MODIFIED IMMUNE CELLS HAVING ENHANCED ANTI-NEOPLASIA ACTIVITY AND IMMUNOSUPPRESSION RESISTANCE

INCOPORATION BY REFERENCE

[0001] This application claims the benefit of U.S. Provisional Application No. 62/793,277 filed on January 16, 2019 and U.S. Provisional Application No. 62/839,870 filed on April 29, 2019.

BACKGROUND OF THE INVENTION

[0002] Autologous and allogeneic immunotherapies are neoplasia treatment approaches in which immune cells expressing chimeric antigen receptors are administered to a subject. To generate an immune cell that expresses a chimeric antigen receptor (CAR), the immune cell is first collected from the subject (autologous) or a donor separate from the subject receiving treatment (allogeneic) and genetically modified to express the chimeric antigen receptor. The resulting cell expresses the chimeric antigen receptor on its cell surface (e.g., CAR T-cell), and upon administration to the subject, the chimeric antigen receptor binds to the marker expressed by the neoplastic cell. This interaction with the neoplasia marker activates the CAR-T cell, which then cell kills the neoplastic cell. But for autologous or allogeneic cell therapy to be effective and efficient, significant conditions and cellular responses, such as T cell signaling inhibition, must be overcome or avoided. For allogeneic cell therapy, graft versus host disease and host rejection of CAR-T cells may provide additional challenges. Editing genes involved in these processes can enhance CAR-T cell function and resistance to immunosuppression or inhibition, but current methodologies for making such edits have the potential to induce large, genomic rearrangements in the CAR-T cell, thereby negatively impacting its efficacy. Thus, there is a significant need for techniques to more precisely modify immune cells, especially CAR-T cells. This application is directed to this and other important needs.

SUMMARY OF THE INVENTION

[0003] As described below, the present invention features genetically modified immune cells having enhanced anti-neoplasia activity, resistance to immune suppression, and decreased risk of eliciting a graft versus host reaction, or host versus graft reaction where host $CD8^+$ T cells recognize a graft as non-self (e.g., where a transplant recipient generates an immune response

against the transplanted organ), or a combination thereof. In one embodiment, a subject having or having a propensity to develop graft versus host disease (GVHD) is administered a CAR-T cell that lacks or has reduced levels of functional TRAC. In one embodiment, a subject having or having a propensity to develop host versus graft disease (HVGD) is administered a CAR-T cell that lacks or has reduced levels of functional beta2 microglobulin (B2M). The present invention also features methods for producing and using these modified immune cells.

[0004] In one aspect, provided herein is a method for producing a modified immune cell with reduced immunogenicity and/or increased anti-neoplasia activity by multiplexed editing, the method comprising: modifying at least four gene sequences or regulatory elements thereof, at a single target nucleobase in each thereof in an immune cell, thereby generating the modified immune cell with reduced immunogenicity and/or increased anti-neoplasia activity.

[0005] In another aspect, provided herein is a method for producing a population of modified immune cells with reduced immunogenicity and/or increased anti-neoplasia activity by multiplexed editing, the method comprising: modifying at least four gene sequences or regulatory elements thereof at a single target nucleobase in each thereof in a population of immune cells, thereby generating the population of modified immune cells with reduced immunogenicity and/or increased anti-neoplasia activity.

[0006] In some embodiments, the at least one of the at least four gene sequences is a checkpoint inhibitor gene sequence, an immune response regulation gene sequence, or an immunogenic gene sequence.

[0007] In some embodiments, the modifying reduces expression of at least one of the at least four gene sequences.

[0008] In some embodiments, the expression of at least one of the at least four genes is reduced by at least 80% as compared to a control cell without the modification.

[0009] In some embodiments, the expression of each one of the at least four genes is reduced by at least 80% as compared to a control cell without the modification.

[0010] In some embodiments, the expression of at least one of the at least four genes is reduced in at least 50% of the population of immune cells.

[0011] In some embodiments, the expression of each one of the at least four genes is reduced in at least 50% of the population of immune cells.

[0012] In some embodiments, the at least four gene sequences comprise a TRAC gene sequence.

[0013] In some embodiments, the at least four gene sequences comprise a check point inhibitor gene sequence.

[0014] In some embodiments, the at least four gene sequences comprise a PDCD1 gene sequence.

[0015] In some embodiments, the at least four gene sequences comprise a T cell marker gene sequence.

[0016] In some embodiments, the at least four gene sequences comprise a CD52 gene sequence.

[0017] In some embodiments, the at least four gene sequences comprises a CD7 gene sequence.

[0018] In some embodiments, the at least four gene sequences comprise a TRAC gene sequence, a PDCD1 gene sequence, a CD52 gene sequence, or a CD7 gene sequence.

[0019] In some embodiments, the at least four sequences comprise a TCR complex gene sequence, a CD7 gene sequence, a CD52 gene sequence ,and a gene sequence selected from the group consisting of CIITA a CD2 gene sequence, a CD4 gene sequence, a CD5 gene sequence, a CD7 gene sequence, a CD30 gene sequence, a CD33 gene sequence, a CD52 gene sequence, a CD70 gene sequence, a B2M gene sequence, and a CIITA gene sequence

[0020] In some embodiments, the at least four gene sequences comprise a gene sequence selected from the group consisting of a CD2 gene sequence, a TRAC gene sequence, a CD3 epsilon gene sequence, a CD3 gamma gene sequence, a CD3 delta gene sequence, a TRBC1 gene sequence, a TRBC2 gene sequence, a CD4 gene sequence, a CD5 gene sequence, a CD7 gene sequence, a CD30 gene sequence, a CD33 gene sequence, a CD52 gene sequence, a CD70 gene sequence, a B2M gene sequence, and a CIITA gene sequence.

[0021] The method of some embodiments described herein comprises modifying five gene sequences or regulatory elements thereof at a single target nucleobase in each thereof in the immune cell.

[0022] The method of some embodiments described herein comprises modifying six gene sequences or regulatory elements thereof at a single target nucleobase in each thereof in the immune cell.

[0023] The method of some embodiments described herein comprises modifying seven gene sequences or regulatory elements thereof at a single target nucleobase in each thereof in the immune cell.

[0024] The method of some embodiments described herein comprises modifying eight gene sequences or regulatory elements thereof at a single target nucleobase in each thereof in the immune cell.

[0025] The method of some embodiments described herein comprises modifying five gene sequences or regulatory elements thereof at a single target nucleobase in each thereof in the population of immune cells.

[0026] The method of some embodiments described herein comprises modifying six gene sequences or regulatory elements thereof at a single target nucleobase in each thereof in the population of immune cells.

[0027] The method of some embodiments described herein comprises modifying seven gene sequences or regulatory elements thereof at a single target nucleobase in each thereof in the population of immune cells.

[0028] The method of some embodiments described herein modifying eight gene sequences or regulatory elements thereof at a single target nucleobase in each thereof in the population of immune cells.

[0029] In some embodiments, the five, six, seven, or eight gene sequences or regulatory elements thereof are selected from the group consisting of a CD2 gene sequence, a TRAC gene sequence, a CD3 epsilon gene sequence, a CD3 gamma gene sequence, a CD3 delta gene sequence, a TRBC1 gene sequence, a TRBC2 gene sequence, a CD4 gene sequence, a CD5 gene sequence, a CD7 gene sequence, a CD30 gene sequence, a CD33 gene sequence, a CD52 gene sequence, a CD70 gene sequence, a B2M gene sequence, and a CIITA gene sequence.

[0030] In some embodiments, the five, six, seven, or eight gene sequences or regulatory elements thereof at comprises a CD3 gene sequence, a CD7 gene sequence, a CD2 gene sequence, a CD5 gene sequence, and a CD52 gene sequence.

[0031] In some embodiments, the modifying comprises deaminating the single target nucleobase.

[0032] In some embodiments, the deaminating is performed by a polypeptide comprising a deaminase.

[0033] In some embodiments, the deaminase is associated with a nucleic acid programmable DNA binding protein (napDNAbp) to form a base editor.

[0034] In some embodiments, the deaminase is fused to the nucleic acid programmable DNA binding protein (napDNAbp).

[0035] In some embodiments, the napDNAbp comprises a Cas9 polypeptide or a portion thereof.

[0036] In some embodiments, the napDNAbp comprises a Cas9 nickase or nuclease dead Cas9.

[0037] In some embodiments, the deaminase is a cytidine deaminase.

[0038] In some embodiments, the single target nucleobase is a cytosine (C) and wherein the modification comprises conversion of the C to a thymine (T).

[0039] In some embodiments, the base editor further comprises a uracil glycosylase inhibitor.

[0040] In some embodiments, the deaminase is an adenosine deaminase.

[0041] In some embodiments, the single target nucleobase is a adenosine (A) and wherein the modification comprises conversion of the A to a guanine (G).

[0042] In some embodiments, the modifying comprises contacting the immune cell with a guide nucleic acid sequences.

[0043] In some embodiments, the modifying comprises contacting the immune cell with at least four guide nucleic acid sequences, wherein each guide nucleic acid sequence targets the napDNAbp to one of the at least four gene sequences or regulatory elements thereof.

[0044] In some embodiments, the guide nucleic acid sequence comprises a sequence selected from guide RNA sequences of table 8A, table 8B, or table 8C.

[0045] In some embodiments, the guide nucleic acid sequence comprises a sequence selected from the group consisting of UUCGUAUCUGUAAAACCAAG, CCUACCUGUCACCAGGACCA, CUCUUACCUGUACCAUAACC, CACCUACCUAAGAACCAUCC, ACUCACGCUGGAUAGCCUCC, ACUCACCCAGCAUCCCCAGC, CACUCACCUUAGCCUGAGCA, and CACGCACCUGGACAGCUGAC.

[0046] In some embodiments, the modifying comprises replacing the single target nucleobase with a different nucleobase by target-primed reverse transcription with a reverse transcriptase and an extended guide nucleic acid sequence.

[0047] In some embodiments, the extended guide nucleic acid sequence comprises a reverse transcription template sequence, a reverse transcription primer binding site, or a combination thereof.

[0048] In some embodiments, the single target nucleobase is in an exon.

[0049] In some embodiments, modifying generates a premature stop codon in the exon.

[0050] In some embodiments, the single target nucleobase is within an exon 1, an exon 2, or an exon 3 of the TRAC gene sequence.

[0051] In some embodiments, the single target nucleobase is within an exon 1, an exon 2, or an exon 5 of the PCDC1 gene sequence.

[0052] In some embodiments, the single target nucleobase is within an exon 1 or an exon 2 of the CD52 gene sequence.

[0053] In some embodiments, the single target nucleobase is within an exon 1, an exon 2, or an exon 3 of the CD7 gene sequence.

[0054] In some embodiments, the single target nucleobase is within an exon 1 or an exon 2 of the B2M gene sequence.

[0055] In some embodiments, the single target nucleobase is within an exon 2, an exon 3, an exon 4, an exon 5, an exon 6, an exon 7, or an exon 8 of the CD5 gene sequence.

[0056] In some embodiments, the single target nucleobase is within an exon 2, an exon 3, an exon 4, or an exon 5 of the CD2 gene sequence.

[0057] In some embodiments, the single target nucleobase is within an exon 1, an exon 2, an exon 4, an exon 7, an exon 8, an exon 9, an exon 10, an exon 11, an exon 12, an exon 14, an exon 15, an exon 18, or an exon 19 of the CIITA gene sequence.

[0058] In some embodiments, the single target nucleobase is in a splice donor site or a splice acceptor site.

[0059] In some embodiments, the single target nucleobase is in an exon 1 splice acceptor site, an exon 1 splice donor site, or an exon 3 splice acceptor site of the TRAC gene sequence.

[0060] In some embodiments, the single target nucleobase is in an exon 1 splice acceptor site, an exon 1 splice donor site, an exon 2 splice acceptor site, an exon 3 splice donor site, an exon 4 splice acceptor site, an exon 4 splice donor site, or an exon 5 splice acceptor site of the PDCD1 gene sequence.

[0061] In some embodiments, the single target nucleobase is in an exon 1 splice donor site, or an exon 2 splice acceptor site of the CD52 gene sequence.

[0062] In some embodiments, the single target nucleobase is in an exon 1 splice donor site, an exon 2 splice donor site, an exon 2 splice acceptor site, or an exon 3 splice acceptor site of the CD7 gene sequence.

[0063] In some embodiments, the single target nucleobase is in an exon 1 splice donor site, an exon 2 splice donor site, an exon 2 splice acceptor site, or an exon 3 splice acceptor site of the B2M gene sequence.

[0064] In some embodiments, the single target nucleobase is in an exon 3 splice donor site of the CD2 gene sequence.

[0065] In some embodiments, the single target nucleobase is in an exon 1 splice donor site, an exon 1 splice acceptor site, an exon 3 splice acceptor site, an exon 3 splice donor site, an exon 4 splice acceptor site, an exon 5 splice donor site, an exon 6 splice acceptor site, an exon 9 splice donor site, an exon 10 splice acceptor site of the CD5 gene sequence.

[0066] In some embodiments, the single target nucleobase is in an exon 1 splice donor site, an exon 7 splice donor site, an exon 8 splice acceptor site, an exon 9 slice donor site, an exon 10 splice acceptor site, an exon 11 splice acceptor site, an exon 14 splice acceptor site, an exon 14 splice donor site, an exon 15 splice donor site, an exon 16 splice acceptor site, an exon 16 splice donor site, an exon 17 splice acceptor site, an exon 17 splice donor site, or an exon 19 splice acceptor site of the CIITACIITA gene sequence.

[0067] In some embodiments, the immune cell is a human cell. In some embodiments, the immune cell is a cytotoxic T cell, a regulatory T cell, a T helper cell, a dendritic cell, a B cell, or a NK cell.

[0068] In some embodiments, the population of immune cells are human cells.

[0069] In some embodiments, the population of immune cells are cytotoxic T cells, regulatory T cells, T helper cells, dendritic cells, B cells, or NK cells.

[0070] In some embodiments, the modifying is ex vivo.

[0071] In some embodiments, the immune cell or the population of immune cells are derived from a single human donor.

[0072] In some embodiments, the method further comprising contacting the immune cell or the population of immune cells with a polynucleotide that encodes an exogenous functional chimeric antigen receptor (CAR) or a functional fragment thereof.

[0073] In some embodiments, contacting the immune cell or the population of immune cells with a lentivirus comprising the polynucleotide that encodes the CAR.

[0074] In some embodiments, contacting the immune cell or the population of immune cells with a napDNAbp and a donor DNA sequence comprising the polynucleotide that encodes the CAR.

[0075] In some embodiments, the napDNAbp is a Cas12b.

[0076] In some embodiments, the CAR specifically binds a marker associated with neoplasia.

[0077] In some embodiments, the neoplasia is a T cell cancer, a B cell cancer, a lymphoma, a leukemia, or a multiple myeloma.

[0078] In some embodiments the CAR specifically binds CD7.

[0079] In some embodiments, the CAR specifically binds BCMA.

[0080] In some embodiments, the immune cell or the population of immune cells comprises no detectable translocation. In some embodiments, at least 50% of the population of immune cells express the CAR. In some embodiments, at least 50% of the population of immune cells are viable. In some embodiments, at least 50% of the population of immune cells expand at least 80% of expansion rate of a population of control cells of a same type without the modification.

[0081] In the method of some embodiments described herein, the modifying generates less than 1% of indels in the immune cell. In some embodiments, the modifying generates less than 5% of non-target edits in the immune cell. In some embodiments, the modifying generates less than 5% of off-target edits in the immune cell.

[0082] In one aspect, provided herein is a modified immune cell produced according to some embodiments described in the preceding paragraphs.

[0083] In one aspect, provided herein is a population of modified immune cells produced according to some embodiments described in the preceding paragraphs.

[0084] In another aspect, provided herein is a modified immune cell with reduced immunogenicity or increased anti-neoplasia activity, wherein the modified immune cell comprises a single target nucleobase modification in each one of at least four gene sequences or regulatory elements thereof. In some embodiments, in the modified immune cell described

above, each one of the at least four gene sequences is a checkpoint inhibitor gene sequence, an immune response regulation gene sequence, or an immunogenic gene sequence.

[0085] In the modified immune cell of the preceding embodiments the at least four gene sequences comprise a TCR complex gene sequence.

[0086] In some embodiments, the at least four gene sequences comprise a TRAC gene sequence. In some embodiments, the at least four gene sequences comprise a check point inhibitor gene sequence. In some embodiments, the at least four gene sequences comprise a PDCD1 gene sequence.

[0087] In some embodiments, the at least four gene sequences comprise a T cell marker gene sequence.

[0088] In some embodiments, the at least four gene sequences comprise CD52 gene sequence.

[0089] In some embodiments, the at least four gene sequences comprises a CD7 gene sequence.

[0090] In some embodiments, the expression of one of the at least four genes is reduced by at least 80% as compared to a control cell without the modification.

[0091] In some embodiments, the expression of each one of the at least four genes is reduced by at least 90% as compared to a control cell without the modification.

[0092] In some embodiments, the immune cell comprises a modification at a single target nucleobase in each one of five gene sequences or regulatory elements thereof, wherein each one of the five gene sequences is a checkpoint inhibitor gene sequence, an immune response regulation gene sequence, or an immunogenic gene sequence.

[0093] In some embodiments, the immune cell comprises a modification at a single target nucleobase in each one of six gene sequences or regulatory elements thereof, wherein each one of the six gene sequences is a checkpoint inhibitor gene sequence, an immune response regulation gene sequence, or an immunogenic gene sequence.

[0094] In some embodiments, the immune cell comprises a modification at a single target nucleobase in each one of seven gene sequences or regulatory elements thereof, wherein each one of the seven gene sequences is a checkpoint inhibitor gene sequence, an immune response regulation gene sequence or an immunogenic gene sequence.

[0095] In some embodiments, the immune cell comprises a modification at a single target nucleobase in each one of eight gene sequences or regulatory elements thereof, wherein each one

of the eight gene sequences is a checkpoint inhibitor gene sequence, an immune response regulation gene sequence, or an immunogenic gene sequence.

[0096] In some embodiments, the expression of at least one of the five, six, seven or eight genes is reduced by at least 90% as compared to a control cell without the modification.

[0097] In some embodiments, the expression of each one of the five, six, seven, or eight genes is reduced by at least 90% as compared to a control cell without the modification.

[0098] In some embodiments, the five, six, seven, or eight gene sequences or regulatory elements thereof comprise a sequence selected from the group consisting of a CD2 gene sequence, a TRAC gene sequence, a CD3 epsilon gene sequence, a CD3 gamma gene sequence, a CD3 delta gene sequence, a TRBC1 gene sequence, a TRBC2 gene sequence, a CD4 gene sequence, a CD5 gene sequence, a CD7 gene sequence, a CD30 gene sequence, a CD33 gene sequence, a CD52 gene sequence, a CD70 gene sequence, a B2M gene sequence, and a CIITA gene sequence.

[0099] In one aspect, provided herein is a modified immune cell comprising a single target nucleobase modification in each one of a CD3 gene sequence, a CD5 gene sequence, a CD52 gene sequence, and a CD7 gene sequence, wherein the modified immune cell exhibits reduced immunogenicity or increased anti-neoplasia activity as compared to a control cell of a same type without the modification.

[00100] In some embodiments, the modified immune cell further comprises a single target nucleobase modification in a CD2 gene sequence, CIITA or a regulatory element of each thereof.

[00101] In some embodiments, the modified immune cell comprises a single target nucleobase modification in a TRAC gene sequence, a CD3 epsilon gene sequence, a CD3 gamma gene sequence, a CD3 delta gene sequence, a TRBC1 gene sequence, or a TRBC2 gene sequence further comprises a single target nucleobase modification in a gene sequence a CD4 gene sequence, a CD30 gene sequence, a CD33 gene sequence, a CD70 gene sequence, a B2M gene sequence, and a CIITA gene sequence or a regulatory element of each thereof.

[00102] In some embodiments, the modified immune cell comprises a single nucleobase modification in each one of a TRAC gene sequence, a PDCD1 gene sequence, a CD52 gene sequence, a CD7 gene sequence, a CD2 gene sequence, a CD5 gene sequence, a CIITA gene sequence, and a B2M gene sequence.

[00103] In some embodiments, the modified immune cell comprises no detectable translocation.

[00104] In some embodiments, the modified immune cell comprises less than 1% of indels.

[00105] In some embodiments, the modified immune cell comprises less than 5% of non-target edits.

[00106] In some embodiments, the modified immune cell comprises less than 5% of off-target edits.

[00107] In some embodiments, the modified immune has increased growth or viability compared to a reference cell. In some embodiments, the reference cell is an immune cell modified with a Cas9 nuclease.

[00108] In some embodiments, the modified immune cell is a mammalian cell.

[00109] In some embodiments, the modified immune cell is a human cell.

[00110] In some embodiments, the modified immune cell is a cytotoxic T cell, a regulatory T cell, a T helper cell, a dendritic cell, a B cell, or a NK cell.

[00111] In some embodiments, the modified the immune cell is in an ex vivo culture.

[00112] In some embodiments, the modified the immune cell is derived from a single human donor.

[00113] In some embodiments, the modified the immune cell further comprises a polynucleotide that encodes an exogenous functional chimeric antigen receptor (CAR) or a functional fragment thereof.

[00114] In some embodiments, the polynucleotide that encodes the CAR is integrated in the genome of the immune cell.

[00115] In some embodiments, the CAR specifically binds a marker associated with neoplasia.

[00116] In some embodiments, the neoplasia is a T cell cancer, a B cell cancer, a lymphoma, a leukemia, or a multiple myeloma.

[00117] In some embodiments, the CAR specifically binds CD7.

[00118] In some embodiments, the CAR specifically binds BCMA.

[00119] In some embodiments, the single target nucleobase is in an exon.

[00120] In some embodiments, the single target nucleobase is within an exon 1, an exon 2, or an exon 3 of the TRAC gene sequence.

[00121] In some embodiments, the single target nucleobase is within an exon 1, an exon 2, or an exon 5 of the PCDC1 gene sequence.

[00122] In some embodiments, the single target nucleobase is within an exon 1 or an exon 2 of the CD52 gene sequence.

[00123] In some embodiments, the single target nucleobase is within an exon 1, an exon 2, or an exon 3 of a CD7 gene sequence.

[00124] In some embodiments, the single target nucleobase is in a splice donor site or a splice acceptor site.

[00125] In some embodiments, the single target nucleobase is in an exon 1 splice acceptor site, an exon 1 splice donor site, or an exon 3 splice acceptor site of the TRAC gene sequence.

[00126] In some embodiments, the single target nucleobase is in an exon 1 splice acceptor site, an exon 1 splice donor site, an exon 2 splice acceptor site, an exon 3 splice donor site, an exon 4 splice acceptor site, an exon 4 splice donor site, or an exon 5 splice acceptor site of the PDCD1 gene sequence.

[00127] In some embodiments, the single target nucleobase is in an exon 1 splice donor site, or an exon 2 splice acceptor site of the CD52 gene sequence.

[00128] In some embodiments, the single target nucleobase is in an exon 1 splice donor site, an exon 2 splice donor site, an exon 2 splice acceptor site, or an exon 3 splice acceptor site of the CD7 gene sequence.

[00129] In one aspect, provided herein is a population of modified immune cells, wherein a plurality of the population of cells comprise a single target nucleobase modification in each one of at least four gene sequences or regulatory elements thereof, and wherein the plurality of the population of cells having the modification exhibit reduced immunogenicity or increased anti-neoplasia activity as compared to a plurality of control cells of a same type without the modification.

[00130] In some embodiments, the plurality of cells comprises at least 50% of the population.

[00131] In some embodiments, each one of the at least four gene sequences is a checkpoint inhibitor gene sequence, an immune response regulation gene sequence, or an immunogenic gene sequence.

[00132] In some embodiments, the at least four gene sequences comprise a TCR component gene sequence, a check point inhibitor gene sequence, or a T cell marker gene sequence.

[00133] In some embodiments, the at least four gene sequences comprise a TRAC gene sequence.

[00134] In some embodiments, the at least four gene sequences comprise a PDCD1 gene sequence.

[00135] In some embodiments, the at least four gene sequences comprise CD52 gene sequence.

[00136] In some embodiments, the at least four gene sequences comprises a CD7 gene sequence.

[00137] In the population of some embodiments, expression of at least one of the at least four genes is reduced by at least 80% in the plurality of cells having the modification as compared to a control cell without the modification

[00138] In the population of some embodiments, expression of each one of the at least four genes is reduced by at least 80% in the plurality of cells having the modification as compared to a control cell without the modification.

[00139] In some embodiments, the plurality of the population comprises a modification at a single target nucleobase in each one of five gene sequences or regulatory elements thereof, wherein each one of the five gene sequences is a checkpoint inhibitor gene sequence, an immune response regulation gene sequence, or an immunogenic gene sequence.

[00140] In some embodiments, the plurality of the population comprises a modification at a single target nucleobase in each one of six gene sequences or regulatory elements thereof, wherein each one of the six sequences is a checkpoint inhibitor gene sequence, an immune response regulation gene sequence, or an immunogenic gene sequence

[00141] In some embodiments, the plurality of the population comprises a modification at a single target nucleobase in each one of seven gene sequences or regulatory elements thereof, wherein each one of the seven gene sequences is a checkpoint inhibitor gene sequence, an immune response regulation gene sequence, or an immunogenic gene sequence.

[00142] In some embodiments, the plurality of the population comprises a modification at a single target nucleobase in each one of eight gene sequences or regulatory elements thereof, wherein each one of the eight gene sequences is a checkpoint inhibitor gene sequence, an immune response regulation gene sequence, or an immunogenic gene sequence.

[00143] In the population of some embodiments, the expression of at least one of the five, six, seven, or eight genes is reduced by at least 90% in the plurality of cells having the modification as compared to a control cell without the modification.

[00144] In the population of some embodiments, the expression of each one of the five, six, seven, or eight genes is reduced by at least 90% in the plurality of cells having the modification as compared to a control cell without the modification.

[00145] In the population of some embodiments, the expression of at least one of the five, six, seven, or eight genes is reduced by at least 90% in the plurality of cells having the modification as compared to a control cell without the modification.

[00146] In some embodiments, the expression of each one of the five, six, seven, or eight genes is reduced by at least 90% in the plurality of cells having the modification as compared to a control cell without the modification.

[00147] In some embodiments, the five, six, seven, or eight gene sequences or regulatory elements thereof are selected from the group consisting of a CD2 gene sequence, a TRAC gene sequence, a CD3 epsilon gene sequence, a CD3 gamma gene sequence, a CD3 delta gene sequence, a TRBC1 gene sequence, a TRBC2 gene sequence, a CD4 gene sequence, a CD5 gene sequence, a CD7 gene sequence, a CD30 gene sequence, a CD33 gene sequence, a CD52 gene sequence, a CD70 gene sequence, a B2M gene sequence, and a CIITA gene sequence.

[00148] In one aspect, provided herein is a population of modified immune cells, wherein a plurality of the population comprise a single target nucleobase modification in each one of a TRAC gene sequence, a PDCD1 gene sequence, a CD52 gene sequence, and a CD7 gene sequence, and wherein the plurality of the population having the modification exhibit reduced immunogenicity or increased anti-neoplasia activity as compared to a plurality of control cells of a same type without the modification.

[00149] In some embodiments, the plurality of the population further comprises a single target nucleobase modification in a CD2 gene sequence, a CD5 gene sequence, a CIITA gene sequence, a B2M gene sequence, or a regulatory element of each thereof. In some embodiments, the plurality of the population further comprises a single target nucleobase modification in a gene sequence of a gene selected from the group consisting of a CD2 gene sequence, a TRAC gene sequence, a CD3 epsilon gene sequence, a CD3 gamma gene sequence, a CD3 delta gene sequence, a TRBC1 gene sequence, a TRBC2 gene sequence, a CD4 gene sequence, a CD5 gene

sequence, a CD7 gene sequence, a CD30 gene sequence, a CD33 gene sequence, a CD52 gene sequence, a CD70 gene sequence, a B2M gene sequence, and a CIITA gene sequence or a regulatory element of each thereof. In some embodiments, the plurality of the population comprises a single nucleobase modification in each one of a TRAC gene sequence, a PDCD1 gene sequence, a CD52 gene sequence, a CD7 gene sequence, a CD2 gene sequence, a CD5 gene sequence, a CIITA gene sequence, and a B2M gene sequence.

[00150] In the population of modified immune cells of some embodiments, the plurality of the population comprises no detectable translocation.

[00151] In the population of modified immune cells of some embodiments, the at least 60% of the population of immune cells are viable. In the population of modified immune cells of some embodiments, the at least 60% of the population of immune cells expand at least 80% of expansion rate of a population of control cells of a same type without the modification. In the population of modified immune cells of some embodiments, the population of immune cells are human cells. In the population of modified immune cells of some embodiments, the population of immune cells are cytotoxic T cells, regulatory T cells, T helper cells, dendritic cells, B cells, or NK cells. In the population of modified immune cells of some embodiments, the population of immune cells are derived from a single human donor. In the population of modified immune cells of some embodiments, the plurality of cells having the modification further comprises a polynucleotide that encodes an exogenous functional chimeric antigen receptor (CAR) or a functional fragment thereof.

[00152] In some embodiments, the at least 50% of the population of immune cells express the CAR.

[00153] In some embodiments, the the CAR specifically binds a marker associated with neoplasia.

[00154] In some embodiments, the neoplasia is a T cell cancer, a B cell cancer, a lymphoma, a leukemia, or a multiple myeloma.

[00155] In some embodiments, the CAR specifically binds CD7.

[00156] In some embodiments, the CAR specifically binds BCMA.

[00157] In some embodiments, the single target nucleobase is in an exon.

[00158] In some embodiments, the single target nucleobase is within an exon 1, an exon 2, or an exon 3 of the TRAC gene sequence.

[00159] In some embodiments, the single target nucleobase is within an exon 1, an exon 2, or an exon 5 of the PCDC1 gene sequence.

[00160] In some embodiments, the single target nucleobase is within an exon 1 or an exon 2 of the CD52 gene sequence.

[00161] In some embodiments, the single target nucleobase is within an exon 1, an exon 2, or an exon 3 of a CD7 gene sequence.

[00162] In the population of modified immune cells of some embodiments, the single target nucleobase is in a splice donor site or a splice acceptor site.

[00163] In some embodiments, the single target nucleobase is in an exon 1 splice acceptor site, an exon 1 splice donor site, or an exon 3 splice acceptor site of the TRAC gene sequence.

[00164] In some embodiments, the single target nucleobase is in an exon 1 splice acceptor site, an exon 1 splice donor site, an exon 2 splice acceptor site, an exon 3 splice donor site, an exon 4 splice acceptor site, an exon 4 splice donor site, or an exon 5 splice acceptor site of the PDCD1 gene sequence.

[00165] In some embodiments, the single target nucleobase is in an exon 1 splice donor site, or an exon 2 splice acceptor site of the CD52 gene sequence.

[00166] In some embodiments, the single target nucleobase is in an exon 1 splice donor site, an exon 2 splice donor site, an exon 2 splice acceptor site, or an exon 3 splice acceptor site of the CD7 gene sequence.

[00167] In one aspect, provided herein is a composition comprising deaminase and a nucleic acid sequence, wherein the guide nucleic acid sequence comprises a sequence selected from the group consisting of UUCGUAUCUGUAAAACCAAG, CCUACCUGUCACCAGGACCA, CUCUUACCUGUACCAUAACC, CACCUACCUAAGAACCAUCC, ACUCACGCUGGAUAGCCUCC, ACUCACCCAGCAUCCCCAGC, CACUCACCUUAGCCUGAGCA, and CACGCACCUGGACAGCUGAC.

[00168] In some embodiments, the deaminase is associated with a nucleic acid programmable DNA binding protein (napDNAbp) to form a base editor.

[00169] In some embodiments, the napDNAbp comprises a Cas9 nickase or nuclease dead Cas9 and wherein the deaminase is a cytidine deaminase.

[00170] In some embodiments, the base editor further comprises a uracil glycosylase inhibitor.

[00171] In some embodiments, the napDNAbp comprises a Cas9 nickase or nuclease dead Cas9 and wherein the deaminase is a adenosine deaminase.

[00172] In one aspect, provided herein is a composition comprising a polymerase and a guide nucleic acid sequence, wherein the guide nucleic acid sequence comprises a sequence selected from the group consisting of the group consisting of UUCGUAUCUGUAAAACCAAG, CCUACCUGUCACCAGGACCA, CUCUUACCUGUACCAUAACC, CACCUACCUAAGAACCAUCC, ACUCACGCUGGAUAGCCUCC, ACUCACCCAGCAUCCCCAGC, CACUCACCUUAGCCUGAGCA, and CACGCACCUGGACAGCUGAC.

[00173] In some embodiments, the polymerase is a reverse transcriptase and wherein the guide nucleic acid sequence is an extended guide nucleic acid sequence comprising a reverse transcription template sequence, a reverse transcription primer binding site, or a combination thereof.

[00174] In one aspect, provided herein is a method for producing a modified immune cell with reduced immunogenicity and/or increased anti-neoplasia activity, the method comprising: a) modifying a single target nucleobase in a first gene sequence or a regulatory element thereof in an immune cell; and b) modifying a second gene sequence or a regulatory element thereof in the immune cell with a Cas12 polypeptide, wherein the Cas12 polypeptide generates a site-specific cleavage in the second gene sequence; wherein each of the first gene and the second gene is a immunogenic gene, a checkpoint inhibitor gene, or an immune response regulation gene, thereby generating a modified immune cell with reduced immunogenicity and/or increased anti-neoplasia activity.

[00175] In some embodiments, the method further comprises expressing an exogenous functional chimeric antigen receptor (CAR) or a functional fragment thereof in the immune cell.

[00176] In some embodiments, a polynucleotide encoding the CAR or the functional fragment thereof is inserted into the site specific cleavage generated by the Cas12 polypeptide.

[00177] In some embodiments, the Cas12 polypeptide is a Cas12b polypeptide.

[00178] In one aspect, provided herein is a method for producing a modified immune cell with reduced immunogenicity and/or increased anti-neoplasia activity, the method comprising:

[00179] a) modifying a single target nucleobase in a first gene sequence or a regulatory element thereof in an immune cell; and b) modifying a second gene sequence or a regulatory

element thereof in the immune cell by inserting an exogenous functional chimeric antigen receptor (CAR) or a functional fragment thereof or an exogenous functional T cell receptor or a functional fragment thereof in the second gene; wherein each of the first gene and the second gene is a immunogenic gene, a checkpoint inhibitor gene, or an immune response regulation gene, thereby generating a modified immune cell with reduced immunogenicity and/or increased anti-neoplasia activity.

[00180] In some embodiments, the step b) further comprises generating a site-specific cleavage in the second gene sequence with a nucleic acid programmable DNA binding protein (napDNAbp).

[00181] In some embodiments, the napDNAbp is a Cas12b.

[00182] In some embodiments, the expression of the first gene is reduced by at least 60% or wherein expression of the second gene is reduced by at least 60% as compared to a control cell of a same type without the modification.

[00183] In some embodiments, the first gene is selected from the group consisting of CD3 epsilon, CD3 gamma, CD3 delta, CD4, TRAC, TRBC1, TRBC2, PDCD1, CD30, CD33, CD7, CD52, B2M, CD70, CIITA, CD2, and CD5.

[00184] In some embodiments, the first gene or the second gene is selected from the group consisting of TRAC, CIITA, CD2, CD5, CD7, and CD52.

[00185] In some embodiments, the second gene is TRAC.

[00186] In some embodiments, the step a) further comprises modifying a single target nucleobase in two other gene sequences or regulatory elements thereof.

[00187] In some embodiments, the step a) further comprises modifying a single target nucleobase in three other gene sequences or regulatory elements thereof.

[00188] In some embodiments, the step a) further comprises modifying a single target nucleobase in four other gene sequences or regulatory elements thereof.

[00189] In some embodiments, the step a) further comprises modifying a single target nucleobase in five other gene sequences or regulatory elements thereof.

[00190] In some embodiments, the step a) further comprises modifying a single target nucleobase in six other gene sequences or regulatory elements thereof.

[00191] In some embodiments, the step a) further comprises modifying a single target nucleobase in seven other gene sequences or regulatory elements thereof.

[00192] In some embodiments, the modifying in step a) comprises deaminating the single target nucleobase with a base editor comprising a deaminase and a nucleic acid programmable DNA binding protein (napDNAbp).

[00193] In some embodiments, the napDNAbp comprises a Cas9 nickase or nuclease dead Cas9.

[00194] In some embodiments, the deaminase is a cytidine deaminase and wherein the modification comprises conversion of a cytidine (C) to a thymine (T).

[00195] In some embodiments, the deaminase is an adenosine deaminase and wherein the modification comprises conversion of an adenine (A) to a guanine (G).

[00196] In some embodiments, the modifying in a) comprises contacting the immune cell with a guide nucleic acid sequence.

[00197] In some embodiments, the guide nucleic acid sequence comprises a sequence selected from the group consisting of UUCGUAUCUGUAAAACCAAG, CCUACCUGUCACCAGGACCA, CUCUUACCUGUACCAUAACC, CACCUACCUAAGAACCAUCC, ACUCACGCUGGAUAGCCUCC, ACUCACCCAGCAUCCCCAGC, CACUCACCUUAGCCUGAGCA, and CACGCACCUGGACAGCUGAC.

[00198] In some embodiments, the modifying in b) comprises contacting the immune cell with a guide nucleic acid sequence.

[00199] In some embodiments, the guide nucleic acid sequence comprises a sequence selected from sequences in Table 1.

[00200] In some embodiments, the modifying in a) comprises replacing the single target nucleobase with a different nucleobase by target-primed reverse transcription with a reverse transcriptase and an extended guide nucleic acid sequence, wherein the extended guide nucleic acid sequence comprises a reverse transcription template sequence, a reverse transcription primer binding site, or a combination thereof.

[00201] In some embodiments, wherein the modifying in a) and b) generates less than 1% indels in the immune cell.

[00202] In some embodiments, the modifying in a) and b) generates less than 5% off target modification in the immune cell.

[00203] In some embodiments, the modifying in a) and b) generate less than 5% non-target modification in the immune cell.

[00204] In some embodiments, the immune cell is a human cell.

[00205] In some embodiments, the immune cell is a cytotoxic T cell, a regulatory T cell, a T helper cell, a dendritic cell, a B cell, or a NK cell.

[00206] In some embodiments, the CAR specifically binds a marker associated with neoplasia.

[00207] In some embodiments, the CAR specifically binds CD7.

[00208] In one aspect, provided herein is a modified immune cell with reduced immunogenicity and/or increased anti-neoplasia activity, wherein the modified immune cell comprises:

[00209] a) a single target nucleobase modification in a first gene sequence or a regulatory element thereof; and b) a modification in a second gene sequence or a regulatory element thereof, wherein the modification is a Cas12 polypeptide generated site-specific cleavage; wherein each of the first gene and the second gene is a immunogenic gene, a checkpoint inhibitor gene, or an immune response regulation gene. In one embodiment, the immune cell further comprises an exogenous functional chimeric antigen receptor (CAR) or a functional fragment thereof.

[00210] In some embodiments, a polynucleotide encoding the CAR or the functional fragment thereof is inserted into the site specific cleavage generated by the Cas12 polypeptide.

[00211] In one aspect, provided herein is a modified immune cell with reduced immunogenicity and/or increased anti-neoplasia activity, the modified immune cell comprising: a) a single target nucleobase modification in a first gene sequence or a regulatory element thereof in an immune cell; and b) a modification in a second gene sequence or a regulatory element thereof, wherein the modification is an insertion of an exogenous chimeric antigen receptor (CAR) or a functional fragment thereof or an exogenous T cell receptor or a functional fragment thereof; wherein each of the first gene and the second gene is a immunogenic gene, a checkpoint inhibitor gene, or immune response regulation gene.

[00212] In some embodiments, the modification in b) is generated by a site-specific cleavage with a Cas12b.

[00213] In some embodiments, expression of the first gene is reduced by at least 60% or wherein expression of the second gene is reduced by at least 60% as compared to a control cell of a same type without the modification.

[00214] In some embodiments, the first gene or the second gene is selected from the group consisting of CD3 epsilon, CD3 gamma, CD3 delta, CD4, TRAC, TRBC1, TRBC2, PDCD1, CD30, CD33, CD7, CD52, B2M, CD70, CIITA, CD2, and CD5.

[00215] In some embodiments, the first gene or the second gene is selected from the group consisting of TRAC, CD2, CD5, CD7, and CD52.

[00216] In some embodiments, the second gene is TRAC.

[00217] In some embodiments, the immune cell further comprises modification in a single target nucleobase in two other gene sequences or regulatory elements thereof.

[00218] In some embodiments, the immune cell further comprises modification in a single target nucleobase in three other gene sequences or regulatory elements thereof.

[00219] In some embodiments, the immune cell further comprises modification in a single target nucleobase in four other gene sequences or regulatory elements thereof.

[00220] In some embodiments, the immune cell further comprises modification in a single target nucleobase in five other gene sequences or regulatory elements thereof.

[00221] In some embodiments, the immune cell further comprises modification in a single target nucleobase in six other gene sequences or regulatory elements thereof.

[00222] In some embodiments, the immune cell further comprises modification in a single target nucleobase in seven other gene sequences or regulatory elements thereof.

[00223] In some embodiments, the modification in a) is generated by a base editor comprising a deaminase and a nucleic acid programmable DNA binding protein (napDNAbp).

[00224] In some embodiments, the deaminase is a cytidine deaminase and the modification comprises conversion of a cytidine (C) to a thymine (T).

[00225] In some embodiments, the deaminase is an adenosine deaminase and wherein the modification comprises conversion of an adenine (A) to a guanine (G).

[00226] In some embodiments, the immune cell comprises less than 1% indels in the genome.

[00227] In some embodiments, the immune cell is a human cell.

[00228] In some embodiments, the immune cell is a cytotoxic T cell, a regulatory T cell, a T helper cell, a dendritic cell, a B cell, or a NK cell.

[00229] In some embodiments, the CAR specifically binds a marker associated with neoplasia.

[00230] In some embodiments, the CAR specifically binds CD7.

[00231] In some embodiments, the modification in b) is an insertion in exon 1 in the TRAC gene sequence.

[00232] In one aspect, provided herein is a population of modified immune cells, wherein a plurality of the population of immune cells comprises: a) a single target nucleobase modification in a first gene sequence or a regulatory element thereof in an immune cell; and b) a modification in a second gene sequence or a regulatory element thereof, wherein the modification is a Cas12 polypeptide generated site-specific cleavage; wherein each of the first gene and the second gene is a immunogenic gene, a checkpoint inhibitor gene, or an immune response regulation gene, and wherein the plurality of the population comprises an exogenous chimeric antigen receptor (CAR) or a functional fragment thereof.

[00233] In some embodiments, a polynucleotide encoding the CAR or the functional fragment thereof is inserted into the site specific cleavage generated by the Cas12 polypeptide.

[00234] In one aspect, provided herein is a population of modified immune cells, wherein a plurality of the population of immune cells comprises: a) a single target nucleobase modification in a first gene sequence or a regulatory element thereof; and b) a modification in a second gene sequence or a regulatory sequence thereof, wherein the modification is an insertion of an exogenous chimeric antigen receptor (CAR) or a functional fragment thereof or an exogenous T cell receptor or a functional fragment thereof; wherein each of the first gene and the second gene is a immunogenic gene, a checkpoint inhibitor gene, or immune response regulation gene, and wherein the plurality of cells with the modification in a) or b) exhibit reduced immunogenicity and/or increased anti-neoplasia activity. In some embodiments, the modification in b) is generated by a site-specific cleavage with a Cas12b. In some embodiments, expression of the first gene is reduced by at least 60% or wherein expression of the second gene is reduced by at least 60% in the plurality of cells with the modification in a) or b) as compared to plurality of control cells of a same type without the modification.

[00235] In some embodiments, the first gene or the second gene is selected from the group consisting of CD3 epsilon, CD3 gamma, CD3 delta, CD4, TRAC, TRBC1, TRBC2, PDCD1, CD30, CD33, CD7, CD52, B2M, CD70, CIITA, CD2, and CD5.

[00236] In some embodiments, the first gene or the second gene is selected from the group consisting of TRAC, CIITA, CD2, CD5, , CD7, and CD52.

[00237] In some embodiments, the first gene is TRAC, CD7, or CD52.

[00238] In some embodiments, the second gene is TRAC.

[00239] In some embodiments, the plurality of cells with the modification in a) or b) further comprises a modification in a single target nucleobase in two other gene sequences or regulatory elements thereof.

[00240] In some embodiments, the plurality of cells with the modification in a) or b) further comprises a single target nucleobase in three, four, five, or six other gene sequences or regulatory elements thereof.

[00241] In some embodiments, the modification in a) is generated by a base editor comprising a deaminase and a nucleic acid programmable DNA binding protein (napDNAbp) to form a base editor.

[00242] In some embodiments, the deaminase is a cytidine deaminase and wherein the modification comprises conversion of a cytidine (C) to a thymine (T).

[00243] In some embodiments, the deaminase is an adenosine deaminase and wherein the modification comprises conversion of an adenine (A) to a guanine (G).

[00244] In some embodiments, the base editor further comprises a uracil glycosylase inhibitor.

[00245] In some embodiments, at least 60% of the population of immune cells are viable.

[00246] In some embodiments, at least 60% of the population of immune cells expand at least 80% of expansion rate of a population of control cells of a same type without the modification.

[00247] In some embodiments, the population of modified immune cells have increased yield of modified immune cells compared to a reference population of cells. In some embodiments, the reference population is a population of immune cells modified with a Cas9 nuclease.

[00248] In some embodiments, the immune cells are a human cells.

[00249] In some embodiments, the immune cells is are cytotoxic T cells, regulatory T cells, T helper cells, dendritic cells, B cells, or NK cells.

[00250] In some embodiments, the CAR specifically binds a marker associated with neoplasia.

[00251] In some embodiments, the CAR specifically binds CD7.

[00252] In some embodiments, the modification in b) is an insertion in exon 1 in the TRAC gene sequence.

[00253] In one aspect, provided herein is a method for producing a modified immune cell with increased anti-neoplasia activity, the method comprising: modifying a single target nucleobase in a Cbl Proto Oncogene B (CBLB) gene sequence or a regulatory element thereof in an immune cell, wherein the modification reduces an activation threshold of the immune cell compared with an immune cell lacking the modification; thereby generating a modified immune cell with increased anti-neoplasia activity.

[00254] In one aspect, provided herein is a composition comprising a modified immune cell with increased anti-neoplasia activity, wherein the modified immune cell comprises: a modification in a single target nucleobase in a Cbl Proto-Oncogene B (CBLB) gene sequence or a regulatory element thereof, wherein the modified immune cell exhibits a reduced activation threshold compared with a control immune cell of a same type without the modification.

[00255] In one aspect, provided herein is a population of immune cells, wherein a plurality of the population of immune cells comprises: a modification in a single target nucleobase in a CBLB gene sequence or a regulatory element thereof, wherein the plurality of the population of the immune cells comprising the modification exhibit a reduced activation threshold compared with an control population of immune cells of a same type without the modification.

[00256] In one aspect, provided herein is a method for producing a population of modified immune cells with increased anti-neoplasia activity, the method comprising: modifying a single target nucleobase in a Cbl Proto Oncogene B (CBLB) gene sequence or a regulatory element thereof in a population of immune cells, wherein at least 50% of the population of immune cells are modified to comprise the single target nucleobase modification.

[00257] In one aspect, provided herein is a composition comprising at least four different guide nucleic acid sequences for base editing. In some embodiments, the composition further comprising a polynucleotide encoding a base editor polypeptide, wherein the base editor polypeptide comprises a nucleic acid programmable DNA binding protein (napDNAbp) and a deaminase. In some embodiments, the polynucleotide encoding the base editor is a mRNA sequence.

[00258] In some embodiments, the deaminase is a cytidine deaminase or an adenosine deaminase.

[00259] In some embodiments, the composition further comprises a base editor polypeptide, wherein the base editor polypeptide comprises a nucleic acid programmable DNA binding protein (napDNAbp) and a deaminase.

[00260] In some embodiments, the deaminase is a cytidine deaminase or an adenosine deaminase.

[00261] In some embodiments, the composition further comprises a lipid nanoparticle.

[00262] In some embodiments, the at least four guide nucleic acid sequences each hybridize with a gene sequence selected from the group consisting of CD2, CD3 epsilon, CD3 gamma, CD3 delta, CD4, CD5, CD7, CD30, CD33, CD52, CD70, and CIITA. In some embodiments, the at least 1, 2, 3, 4, 5, 6, 7, 8, or more genes or regulatory elements thereof are selected from CD2, CD3 epsilon, CD3 gamma, CD3 delta, CD4, CD5, CD7, CD30, CD33, CD52, CD70, and CIITA.

[00263] In some embodiments, the at least 1, 2, 3, 4, 5, 6, 7, 8, or more genes or regulatory elements thereof comprise one or more genes selected from CD2, CD3 epsilon, CD3 gamma, CD3 delta, CD4, CD5, CD7, CD30, CD33, CD52, CD70, and CIITA. In some embodiments, the at least 1, 2, 3, 4, 5, 6, 7, 8, or more genes or regulatory elements thereof are selected from ACAT1, ACLY, ADORA2A, AXL, B2M , BATF, BCL2L11, BTLA, CAMK2D, cAMP, CASP8, Cblb, CCR5, CD2, CD3D, CD3E, CD3G, CD4, CD5, CD7, CD8A, CD33, CD38, CD52, CD70, CD82, CD86, CD96, CD123, CD160, CD244, CD276, CDK8, CDKN1B, Chi3l1, CIITA, CISH, CSF2CSK, CTLA-4, CUL3, Cyp11a1, DCK, DGKA, DGKZ, DHX37, ELOB(TCEB2), ENTPD1 (CD39), FADD, FAS, GATA3, IL6, IL6R, IL10, IL10RA, IRF4, IRF8, JUNB, Lag3, , LAIR-1 (CD305), LDHA, LIF, LYN, MAP4K4, MAPK14, MCJ, MEF2D, MGAT5, NR4A1, NR4A2, NR4A3, NT5E (CD73), ODC1, OTULINL (FAM105A), PAG1, PDCD1, PDIA3, PHD1 (EGLN2), PHD2 (EGLN1), PHD3 (EGLN3), PIK3CD, PIKFYVE, PPARa, PPARd, PRDMI1, PRKACA, PTEN, PTPN2, PTPN6, PTPN11, PVRIG (CD112R), RASA2, RFXANK, SELPG/PSGL1, SIGLEC15, SLA, SLAMF7, SOCS1, Spry1, Spry2, STK4, SUV39, H1TET2, TGFbRII, TIGIT, Tim-3, TMEM222, TNFAIP3, TNFRSF8 (CD30), TNFRSF10B, TOX, TOX2, , TRAC, TRBC1, TRBC2, UBASH3A, VHL, VISTA, In some embodiments, the at least four guide nucleic acid sequences each hybridize with a gene sequence selected from the group consisting of CD3epsilon, CD3 delta, CD3 gamma, TRAC, TRBC1, and TRBC2, CD2, CD5, CD7, CD52, CD70, and CIITA.

[00264] In some embodiments, the at least four guide nucleic acid sequences comprise a sequence selected from the group consisting of UUCGUAUCUGUAAAACCAAG, CCUACCUGUCACCAGGACCA, CUCUUACCUGUACCAUAACC, CACCUACCUAAGAACCAUCC, ACUCACGCUGGAUAGCCUCC, ACUCACCCAGCAUCCCCAGC, CACUCACCUUAGCCUGAGCA, and CACGCACCUGGACAGCUGAC.

[00265] In one aspect, provided herein is an immune cell comprising the composition of some of the embodiments described above, wherein the composition is introduced into the immune cell with electroporation.

[00266] In one aspect, provided herein is an immune cell comprising the composition of some of the embodiments described above, wherein the composition is introduced into the immune cell with electroporation, nucleofection, viral transduction, or a combination thereof.

[00267] Other features and advantages of the invention will be apparent from the detailed description, and from the claims.

Definitions

[00268] Unless defined otherwise, all technical and scientific terms used herein have the meaning commonly understood by a person skilled in the art to which this invention belongs. The following references provide one of skill with a general definition of many of the terms used in this invention: Singleton et al., Dictionary of Microbiology and Molecular Biology (2nd ed. 1994); The Cambridge Dictionary of Science and Technology (Walker ed., 1988); The Glossary of Genetics, 5th Ed., R. Rieger et al. (eds.), Springer Verlag (1991); and Hale & Marham, The Harper Collins Dictionary of Biology (1991). As used herein, the following terms have the meanings ascribed to them below, unless specified otherwise.

[00269] By "adenosine deaminase" is meant a polypeptide or fragment thereof capable of catalyzing the hydrolytic deamination of adenine or adenosine. In some embodiments, the deaminase or deaminase domain is an adenosine deaminase catalyzing the hydrolytic deamination of adenosine to inosine or deoxyadenosine to deoxyinosine. In some embodiments, the adenosine deaminase catalyzes the hydrolytic deamination of adenine or adenosine in deoxyribonucleic acid (DNA). The adenosine deaminases (e.g., engineered adenosine deaminases, evolved adenosine deaminases) provided herein may be from any organism, such as

a bacterium. In some embodiments, the deaminase or deaminase domain is a variant of a naturally-occurring deaminase from an organism. In some embodiments, the deaminase or deaminase domain does not occur in nature. For example, in some embodiments, the deaminase or deaminase domain is at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75% at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5% identical to a naturally-occurring deaminase. In some embodiments, the adenosine deaminase is from a bacterium, such as, *E. coli*, *S. aureus, S. typhi, S. putrefaciens, H. influenzae, or C. crescentus*. In some embodiments, the adenosine deaminase is a TadA deaminase. In some embodiments, the TadA deaminase is an *E. coli* TadA (ecTadA) deaminase or a fragment thereof.

[00270] For example, the truncated ecTadA may be missing one or more N-terminal amino acids relative to a full-length ecTadA. In some embodiments, the truncated ecTadA may be missing 1, 2, 3, 4, 5 ,6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 6, 17, 18, 19, or 20 N-terminal amino acid residues relative to the full length ecTadA. In some embodiments, the truncated ecTadA may be missing 1, 2, 3, 4, 5 ,6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 6, 17, 18, 19, or 20 C-terminal amino acid residues relative to the full length ecTadA. In some embodiments, the ecTadA deaminase does not comprise an N-terminal methionine. In some embodiments, the TadA deaminase is an N-terminal truncated TadA. In particular embodiments, the TadA is any one of the TadAs described in PCT/US2017/045381, which is incorporated herein by reference in its entirety.

[00271] In certain embodiments, the adenosine deaminase comprises the amino acid sequence:
MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPT
AHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKT
GAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTD,
which is termed "the TadA reference sequence."

[00272] In some embodiments the TadA deaminase is a full-length *E. coli* TadA deaminase. For example, in certain embodiments, the adenosine deaminase comprises the amino acid sequence:
MRRAFITGVFFLSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEG
WNRPIGRHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIG
RVVFGARDAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEI
KAQKKAQSSTD

[00273] It should be appreciated, however, that additional adenosine deaminases useful in the present application would be apparent to the skilled artisan and are within the scope of this disclosure. For example, the adenosine deaminase may be a homolog of adenosine deaminase acting on tRNA (AD AT). Exemplary AD AT homologs include, without limitation:

[00274] *Staphylococcus aureus* TadA:

MGSHMTNDIYFMTLAIEEAKKAAQLGEVPIGAIITKDDEVIARAHNLRETLQQPTAH AEHIAIERAAKVLGSWRLEGCTLYVTLEPCVMCAGTIVMSRIPRVVYGADDPKGGCS GS LMNLLQQS NFNHRAIVDKG VLKE AC S TLLTTFFKNLRANKKS TN

[00275] *Bacillus subtilis* TadA:

MTQDELYMKEAIKEAKKAEEKGEVPIGAVLVINGEIIARAHNLRETEQRSIAHAEML VIDEACKALGTWRLEGATLYVTLEPCPMCAGAVVLSRVEKVVFGAFDPKGGC S GTLMN LLQEERFNHQAEVVSGVLEEECGGMLSAFFRELRKKKKAARKNLSE

[00276] *Salmonella typhimurium* (*S. typhimurium*) TadA:

MPPAFITGVTSLSDVELDHEYWMRHALTLAKRAWDEREVPVGAVLVHNHRVIGEG WNRPIGRHDPTAHAEIMALRQGGLVLQNYRLLDTTLYVTLEPCVMCAGAMVHSRIG RVVFGARDAKTGAAGSLIDVLHHPGMNHRVEIIEGVLRDECATLLSDFFRMRRQEIK ALKKADRAEGAGPAV

[00277] *Shewanella putrefaciens* (*S. putrefaciens*) TadA:

MDE YWMQVAMQM AEKAEAAGE VPVGA VLVKDGQQIATGYNLS IS QHDPT AHAEI LCLRSAGKKLENYRLLDATLYITLEPCAMCAGAMVHSRIARVVYGARDEKTGAAGT VVNLLQHPAFNHQVEVTSGVLAEACSAQLSRFFKRRRDEKKALKLAQRAQQGIE

[00278] *Haemophilus influenzae* F3031 (*H. influenzae*) TadA:

MDAAKVRSEFDEKMMRYALELADKAEALGEIPVGAVLVDDARNIIGEGWNLSIVQS DPT AH AEIIALRNG AKNIQN YRLLNS TLY VTLEPCTMC AG AILHS RIKRLVFG AS D YK TGAIGSRFHFFDDYKMNHTLEITSGVLAEECSQKLSTFFQKRREEKKIEKALLKSLSD K

[00279] *Caulobacter crescentus* (*C. crescentus*) TadA:

MRTDESEDQDHRMMRLALDAARAAAEAGETPVGAVILDPSTGEVIATAGNGPIAAH
DPTAHAEIAAMRAAAAKLGNYRLTDLTLVVTLEPCAMCAGAISHARIGRVVFGADD
PKGGAVVHGPKFFAQPTCHWRPEVTGGVLADESADLLRGFFRARRKAKI

[00280] *Geobacter sulfurreducens* (*G. sulfurreducens*) TadA:

MSSLKKTPIRDDAYWMGKAIREAAKAAARDEVPIGAVIVRDGAVIGRGHNLREGSN
DPSAHAEMIAIRQAARRSANWRLTGATLYVTLEPCLMCMGAIILARLERVVFGCYDP
KGGAAGSLYDLSADPRLNHQVRLSPGVCQEECGTMLSDFFRDLRRRKKAKATPALF
IDERKVPPEP

[00281] TadA7.10

MSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTAH
AEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAG
SLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRQVFNAQKKAQSSTD

[00282] By "agent" is meant any small molecule chemical compound, antibody, nucleic acid molecule, or polypeptide, or fragments thereof.

[00283] By "alteration" is meant a change in the structure, expression levels or activity of a gene or polypeptide as detected by standard art known methods such as those described herein. As used herein, an alteration (e.g., increase or decrease) includes a 10% change in expression levels, a 25% change, a 40% change, and a 50% or greater change in expression levels.

[00284] "Allogeneic," as used herein, refers to cells of the same species that differ genetically to the cell in comparison.

[00285] By "analog" is meant a molecule that is not identical, but has analogous functional or structural features. For example, a polypeptide analog retains the biological activity of a corresponding naturally-occurring polypeptide, while having certain sequence modifications that enhance the analog's function relative to a naturally occurring polypeptide. Such modifications could increase the analog's protease resistance, membrane permeability, or half-life, without altering, for example, polynucleotide binding activity. In another example, a polynucleotide analog retains the biological activity of a corresponding naturally-occurring polynucleotide while having certain modifications that enhance the analog's function relative to a naturally occurring polynucleotide. Such modifications could increase the polynucleotide's affinity for DNA, half-life, and/or nuclease resistance, an analog may include an unnatural nucleotide or amino acid.

[00286] By “anti-neoplasia activity” is meant preventing or inhibiting the maturation and/or proliferation of neoplasms.

[00287] "Autologous," as used herein, refers to cells from the same subject.

[00288] By “B cell maturation antigen, or tumor necrosis factor receptor superfamily member 17 polypeptide, (BCMA)” is meant a protein having at least about 85% amino acid sequence identify to NCBI Accession No. NP_001183 or a fragment thereof that is expressed on mature B lymphocytes. An exemplary BCMA polypeptide sequence is provided below.

[00289] >NP_001183.2 tumor necrosis factor receptor superfamily member 17 [Homo sapiens]
MLQMAGQCSQNEYFDSLLHACIPCQLRCSSNTPPLTCQRYCNASVTNSVKGTNAILWTC
LGLSLIISLAVFVLMFLLRKINSEPLKDEFKNTGSGLLGMANIDLEKSRTGDEIILPRGLEY
TVEECTCEDCIKSKPKVDSDHCFPLPAMEEGATILVTTKTNDYCKSLPAALSATEIEKSIS
AR

[00290] This antigen can be targeted in relapsed or refractory multiple myeloma and other hematological neoplasia therapies.

[00291] By “B cell maturation antigen, or tumor necrosis factor receptor superfamily member 17, (BCMA) polynucleotide” is meant a nucleic acid molecule encoding a BCMA polypeptide. The BCMA gene encodes a cell surface receptor that recognizes B cell activating factor. An exemplary B2M polynucleotide sequence is provided below.

[00292] >NM_001192.2 Homo sapiens TNF receptor superfamily member 17 (TNFRSF17), mRNA
AAGACTCAAACTTAGAAACTTGAATTAGATGTGGTATTCAAATCCTTAGCTGCCGCG
AAGACACAGACAGCCCCCGTAAGAACCCACGAAGCAGGCGAAGTTCATTGTTCTCA
ACATTCTAGCTGCTCTTGCTGCATTTGCTCTGGAATTCTTGTAGAGATATTACTTGTC
CTTCCAGGCTGTTCTTTCTGTAGCTCCCTTGTTTTCTTTTTGTGATCATGTTGCAGATG
GCTGGGCAGTGCTCCCAAAATGAATATTTTGACAGTTTGTTGCATGCTTGCATACCTT
GTCAACTTCGATGTTCTTCTAATACTCCTCCTCTAACATGTCAGCGTTATTGTAATGC
AAGTGTGACCAATTCAGTGAAAGGAACGAATGCGATTCTCTGGACCTGTTTGGGACT
GAGCTTAATAATTTCTTTGGCAGTTTTCGTGCTAATGTTTTTGCTAAGGAAGATAAAC
TCTGAACCATTAAAGGACGAGTTTAAAAACACAGGATCAGGTCTCCTGGGCATGGC
TAACATTGACCTGGAAAAGAGCAGGACTGGTGATGAAATTATTCTTCCGAGAGGCC
TCGAGTACACGGTGGAAGAATGCACCTGTGAAGACTGCATCAAGAGCAAACCGAAG

GTCGACTCTGACCATTGCTTTCCACTCCCAGCTATGGAGGAAGGCGCAACCATTCTT
GTCACCACGAAAACGAATGACTATTGCAAGAGCCTGCCAGCTGCTTTGAGTGCTACG
GAGATAGAGAAATCAATTTCTGCTAGGTAATTAACCATTTCGACTCGAGCAGTGCCA
CTTTAAAAATCTTTTGTCAGAATAGATGATGTGTCAGATCTCTTTAGGATGACTGTAT
TTTTCAGTTGCCGATACAGCTTTTTGTCCTCTAACTGTGGAAACTCTTTATGTTAGAT
ATATTTCTCTAGGTTACTGTTGGGAGCTTAATGGTAGAAACTTCCTTGGTTTCATGAT
TAAACTCTTTTTTTTCCTGA

[00293] By "base editor (BE)," or "nucleobase editor (NBE)" is meant an agent that binds a polynucleotide and has nucleobase modifying activity. In one embodiment, the agent binds the polynucleotide at a specific sequence using a nucleic acid programmable DNA binding protein. In another embodiment, the base editor is an enzyme capable of modifying a cytidine base within a nucleic acid molecule (e.g., DNA). In some embodiments, the base editor is capable of deaminating a base within a nucleic acid molecule. In some embodiments, the base editor is capable of deaminating a base within a DNA molecule. In some embodiments, the base editor is capable of deaminating a cytidine in DNA. In some embodiments, the base editor is a fusion protein comprising a cytidine deaminase or an adenosine deaminase. In some embodiments, the base editor is a Cas9 protein fused to a cytidine deaminase or an adenosine deaminase. In some embodiments, the base editor is a Cas9 nickase (nCas9) fused to a cytidine deaminase or an adenosine deaminase. In some embodiments, the base editor is fused to an inhibitor of base excision repair, for example, a UGI domain. In some embodiments, the fusion protein comprises a Cas9 nickase fused to a deaminase and an inhibitor of base excision repair, such as a UGI domain. In some embodiments, the cytidine deaminase or an or an adenosine deaminase nucleobase editor polypeptide comprising the following domains A-B:

NH_2-[A-B]-COOH,

[00294] wherein A comprises a cytidine deaminase domain, an adenosine deaminase domain or an active fragment thereof, and wherein B comprises one or more domains having nucleic acid sequence specific binding activity. In one embodiment, the cytidine or adenosine deaminase Nucleobase Editor polypeptide of the previous aspect contains:

[00295] NH_2-[A_n-B_o]-COOH, wherein A comprises: a cytidine deaminase domain, an adenosine deaminase domain, or an active fragment thereof, wherein n is an integer: 1, 2, 3, 4, or 5; and wherein B comprises a domain having nucleic acid sequence specific binding activity; and

wherein o is an integer: 1, 2, 3, 4, or 5. In one embodiment, the polypeptide contains one or more nuclear localization sequences. In one embodiment, the polypeptide contains at least one of said nuclear localization sequences is at the N-terminus or C-terminus. In one embodiment, the polypeptide contains the nuclear localization signal is a bipartite nuclear localization signal. In one embodiment, the polypeptide contains one or more domains linked by a linker.

[00296] In some embodiments, the base editor is a cytidine base editor (CBE). In some embodiments, the base editor is an adenosine base editor (ABE). In some embodiments, the base editor is an adenosine base editor (ABE) and a cytidine base editor (CBE). In some embodiments, the base editor is a nuclease-inactive Cas9 (dCas9) fused to an adenosine deaminase. In some embodiments, the Cas9 is a circular permutant Cas9 (e.g., spCas9 or saCas9). Circular permutant Cas9s are known in the art and described, for example, in Oakes et al., Cell 176, 254–267, 2019. In some embodiments, the base editor is fused to an inhibitor of base excision repair, for example, a UGI domain, or a dISN domain. In some embodiments, the fusion protein comprises a Cas9 nickase fused to a deaminase and an inhibitor of base excision repair, such as a UGI or dISN domain. In other embodiments the base editor is an abasic base editor.

[00297] In some embodiments, an adenosine deaminase is evolved from TadA. In some embodiments, the polynucleotide programmable DNA binding domain is a CRISPR associated (*e.g.*, Cas or Cpf1) enzyme. In some embodiments, the base editor is a catalytically dead Cas9 (dCas9) fused to a deaminase domain. In some embodiments, the base editor is a Cas9 nickase (nCas9) fused to a deaminase domain. In some embodiments, the base editor is fused to an inhibitor of base excision repair (BER). In some embodiments, the inhibitor of base excision repair is a uracil DNA glycosylase inhibitor (UGI). In some embodiments, the inhibitor of base excision repair is an inosine base excision repair inhibitor. Details of base editors are described in International PCT Application Nos. PCT/2017/045381 (WO2018/027078) and PCT/US2016/058344 (WO2017/070632), each of which is incorporated herein by reference for its entirety. Also see Komor, A.C., *et al.*, "Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage" Nature 533, 420-424 (2016); Gaudelli, N.M., *et al.*, "Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage" Nature 551, 464-471 (2017); Komor, A.C., *et al.*, "Improved base excision repair inhibition and bacteriophage Mu Gam protein yields C:G-to-T:A base editors with higher efficiency and

product purity" Science Advances 3:eaao4774 (2017), and Rees, H.A., *et al.*, "Base editing: precision chemistry on the genome and transcriptome of living cells." Nat Rev Genet. 2018 Dec;19(12):770-788. doi: 10.1038/s41576-018-0059-1, the entire contents of which are hereby incorporated by reference.

[00298] In some embodiments, base editors are generated by cloning an adenosine deaminase variant (*e.g.*, TadA*7.10) into a scaffold that includes a circular permutant Cas9 (*e.g.*, spCAS9) and a bipartite nuclear localization sequence. Circular permutant Cas9s are known in the art and described, for example, in Oakes *et al.*, Cell 176, 254–267, 2019. Exemplary circular permutant sequences are set forth below, in which the bold sequence indicates sequence derived from Cas9, the italics sequence denotes a linker sequence, and the underlined sequence denotes a bipartite nuclear localization sequence.

[00299] **CP5** (with MSP "NGC=Pam Variant with mutations Regular Cas9 likes NGG" PID=Protein Interacting Domain and "D10A" nickase):

[00300] **EIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGR DFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGG FMQPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKE VKKDLIIKLPKYSLFELENGRKRMLASAKFLQKGNELALPSKYVNFLYLASHYEKL KGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPI REQAENIIHLFTLTNLGAPRAFKYFDTTIARKEYRSTKEVLDATLIHQSITGLYETRI DLSQLGGD***GGSGGSGGSGGSGGSGGSGGM*****DKKYSIGLAIGTNSVGWAVITDEYKVPS KKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEI FSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKK LVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNF DLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEIT KAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQ EEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVD KGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPA FLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTY HDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQL**

KRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKED IQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEM ARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQN GRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEE VVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITK HVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAH DAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEGADKRTADGSEFE SPKKKRKV*

[00301] The nucleobase components and the polynucleotide programmable nucleotide binding component of a base editor system may be associated with each other covalently or non-covalently. For example, in some embodiments, the deaminase domain can be targeted to a target nucleotide sequence by a polynucleotide programmable nucleotide binding domain. In some embodiments, a polynucleotide programmable nucleotide binding domain can be fused or linked to a deaminase domain. In some embodiments, a polynucleotide programmable nucleotide binding domain can target a deaminase domain to a target nucleotide sequence by non-covalently interacting with or associating with the deaminase domain. For example, in some embodiments, the nucleobase editing component, e.g., the deaminase component can comprise an additional heterologous portion or domain that is capable of interacting with, associating with, or capable of forming a complex with an additional heterologous portion or domain that is part of a polynucleotide programmable nucleotide binding domain. In some embodiments, the additional heterologous portion may be capable of binding to, interacting with, associating with, or forming a complex with a polypeptide. In some embodiments, the additional heterologous portion may be capable of binding to, interacting with, associating with, or forming a complex with a polynucleotide. In some embodiments, the additional heterologous portion may be capable of binding to a guide polynucleotide. In some embodiments, the additional heterologous portion may be capable of binding to a polypeptide linker. In some embodiments, the additional heterologous portion may be capable of binding to a polynucleotide linker. The additional heterologous portion may be a protein domain. In some embodiments, the additional heterologous portion may be a K Homology (KH) domain, a MS2 coat protein domain, a PP7 coat protein domain, a SfMu Com coat protein domain, a steril alpha motif, a telomerase Ku

binding motif and Ku protein, a telomerase Sm7 binding motif and Sm7 protein, or a RNA recognition motif.

[00302] A base editor system may further comprise a guide polynucleotide component. It should be appreciated that components of the base editor system may be associated with each other via covalent bonds, noncovalent interactions, or any combination of associations and interactions thereof. In some embodiments, a deaminase domain can be targeted to a target nucleotide sequence by a guide polynucleotide. For example, in some embodiments, the nucleobase editing component of the base editor system, e.g., the deaminase component, can comprise an additional heterologous portion or domain (e.g., polynucleotide binding domain such as an RNA or DNA binding protein) that is capable of interacting with, associating with, or capable of forming a complex with a portion or segment (e.g., a polynucleotide motif) of a guide polynucleotide. In some embodiments, the additional heterologous portion or domain (e.g., polynucleotide binding domain such as an RNA or DNA binding protein) can be fused or linked to the deaminase domain. In some embodiments, the additional heterologous portion may be capable of binding to, interacting with, associating with, or forming a complex with a polypeptide. In some embodiments, the additional heterologous portion may be capable of binding to, interacting with, associating with, or forming a complex with a polynucleotide. In some embodiments, the additional heterologous portion may be capable of binding to a guide polynucleotide. In some embodiments, the additional heterologous portion may be capable of binding to a polypeptide linker. In some embodiments, the additional heterologous portion may be capable of binding to a polynucleotide linker. The additional heterologous portion may be a protein domain. In some embodiments, the additional heterologous portion may be a K Homology (KH) domain, a MS2 coat protein domain, a PP7 coat protein domain, a SfMu Com coat protein domain, a sterile alpha motif, a telomerase Ku binding motif and Ku protein, a telomerase Sm7 binding motif and Sm7 protein, or a RNA recognition motif.

[00303] In some embodiments, a base editor system can further comprise an inhibitor of base excision repair (BER) component. It should be appreciated that components of the base editor system may be associated with each other via covalent bonds, noncovalent interactions, or any combination of associations and interactions thereof. The inhibitor of BER component may comprise a base excision repair inhibitor. In some embodiments, the inhibitor of base excision repair can be a uracil DNA glycosylase inhibitor (UGI). In some embodiments, the inhibitor of

base excision repair can be an inosine base excision repair inhibitor. In some embodiments, the inhibitor of base excision repair can be targeted to the target nucleotide sequence by the polynucleotide programmable nucleotide binding domain. In some embodiments, a polynucleotide programmable nucleotide binding domain can be fused or linked to an inhibitor of base excision repair. In some embodiments, a polynucleotide programmable nucleotide binding domain can be fused or linked to a deaminase domain and an inhibitor of base excision repair. In some embodiments, a polynucleotide programmable nucleotide binding domain can target an inhibitor of base excision repair to a target nucleotide sequence by non-covalently interacting with or associating with the inhibitor of base excision repair. For example, in some embodiments, the inhibitor of base excision repair component can comprise an additional heterologous portion or domain that is capable of interacting with, associating with, or capable of forming a complex with an additional heterologous portion or domain that is part of a polynucleotide programmable nucleotide binding domain. In some embodiments, the inhibitor of base excision repair can be targeted to the target nucleotide sequence by the guide polynucleotide. For example, in some embodiments, the inhibitor of base excision repair can comprise an additional heterologous portion or domain (e.g., polynucleotide binding domain such as an RNA or DNA binding protein) that is capable of interacting with, associating with, or capable of forming a complex with a portion or segment (e.g., a polynucleotide motif) of a guide polynucleotide. In some embodiments, the additional heterologous portion or domain of the guide polynucleotide (e.g., polynucleotide binding domain such as an RNA or DNA binding protein) can be fused or linked to the inhibitor of base excision repair. In some embodiments, the additional heterologous portion may be capable of binding to, interacting with, associating with, or forming a complex with a polynucleotide. In some embodiments, the additional heterologous portion may be capable of binding to a guide polynucleotide. In some embodiments, the additional heterologous portion may be capable of binding to a polypeptide linker. In some embodiments, the additional heterologous portion may be capable of binding to a polynucleotide linker. The additional heterologous portion may be a protein domain. In some embodiments, the additional heterologous portion may be a K Homology (KH) domain, a MS2 coat protein domain, a PP7 coat protein domain, a SfMu Com coat protein domain, a sterile alpha motif, a telomerase Ku binding motif and Ku protein, a telomerase Sm7 binding motif and Sm7 protein, or a RNA recognition motif. By “base editing activity” is meant acting to

chemically alter a base within a polynucleotide. In one embodiment, a first base is converted to a second base. In one embodiment, the base editing activity is cytidine deaminase activity, e.g., converting target C•G to T•A. In another embodiment, the base editing activity is adenosine deaminase activity, e.g., converting A•T to G•C.

[00304] By "beta-2 microglobulin (B2M) polypeptide" is meant a protein having at least about 85% amino acid sequence identity to UniProt Accession No. P61769 or a fragment thereof and having immunomodulatory activity. An exemplary B2M polypeptide sequence is provided below.

>sp|P61769|B2MG_HUMAN Beta-2-microglobulin OS=Homo sapiens OX=9606 GN=B2M PE=1 SV=1

MSRSVALAVLALLSLSGLEAIQRTPKIQVYSRHPAENGKSNFLNCYVSGFHPSDIEVDLL KNGERIEKVEHSDLSFSKDWSFYLLYYTEFTPTEKDEYACRVNHVTLSQPKIVKWDRD M

[00305] By "beta-2-microglobulin (B2M) polynucleotide" is meant a nucleic acid molecule encoding a B2M polypeptide. The beta-2-microglobulin gene encodes a serum protein associated with the major histocompatibility complex. B2M is involved in non-self recognition by host CD8+ T cells. An exemplary B2M polynucleotide sequence is provided below.

>DQ217933.1 Homo sapiens beta-2-microglobin (B2M) gene, complete cds

CATGTCATAAATGGTAAGTCCAAGAAAAATACAGGTATTCCCCCCCAAAGAAAACT
GTAAAATCGACTTTTTTCTATCTGTACTGTTTTTTATTGGTTTTTAAATTGGTTTTCCA
AGTGAGTAAATCAGAATCTATCTGTAATGGATTTTAAATTTAGTGTTTCTCTGTGATG
TAGTAAACAAGAAACTAGAGGCAAAAATAGCCCTGTCCCTTGCTAAACTTCTAAGG
CACTTTTCTAGTACAACTCAACACTAACATTTCAGGCCTTTAGTGCCTTATATGAGTT
TTTAAAAGGGGGAAAAGGGAGGGAGCAAGAGTGTCTTAACTCATACATTTAGGCAT
AACAATTATTCTCATATTTTAGTTATTGAGAGGGCTGGTAGAAAAACTAGGTAAATA
ATATTAATAATTATAGCGCTTATTAAACACTACAGAACACTTACTATGTACCAGGCA
TTGTGGGAGGCTCTCTCTTGTGCATTATCTCATTTCATTAGGTCCATGGAGAGTATTG
CATTTTCTTAGTTTAGGCATGGCCTCCACAATAAAGATTATCAAAAGCCTAAAAATA
TGTAAAAGAAACCTAGAAGTTATTTGTTGTGCTCCTTGGGGAAGCTAGGCAAATCCT
TTCAACTGAAAACCATGGTGACTTCCAAGATCTCTGCCCCTCCCCATCGCCATGGTC
CACTTCCTCTTCTCACTGTTCCTCTTAGAAAAGATCTGTGGACTCCACCACCACGAA

ATGGCGGCACCTTATTTATGGTCACTTTAGAGGGTAGGTTTTCTTAATGGGTCTGCCT
GTCATGTTTAACGTCCTTGGCTGGGTCCAAGGCAGATGCAGTCCAAACTCTCACTAA
AATTGCCGAGCCCTTTGTCTTCCAGTGTCTAAAATATTAATGTCAATGGAATCAGGC
CAGAGTTTGAATTCTAGTCTCTTAGCCTTTGTTTCCCCTGTCCATAAAATGAATGGGG
GTAATTCTTTCCTCCTACAGTTTATTTATATATTCACTAATTCATTCATTCATCCATCC
ATTCGTTCATTCGGTTTACTGAGTACCTACTATGTGCCAGCCCCTGTTCTAGGGTGGA
AACTAAGAGAATGATGTACCTAGAGGGCGCTGGAAGCTCTAAAGCCCTAGCAGTTA
CTGCTTTTACTATTAGTGGTCGTTTTTTTCTCCCCCCCCGCCCCCCGACAAATCAACAG
AACAAAGAAAATTACCTAAACAGCAAGGACATAGGGAGGAACTTCTTGGCACAGAA
CTTTCCAAACACTTTTTCCTGAAGGGATACAAGAAGCAAGAAAGGTACTCTTTCACT
AGGACCTTCTCTGAGCTGTCCTCAGGATGCTTTTGGGACTATTTTTCTTACCCAGAGA
ATGGAGAAACCCTGCAGGGAATTCCCAAGCTGTAGTTATAAACAGAAGTTCTCCTTC
TGCTAGGTAGCATTCAAAGATCTTAATCTTCTGGGTTTCCGTTTTCTCGAATGAAAAA
TGCAGGTCCGAGCAGTTAACTGGCTGGGGCACCATTAGCAAGTCACTTAGCATCTCT
GGGGCCAGTCTGCAAAGCGAGGGGGCAGCCTTAATGTGCCTCCAGCCTGAAGTCCT
AGAATGAGCGCCCGGTGTCCCAAGCTGGGGCGCGCACCCCAGATCGGAGGGCGCCG
ATGTACAGACAGCAAACTCACCCAGTCTAGTGCATGCCTTCTTAAACATCACGAGAC
TCTAAGAAAAGGAAACTGAAAACGGGAAAGTCCCTCTCTCTAACCTGGCACTGCGT
CGCTGGCTTGGAGACAGGTGACGGTCCCTGCGGGCCTTGTCCTGATTGGCTGGGCAC
GCGTTTAATATAAGTGGAGGCGTCGCGCTGGCGGGCATTCCTGAAGCTGACAGCATT
CGGGCCGAGATGTCTCGCTCCGTGGCCTTAGCTGTGCTCGCGCTACTCTCTCTTTCTG
GCCTGGAGGCTATCCAGCGTGAGTCTCTCCTACCCTCCCGCTCTGGTCCTTCCTCTCC
CGCTCTGCACCCTCTGTGGCCCTCGCTGTGCTCTCTCGCTCCGTGACTTCCCTTCTCC
AAGTTCTCCTTGGTGGCCCGCCGTGGGGCTAGTCCAGGGCTGGATCTCGGGGAAGCG
GCGGGGTGGCCTGGGAGTGGGGAAGGGGGTGCGCACCCGGGACGCGCGCTACTTGC
CCCTTTCGGCGGGGAGCAGGGGAGACCTTTGGCCTACGGCGACGGGAGGGTCGGGA
CAAAGTTTAGGGCGTCGATAAGCGTCAGAGCGCCGAGGTTGGGGGAGGGTTTCTCT
TCCGCTCTTTCGCGGGGCCTCTGGCTCCCCCAGCGCAGCTGGAGTGGGGGACGGGTA
GGCTCGTCCCAAAGGCGCGGCGCTGAGGTTTGTGAACGCGTGGAGGGGCGCTTGGG
GTCTGGGGGAGGCGTCGCCCGGGTAAGCCTGTCTGCTGCGGCTCTGCTTCCCTTAGA
CTGGAGAGCTGTGGACTTCGTCTAGGCGCCCGCTAAGTTCGCATGTCCTAGCACCTC

TGGGTCTATGTGGGGCCACACCGTGGGGAGGAAACAGCACGCGACGTTTGTAGAAT
GCTTGGCTGTGATACAAAGCGGTTTCGAATAATTAACTTATTTGTTCCCATCACATGT
CACTTTTAAAAAATTATAAGAACTACCCGTTATTGACATCTTTCTGTGTGCCAAGGA
CTTTATGTGCTTTGCGTCATTTAATTTTGAAAACAGTTATCTTCCGCCATAGATAACT
ACTATGGTTATCTTCTGCCTCTCACAGATGAAGAAACTAAGGCACCGAGATTTTAAG
AAACTTAATTACACAGGGGATAAATGGCAGCAATCGAGATTGAAGTCAAGCCTAAC
CAGGGCTTTTGCGGGAGCGCATGCCTTTTGGCTGTAATTCGTGCATTTTTTTTTAAGA
AAAACGCCTGCCTTCTGCGTGAGATTCTCCAGAGCAAACTGGGCGGCATGGGCCCT
GTGGTCTTTTCGTACAGAGGGCTTCCTCTTTGGCTCTTTGCCTGGTTGTTTCCAAGAT
GTACTGTGCCTCTTACTTTCGGTTTTGAAAACATGAGGGGGTTGGGCGTGGTAGCTT
ACGCCTGTAATCCCAGCACTTAGGGAGGCCGAGGCGGGAGGATGGCTTGAGGTCCG
TAGTTGAGACCAGCCTGGCCAACATGGTGAAGCCTGGTCTCTACAAAAAATAATAA
CAAAAATTAGCCGGGTGTGGTGGCTCGTGCCTGTGGTCCCAGCTGCTCCGGTGGCTG
AGGCGGGAGGATCTCTTGAGCTTAGGCTTTTGAGCTATCATGGCGCCAGTGCACTCC
AGCGTGGGCAACAGAGCGAGACCCTGTCTCTCAAAAAAGAAAAAAAAAAAAAAAG
AAAGAGAAAAGAAAAGAAAGAAAGAAGTGAAGGTTTGTCAGTCAGGGGAGCTGTA
AAACCATTAATAAAGATAATCCAAGATGGTTACCAAGACTGTTGAGGACGCCAGAG
ATCTTGAGCACTTTCTAAGTACCTGGCAATACACTAAGCGCGCTCACCTTTTCCTCTG
GCAAAACATGATCGAAAGCAGAATGTTTTGATCATGAGAAAATTGCATTTAATTTGA
ATACAATTTATTTACAACATAAAGGATAATGTATATATCACCACCATTACTGGTATT
TGCTGGTTATGTTAGATGTCATTTTAAAAAATAACAATCTGATATTTAAAAAAAAAT
CTTATTTTGAAAATTTCCAAAGTAATACATGCCATGCATAGACCATTTCTGGAAGAT
ACCACAAGAAACATGTAATGATGATTGCCTCTGAAGGTCTATTTTCCTCCTCTGACC
TGTGTGTGGGTTTTGTTTTTGTTTTACTGTGGGCATAAATTAATTTTTCAGTTAAGTTT
TGGAAGCTTAAATAACTCTCCAAAAGTCATAAAGCCAGTAACTGGTTGAGCCCAAA
TTCAAACCCAGCCTGTCTGATACTTGTCCTCTTCTTAGAAAAGATTACAGTGATGCTC
TCACAAAATCTTGCCGCCTTCCCTCAAACAGAGAGTTCCAGGCAGGATGAATCTGTG
CTCTGATCCCTGAGGCATTTAATATGTTCTTATTATTAGAAGCTCAGATGCAAAGAG
CTCTCTTAGCTTTTAATGTTATGAAAAAAATCAGGTCTTCATTAGATTCCCCAATCCA
CCTCTTGATGGGGCTAGTAGCCTTTCCTTAATGATAGGGTGTTTCTAGAGAGATATA
TCTGGTCAAGGTGGCCTGGTACTCCTCCTTCTCCCCACAGCCTCCCAGACAAGGAGG

AGTAGCTGCCTTTTAGTGATCATGTACCCTGAATATAAGTGTATTTAAAAGAATTTT
ATACACATATATTTAGTGTCAATCTGTATATTTAGTAGCACTAACACTTCTCTTCATT
TTCAATGAAAAATATAGAGTTTATAATATTTTCTTCCCACTTCCCCATGGATGGTCTA
GTCATGCCTCTCATTTTGGAAAGTACTGTTTCTGAAACATTAGGCAATATATTCCCAA
CCTGGCTAGTTTACAGCAATCACCTGTGGATGCTAATTAAAACGCAAATCCCACTGT
CACATGCATTACTCCATTTGATCATAATGGAAAGTATGTTCTGTCCCATTTGCCATAG
TCCTCACCTATCCCTGTTGTATTTTATCGGGTCCAACTCAACCATTTAAGGTATTTGC
CAGCTCTTGTATGCATTTAGGTTTTGTTTCTTTGTTTTTTAGCTCATGAAATTAGGTAC
AAAGTCAGAGAGGGGTCTGGCATATAAAACCTCAGCAGAAATAAAGAGGTTTTGTT
GTTTGGTAAGAACATACCTTGGGTTGGTTGGGCACGGTGGCTCGTGCCTGTAATCCC
AACACTTTGGGAGGCCAAGGCAGGCTGATCACTTGAAGTTGGGAGTTCAAGACCAG
CCTGGCCAACATGGTGAAATCCCGTCTCTACTGAAAATACAAAAATTAACCAGGCAT
GGTGGTGTGTGCCTGTAGTCCCAGGAATCACTTGAACCCAGGAGGCGGAGGTTGCA
GTGAGCTGAGATCTCACCACTGCACACTGCACTCCAGCCTGGGCAATGGAATGAGA
TTCCATCCCAAAAAATAAAAAAATAAAAAAATAAAGAACATACCTTGGGTTGATCC
ACTTAGGAACCTCAGATAATAACATCTGCCACGTATAGAGCAATTGCTATGTCCCAG
GCACTCTACTAGACACTTCATACAGTTTAGAAAATCAGATGGGTGTAGATCAAGGCA
GGAGCAGGAACCAAAAAGAAAGGCATAAACATAAGAAAAAAAATGGAAGGGGTGG
AAACAGAGTACAATAACATGAGTAATTTGATGGGGGCTATTATGAACTGAGAAATG
AACTTTGAAAAGTATCTTGGGGCCAAATCATGTAGACTCTTGAGTGATGTGTTAAGG
AATGCTATGAGTGCTGAGAGGGCATCAGAAGTCCTTGAGAGCCTCCAGAGAAAGGC
TCTTAAAAATGCAGCGCAATCTCCAGTGACAGAAGATACTGCTAGAAATCTGCTAG
AAAAAAAACAAAAAAGGCATGTATAGAGGAATTATGAGGGAAAGATACCAAGTCA
CGGTTTATTCTTCAAAATGGAGGTGGCTTGTTGGGAAGGTGGAAGCTCATTTGGCCA
GAGTGGAAATGGAATTGGGAGAAATCGATGACCAAATGTAAACACTTGGTGCCTGA
TATAGCTTGACACCAAGTTAGCCCCAAGTGAAATACCCTGGCAATATTAATGTGTCT
TTTCCCGATATTCCTCAGGTACTCCAAAGATTCAGGTTTACTCACGTCATCCAGCAG
AGAATGGAAAGTCAAATTTCCTGAATTGCTATGTGTCTGGGTTTCATCCATCCGACA
TTGAAGTTGACTTACTGAAGAATGGAGAGAGAATTGAAAAAGTGGAGCATTCAGAC
TTGTCTTTCAGCAAGGACTGGTCTTTCTATCTCTTGTACTACACTGAATTCACCCCCA
CTGAAAAAGATGAGTATGCCTGCCGTGTGAACCATGTGACTTTGTCACAGCCCAAGA

TAGTTAAGTGGGGTAAGTCTTACATTCTTTTGTAAGCTGCTGAAAGTTGTGTATGAG
TAGTCATATCATAAAGCTGCTTTGATATAAAAAAGGTCTATGGCCATACTACCCTGA
ATGAGTCCCATCCCATCTGATATAAACAATCTGCATATTGGGATTGTCAGGGAATGT
TCTTAAAGATCAGATTAGTGGCACCTGCTGAGATACTGATGCACAGCATGGTTTCTG
AACCAGTAGTTTCCCTGCAGTTGAGCAGGGAGCAGCAGCAGCACTTGCACAAATAC
ATATACACTCTTAACACTTCTTACCTACTGGCTTCCTCTAGCTTTTGTGGCAGCTTCA
GGTATATTTAGCACTGAACGAACATCTCAAGAAGGTATAGGCCTTTGTTTGTAAGTC
CTGCTGTCCTAGCATCCTATAATCCTGGACTTCTCCAGTACTTTCTGGCTGGATTGGT
ATCTGAGGCTAGTAGGAAGGGCTTGTTCCTGCTGGGTAGCTCTAAACAATGTATTCA
TGGGTAGGAACAGCAGCCTATTCTGCCAGCCTTATTTCTAACCATTTTAGACATTTGT
TAGTACATGGTATTTTAAAAGTAAAACTTAATGTCTTCCTTTTTTTTCTCCACTGTCTT
TTTCATAGATCGAGACATGTAAGCAGCATCATGGAGGTAAGTTTTTGACCTTGAGAA
AATGTTTTTGTTTCACTGTCCTGAGGACTATTTATAGACAGCTCTAACATGATAACCC
TCACTATGTGGAGAACATTGACAGAGTAACATTTTAGCAGGGAAAGAAGAATCCTA
CAGGGTCATGTTCCCTTCTCCTGTGGAGTGGCATGAAGAAGGTGTATGGCCCCAGGT
ATGGCCATATTACTGACCCTCTACAGAGAGGGCAAAGGAACTGCCAGTATGGTATT
GCAGGATAAAGGCAGGTGGTTACCCACATTACCTGCAAGGCTTTGATCTTTCTTCTG
CCATTTCCACATTGGACATCTCTGCTGAGGAGAGAAAATGAACCACTCTTTTCCTTT
GTATAATGTTGTTTTATTCTTCAGACAGAAGAGAGGAGTTATACAGCTCTGCAGACA
TCCCATTCCTGTATGGGGACTGTGTTTGCCTCTTAGAGGTTCCCAGGCCACTAGAGG
AGATAAAGGGAAACAGATTGTTATAACTTGATATAATGATACTATAATAGATGTAA
CTACAAGGAGCTCCAGAAGCAAGAGAGAGGGAGGAACTTGGACTTCTCTGCATCTT
TAGTTGGAGTCCAAAGGCTTTTCAATGAAATTCTACTGCCCAGGGTACATTGATGCT
GAAACCCCATTCAAATCTCCTGTTATATTCTAGAACAGGGAATTGATTTGGGAGAGC
ATCAGGAAGGTGGATGATCTGCCCAGTCACACTGTTAGTAAATTGTAGAGCCAGGA
CCTGAACTCTAATATAGTCATGTGTTACTTAATGACGGGGACATGTTCTGAGAAATG
CTTACACAAACCTAGGTGTTGTAGCCTACTACACGCATAGGCTACATGGTATAGCCT
ATTGCTCCTAGACTACAAACCTGTACAGCCTGTTACTGTACTGAATACTGTGGGCAG
TTGTAACACAATGGTAAGTATTTGTGTATCTAAACATAGAAGTTGCAGTAAAAATAT
GCTATTTTAATCTTATGAGACCACTGTCATATATACAGTCCATCATTGACCAAAACA
TCATATCAGCATTTTTTCTTCTAAGATTTTGGGAGCACCAAAGGGATACACTAACAG

GATATACTCTTTATAATGGGTTTGGAGAACTGTCTGCAGCTACTTCTTTTAAAAAGGT
GATCTACACAGTAGAAATTAGACAAGTTTGGTAATGAGATCTGCAATCCAAATAAA
ATAAATTCATTGCTAACCTTTTTCTTTTCTTTTCAGGTTTGAAGATGCCGCATTTGGA
TTGGATGAATTCCAAATTCTGCTTGCTTGCTTTTTAATATTGATATGCTTATACACTT
ACACTTTATGCACAAAATGTAGGGTTATAATAATGTTAACATGGACATGATCTTCTT
TATAATTCTACTTTGAGTGCTGTCTCCATGTTTGATGTATCTGAGCAGGTTGCTCCAC
AGGTAGCTCTAGGAGGGCTGGCAACTTAGAGGTGGGGAGCAGAGAATTCTCTTATC
CAACATCAACATCTTGGTCAGATTTGAACTCTTCAATCTCTTGCACTCAAAGCTTGTT
AAGATAGTTAAGCGTGCATAAGTTAACTTCCAATTTACATACTCTGCTTAGAATTTG
GGGGAAAATTTAGAAATATAATTGACAGGATTATTGGAAATTTGTTATAATGAATGA
AACATTTTGTCATATAAGATTCATATTTACTTCTTATACATTTGATAAAGTAAGGCAT
GGTTGTGGTTAATCTGGTTTATTTTTGTTCCACAAGTTAAATAAATCATAAAACTTGA
TGTGTTATCTCTTATATCTCACTCCCACTATTACCCCTTTATTTTCAAACAGGGAAAC
AGTCTTCAAGTTCCACTTGGTAAAAAATGTGAACCCCTTGTATATAGAGTTTGGCTC
ACAGTGTAAAGGGCCTCAGTGATTCACATTTTCCAGATTAGGAATCTGATGCTCAAA
GAAGTTAAATGGCATAGTTGGGGTGACACAGCTGTCTAGTGGGAGGCCAGCCTTCT
ATATTTTAGCCAGCGTTCTTTCCTGCGGGCCAGGTCATGAGGAGTATGCAGACTCTA
AGAGGGAGCAAAAGTATCTGAAGGATTTAATATTTTAGCAAGGAATAGATATACAA
TCATCCCTTGGTCTCCCTGGGGGATTGGTTTCAGGACCCCTTCTTGGACACCAAATCT
ATGGATATTTAAGTCCCTTCTATAAAATGGTATAGTATTTGCATATAACCTATCCACA
TCCTCCTGTATACTTTAAATCATTTCTAGATTACTTGTAATACCTAATACAATGTAAA
TGCTATGCAAATAGTTGTTATTGTTTAAGGAATAATGACAAGAAAAAAAAGTCTGTA
CATGCTCAGTAAAGACACAACCATCCCTTTTTTTCCCCAGTGTTTTTGATCCATGGTT
TGCTGAATCCACAGATGTGGAGCCCCTGGATACGGAAGGCCCGCTGTACTTTGAATG
ACAAATAACAGATTTAAA

[00306] The term “Cas9” or “Cas9 domain” refers to an RNA-guided nuclease comprising a Cas9 protein, or a fragment thereof (*e.g.*, a protein comprising an active, inactive, or partially active DNA cleavage domain of Cas9, and/or the gRNA binding domain of Cas9). A Cas9 nuclease is also referred to sometimes as a casn1 nuclease or a CRISPR (“clustered regularly interspaced short palindromic repeat”)-associated nuclease. CRISPR is an adaptive immune system that provides protection against mobile genetic elements (viruses, transposable elements

and conjugative plasmids). CRISPR clusters contain spacers, sequences complementary to antecedent mobile elements, and target invading nucleic acids. CRISPR clusters are transcribed and processed into CRISPR RNA (crRNA). In type II CRISPR systems correct processing of pre-crRNA requires a trans-encoded small RNA (tracrRNA), endogenous ribonuclease 3 (rnc) and a Cas9 protein. The tracrRNA serves as a guide for ribonuclease 3-aided processing of pre-crRNA. Subsequently, Cas9/crRNA/tracrRNA endonucleolytically cleaves linear or circular dsDNA target complementary to the spacer. The target strand not complementary to crRNA is first cut endonucleolytically, then trimmed 3′-5′ exonucleolytically. In nature, DNA-binding and cleavage typically requires protein and both RNAs. However, single guide RNAs ("sgRNA", or simply "gNRA") can be engineered so as to incorporate aspects of both the crRNA and tracrRNA into a single RNA species. See, e.g., Jinek M., Chylinski K., Fonfara I., Hauer M., Doudna J.A., Charpentier E. *Science* 337:816-821(2012), the entire contents of which is hereby incorporated by reference. Cas9 recognizes a short motif in the CRISPR repeat sequences (the PAM or protospacer adjacent motif) to help distinguish self versus non-self. Cas9 nuclease sequences and structures are well known to those of skill in the art (see, e.g., "Complete genome sequence of an M1 strain of Streptococcus pyogenes." Ferretti *et al.*, J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K., Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P., Qian Y., Jia H.G., Najar F.Z., Ren Q., Zhu H., Song L., White J., Yuan X., Clifton S.W., Roe B.A., McLaughlin R.E., Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001); "CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III." Deltcheva E., Chylinski K., Sharma C.M., Gonzales K., Chao Y., Pirzada Z.A., Eckert M.R., Vogel J., Charpentier E., Nature 471:602-607(2011); and "A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity." Jinek M., Chylinski K., Fonfara I., Hauer M., Doudna J.A., Charpentier E. *Science* 337:816-821(2012), the entire contents of each of which are incorporated herein by reference). Cas9 orthologs have been described in various species, including, but not limited to, *S. pyogenes* and *S. thermophilus*. Additional suitable Cas9 nucleases and sequences will be apparent to those of skill in the art based on this disclosure, and such Cas9 nucleases and sequences include Cas9 sequences from the organisms and loci disclosed in Chylinski, Rhun, and Charpentier, "The tracrRNA and Cas9 families of type II CRISPR-Cas immunity systems" (2013) RNA Biology 10:5, 726-737; the entire contents of

which are incorporated herein by reference. In some embodiments, a Cas9 nuclease has an inactive (*e.g.*, an inactivated) DNA cleavage domain, that is, the Cas9 is a nickase.

[00307] A nuclease-inactivated Cas9 protein may interchangeably be referred to as a "dCas9" protein (for nuclease-"dead" Cas9). Methods for generating a Cas9 protein (or a fragment thereof) having an inactive DNA cleavage domain are known (See, *e.g.*, Jinek *et al., Science.* 337:816-821(2012); Qi *et al.,* "Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression" (2013) *Cell.* 28;152(5):1173-83, the entire contents of each of which are incorporated herein by reference). For example, the DNA cleavage domain of Cas9 is known to include two subdomains, the HNH nuclease subdomain and the RuvC1 subdomain. The HNH subdomain cleaves the strand complementary to the gRNA, whereas the RuvC1 subdomain cleaves the non-complementary strand. Mutations within these subdomains can silence the nuclease activity of Cas9. For example, the mutations D10A and H840A completely inactivate the nuclease activity of *S. pyogenes* Cas9 (Jinek *et al., Science.* 337:816-821(2012); Qi *et al., Cell.* 28;152(5):1173-83 (2013)). In some embodiments, proteins comprising fragments of Cas9 are provided. For example, in some embodiments, a protein comprises one of two Cas9 domains: (1) the gRNA binding domain of Cas9; or (2) the DNA cleavage domain of Cas9. In some embodiments, proteins comprising Cas9 or fragments thereof are referred to as "Cas9 variants." A Cas9 variant shares homology to Cas9, or a fragment thereof. For example, a Cas9 variant is at least about 70% identical, at least about 80% identical, at least about 90% identical, at least about 95% identical, at least about 96% identical, at least about 97% identical, at least about 98% identical, at least about 99% identical, at least about 99.5% identical, or at least about 99.9% identical to wild type Cas9. In some embodiments, the Cas9 variant may have 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 21, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50 or more amino acid changes compared to wild type Cas9. In some embodiments, the Cas9 variant comprises a fragment of Cas9 (*e.g.*, a gRNA binding domain or a DNA-cleavage domain), such that the fragment is at least about 70% identical, at least about 80% identical, at least about 90% identical, at least about 95% identical, at least about 96% identical, at least about 97% identical, at least about 98% identical, at least about 99% identical, at least about 99.5% identical, or at least about 99.9% identical to the corresponding fragment of wild type Cas9. In some embodiments, the fragment is at least 30%, at least 35%, at least 40%, at least 45%, at least

50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95% identical, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5% of the amino acid length of a corresponding wild type Cas9.

[00308] In some embodiments, the fragment is at least 100 amino acids in length. In some embodiments, the fragment is at least 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1150, 1200, 1250, or at least 1300 amino acids in length. In some embodiments, wild type Cas9 corresponds to Cas9 from *Streptococcus pyogenes* (NCBI Reference Sequence: NC_017053.1, nucleotide and amino acid sequences as follows).

ATGGATAAGAAATACTCAATAGGCTTAGATATCGGCACAAATAGCGTCGGATGGGC
GGTGATCACTGATGATTATAAGGTTCCGTCTAAAAAGTTCAAGGTTCTGGGAAATAC
AGACCGCCACAGTATCAAAAAAAATCTTATAGGGGCTCTTTTATTTGGCAGTGGAGA
GACAGCGGAAGCGACTCGTCTCAAACGGACAGCTCGTAGAAGGTATACACGTCGGA
AGAATCGTATTTGTTATCTACAGGAGATTTTTTCAAATGAGATGGCGAAAGTAGATG
ATAGTTTCTTTCATCGACTTGAAGAGTCTTTTTTGGTGGAAGAAGACAAGAAGCATG
AACGTCATCCTATTTTTGGAAATATAGTAGATGAAGTTGCTTATCATGAGAAATATC
CAACTATCTATCATCTGCGAAAAAAATTGGCAGATTCTACTGATAAAGCGGATTTGC
GCTTAATCTATTTGGCCTTAGCGCATATGATTAAGTTTCGTGGTCATTTTTTGATTGA
GGGAGATTTAAATCCTGATAATAGTGATGTGGACAAACTATTTATCCAGTTGGTACA
AATCTACAATCAATTATTTGAAGAAAACCCTATTAACGCAAGTAGAGTAGATGCTAA
AGCGATTCTTTCTGCACGATTGAGTAAATCAAGACGATTAGAAAATCTCATTGCTCA
GCTCCCCGGTGAGAAGAGAAATGGCTTGTTTGGGAATCTCATTGCTTTGTCATTGGG
ATTGACCCCTAATTTTAAATCAAATTTTGATTTGGCAGAAGATGCTAAATTACAGCT
TTCAAAAGATACTTACGATGATGATTTAGATAATTTATTGGCGCAAATTGGAGATCA
ATATGCTGATTTGTTTTTGGCAGCTAAGAATTTATCAGATGCTATTTTACTTTCAGAT
ATCCTAAGAGTAAATAGTGAAATAACTAAGGCTCCCCTATCAGCTTCAATGATTAAG
CGCTACGATGAACATCATCAAGACTTGACTCTTTTAAAAGCTTTAGTTCGACAACAA
CTTCCAGAAAAGTATAAAGAAATCTTTTTTGATCAATCAAAAAACGGATATGCAGGT
TATATTGATGGGGGAGCTAGCCAAGAAGAATTTTATAAATTTATCAAACCAATTTTA
GAAAAAATGGATGGTACTGAGGAATTATTGGTGAAACTAAATCGTGAAGATTTGCT

GCGCAAGCAACGGACCTTTGACAACGGCTCTATTCCCCATCAAATTCACTTGGGTGA
GCTGCATGCTATTTTGAGAAGACAAGAAGACTTTTATCCATTTTTAAAAGACAATCG
TGAGAAGATTGAAAAAATCTTGACTTTTCGAATTCCTTATTATGTTGGTCCATTGGCG
CGTGGCAATAGTCGTTTTGCATGGATGACTCGGAAGTCTGAAGAAACAATTACCCCA
TGGAATTTTGAAGAAGTTGTCGATAAAGGTGCTTCAGCTCAATCATTTATTGAACGC
ATGACAAACTTTGATAAAAATCTTCCAAATGAAAAAGTACTACCAAAACATAGTTTG
CTTTATGAGTATTTTACGGTTTATAACGAATTGACAAAGGTCAAATATGTTACTGAG
GGAATGCGAAAACCAGCATTTCTTTCAGGTGAACAGAAGAAAGCCATTGTTGATTTA
CTCTTCAAAACAAATCGAAAAGTAACCGTTAAGCAATTAAAAGAAGATTATTTCAA
AAAAATAGAATGTTTTGATAGTGTTGAAATTTCAGGAGTTGAAGATAGATTTAATGC
TTCATTAGGCGCCTACCATGATTTGCTAAAAATTATTAAAGATAAAGATTTTTTGGA
TAATGAAGAAAATGAAGATATCTTAGAGGATATTGTTTTAACATTGACCTTATTTGA
AGATAGGGGGATGATTGAGGAAAGACTTAAAACATATGCTCACCTCTTTGATGATA
AGGTGATGAAACAGCTTAAACGTCGCCGTTATACTGGTTGGGGACGTTTGTCTCGAA
AATTGATTAATGGTATTAGGGATAAGCAATCTGGCAAAACAATATTAGATTTTTTGA
AATCAGATGGTTTTGCCAATCGCAATTTTATGCAGCTGATCCATGATGATAGTTTGA
CATTTAAAGAAGATATTCAAAAAGCACAGGTGTCTGGACAAGGCCATAGTTTACAT
GAACAGATTGCTAACTTAGCTGGCAGTCCTGCTATTAAAAAAGGTATTTTACAGACT
GTAAAAATTGTTGATGAACTGGTCAAAGTAATGGGGCATAAGCCAGAAAATATCGT
TATTGAAATGGCACGTGAAAATCAGACAACTCAAAAGGGCCAGAAAAATTCGCGAG
AGCGTATGAAACGAATCGAAGAAGGTATCAAAGAATTAGGAAGTCAGATTCTTAAA
GAGCATCCTGTTGAAAATACTCAATTGCAAAATGAAAAGCTCTATCTCTATTATCTA
CAAAATGGAAGAGACATGTATGTGGACCAAGAATTAGATATTAATCGTTTAAGTGA
TTATGATGTCGATCACATTGTTCCACAAAGTTTCATTAAAGACGATTCAATAGACAA
TAAGGTACTAACGCGTTCTGATAAAAATCGTGGTAAATCGGATAACGTTCCAAGTGA
AGAAGTAGTCAAAAAGATGAAAAACTATTGGAGACAACTTCTAAACGCCAAGTTAA
TCACTCAACGTAAGTTTGATAATTTAACGAAAGCTGAACGTGGAGGTTTGAGTGAAC
TTGATAAAGCTGGTTTTATCAAACGCCAATTGGTTGAAACTCGCCAAATCACTAAGC
ATGTGGCACAAATTTTGGATAGTCGCATGAATACTAAATACGATGAAAATGATAAA
CTTATTCGAGAGGTTAAAGTGATTACCTTAAAATCTAAATTAGTTTCTGACTTCCGA
AAAGATTTCCAATTCTATAAAGTACGTGAGATTAACAATTACCATCATGCCCATGAT

GCGTATCTAAATGCCGTCGTTGGAACTGCTTTGATTAAGAAATATCCAAAACTTGAA
TCGGAGTTTGTCTATGGTGATTATAAAGTTTATGATGTTCGTAAAATGATTGCTAAGT
CTGAGCAAGAAATAGGCAAAGCAACCGCAAAATATTTCTTTTACTCTAATATCATGA
ACTTCTTCAAAACAGAAATTACACTTGCAAATGGAGAGATTCGCAAACGCCCTCTAA
TCGAAACTAATGGGGAAACTGGAGAAATTGTCTGGGATAAAGGGCGAGATTTTGCC
ACAGTGCGCAAAGTATTGTCCATGCCCCAAGTCAATATTGTCAAGAAAACAGAAGT
ACAGACAGGCGGATTCTCCAAGGAGTCAATTTTACCAAAAAGAAATTCGGACAAGC
TTATTGCTCGTAAAAAAGACTGGGATCCAAAAAAATATGGTGGTTTTGATAGTCCAA
CGGTAGCTTATTCAGTCCTAGTGGTTGCTAAGGTGGAAAAAGGGAAATCGAAGAAG
TTAAAATCCGTTAAAGAGTTACTAGGGATCACAATTATGGAAAGAAGTTCCTTTGAA
AAAAATCCGATTGACTTTTTAGAAGCTAAAGGATATAAGGAAGTTAAAAAAGACTT
AATCATTAAACTACCTAAATATAGTCTTTTTGAGTTAGAAAACGGTCGTAAACGGAT
GCTGGCTAGTGCCGGAGAATTACAAAAAGGAAATGAGCTGGCTCTGCCAAGCAAAT
ATGTGAATTTTTTATATTTAGCTAGTCATTATGAAAAGTTGAAGGGTAGTCCAGAAG
ATAACGAACAAAAACAATTGTTTGTGGAGCAGCATAAGCATTATTTAGATGAGATT
ATTGAGCAAATCAGTGAATTTTCTAAGCGTGTTATTTTAGCAGATGCCAATTTAGAT
AAAGTTCTTAGTGCATATAACAAACATAGAGACAAACCAATACGTGAACAAGCAGA
AAATATTATTCATTTATTTACGTTGACGAATCTTGGAGCTCCCGCTGCTTTTAAATAT
TTTGATACAACAATTGATCGTAAACGATATACGTCTACAAAAGAAGTTTTAGATGCC
ACTCTTATCCATCAATCCATCACTGGTCTTTATGAAACACGCATTGATTTGAGTCAGC
TAGGAGGTGACTGA

[00309]

MDKKYSIGLDIGTNSVGWAVITDDYKVPSKKFKVLGNTDRHSIKKNLIGALLFGSGETA
EATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIF
GNIVDEVAYHEKYPTIYHLRKKLADSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNS
DVDKLFIQLVQIYNQLFEENPINASRVDAKAILSARLSKSRRLENLIAQLPGEKRNGLFGN
LIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDA
ILLSDILRVNSEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYA
GYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELH
AILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEV
VDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAF

LSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGAYHDLL
KIIKDKDFLDNEENEDILEDIVLTLTLFEDRGMIEERLKTYAHLFDDKVMKQLKRRRYTG
WGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQG
HSLHEQIANLAGSPAIKKGILQTVKIVDELVKVMGHKPENIVIEMARENQTTQKGQKNS
RERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDY
DVDHIVPQSFIKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQ
RKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREV
KVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGD
YKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIV
WDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKY
GGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEV
KKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSP
EDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENII
HLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD

[00310] (single underline: HNH domain; double underline: RuvC domain)

[00311] In some embodiments, wild type Cas9 corresponds to, or comprises the following nucleotide and/or amino acid sequences:

ATGGATAAAAAGTATTCTATTGGTTTAGACATCGGCACTAATTCCGTTGGATGGGCT
GTCATAACCGATGAATACAAAGTACCTTCAAAGAAATTTAAGGTGTTGGGGAACAC
AGACCGTCATTCGATTAAAAAGAATCTTATCGGTGCCCTCCTATTCGATAGTGGCGA
AACGGCAGAGGCGACTCGCCTGAAACGAACCGCTCGGAGAAGGTATACACGTCGCA
AGAACCGAATATGTTACTTACAAGAAATTTTTAGCAATGAGATGGCCAAAGTTGAC
GATTCTTTCTTTCACCGTTTGGAAGAGTCCTTCCTTGTCGAAGAGGACAAGAAACAT
GAACGGCACCCCATCTTTGGAAACATAGTAGATGAGGTGGCATATCATGAAAAGTA
CCCAACGATTTATCACCTCAGAAAAAAGCTAGTTGACTCAACTGATAAAGCGGACCT
GAGGTTAATCTACTTGGCTCTTGCCCATATGATAAAGTTCCGTGGGCACTTTCTCATT
GAGGGTGATCTAAATCCGGACAACTCGGATGTCGACAAACTGTTCATCCAGTTAGTA
CAAACCTATAATCAGTTGTTTGAAGAGAACCCTATAAATGCAAGTGGCGTGGATGC
GAAGGCTATTCTTAGCGCCCGCCTCTCTAAATCCCGACGGCTAGAAAACCTGATCGC

ACAATTACCCGGAGAGAAGAAAAATGGGTTGTTCGGTAACCTTATAGCGCTCTCACT
AGGCCTGACACCAAATTTTAAGTCGAACTTCGACTTAGCTGAAGATGCCAAATTGCA
GCTTAGTAAGGACACGTACGATGACGATCTCGACAATCTACTGGCACAAATTGGAG
ATCAGTATGCGGACTTATTTTTGGCTGCCAAAAACCTTAGCGATGCAATCCTCCTAT
CTGACATACTGAGAGTTAATACTGAGATTACCAAGGCGCCGTTATCCGCTTCAATGA
TCAAAAGGTACGATGAACATCACCAAGACTTGACACTTCTCAAGGCCCTAGTCCGTC
AGCAACTGCCTGAGAAATATAAGGAAATATTCTTTGATCAGTCGAAAAACGGGTAC
GCAGGTTATATTGACGGCGGAGCGAGTCAAGAGGAATTCTACAAGTTTATCAAACC
CATATTAGAGAAGATGGATGGGACGGAAGAGTTGCTTGTAAAACTCAATCGCGAAG
ATCTACTGCGAAAGCAGCGGACTTTCGACAACGGTAGCATTCCACATCAAATCCACT
TAGGCGAATTGCATGCTATACTTAGAAGGCAGGAGGATTTTTATCCGTTCCTCAAAG
ACAATCGTGAAAAGATTGAGAAAATCCTAACCTTTCGCATACCTTACTATGTGGGAC
CCCTGGCCCGAGGGAACTCTCGGTTCGCATGGATGACAAGAAAGTCCGAAGAAACG
ATTACTCCATGGAATTTTGAGGAAGTTGTCGATAAAGGTGCGTCAGCTCAATCGTTC
ATCGAGAGGATGACCAACTTTGACAAGAATTTACCGAACGAAAAAGTATTGCCTAA
GCACAGTTTACTTTACGAGTATTTCACAGTGTACAATGAACTCACGAAAGTTAAGTA
TGTCACTGAGGGCATGCGTAAACCCGCCTTTCTAAGCGGAGAACAGAAGAAAGCAA
TAGTAGATCTGTTATTCAAGACCAACCGCAAAGTGACAGTTAAGCAATTGAAAGAG
GACTACTTTAAGAAAATTGAATGCTTCGATTCTGTCGAGATCTCCGGGGTAGAAGAT
CGATTTAATGCGTCACTTGGTACGTATCATGACCTCCTAAAGATAATTAAAGATAAG
GACTTCCTGGATAACGAAGAGAATGAAGATATCTTAGAAGATATAGTGTTGACTCTT
ACCCTCTTTGAAGATCGGGAAATGATTGAGGAAAGACTAAAAACATACGCTCACCT
GTTCGACGATAAGGTTATGAAACAGTTAAAGAGGCGTCGCTATACGGGCTGGGGAC
GATTGTCGCGGAAACTTATCAACGGGATAAGAGACAAGCAAAGTGGTAAAACTATT
CTCGATTTTCTAAAGAGCGACGGCTTCGCCAATAGGAACTTTATGCAGCTGATCCAT
GATGACTCTTTAACCTTCAAAGAGGATATACAAAAGGCACAGGTTTCCGGACAAGG
GGACTCATTGCACGAACATATTGCGAATCTTGCTGGTTCGCCAGCCATCAAAAAGGG
CATACTCCAGACAGTCAAAGTAGTGGATGAGCTAGTTAAGGTCATGGGACGTCACA
AACCGGAAAACATTGTAATCGAGATGGCACGCGAAAATCAAACGACTCAGAAGGG
GCAAAAAAACAGTCGAGAGCGGATGAAGAGAATAGAAGAGGGTATTAAAGAACTG
GGCAGCCAGATCTTAAAGGAGCATCCTGTGGAAAATACCCAATTGCAGAACGAGAA

ACTTTACCTCTATTACCTACAAAATGGAAGGGACATGTATGTTGATCAGGAACTGGA
CATAAACCGTTTATCTGATTACGACGTCGATCACATTGTACCCCAATCCTTTTTGAAG
GACGATTCAATCGACAATAAAGTGCTTACACGCTCGGATAAGAACCGAGGGAAAAG
TGACAATGTTCCAAGCGAGGAAGTCGTAAAGAAAATGAAGAACTATTGGCGGCAGC
TCCTAAATGCGAAACTGATAACGCAAAGAAAGTTCGATAACTTAACTAAAGCTGAG
AGGGGTGGCTTGTCTGAACTTGACAAGGCCGGATTTATTAAACGTCAGCTCGTGGAA
ACCCGCCAAATCACAAAGCATGTTGCACAGATACTAGATTCCCGAATGAATACGAA
ATACGACGAGAACGATAAGCTGATTCGGGAAGTCAAAGTAATCACTTTAAAGTCAA
AATTGGTGTCGGACTTCAGAAAGGATTTTCAATTCTATAAAGTTAGGGAGATAAATA
ACTACCACCATGCGCACGACGCTTATCTTAATGCCGTCGTAGGGACCGCACTCATTA
AGAAATACCCGAAGCTAGAAAGTGAGTTTGTGTATGGTGATTACAAAGTTTATGAC
GTCCGTAAGATGATCGCGAAAAGCGAACAGGAGATAGGCAAGGCTACAGCCAAAT
ACTTCTTTTATTCTAACATTATGAATTTCTTTAAGACGGAAATCACTCTGGCAAACGG
AGAGATACGCAAACGACCTTTAATTGAAACCAATGGGGAGACAGGTGAAATCGTAT
GGGATAAGGGCCGGGACTTCGCGACGGTGAGAAAAGTTTTGTCCATGCCCCAAGTC
AACATAGTAAAGAAAACTGAGGTGCAGACCGGAGGGTTTTCAAAGGAATCGATTCT
TCCAAAAAGGAATAGTGATAAGCTCATCGCTCGTAAAAAGGACTGGGACCCGAAAA
AGTACGGTGGCTTCGATAGCCCTACAGTTGCCTATTCTGTCCTAGTAGTGGCAAAAG
TTGAGAAGGGAAAATCCAAGAAACTGAAGTCAGTCAAAGAATTATTGGGGATAACG
ATTATGGAGCGCTCGTCTTTTGAAAAGAACCCCATCGACTTCCTTGAGGCGAAAGGT
TACAAGGAAGTAAAAAAGGATCTCATAATTAAACTACCAAAGTATAGTCTGTTTGA
GTTAGAAAATGGCCGAAAACGGATGTTGGCTAGCGCCGGAGAGCTTCAAAAGGGGA
ACGAACTCGCACTACCGTCTAAATACGTGAATTTCCTGTATTTAGCGTCCCATTACG
AGAAGTTGAAAGGTTCACCTGAAGATAACGAACAGAAGCAACTTTTTGTTGAGCAG
CACAAACATTATCTCGACGAAATCATAGAGCAAATTTCGGAATTCAGTAAGAGAGT
CATCCTAGCTGATGCCAATCTGGACAAAGTATTAAGCGCATACAACAAGCACAGGG
ATAAACCCATACGTGAGCAGGCGGAAAATATTATCCATTTGTTTACTCTTACCAACC
TCGGCGCTCCAGCCGCATTCAAGTATTTTGACACAACGATAGATCGCAAACGATACA
CTTCTACCAAGGAGGTGCTAGACGCGACACTGATTCACCAATCCATCACGGGATTAT
ATGAAACTCGGATAGATTTGTCACAGCTTGGGGGTGACGGATCCCCCAAGAAGAAG

AGGAAAGTCTCGAGCGACTACAAAGACCATGACGGTGATTATAAAGATCATGACAT
CGATTACAAGGATGACGATGACAAGGCTGCAGGA

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETA
EATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIF
GNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNS
DVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFG
NLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSD
AILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGY
AGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGEL
HAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEE
VVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPA
FLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLL
KIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTG
WGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQG
DSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKN
SRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSD
YDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLIT
QRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIRE
VKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYG
DYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEI
VWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKK
YGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKE
VKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGS
PEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENI
IHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD

[00312] (single underline: HNH domain; double underline: RuvC domain)

[00313] In some embodiments, wild type Cas9 corresponds to Cas9 from *Streptococcus pyogenes* (NCBI Reference Sequence: NC_002737.2 (nucleotide sequence as follows); and Uniprot Reference Sequence: Q99ZW2 (amino acid sequence as follows).

ATGGATAAGAAATACTCAATAGGCTTAGATATCGGCACAAATAGCGTCGGATGGGC
GGTGATCACTGATGAATATAAGGTTCCGTCTAAAAAGTTCAAGGTTCTGGGAAATAC
AGACCGCCACAGTATCAAAAAAAATCTTATAGGGGCTCTTTTATTTGACAGTGGAGA
GACAGCGGAAGCGACTCGTCTCAAACGGACAGCTCGTAGAAGGTATACACGTCGGA
AGAATCGTATTTGTTATCTACAGGAGATTTTTTCAAATGAGATGGCGAAAGTAGATG
ATAGTTTCTTTCATCGACTTGAAGAGTCTTTTTTGGTGGAAGAAGACAAGAAGCATG
AACGTCATCCTATTTTTGGAAATATAGTAGATGAAGTTGCTTATCATGAGAAATATC
CAACTATCTATCATCTGCGAAAAAAATTGGTAGATTCTACTGATAAAGCGGATTTGC
GCTTAATCTATTTGGCCTTAGCGCATATGATTAAGTTTCGTGGTCATTTTTTGATTGA
GGGAGATTTAAATCCTGATAATAGTGATGTGGACAAACTATTTATCCAGTTGGTACA
AACCTACAATCAATTATTTGAAGAAAACCCTATTAACGCAAGTGGAGTAGATGCTA
AAGCGATTCTTTCTGCACGATTGAGTAAATCAAGACGATTAGAAAATCTCATTGCTC
AGCTCCCCGGTGAGAAGAAAAATGGCTTATTTGGGAATCTCATTGCTTTGTCATTGG
GTTTGACCCCTAATTTTAAATCAAATTTTGATTTGGCAGAAGATGCTAAATTACAGC
TTTCAAAAGATACTTACGATGATGATTTAGATAATTTATTGGCGCAAATTGGAGATC
AATATGCTGATTTGTTTTTGGCAGCTAAGAATTTATCAGATGCTATTTTACTTTCAGA
TATCCTAAGAGTAAATACTGAAATAACTAAGGCTCCCCTATCAGCTTCAATGATTAA
ACGCTACGATGAACATCATCAAGACTTGACTCTTTTAAAAGCTTTAGTTCGACAACA
ACTTCCAGAAAAGTATAAAGAAATCTTTTTTGATCAATCAAAAAACGGATATGCAG
GTTATATTGATGGGGGAGCTAGCCAAGAAGAATTTTATAAATTTATCAAACCAATTT
TAGAAAAAATGGATGGTACTGAGGAATTATTGGTGAAACTAAATCGTGAAGATTTG
CTGCGCAAGCAACGGACCTTTGACAACGGCTCTATTCCCCATCAAATTCACTTGGGT
GAGCTGCATGCTATTTTGAGAAGACAAGAAGACTTTTATCCATTTTTAAAAGACAAT
CGTGAGAAGATTGAAAAAATCTTGACTTTTCGAATTCCTTATTATGTTGGTCCATTGG
CGCGTGGCAATAGTCGTTTTGCATGGATGACTCGGAAGTCTGAAGAAACAATTACCC
CATGGAATTTTGAAGAAGTTGTCGATAAAGGTGCTTCAGCTCAATCATTTATTGAAC
GCATGACAAACTTTGATAAAAATCTTCCAAATGAAAAAGTACTACCAAAACATAGT
TTGCTTTATGAGTATTTTACGGTTTATAACGAATTGACAAAGGTCAAATATGTTACTG
AAGGAATGCGAAAACCAGCATTTCTTTCAGGTGAACAGAAGAAAGCCATTGTTGAT
TTACTCTTCAAAACAAATCGAAAAGTAACCGTTAAGCAATTAAAAGAAGATTATTTC

AAAAAAATAGAATGTTTTGATAGTGTTGAAATTTCAGGAGTTGAAGATAGATTTAAT
GCTTCATTAGGTACCTACCATGATTTGCTAAAAATTATTAAAGATAAAGATTTTTTG
GATAATGAAGAAAATGAAGATATCTTAGAGGATATTGTTTTAACATTGACCTTATTT
GAAGATAGGGAGATGATTGAGGAAAGACTTAAAACATATGCTCACCTCTTTGATGA
TAAGGTGATGAAACAGCTTAAACGTCGCCGTTATACTGGTTGGGGACGTTTGTCTCG
AAAATTGATTAATGGTATTAGGGATAAGCAATCTGGCAAAACAATATTAGATTTTTT
GAAATCAGATGGTTTTGCCAATCGCAATTTTATGCAGCTGATCCATGATGATAGTTT
GACATTTAAAGAAGACATTCAAAAAGCACAAGTGTCTGGACAAGGCGATAGTTTAC
ATGAACATATTGCAAATTTAGCTGGTAGCCCTGCTATTAAAAAAGGTATTTTACAGA
CTGTAAAAGTTGTTGATGAATTGGTCAAAGTAATGGGGCGGCATAAGCCAGAAAAT
ATCGTTATTGAAATGGCACGTGAAAATCAGACAACTCAAAAGGGCCAGAAAAATTC
GCGAGAGCGTATGAAACGAATCGAAGAAGGTATCAAAGAATTAGGAAGTCAGATTC
TTAAAGAGCATCCTGTTGAAAATACTCAATTGCAAAATGAAAAGCTCTATCTCTATT
ATCTCCAAAATGGAAGAGACATGTATGTGGACCAAGAATTAGATATTAATCGTTTAA
GTGATTATGATGTCGATCACATTGTTCCACAAAGTTTCCTTAAAGACGATTCAATAG
ACAATAAGGTCTTAACGCGTTCTGATAAAAATCGTGGTAAATCGGATAACGTTCCAA
GTGAAGAAGTAGTCAAAAAGATGAAAAACTATTGGAGACAACTTCTAAACGCCAAG
TTAATCACTCAACGTAAGTTTGATAATTTAACGAAAGCTGAACGTGGAGGTTTGAGT
GAACTTGATAAAGCTGGTTTTATCAAACGCCAATTGGTTGAAACTCGCCAAATCACT
AAGCATGTGGCACAAATTTTGGATAGTCGCATGAATACTAAATACGATGAAAATGA
TAAACTTATTCGAGAGGTTAAAGTGATTACCTTAAAATCTAAATTAGTTTCTGACTTC
CGAAAAGATTTCCAATTCTATAAAGTACGTGAGATTAACAATTACCATCATGCCCAT
GATGCGTATCTAAATGCCGTCGTTGGAACTGCTTTGATTAAGAAATATCCAAAACTT
GAATCGGAGTTTGTCTATGGTGATTATAAAGTTTATGATGTTCGTAAAATGATTGCT
AAGTCTGAGCAAGAAATAGGCAAAGCAACCGCAAAATATTTCTTTTACTCTAATATC
ATGAACTTCTTCAAAACAGAAATTACACTTGCAAATGGAGAGATTCGCAAACGCCCT
CTAATCGAAACTAATGGGGAAACTGGAGAAATTGTCTGGGATAAAGGGCGAGATTT
TGCCACAGTGCGCAAAGTATTGTCCATGCCCCAAGTCAATATTGTCAAGAAAACAG
AAGTACAGACAGGCGGATTCTCCAAGGAGTCAATTTTACCAAAAAGAAATTCGGAC
AAGCTTATTGCTCGTAAAAAAGACTGGGATCCAAAAAAATATGGTGGTTTTGATAGT
CCAACGGTAGCTTATTCAGTCCTAGTGGTTGCTAAGGTGGAAAAAGGGAAATCGAA

GAAGTTAAAATCCGTTAAAGAGTTACTAGGGATCACAATTATGGAAAGAAGTTCCTT
TGAAAAAAATCCGATTGACTTTTTAGAAGCTAAAGGATATAAGGAAGTTAAAAAAG
ACTTAATCATTAAACTACCTAAATATAGTCTTTTTGAGTTAGAAAACGGTCGTAAAC
GGATGCTGGCTAGTGCCGGAGAATTACAAAAAGGAAATGAGCTGGCTCTGCCAAGC
AAATATGTGAATTTTTTATATTTAGCTAGTCATTATGAAAAGTTGAAGGGTAGTCCA
GAAGATAACGAACAAAAACAATTGTTTGTGGAGCAGCATAAGCATTATTTAGATGA
GATTATTGAGCAAATCAGTGAATTTTCTAAGCGTGTTATTTTAGCAGATGCCAATTT
AGATAAAGTTCTTAGTGCATATAACAAACATAGAGACAAACCAATACGTGAACAAG
CAGAAAATATTATTCATTTATTTACGTTGACGAATCTTGGAGCTCCCGCTGCTTTTAA
ATATTTTGATACAACAATTGATCGTAAACGATATACGTCTACAAAAGAAGTTTTAGA
TGCCACTCTTATCCATCAATCCATCACTGGTCTTTATGAAACACGCATTGATTTGAGT
CAGCTAGGAGGTGACTGA

[00314] MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLF
DSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKK
HERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEG
DLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGE
KKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFL
AAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFF
DQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPH
QIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETI
TPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTE
GMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASL
GTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQL
KRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKA
QVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTT
QKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQEL
DINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQL
LNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDE
NDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKL
ESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIET

NGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKK
DWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLE
AKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY
EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPI
REQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQ
LGGD (single underline: HNH domain; double underline: RuvC domain)

[00315] In some embodiments, Cas9 refers to Cas9 from: *Corynebacterium ulcerans* (NCBI Refs: NC_015683.1, NC_017317.1); *Corynebacterium diphtheria* (NCBI Refs: NC_016782.1, NC_016786.1); *Spiroplasma syrphidicola* (NCBI Ref: NC_021284.1); *Prevotella intermedia* (NCBI Ref: NC_017861.1); *Spiroplasma taiwanense* (NCBI Ref: NC_021846.1); *Streptococcus iniae* (NCBI Ref: NC_021314.1); *Belliella baltica* (NCBI Ref: NC_018010.1); *Psychroflexus torquisI* (NCBI Ref: NC_018721.1); *Streptococcus thermophilus* (NCBI Ref: YP_820832.1), *Listeria innocua* (NCBI Ref: NP_472073.1), *Campylobacter jejuni* (NCBI Ref: YP_002344900.1) or *Neisseria. meningitidis* (NCBI Ref: YP_002342100.1) or to a Cas9 from any other organism.

[00316] In some embodiments, dCas9 corresponds to, or comprises in part or in whole, a Cas9 amino acid sequence having one or more mutations that inactivate the Cas9 nuclease activity. For example, in some embodiments, a dCas9 domain comprises D10A and an H840A mutation or corresponding mutations in another Cas9. In some embodiments, the dCas9 comprises the amino acid sequence of dCas9 (D10A and H840A):

[00317] MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLF
DSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKK
HERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEG
DLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGE
KKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFL
AAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFF
DQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPH
QIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETI
TPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTE

GMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASL
GTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQL
KRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKA
QVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTT
QKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQEL
DINRLSDYDVDAIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQL
LNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDE
NDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKL
ESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIET
NGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKK
DWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLE
AKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY
EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPI
REQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQ
LGGD (single underline: HNH domain; double underline: RuvC domain).

[00318] In some embodiments, the Cas9 domain comprises a D10A mutation, while the residue at position 840 remains a histidine in the amino acid sequence provided above, or at corresponding positions in any of the amino acid sequences provided herein.

[00319] In other embodiments, dCas9 variants having mutations other than D10A and H840A are provided, which, e.g., result in nuclease inactivated Cas9 (dCas9). Such mutations, by way of example, include other amino acid substitutions at D10 and H840, or other substitutions within the nuclease domains of Cas9 (e.g., substitutions in the HNH nuclease subdomain and/or the RuvC1 subdomain).

[00320] In some embodiments, variants or homologues of dCas9 are provided which are at least about 70% identical, at least about 80% identical, at least about 90% identical, at least about 95% identical, at least about 98% identical, at least about 99% identical, at least about 99.5% identical, or at least about 99.9% identical. In some embodiments, variants of dCas9 are provided having amino acid sequences which are shorter, or longer, by about 5 amino acids, by about 10 amino acids, by about 15 amino acids, by about 20 amino acids, by about 25 amino acids, by about 30 amino acids, by about 40 amino acids, by about 50 amino acids, by about 75 amino acids, by about 100 amino acids or more.

[00321] In some embodiments, Cas9 fusion proteins as provided herein comprise the full-length amino acid sequence of a Cas9 protein, e.g., one of the Cas9 sequences provided herein. In other embodiments, however, fusion proteins as provided herein do not comprise a full-length Cas9 sequence, but only a fragment thereof. For example, in some embodiments, a Cas9 fusion protein provided herein comprises a Cas9 fragment, wherein the fragment binds crRNA and tracrRNA or sgRNA, but does not comprise a functional nuclease domain, *e.g.*, in that it comprises only a truncated version of a nuclease domain or no nuclease domain at all.

[00322] Exemplary amino acid sequences of suitable Cas9 domains and Cas9 fragments are provided herein, and additional suitable sequences of Cas9 domains and fragments will be apparent to those of skill in the art.

[00323] In some embodiments, Cas9 refers to Cas9 from: *Corynebacterium ulcerans* (NCBI Refs: NC_015683.1, NC_017317.1); *Corynebacterium diphtheria* (NCBI Refs: NC_016782.1, NC_016786.1); *Spiroplasma syrphidicola* (NCBI Ref: NC_021284.1); *Prevotella intermedia* (NCBI Ref: NC_017861.1); *Spiroplasma taiwanense* (NCBI Ref: NC_021846.1); *Streptococcus iniae* (NCBI Ref: NC_021314.1); *Belliella baltica* (NCBI Ref: NC_018010.1); *Psychroflexus torquisI* (NCBI Ref: NC_018721.1); *Streptococcus thermophilus* (NCBI Ref: YP_820832.1); *Listeria innocua* (NCBI Ref: NP_472073.1); *Campylobacter jejuni* (NCBI Ref: YP_002344900.1); or *Neisseria. meningitidis* (NCBI Ref: YP_002342100.1).

[00324] It should be appreciated that additional Cas9 proteins (*e.g.,* a nuclease dead Cas9 (dCas9), a Cas9 nickase (nCas9), or a nuclease active Cas9), including variants and homologs thereof, are within the scope of this disclosure. Exemplary Cas9 proteins include, without limitation, those provided below. In some embodiments, the Cas9 protein is a nuclease dead Cas9 (dCas9). In some embodiments, the Cas9 protein is a Cas9 nickase (nCas9). In some embodiments, the Cas9 protein is a nuclease active Cas9.

[00325] Exemplary catalytically inactive Cas9 (dCas9):
DKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEA
TRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGN
IVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDV
DKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLI
ALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAIL

LSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAG
YIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAI
LRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVV
DKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLS
GEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKII
KDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWG
RLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSL
HEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRE
RMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDV
DAIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRK
FDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVI
TLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYK
VYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWD
KGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGG
FDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKK
DLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPED
NEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHL
FTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD

[00326] Exemplary catalytically Cas9 nickase (nCas9):
DKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEA
TRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGN
IVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDV
DKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLI
ALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAIL
LSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAG
YIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAI
LRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVV
DKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLS
GEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKII
KDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWG

RLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSL
HEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRE
RMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDV
DHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRK
FDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVI
TLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYK
VYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWD
KGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGG
FDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKK
DLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPED
NEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHL
FTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD

[00327] Exemplary catalytically active Cas9:

DKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEA
TRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGN
IVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDV
DKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLI
ALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAIL
LSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAG
YIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAI
LRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVV
DKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLS
GEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKII
KDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWG
RLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSL
HEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRE
RMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDV
DHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRK
FDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVI
TLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYK

VYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWD
KGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGG
FDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKK
DLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPED
NEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHL
FTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD.

[00328] In some embodiments, Cas9 refers to a Cas9 from archaea (e.g. nanoarchaea), which constitute a domain and kingdom of single-celled prokaryotic microbes. In some embodiments, Cas9 refers to CasX or CasY, which have been described in, for example, Burstein et al., "New CRISPR-Cas systems from uncultivated microbes." Cell Res. 2017 Feb 21. doi: 10.1038/cr.2017.21, the entire contents of which is hereby incorporated by reference. Using genome-resolved metagenomics, a number of CRISPR-Cas systems were identified, including the first reported Cas9 in the archaeal domain of life. This divergent Cas9 protein was found in little- studied nanoarchaea as part of an active CRISPR-Cas system. In bacteria, two previously unknown systems were discovered, CRISPR-CasX and CRISPR-CasY, which are among the most compact systems yet discovered. In some embodiments, Cas9 refers to CasX, or a variant of CasX. In some embodiments, Cas9 refers to a CasY, or a variant of CasY. It should be appreciated that other RNA-guided DNA binding proteins may be used as a nucleic acid programmable DNA binding protein (napDNAbp), and are within the scope of this disclosure.

[00329] In some embodiments, the nucleic acid programmable DNA binding protein (napDNAbp) or any of the fusion proteins provided herein may be a CasX or CasY protein. In some embodiments, the napDNAbp is a CasX protein. In some embodiments, the napDNAbp is a CasY protein. In some embodiments, the napDNAbp comprises an amino acid sequence that is at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at ease 99.5% identical to a naturally-occurring CasX or CasY protein. In some embodiments, the napDNAbp is a naturally-occurring CasX or CasY protein. In some embodiments, the napDNAbp comprises an amino acid sequence that is at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at ease 99.5% identical to any CasX or CasY protein described herein. It should be appreciated that CasX and CasY from other bacterial species may also be used in accordance with the present disclosure.

[00330] CasX (uniprot.org/uniprot/F0NN87; uniprot.org/uniprot/F0NH53)

[00331] >tr|F0NN87|F0NN87_SULIH CRISPR-associated Casx protein OS = *Sulfolobus islandicus* (strain HVE10/4) GN = SiH_0402 PE=4 SV=1

MEVPLYNIFGDNYIIQVATEAENSTIYNNKVEIDDEELRNVLNLAYKIAKNNEDAAAERR
GKAKKKKGEEGETTTSNIILPLSGNDKNPWTETLKCYNFPTTVALSEVFKNFSQVKECEE
VSAPSFVKPEFYEFGRSPGMVERTRRVKLEVEPHYLIIAAAGWVLTRLGKAKVSEGDYV
GVNVFTPTRGILYSLIQNVNGIVPGIKPETAFGLWIARKVVSSVTNPNVSVVRIYTISDAV
GQNPTTINGGFSIDLTKLLEKRYLLSERLEAIARNALSISSNMRERYIVLANYIYEYLTG
SKRLEDLLYFANRDLIMNLNSDDGKVRDLKLISAYVNGELIRGEG

[00332] >tr|F0NH53|F0NH53_SULIR CRISPR associated protein, Casx OS = *Sulfolobus islandicus* (strain REY15A) GN=SiRe_0771 PE=4 SV=1

MEVPLYNIFGDNYIIQVATEAENSTIYNNKVEIDDEELRNVLNLAYKIAKNNEDAAAERR
GKAKKKKGEEGETTTSNIILPLSGNDKNPWTETLKCYNFPTTVALSEVFKNFSQVKECEE
VSAPSFVKPEFYKFGRSPGMVERTRRVKLEVEPHYLIMAAAGWVLTRLGKAKVSEGDY
VGVNVFTPTRGILYSLIQNVNGIVPGIKPETAFGLWIARKVVSSVTNPNVSVVSIYTISDA
VGQNPTTINGGFSIDLTKLLEKRDLLSERLEAIARNALSISSNMRERYIVLANYIYEYLTGS
KRLEDLLYFANRDLIMNLNSDDGKVRDLKLISAYVNGELIRGEG

[00333] CasY (ncbi.nlm.nih.gov/protein/APG80656.1)

[00334] >APG80656.1 CRISPR-associated protein CasY [uncultured Parcubacteria group bacterium]

MSKRHPRISGVKGYRLHAQRLEYTGKSGAMRTIKYPLYSSPSGGRTVPREIVSAINDDY
VGLYGLSNFDDLYNAEKRNEEKVYSVLDFWYDCVQYGAVFSYTAPGLLKNVAEVRGG
SYELTKTLKGSHLYDELQIDKVIKFLNKKEISRANGSLDKLKKDIIDCFKAEYRERHKDQ
CNKLADDIKNAKKDAGASLGERQKKLFRDFFGISEQSENDKPSFTNPLNLTCCLLPFDTV
NNNRNRGEVLFNKLKEYAQKLDKNEGSLEMWEYIGIGNSGTAFSNFLGEGFLGRLREN
KITELKKAMMDITDAWRGQEQEEELEKRLRILAALTIKLREPKFDNHWGGYRSDINGKL
SSWLQNYINQTVKIKEDLKGHKKDLKKAKEMINRFGESDTKEEAVVSSLLESIEKIVPDD
SADDEKPDIPAIAIYRRFLSDGRLTLNRFVQREDVQEALIKERLEAEKKKKPKKRKKKSD

AEDEKETIDFKELFPHLAKPLKLVPNFYGDSKRELYKKYKNAAIYTDALWKAVEKIYKS
AFSSSLKNSFFDTDFDKDFFIKRLQKIFSVYRRFNTDKWKPIVKNSFAPYCDIVSLAENEV
LYKPKQSRSRKSAAIDKNRVRLPSTENIAKAGIALARELSVAGFDWKDLLKKEEHEEYID
LIELHKTALALLLAVTETQLDISALDFVENGTVKDFMKTRDGNLVLEGRFLEMFSQSIVF
SELRGLAGLMSRKEFITRSAIQTMNGKQAELLYIPHEFQSAKITTPKEMSRAFLDLAPAEF
ATSLEPESLSEKSLLKLKQMRYYPHYFGYELTRTGQGIDGGVAENALRLEKSPVKKREIK
CKQYKTLGRGQNKIVLYVRSSYYQTQFLEWFLHRPKNVQTDVAVSGSFLIDEKKVKTR
WNYDALTVALEPVSGSERVFVSQPFTIFPEKSAEEEGQRYLGIDIGEYGIAYTALEITGDS
AKILDQNFISDPQLKTLREEVKGLKLDQRRGTFAMPSTKIARIRESLVHSLRNRIHHLALK
HKAKIVYELEVSRFEEGKQKIKKVYATLKKADVYSEIDADKNLQTTVWGKLAVASEISA
SYTSQFCGACKKLWRAEMQVDETITTQELIGTVRVIKGGTLIDAIKDFMRPPIFDENDTPF
PKYRDFCDKHHISKKMRGNSCLFICPFCRANADADIQASQTIALLRYVKEEKKVEDYFE
RFRKLKN IKVLGQMKKI

[00335] The term "Cas12b" or "Cas12b domain" refers to an RNA-guided nuclease comprising a Cas12b/C2c1 protein, or a fragment thereof (*e.g.*, a protein comprising an active, inactive, or partially active DNA cleavage domain of Cas12b, and/or the gRNA binding domain of Cas12b). contents of each of which are incorporated herein by reference). Cas12b orthologs have been described in various species, including, but not limited to, *Alicyclobacillus acidoterrestris, Alicyclobacillus acidophilus* (Teng et al., Cell Discov. 2018 Nov 27;4:63), *Bacillus hisashi,* and *Bacillus sp. V3-13*. Additional suitable Cas12b nucleases and sequences will be apparent to those of skill in the art based on this disclosure.

[00336] In some embodiments, proteins comprising Cas12b or fragments thereof are referred to as "Cas12b variants." A Cas12b variant shares homology to Cas12b, or a fragment thereof. For example, a Cas12b variant is at least about 70% identical, at least about 80% identical, at least about 90% identical, at least about 95% identical, at least about 96% identical, at least about 97% identical, at least about 98% identical, at least about 99% identical, at least about 99.5% identical, or at least about 99.9% identical to wild type Cas12b. In some embodiments, the Cas12b variant may have 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 21, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50 or more amino acid changes compared to wild type Cas12b. In some embodiments, the

Cas12b variant comprises a fragment of Cas12b (*e.g.*, a gRNA binding domain or a DNA-cleavage domain), such that the fragment is at least about 70% identical, at least about 80% identical, at least about 90% identical, at least about 95% identical, at least about 96% identical, at least about 97% identical, at least about 98% identical, at least about 99% identical, at least about 99.5% identical, or at least about 99.9% identical to the corresponding fragment of wild type Cas12b. In some embodiments, the fragment is at least 30%, at least 35%, at least 40%, at least 45%, at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95% identical, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5% of the amino acid length of a corresponding wild type Cas12b. Exemplary Cas12b polypeptides are listed below.

[00337] Cas12b/C2c1 (uniprot.org/uniprot/T0D7A2#2)

[00338] sp|T0D7A2|C2C1_ALIAG CRISPR-associated endo- nuclease C2c1 OS = *Alicyclobacillus acido- terrestris* (strain ATCC 49025 / DSM 3922/ CIP 106132 / NCIMB 13137/GD3B) GN=c2c1 PE=1 SV=1

MAVKSIKVKLRLDDMPEIRAGLWKLHKEVNAGVRYYTEWLSLLRQENLYRRSPNGDG
EQECDKTAEECKAELLERLRARQVENGHRGPAGSDDELLQLARQLYELLVPQAIGAKG
DAQQIARKFLSPLADKDAVGGLGIAKAGNKPRWVRMREAGEPGWEEEKEKAETRKSA
DRTADVLRALADFGLKPLMRVYTDSEMSSVEWKPLRKGQAVRTWDRDMFQQAIERM
MSWESWNQRVGQEYAKLVEQKNRFEQKNFVGQEHLVHLVNQLQQDMKEASPGLESK
EQTAHYVTGRALRGSDKVFEKWGKLAPDAPFDLYDAEIKNVQRRNTRRFGSHDLFAKL
AEPEYQALWREDASFLTRYAVYNSILRKLNHAKMFATFTLPDATAHPIWTRFDKLGGN
LHQYTFLFNEFGERRHAIRFHKLLKVENGVAREVDDVTVPISMSEQLDNLLPRDPNEPIA
LYFRDYGAEQHFTGEFGGAKIQCRRDQLAHMHRRRGARDVYLNVSVRVQSQSEARGE
RRPPYAAVFRLVGDNHRAFVHFDKLSDYLAEHPDDGKLGSEGLLSGLRVMSVDLGLRT
SASISVFRVARKDELKPNSKGRVPFFFPIKGNDNLVAVHERSQLLKLPGETESKDLRAIRE
ERQRTLRQLRTQLAYLRLLVRCGSEDVGRRERSWAKLIEQPVDAANHMTPDWREAFEN
ELQKLKSLHGICSDKEWMDAVYESVRRVWRHMGKQVRDWRKDVRSGERPKIRGYAK
DVVGGNSIEQIEYLERQYKFLKSWSFFGKVSGQVIRAEKGSRFAITLREHIDHAKEDRLK
KLADRIIMEALGYVYALDERGKGKWVAKYPPCQLILLEELSEYQFNNDRPPSENNQLM
QWSHRGVFQELINQAQVHDLLVGTMYAAFSSRFDARTGAPGIRCRRVPARCTQEHNPE
PFPWWLNKFVVEHTLDACPLRADDLIPTGEGEIFVSPFSAEEGDFHQIHADLNAAQNLQ

QRLWSDFDISQIRLRCDWGEVDGELVLIPRLTGKRTADSYSNKVFYTNTGVTYYERERG
KKRRKVFAQEKLSEEEAELLVEADEAREKSVVLMRDPSGIINRGNWTRQKEFWSMV
NQRIEGYLVKQIRSRVPLQDSACENTGDI

[00339] AacCas12b (Alicyclobacillus acidiphilus) - WP_067623834

MAVKSMKVKLRLDNMPEIRAGLWKLHTEVNAGVRYYTEWLSLLRQENLYRRSPNGDG
EQECYKTAEECKAELLERLRARQVENGHCGPAGSDDELLQLARQLYELLVPQAIGAKG
DAQQIARKFLSPLADKDAVGGLGIAKAGNKPRWVRMREAGEPGWEEEKAKAEARKST
DRTADVLRALADFGLKPLMRVYTDSDMSSVQWKPLRKGQAVRTWDRDMFQQAIERM
MSWESWNQRVGEAYAKLVEQKSRFEQKNFVGQEHLVQLVNQLQQDMKEASHGLESK
EQTAHYLTGRALRGSDKVFEKWEKLDPDAPFDLYDTEIKNVQRRNTRRFGSHDLFAKL
AEPKYQALWREDASFLTRYAVYNSIVRKLNHAKMFATFTLPDATAHPIWTRFDKLGGN
LHQYTFLFNEFGEGRHAIRFQKLLTVEDGVAKEVDDVTVPISMSAQLDDLLPRDPHELV
ALYFQDYGAEQHLAGEFGGAKIQYRRDQLNHLHARRGARDVYLNLSVRVQSQSEARG
ERRPPYAAVFRLVGDNHRAFVHFDKLSDYLAEHPDDGKLGSEGLLSGLRVMSVDLGLR
TSASISVFRVARKDELKPNSEGRVPFCFPIEGNENLVAVHERSQLLKLPGETESKDLRAIR
EERQRTLRQLRTQLAYLRLLVRCGSEDVGRRERSWAKLIEQPMDANQMTPDWREAFED
ELQKLKSLYGICGDREWTEAVYESVRRVWRHMGKQVRDWRKDVRSGERPKIRGYQKD
VVGGNSIEQIEYLERQYKFLKSWSFFGKVSGQVIRAEKGSRFAITLREHIDHAKEDRLKK
LADRIIMEALGYVYALDDERGKGKWVAKYPPCQLILLEELSEYQFNNDRPPSENNQLM
QWSHRGVFQELLNQAQVHDLLVGTMYAAFSSRFDARTGAPGIRCRRVPARCAREQNPE
PFPWWLNKFVAEHKLDGCPLRADDLIPTGEGEFFVSPFSAEEGDFHQIHADLNAAQNLQ
RRLWSDFDISQIRLRCDWGEVDGEPVLIPRTTGKRTADSYGNKVFYTKTGVTYYERERG
KKRRKVFAQEELSEEEAELLVEADEAREKSVVLMRDPSGIINRGDWTRQKEFWSMVNQ
RIEGYLVKQIRSRVRLQESACENTGDI

[00340] BhCas12b (Bacillus hisashii) NCBI Reference Sequence: WP_095142515

MAPKKKRKVGIHGVPAAATRSFILKIEPNEEVKKGLWKTHEVLNHGIAYYMNILKLIRQ
EAIYEHHEQDPKNPKKVSKAEIQAELWDFVLKMQKCNSFTHEVDKDEVFNILRELYEEL
VPSSVEKKGEANQLSNKFLYPLVDPNSQSGKGTASSGRKPRWYNLKIAGDPSWEEEKK
KWEEDKKKDPLAKILGKLAEYGLIPLFIPYTDSNEPIVKEIKWMEKSRNQSVRRLDKDM
FIQALERFLSWESWNLKVKEEYEKVEKEYKTLEERIKEDIQALKALEQYEKERQEQLLR

DTLNTNEYRLSKRGLRGWREIIQKWLKMDENEPSEKYLEVFKDYQRKHPREAGDYSVY
EFLSKKENHFIWRNHPEYPYLYATFCEIDKKKKDAKQQATFTLADPINHPLWVRFEERS
GSNLNKYRILTEQLHTEKLKKKLTVQLDRLIYPTESGGWEEKGKVDIVLLPSRQFYNQIF
LDIEEKGKHAFTYKDESIKFPLKGTLGGARVQFDRDHLRRYPHKVESGNVGRIYFNMTV
NIEPTESPVSKSLKIHRDDFPKVVNFKPKELTEWIKDSKGKKLKSGIESLEIGLRVMSIDL
GQRQAAAASIFEVVDQKPDIEGKLFFPIKGTELYAVHRASFNIKLPGETLVKSREVLRKA
REDNLKLMNQKLNFLRNVLHFQQFEDITEREKRVTKWISRQENSDVPLVYQDELIQIREL
MYKPYKDWVAFLKQLHKRLEVEIGKEVKHWRKSLSDGRKGLYGISLKNIDEIDRTRKF
LLRWSLRPTEPGEVRRLEPGQRFAIDQLNHLNALKEDRLKKMANTIIMHALGYCYDVR
KKKWQAKNPACQIILFEDLSNYNPY**E**ERSRFENSKLM**K**WSRREIPRQVALQGEIYGLQV
GEVGAQFSSRFHAKTGSPGIRC**S**VVTKEKLQDNRFFKNLQREGRLTLDKIAVLKEGDLY
PDKGGEKFISLSKDRKCVTTHADINAAQNLQKRFWTRTHGFYKVYCKAYQVDGQTVYI
PESKDQKQKIIEEFGEGYFILKDGVYEWVNAGKLKIKKGSSKQSSSELVDSDILKDSFDL
ASELKGEKLMLYRDPSGNVFPSDKWMAAGVFFGKLERILISKLTNQYSISTIEDDSSKQS
MKRPAATKKAGQAKKKK

including the variant termed BvCas12b V4 (S893R/K846R/E837G changes rel. to wt above)

[00341] BvCas12b (Bacillus sp. V3-13) NCBI Reference Sequence: WP_101661451.1

MAIRSIKLKMKTNSGTDSIYLRKALWRTHQLINEGIAYYMNLLTLYRQEAIGDKTKEAY
QAELINIIRNQQRNNGSSEEHGSDQEILALLRQLYELIIPSSIGESGDANQLGNKFLYPLVD
PNSQSGKGTSNAGRKPRWKRLKEEGNPDWELEKKKDEERKAKDPTVKIFDNLNKYGLL
PLFPLFTNIQKDIEWLPLGKRQSVRKWDKDMFIQAIERLLSWESWNRRVADEYKQLKEK
TESYYKEHLTGGEEWIEKIRKFEKERNMELEKNAFAPNDGYFITSRQIRGWDRVYEKWS
KLPESASPEELWKVVAEQQNKMSEGFGDPKVFSFLANRENRDIWRGHSERIYHIAAYNG
LQKKLSRTKEQATFTLPDAIEHPLWIRYESPGGTNLNLFKLEEKQKKNYYVTLSKIIWPS
EEKWIEKENIEIPLAPSIQFNRQIKLKQHVKGKQEISFSDYSSRISLDGVLGGSRIQFNRKYI
KNHKELLGEGDIGPVFFNLVVDVAPLQETRNGRLQSPIGKALKVISSDFSKVIDYKPKEL
MDWMNTGSASNSFGVASLLEGMRVMSIDMGQRTSASVSIFEVVKELPKDQEQKLFYSI
NDTELFAIHKRSFLLNLPGEVVTKNNKQQRQERRKKRQFVRSQIRMLANVLRLETKKTP
DERKKAIHKLMEIVQSYDSWTASQKEVWEKELNLLTNMAAFNDEIWKESLVELHHRIE
PYVGQIVSKWRKGLSEGRKNLAGISMWNIDELEDTRRLLISWSKRSRTPGEANRIETDEP

FGSSLLQHIQNVKDDRLKQMANLIIMTALGFKYDKEEKDRYKRWKETYPACQIILFENL
NRYLFNLDRSRRENSRLMKWAHRSIPRTVSMQGEMFGLQVGDVRSEYSSRFHAKTGAP
GIRCHALTEEDLKAGSNTLKRLIEDGFINESELAYLKKGDIIPSQGGELFVTLSKRYKKDS
DNNELTVIHADINAAQNLQKRFWQQNSEVYRVPCQLARMGEDKLYIPKSQTETIKKYFG
KGSFVKNNTEQEVYKWEKSEKMKIKTDTTFDLQDLDGFEDISKTIELAQEQQKKYLTMF
RDPSGYFFNNETWRPQKEYWSIVNNIIKSCLKKKILSNKVEL

[00342] By "Cbl proto-oncogene B (CBLB) polypeptide" is meant a protein having at least about 85% amino acid sequence identity to GenBank Accession No. ABC86700.1 or a fragment thereof that is involved in the regulation of immune responses. An exemplary CBLB polypeptide sequence is provided below.

[00343] >ABC86700.1 CBL-B [Homo sapiens]
MANSMNGRNPGGRGGNPRKGRILGIIDAIQDAVGPPKQAAADRRTVEKTWKLMDKVV
RLCQNPKLQLKNSPPYILDILPDTYQHLRLILSKYDDNQKLAQLSENEYFKIYIDSLMKKS
KRAIRLFKEGKERMYEEQSQDRRNLTKLSLIFSHMLAEIKAIFPNGQFQGDNFRITKADA
AEFWRKFFGDKTIVPWKVFRQCLHEVHQISSGLEAMALKSTIDLTCNDYISVFEFDIFTR
LFQPWGSILRNWNFLAVTHPGYMAFLTYDEVKARLQKYSTKPGSYIFRLSCTRLGQWAI
GYVTGDGNILQTIPHNKPLFQALIDGSREGFYLYPDGRSYNPDLTGLCEPTPHDHIKVTQ
EQYELYCEMGSTFQLCKICAENDKDVKIEPCGHLMCTSCLTAWQESDGQGCPFCRCEIK
GTEPIIVDPFDPRDEGSRCCSIIDPFGMPMLDLDDDDDREESLMMNRLANVRKCTDRQN
SPVTSPGSSPLAQRRKPQPDPLQIPHLSLPPVPPRLDLIQKGIVRSPCGSPTGSPKSSPCMV
RKQDKPLPAPPPPLRDPPPPPPERPPPIPPDNRLSRHIHHVESVPSRDPPMPLEAWCPRDVF
GTNQLVGCRLLGEGSPKPGITASSNVNGRHSRVGSDPVLMRKHRRHDLPLEGAKVFSN
GHLGSEEYDVPPRLSPPPPVTTLLPSIKCTGPLANSLSEKTRDPVEEDDDEYKIPSSHPVSL
NSQPSHCHNVKPPVRSCDNGHCMLNGTHGPSSEKKSNIPDLSIYLKGDVFDSASDPVPLP
PARPPTRDNPKHGSSLNRTPSDYDLLIPPLGEDAFDALPPSLPPPPPPARHSLIEHSKPPGSS
SRPSSGQDLFLLPSDPFVDLASGQVPLPPARRLPGENVKTNRTSQDYDQLPSCSDGSQAP
ARPPKPRPRRTAPEIHHRKPHGPEAALENVDAKIAKLMGEGYAFEEVKRALEIAQNNVE
VARSILREFAFPPPVSPRLNL

By "Cbl proto-oncogene B (CBLB) polynucleotide" is meant a nucleic acid molecule encoding a CBLB polypeptide. The CBLB gene encodes an E3 ubiquitin ligase. An exemplary CBLB

nucleic acid sequence is provided below. Additional exemplary CBLB genomic sequences are indicated in NCBI Reference Sequence: NC_000003.12, or transcript reference NM_001321813.1.

>DQ349203.1 Homo sapiens CBL-B mRNA, complete cds

ATGGCAAACTCAATGAATGGCAGAAACCCTGGTGGTCGAGGAGGAAATCCCCGAAA
AGGTCGAATTTTGGGTATTATTGATGCTATTCAGGATGCAGTTGGACCCCCTAAGCA
AGCTGCCGCAGATCGCAGGACCGTGGAGAAGACTTGGAAGCTCATGGACAAAGTGG
TAAGACTGTGCCAAAATCCCAAACTTCAGTTGAAAAATAGCCCACCATATATACTTG
ATATTTTGCCTGATACATATCAGCATTTACGACTTATATTGAGTAAATATGATGACA
ACCAGAAACTTGCCCAACTCAGTGAGAATGAGTACTTTAAAATCTACATTGATAGCC
TTATGAAAAAGTCAAAACGGGCAATAAGACTCTTTAAAGAAGGCAAGGAGAGAATG
TATGAAGAACAGTCACAGGACAGACGAAATCTCACAAAACTGTCCCTTATCTTCAGT
CACATGCTGGCAGAAATCAAAGCAATCTTTCCCAATGGTCAATTCCAGGGAGATAA
CTTTCGTATCACAAAAGCAGATGCTGCTGAATTCTGGAGAAAGTTTTTTGGAGACAA
AACTATCGTACCATGGAAAGTATTCAGACAGTGCCTTCATGAGGTCCACCAGATTAG
CTCTGGCCTGGAAGCAATGGCTCTAAAATCAACAATTGATTTAACTTGCAATGATTA
CATTTCAGTTTTTGAATTTGATATTTTTACCAGGCTGTTTCAGCCTTGGGGCTCTATTT
TGCGGAATTGGAATTTCTTAGCTGTGACACATCCAGGTTACATGGCATTTCTCACAT
ATGATGAAGTTAAAGCACGACTACAGAAATATAGCACCAAACCCGGAAGCTATATT
TTCCGGTTAAGTTGCACTCGATTGGGACAGTGGGCCATTGGCTATGTGACTGGGGAT
GGGAATATCTTACAGACCATACCTCATAACAAGCCCTTATTTCAAGCCCTGATTGAT
GGCAGCAGGGAAGGATTTTATCTTTATCCTGATGGGAGGAGTTATAATCCTGATTTA
ACTGGATTATGTGAACCTACACCTCATGACCATATAAAAGTTACACAGGAACAATAT
GAATTATATTGTGAAATGGGCTCCACTTTTCAGCTCTGTAAGATTTGTGCAGAGAAT
GACAAAGATGTCAAGATTGAGCCTTGTGGGCATTTGATGTGCACCTCTTGCCTTACG
GCATGGCAGGAGTCGGATGGTCAGGGCTGCCCTTTCTGTCGTTGTGAAATAAAAGG
AACTGAGCCCATAATCGTGGACCCCTTTGATCCAAGAGATGAAGGCTCCAGGTGTTG
CAGCATCATTGACCCCTTTGGCATGCCGATGCTAGACTTGGACGACGATGATGATCG
TGAGGAGTCCTTGATGATGAATCGGTTGGCAAACGTCCGAAAGTGCACTGACAGGC
AGAACTCACCAGTCACATCACCAGGATCCTCTCCCCTTGCCCAGAGAAGAAAGCCA
CAGCCTGACCCACTCCAGATCCCACATCTAAGCCTGCCACCCGTGCCTCCTCGCCTG

GATCTAATTCAGAAAGGCATAGTTAGATCTCCCTGTGGCAGCCCAACGGGTTCACCA
AAGTCTTCTCCTTGCATGGTGAGAAAACAAGATAAACCACTCCCAGCACCACCTCCT
CCCTTAAGAGATCCTCCTCCACCGCCACCTGAAAGACCTCCACCAATCCCACCAGAC
AATAGACTGAGTAGACACATCCATCATGTGGAAAGCGTGCCTTCCAGAGACCCGCC
AATGCCTCTTGAAGCATGGTGCCCTCGGGATGTGTTTGGGACTAATCAGCTTGTGGG
ATGTCGACTCCTAGGGGAGGGCTCTCCAAAACCTGGAATCACAGCGAGTTCAAATG
TCAATGGAAGGCACAGTAGAGTGGGCTCTGACCCAGTGCTTATGCGGAAACACAGA
CGCCATGATTTGCCTTTAGAAGGAGCTAAGGTCTTTTCCAATGGTCACCTTGGAAGT
GAAGAATATGATGTTCCTCCCCGGCTTTCTCCTCCTCCTCCAGTTACCACCCTCCTCC
CTAGCATAAAGTGTACTGGTCCGTTAGCAAATTCTCTTTCAGAGAAAACAAGAGACC
CAGTAGAGGAAGATGATGATGAATACAAGATTCCTTCATCCCACCCTGTTTCCCTGA
ATTCACAACCATCTCATTGTCATAATGTAAAACCTCCTGTTCGGTCTTGTGATAATGG
TCACTGTATGCTGAATGGAACACATGGTCCATCTTCAGAGAAGAAATCAAACATCCC
TGACTTAAGCATATATTTAAAGGGAGATGTTTTTGATTCAGCCTCTGATCCCGTGCC
ATTACCACCTGCCAGGCCTCCAACTCGGGACAATCCAAAGCATGGTTCTTCACTCAA
CAGGACGCCCTCTGATTATGATCTTCTCATCCCTCCATTAGGTGAAGATGCTTTTGAT
GCCCTCCCTCCATCTCTCCCACCTCCCCCACCTCCTGCAAGGCATAGTCTCATTGAAC
ATTCAAAACCTCCTGGCTCCAGTAGCCGGCCATCCTCAGGACAGGATCTTTTTCTTCT
TCCTTCAGATCCCTTTGTTGATCTAGCAAGTGGCCAAGTTCCTTTGCCTCCTGCTAGA
AGGTTACCAGGTGAAAATGTCAAAACTAACAGAACATCACAGGACTATGATCAGCT
TCCTTCATGTTCAGATGGTTCACAGGCACCAGCCAGACCCCCTAAACCACGACCGCG
CAGGACTGCACCAGAAATTCACCACAGAAAACCCCATGGGCCTGAGGCGGCATTGG
AAAATGTCGATGCAAAAATTGCAAAACTCATGGGAGAGGGTTATGCCTTTGAAGAG
GTGAAGAGAGCCTTAGAGATAGCCCAGAATAATGTCGAAGTTGCCCGGAGCATCCT
CCGAGAATTTGCCTTCCCTCCTCCAGTATCCCCACGTCTAAATCTATAG

[00344] By “chimeric antigen receptor” is meant a synthetic receptor comprising an extracellular antigen binding domain, a transmembrane domain, and an intracellular signaling domain that confers specificity for an antigen onto an immune cell.

[00345] In this disclosure, “comprises,” “comprising,” “containing” and “having” and the like can have the meaning ascribed to them in U.S. Patent law and can mean “ includes,” “including,” and the like; “consisting essentially of” or “consists essentially” likewise has the meaning

ascribed in U.S. Patent law and the term is open-ended, allowing for the presence of more than that which is recited so long as basic or novel characteristics of that which is recited is not changed by the presence of more than that which is recited, but excludes prior art embodiments.

[00346] By "cluster of differentiation 2 (CD2)" is meant a protein having at least about 85% amino acid sequence identity to NCBI Accession No. **NP_001315538.1** or fragment thereof and having immunomodulatory activity. An exemplary amino acid sequence is provided below.

>NP_001315538.1 T-cell surface antigen CD2 isoform 1 precursor [Homo sapiens]
MSFPCKFVASFLLIFNVSSKGAVSKEITNALETWGALGQDINLDIPSFQMSDDIDDIKWE
KTSDKKKIAQFRKEKETFKEKDTYKLFKNGTLKIKHLKTDDQDIYKVSIYDTKGKNVLE
KIFDLKIQERVSKPKISWTCINTTLTCEVMNGTDPELNLYQDGKHLKLSQRVITHKWTTS
LSAKFKCTAGNKVSKESSVEPVSCPGGSILGQSNGLSAWTPPSHPTSLPFAEKGLDIYLII
GICGGGSLLMVFVALLVFYITKRKKQRSRRNDEELETRAHRVATEERGRKPHQIPASTPQ
NPATSQHPPPPPGHRSQAPSHRPPPPGHRVQHQPQKRPPAPSGTQVHQQKGPPLPRPRVQ
PKPPHGAAENSLSPSSN

[00347] By "cluster of differentiation 2 (CD2)" is meant a nucleic acid encoding a CD2 polypeptide. An exemplary CD2 nucleic acid sequence is provided below. >NM_001328609.2 Homo sapiens CD2 molecule (CD2), transcript variant 1, mRNA
AGTCTCACTTCAGTTCCTTTTGCATGAAGAGCTCAGAATCAAAAGAGGAAACCAACC
CCTAAGATGAGCTTTCCATGTAAATTTGTAGCCAGCTTCCTTCTGATTTTCAATGTTT
CTTCCAAAGGTGCAGTCTCCAAAGAGATTACGAATGCCTTGGAAACCTGGGGTGCCT
TGGGTCAGGACATCAACTTGGACATTCCTAGTTTTCAAATGAGTGATGATATTGACG
ATATAAAATGGGAAAAAACTTCAGACAAGAAAAAGATTGCACAATTCAGAAAAGA
GAAAGAGACTTTCAAGGAAAAAGATACATATAAGCTATTTAAAAATGGAACTCTGA
AAATTAAGCATCTGAAGACCGATGATCAGGATATCTACAAGGTATCAATATATGAT
ACAAAAGGAAAAAATGTGTTGGAAAAAATATTTGATTTGAAGATTCAAGAGAGGGT
CTCAAAACCAAAGATCTCCTGGACTTGTATCAACACAACCCTGACCTGTGAGGTAAT
GAATGGAACTGACCCCGAATTAAACCTGTATCAAGATGGGAAACATCTAAAACTTT
CTCAGAGGGTCATCACACACAAGTGGACCACCAGCCTGAGTGCAAAATTCAAGTGC
ACAGCAGGGAACAAAGTCAGCAAGGAATCCAGTGTCGAGCCTGTCAGCTGTCCAGG
AGGCAGCATCCTTGGCCAGAGTAATGGGCTCTCTGCCTGGACCCCTCCCAGCCATCC
CACTTCTCTTCCTTTTGCAGAGAAAGGTCTGGACATCTATCTCATCATTGGCATATGT

GGAGGAGGCAGCCTCTTGATGGTCTTTGTGGCACTGCTCGTTTTCTATATCACCAAA
AGGAAAAAACAGAGGAGTCGGAGAAATGATGAGGAGCTGGAGACAAGAGCCCACA
GAGTAGCTACTGAAGAAAGGGGCCGGAAGCCCCACCAAATTCCAGCTTCAACCCCT
CAGAATCCAGCAACTTCCCAACATCCTCCTCCACCACCTGGTCATCGTTCCCAGGCA
CCTAGTCATCGTCCCCCGCCTCCTGGACACCGTGTTCAGCACCAGCCTCAGAAGAGG
CCTCCTGCTCCGTCGGGCACACAAGTTCACCAGCAGAAAGGCCCGCCCCTCCCCAGA
CCTCGAGTTCAGCCAAAACCTCCCCATGGGGCAGCAGAAAACTCATTGTCCCCTTCC
TCTAATTAAAAAAGATAGAAACTGTCTTTTTCAATAAAAAGCACTGTGGATTTCTGC
CCTCCTGATGTGCATATCCGTACTTCCATGAGGTGTTTTCTGTGTGCAGAACATTGTC
ACCTCCTGAGGCTGTGGGCCACAGCCACCTCTGCATCTTCGAACTCAGCCATGTGGT
CAACATCTGGAGTTTTTGGTCTCCTCAGAGAGCTCCATCACACCAGTAAGGAGAAGC
AATATAAGTGTGATTGCAAGAATGGTAGAGGACCGAGCACAGAAATCTTAGAGATT
TCTTGTCCCCTCTCAGGTCATGTGTAGATGCGATAAATCAAGTGATTGGTGTGCCTG
GGTCTCACTACAAGCAGCCTATCTGCTTAAGAGACTCTGGAGTTTCTTATGTGCCCT
GGTGGACACTTGCCCACCATCCTGTGAGTAAAAGTGAAATAAAAGCTTTGACTAGA

[00348] By "cluster of differentiation 3 epsilon (CD3e or CD3 epsilon)" is meant a protein having at least about 85% amino acid sequence identity to NCBI Accession No. **NP_000724.1** or fragment thereof and having immunomodulatory activity. An exemplary amino acid sequence is provided below.

>NP_000724.1 T-cell surface glycoprotein CD3 epsilon chain precursor [Homo sapiens]
MQSGTHWRVLGLCLLSVGVWGQDGNEEMGGITQTPYKVSISGTTVILTCPQYPGSEILW
QHNDKNIGGDEDDKNIGSDEDHLSLKEFSELEQSGYYVCYPRGSKPEDANFYLYLRARV
CENCMEMDVMSVATIVIVDICITGGLLLLVYYWSKNRKAKAKPVTRGAGAGGRQRGQ
NKERPPPVPNPDYEPIRKGQRDLYSGLNQRRI

[00349] By "cluster of differentiation 3 epsilon (CD3e or CD3 epsilon)" is meant a nucleic acid encoding a CD3e polypeptide. An exemplary CD3e nucleic acid sequence is provided below.

>NM_000733.4 Homo sapiens CD3e molecule (CD3E), mRNA
AGAAACCCTCCTCCCCTCCCAGCCTCAGGTGCCTGCTTCAGAAAATGAAGTAGTAAG
TCTGCTGGCCTCCGCCATCTTAGTAAAGTAACAGTCCCATGAAACAAAGATGCAGTC
GGGCACTCACTGGAGAGTTCTGGGCCTCTGCCTCTTATCAGTTGGCGTTTGGGGGCA

AGATGGTAATGAAGAAATGGGTGGTATTACACAGACACCATATAAAGTCTCCATCT
CTGGAACCACAGTAATATTGACATGCCCTCAGTATCCTGGATCTGAAATACTATGGC
AACACAATGATAAAAACATAGGCGGTGATGAGGATGATAAAAACATAGGCAGTGAT
GAGGATCACCTGTCACTGAAGGAATTTTCAGAATTGGAGCAAAGTGGTTATTATGTC
TGCTACCCCAGAGGAAGCAAACCAGAAGATGCGAACTTTTATCTCTACCTGAGGGC
AAGAGTGTGTGAGAACTGCATGGAGATGGATGTGATGTCGGTGGCCACAATTGTCA
TAGTGGACATCTGCATCACTGGGGGCTTGCTGCTGCTGGTTTACTACTGGAGCAAGA
ATAGAAAGGCCAAGGCCAAGCCTGTGACACGAGGAGCGGGTGCTGGCGGCAGGCA
AAGGGGACAAAACAAGGAGAGGCCACCACCTGTTCCCAACCCAGACTATGAGCCCA
TCCGGAAAGGCCAGCGGGACCTGTATTCTGGCCTGAATCAGAGACGCATCTGACCC
TCTGGAGAACACTGCCTCCCGCTGGCCCAGGTCTCCTCTCCAGTCCCCCTGCGACTC
CCTGTTTCCTGGGCTAGTCTTGGACCCCACGAGAGAGAATCGTTCCTCAGCCTCATG
GTGAACTCGCGCCCTCCAGCCTGATCCCCCGCTCCCTCCTCCCTGCCTTCTCTGCTGG
TACCCAGTCCTAAAATATTGCTGCTTCCTCTTCCTTTGAAGCATCATCAGTAGTCACA
CCCTCACAGCTGGCCTGCCCTCTTGCCAGGATATTTATTTGTGCTATTCACTCCCTTC
CCTTTGGATGTAACTTCTCCGTTCAGTTCCCTCCTTTTCTTGCATGTAAGTTGTCCCCC
ATCCCAAAGTATTCCATCTACTTTTCTATCGCCGTCCCCTTTTGCAGCCCTCTCTGGG
GATGGACTGGGTAAATGTTGACAGAGGCCCTGCCCCGTTCACAGATCCTGGCCCTGA
GCCAGCCCTGTGCTCCTCCCTCCCCCAACACTCCCTACCAACCCCCTAATCCCCTACT
CCCTCCACCCCCCCTCCACTGTAGGCCACTGGATGGTCATTTGCATCTCCGTAAATGT
GCTCTGCTCCTCAGCTGAGAGAGAAAAAAATAAACTGTATTTGGCTGCAA

[00350] By "cluster of differentiation 3 gamma (CD3g or CD3 gamma) is meant a protein having at least about 85% amino acid sequence identity to NCBI Accession No. **NP_000064.1** or fragment thereof and having immunomodulatory activity. An exemplary amino acid sequence is provided below.

>NP_000064.1 T-cell surface glycoprotein CD3 gamma chain precursor [Homo sapiens]
MEQGKGLAVLILAIILLQGTLAQSIKGNHLVKVYDYQEDGSVLLTCDAEAKNITWFKDG
KMIGFLTEDKKKWNLGSNAKDPRGMYQCKGSQNKSKPLQVYYRMCQNCIELNAATIS
GFLFAEIVSIFVLAVGVYFIAGQDGVRQSRASDKQTLLPNDQLYQPLKDREDDQYSHLQ
GNQLRRN

[00351] By "cluster of differentiation 3 gamma (CD3g or CD3 gamma)" is meant a nucleic acid encoding a CD3g polypeptide. An exemplary CD3g nucleic acid sequence is provided below.

>NM_000073.3 Homo sapiens CD3g molecule (CD3G), mRNA

AGTCTAGCTGCTGCACAGGCTGGCTGGCTGGCTGGCTGCTAAGGGCTGCTCCACGCT
TTTGCCGGAGGACAGAGACTGACATGGAACAGGGGAAGGGCCTGGCTGTCCTCATC
CTGGCTATCATTCTTCTTCAAGGTACTTTGGCCCAGTCAATCAAAGGAAACCACTTG
GTTAAGGTGTATGACTATCAAGAAGATGGTTCGGTACTTCTGACTTGTGATGCAGAA
GCCAAAAATATCACATGGTTTAAAGATGGGAAGATGATCGGCTTCCTAACTGAAGA
TAAAAAAAAAATGGAATCTGGGAAGTAATGCCAAGGACCCTCGAGGGATGTATCAGT
GTAAAGGATCACAGAACAAGTCAAAACCACTCCAAGTGTATTACAGAATGTGTCAG
AACTGCATTGAACTAAATGCAGCCACCATATCTGGCTTTCTCTTTGCTGAAATCGTC
AGCATTTTCGTCCTTGCTGTTGGGGTCTACTTCATTGCTGGACAGGATGGAGTTCGCC
AGTCGAGAGCTTCAGACAAGCAGACTCTGTTGCCCAATGACCAGCTCTACCAGCCCC
TCAAGGATCGAGAAGATGACCAGTACAGCCACCTTCAAGGAAACCAGTTGAGGAGG
AATTGAACTCAGGACTCAGAGTAGTCCAGGTGTTCTCCTCCTATTCAGTTCCCAGAA
TCAAAGCAATGCATTTTGGAAAGCTCCTAGCAGAGAGACTTTCAGCCCTAAATCTAG
ACTCAAGGTTCCCAGAGATGACAAATGGAGAAGAAAGGCCATCAGAGCAAATTTGG
GGGTTTCTCAAATAAAATAAAAATAAAAACAAATACTGTGTTTCAGAAGCGCCACC
TATTGGGGAAAATTGTAAAAGAAAAATGAAAAGATCAAATAACCCCCTGGATTTGA
ATATAATTTTTTGTGTTGTAATTTTTATTTCGTTTTTGTATAGGTTATAATTCACATGG
CTCAAATATTCAGTGAAAGCTCTCCCTCCACCGCCATCCCCTGCTACCCAGTGACCC
TGTTGCCCTCTTCAGAGACAAATTAGTTTCTCTTTTTTTTTTTTTTTTTTTTTTTGA
GACAGTCTGGCTCTGTCACCCAGGCTGAAATGCAGTGGCACCATCTCGGCTCACTGC
AACCTCTGCCTCCTGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCCGGGCAGCTGGG
ATTACAGGCACACACTACCACACCTGGCTAATTTTTGTATTTTTAGTAGAGACAGGG
TTTTGCTCTGTTGGCCAAGCTGGTCTCGAACTCCTGACCTCAAGTGATCCGCCCGCCT
CAGCCTCCCAAAGTGCTGGGATTACAGGTGTGAGCCACCATGCCTGGTCTTAAAACC
AGTTTCTTATATATCTCTCTGGAGGTATTCTAGGCATATATGAGCACATTCTCAAGTA
CATATTATCCTCCCTTCCCCTATCTTTTAGACAAATGATATCAAACTATACATCTTGT
GAGATTATTGCATACCATTATATGAAGATACCATTATATCCTTTTTAATGCAACCATA

TTGTACAAATAGACTATGATTTATTTAACCTGTTATCTATCAGTGGATATTTAAGTTG
GTAGTTGGTTCCAATCTTTTGCTCTTACAACAATTCTGCAATGACTAACATTGTATAA
ATATCATTTTTAAAAATAATTGCATTGAAGCATAATGTACATGCCATAAAATCCACC
CATCTTAAGTGATTTCACCTGTTCTCAGAAATTTTTAGTAAATTTAACTAATTGTACA
GCCATTACCATAATCCAGCTTTAGGACATTTTCTTTTTTTTCTTTTCTTTTCTTTTTTTT
CTTTTTTTTTTTTTTTTGAAGTGGAATCTTGCTCTGTGGCCCAGGCTGGAGTGCAGTG
GCGCGATCTCAGCTCACTGCAACCTCCACCTCCTGGGTTCAAGCGATTCTCTTGCCTT
GGCCTCCCGAGTAGCTGAGACTACAGGCACATGCCACCACGCCCAGCTCATTTTTTG
TGTATTTAGTATTTGTGTATCTAGTATTTGTGTACTTAGTAGAGACAGGGTTTCACCA
TGTTGGCCAGGCTGGTCTCCAATTCCTGACCTCAGGCGATCCACCCGCCTTGACCTC
CCAAAGTGCTGGGATTACAGGTGTGAGCCACCGCGCCAGGCCCGTAACTGTATTTTA
ATATAGCCATTCTATGGATTTAATATGGTATTTTATTATGGCCTTAATTTGCATTTCC
CTAGATACTAACCATGCTGAGTGTCCTGTCTTGTGTTTATTAACCATTCATATATTTT
TAGTGAAATGTGTATCAAATCTTTTGCCCATTTTTAAGTTGACTTATTTGTTTGTCTTC
TTACTATTGGGTTGCATATGTTTTTGATATAAGTCCTTTATCAGATATATGATTTGGA
AATATTTTCTACCAATCTGTGGTTTGTTTTTCTTAATGGTGTCTTTTGAAGTGCAAAA
GGTTTGAATTTTGAAGTACATTTTATTGATTTTTTCTTCTATATATTGTGCTTTTGGTA
TCATGTCTAATAAATCTTTACCAAACCCACAGTTACAAAGATTTTCTCCTGTCTTCTT
TTTATACTTTTTACAGCTTTATGGTTTTAGCTCTAACAATAAATGTGATTTTGAACAT
ACATAAGACTATTTGTAACAAACACAAATAAATTGAATTGTTGGGCA

[00352] By "cluster of differentiation 3 delta (CD3d or CD3 delta) is meant a protein having at least about 85% amino acid sequence identity to NCBI Accession No. **NP_000723.1** or fragment thereof and having immunomodulatory activity. An exemplary amino acid sequence is provided below.

>NP_000723.1 T-cell surface glycoprotein CD3 delta chain isoform A precursor [Homo sapiens]
MEHSTFLSGLVLATLLSQVSPFKIPIEELEDRVFVNCNTSITWVEGTVGTLLSDITRLDLG
KRILDPRGIYRCNGTDIYKDKESTVQVHYRMCQSCVELDPATVAGIIVTDVIATLLLALG
VFCFAGHETGRLSGAADTQALLRNDQVYQPLRDRDDAQYSHLGGNWARNK

[00353] By "cluster of differentiation 3 delta (CD3d or CD3 delta)" is meant a nucleic acid encoding a CD3d polypeptide. An exemplary CD3d nucleic acid sequence is provided below.

>NM_000732.4 Homo sapiens CD3d molecule (CD3D), transcript variant 1, mRNA

AGAGAAGCAGACATCTTCTAGTTCCTCCCCCACTCTCCTCTTTCCGGTACCTGTGAGT
CAGCTAGGGGAGGGCAGCTCTCACCCAGGCTGATAGTTCGGTGACCTGGCTTTATCT
ACTGGATGAGTTCCGCTGGGAGATGGAACATAGCACGTTTCTCTCTGGCCTGGTACT
GGCTACCCTTCTCTCGCAAGTGAGCCCCTTCAAGATACCTATAGAGGAACTTGAGGA
CAGAGTGTTTGTGAATTGCAATACCAGCATCACATGGGTAGAGGGAACGGTGGGAA
CACTGCTCTCAGACATTACAAGACTGGACCTGGGAAAACGCATCCTGGACCCACGA
GGAATATATAGGTGTAATGGGACAGATATATACAAGGACAAAGAATCTACCGTGCA
AGTTCATTATCGAATGTGCCAGAGCTGTGTGGAGCTGGATCCAGCCACCGTGGCTGG
CATCATTGTCACTGATGTCATTGCCACTCTGCTCCTTGCTTTGGGAGTCTTCTGCTTT
GCTGGACATGAGACTGGAAGGCTGTCTGGGGCTGCCGACACACAAGCTCTGTTGAG
GAATGACCAGGTCTATCAGCCCCTCCGAGATCGAGATGATGCTCAGTACAGCCACCT
TGGAGGAAACTGGGCTCGGAACAAGTGAACCTGAGACTGGTGGCTTCTAGAAGCAG
CCATTACCAACTGTACCTTCCCTTCTTGCTCAGCCAATAAATATATCCTCTTTCACTC
AGAAAAAAAAAAAAAAAAAAAAAAAAAAAA

[00354] By "cluster of differentiation 4 (CD4)" is meant a protein having at least about 85% amino acid sequence identity to NCBI Accession No. **NP_000607.1** or fragment thereof and having immunomodulatory activity. An exemplary amino acid sequence is provided below.

>NP_000607.1 T-cell surface glycoprotein CD4 isoform 1 precursor [Homo sapiens]
MNRGVPFRHLLLVLQLALLPAATQGKKVVLGKKGDTVELTCTASQKKSIQFHWKNSNQ
IKILGNQGSFLTKGPSKLNDRADSRRSLWDQGNFPLIIKNLKIEDSDTYICEVEDQKEEVQ
LLVFGLTANSDTHLLQGQSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLSVSQLELQDS
GTWTCTVLQNQKKVEFKIDIVVLAFQKASSIVYKKEGEQVEFSFPLAFTVEKLTGSGEL
WWQAERASSSKSWITFDLKNKEVSVKRVTQDPKLQMGKKLPLHLTLPQALPQYAGSG
NLTLALEAKTGKLHQEVNLVVMRATQLQKNLTCEVWGPTSPKLMLSLKLENKEAKVS
KREKAVWVLNPEAGMWQCLLSDSGQVLLESNIKVLPTWSTPVQPMALIVLGGVAGLLL
FIGLGIFFCVRCRHRRRQAERMSQIKRLLSEKKTCQCPHRFQKTCSPI

[00355] By "cluster of differentiation 4 (CD4)" is meant a nucleic acid encoding a CD4 polypeptide. An exemplary CD4 nucleic acid sequence is provided below.

>NM_000616.5 Homo sapiens CD4 molecule (CD4), transcript variant 1, mRNA
CTCTCTTCATTTAAGCACGACTCTGCAGAAGGAACAAAGCACCCTCCCCACTGGGCT
CCTGGTTGCAGAGCTCCAAGTCCTCACACAGATACGCCTGTTTGAGAAGCAGCGGG

CAAGAAAGACGCAAGCCCAGAGGCCCTGCCATTTCTGTGGGCTCAGGTCCCTACTG
GCTCAGGCCCCTGCCTCCCTCGGCAAGGCCACAATGAACCGGGGAGTCCCTTTTAGG
CACTTGCTTCTGGTGCTGCAACTGGCGCTCCTCCCAGCAGCCACTCAGGGAAAGAAA
GTGGTGCTGGGCAAAAAAGGGGATACAGTGGAACTGACCTGTACAGCTTCCCAGAA
GAAGAGCATACAATTCCACTGGAAAAACTCCAACCAGATAAAGATTCTGGGAAATC
AGGGCTCCTTCTTAACTAAAGGTCCATCCAAGCTGAATGATCGCGCTGACTCAAGAA
GAAGCCTTTGGGACCAAGGAAACTTTCCCCTGATCATCAAGAATCTTAAGATAGAA
GACTCAGATACTTACATCTGTGAAGTGGAGGACCAGAAGGAGGAGGTGCAATTGCT
AGTGTTCGGATTGACTGCCAACTCTGACACCCACCTGCTTCAGGGGCAGAGCCTGAC
CCTGACCTTGGAGAGCCCCCCTGGTAGTAGCCCCTCAGTGCAATGTAGGAGTCCAAG
GGGTAAAAACATACAGGGGGGGAAGACCCTCTCCGTGTCTCAGCTGGAGCTCCAGG
ATAGTGGCACCTGGACATGCACTGTCTTGCAGAACCAGAAGAAGGTGGAGTTCAAA
ATAGACATCGTGGTGCTAGCTTTCCAGAAGGCCTCCAGCATAGTCTATAAGAAAGA
GGGGGAACAGGTGGAGTTCTCCTTCCCACTCGCCTTTACAGTTGAAAAGCTGACGGG
CAGTGGCGAGCTGTGGTGGCAGGCGGAGAGGGCTTCCTCCTCCAAGTCTTGGATCAC
CTTTGACCTGAAGAACAAGGAAGTGTCTGTAAAACGGGTTACCCAGGACCCTAAGC
TCCAGATGGGCAAGAAGCTCCCGCTCCACCTCACCCTGCCCCAGGCCTTGCCTCAGT
ATGCTGGCTCTGGAAACCTCACCCTGGCCCTTGAAGCGAAAACAGGAAAGTTGCAT
CAGGAAGTGAACCTGGTGGTGATGAGAGCCACTCAGCTCCAGAAAAATTTGACCTG
TGAGGTGTGGGGACCCACCTCCCCTAAGCTGATGCTGAGTTTGAAACTGGAGAACA
AGGAGGCAAAGGTCTCGAAGCGGGAGAAGGCGGTGTGGGTGCTGAACCCTGAGGC
GGGGATGTGGCAGTGTCTGCTGAGTGACTCGGGACAGGTCCTGCTGGAATCCAACA
TCAAGGTTCTGCCCACATGGTCCACCCCGGTGCAGCCAATGGCCCTGATTGTGCTGG
GGGGCGTCGCCGGCCTCCTGCTTTTCATTGGGCTAGGCATCTTCTTCTGTGTCAGGTG
CCGGCACCGAAGGCGCCAAGCAGAGCGGATGTCTCAGATCAAGAGACTCCTCAGTG
AGAAGAAGACCTGCCAGTGTCCTCACCGGTTTCAGAAGACATGTAGCCCCATTTGAG
GCACGAGGCCAGGCAGATCCCACTTGCAGCCTCCCCAGGTGTCTGCCCCGCGTTTCC
TGCCTGCGGACCAGATGAATGTAGCAGATCCCCAGCCTCTGGCCTCCTGTTCGCCTC
CTCTACAATTTGCCATTGTTTCTCCTGGGTTAGGCCCCGGCTTCACTGGTTGAGTGTT
GCTCTCTAGTTTCCAGAGGCTTAATCACACCGTCCTCCACGCCATTTCCTTTTCCTTC
AAGCCTAGCCCTTCTCTCATTATTTCTCTCTGACCCTCTCCCCACTGCTCATTTGGAT

CCCAGGGGAGTGTTCAGGGCCAGCCCTGGCTGGCATGGAGGGTGAGGCTGGGTGTC
TGGAAGCATGGAGCATGGGACTGTTCTTTTACAAGACAGGACCCTGGGACCACAGA
GGGCAGGAACTTGCACAAAATCACACAGCCAAGCCAGTCAAGGATGGATGCAGATC
CAGAGGTTTCTGGCAGCCAGTACCTCCTGCCCCATGCTGCCCGCTTCTCACCCTATGT
GGGTGGGACCACAGACTCACATCCTGACCTTGCACAAACAGCCCCTCTGGACACAG
CCCCATGTACACGGCCTCAAGGGATGTCTCACATCCTCTGTCTATTTGAGACTTAGA
AAAATCCTACAAGGCTGGCAGTGACAGAACTAAGATGATCATCTCCAGTTTATAGA
CCAGAACCAGAGCTCAGAGAGGCTAGATGATTGATTACCAAGTGCCGGACTAGCAA
GTGCTGGAGTCGGGACTAACCCAGGTCCCTTGTCCCAAGTTCCACTGCTGCCTCTTG
AATGCAGGGACAAATGCCACACGGCTCTCACCAGTGGCTAGTGGTGGGTACTCAAT
GTGTACTTTTGGGTTCACAGAAGCACAGCACCCATGGGAAGGGTCCATCTCAGAGA
ATTTACGAGCAGGGATGAAGGCCTCCCTGTCTAAAATCCCTCCTTCATCCCCCGCTG
GTGGCAGAATCTGTTACCAGAGGACAAAGCCTTTGGCTCTTCTAATCAGAGCGCAAG
CTGGGAGCACAGGCACTGCAGGAGAGAATGCCCAGTGACCAGTCACTGACCCTGTG
CAGAACCTCCTGGAAGCGAGCTTTGCTGGGAGAGGGGGTAGCTAGCCTGAGAGGGA
ACCCTCTAAGGGACCTCAAAGGTGATTGTGCCAGGCTCTGCGCCTGCCCCACACCCT
CCCTTACCCTCCTCCAGACCATTCAGGACACAGGGAAATCAGGGTTACAAATCTTCT
TGATCCACTTCTCTCAGGATCCCCTCTCTTCCTACCCTTCCTCACCACTTCCCTCAGTC
CCAACTCCTTTTCCCTATTTCCTTCTCCTCCTGTCTTTAAAGCCTGCCTCTTCCAGGAA
GACCCCCCTATTGCTGCTGGGGCTCCCCATTTGCTTACTTTGCATTTGTGCCCACTCT
CCACCCCTGCTCCCCTGAGCTGAAATAAAAATACAATAAACTTAC

[00356] By "cluster of differentiation 5 (CD5)" is meant a protein having at least about 85% amino acid sequence identity to NCBI Accession No. **NP_001333385.1** or fragment thereof and having immunomodulatory activity. An exemplary amino acid sequence is provided below.

>NP_001333385.1 T-cell surface glycoprotein CD5 isoform 2 [Homo sapiens]
MVCSQSWGRSSKQWEDPSQASKVCQRLNCGVPLSLGPFLVTYTPQSSIICYGQLGSFSN
CSHSRNDMCHSLGLTCLEPQKTTPPTTRPPPTTTPEPTAPPRLQLVAQSGGQHCAGVVEF
YSGSLGGTISYEAQDKTQDLENFLCNNLQCGSFLKHLPETEAGRAQDPGEPREHQPLPIQ
WKIQNSSCTSLEHCFRKIKPQKSGRVLALLCSGFQPKVQSRLVGGSSICEGTVEVRQGAQ
WAALCDSSSARSSLRWEEVCREQQCGSVNSYRVLDAGDPTSRGLFCPHQKLSQCHELW
ERNSYCKKVFVTCQDPNPAGLAAGTVASIILALVLLVVLLVVCGPLAYKKLVKKFRQK

KQRQWIGPTGMNQNMSFHRNHTATVRSHAENPTASHVDNEYSQPPRNSHLSAYPALEG
ALHRSSMQPDNSSDSDYDLHGAQRL

[00357] By "cluster of differentiation 5 (CD5)" is meant a nucleic acid encoding a CD5 polypeptide. An exemplary CD5 nucleic acid sequence is provided below. >NM_001346456.1 Homo sapiens CD5 molecule (CD5), transcript variant 2, mRNA

GAGTCTTGCTGATGCTCCCGGCTGAATAAACCCCTTCCTTCTTTAACTTGGTGTCTGA
GGGGTTTTGTCTGTGGCTTGTCCTGCTACATTTCTTGGTTCCCTGACCAGGAAGCAAA
GTGATTAACGGACAGTTGAGGCAGCCCCTTAGGCAGCTTAGGCCTGCCTTGTGGAGC
ATCCCCGCGGGGAACTCTGGCCAGCTTGAGCGACACGGATCCTCAGAGCGCTCCCA
GGTAGGCAATTGCCCCAGTGGAATGCCTCGTCAGAGCAGTGCATGGCAGGCCCCTG
TGGAGGATCAACGCAGTGGCTGAACACAGGGAAGGAACTGGCACTTGGAGTCCGGA
CAACTGAAACTTGTCGCTTCCTGCCTCGGACGGCTCAGCTGGTATGACCCAGATTTC
CAGGCAAGGCTCACCCGTTCCAACTCGAAGTGCCAGGGCCAGCTGGAGGTCTACCT
CAAGGACGGATGGCACATGGTTTGCAGCCAGAGCTGGGGCCGGAGCTCCAAGCAGT
GGGAGGACCCCAGTCAAGCGTCAAAAGTCTGCCAGCGGCTGAACTGTGGGGTGCCC
TTAAGCCTTGGCCCCTTCCTTGTCACCTACACACCTCAGAGCTCAATCATCTGCTACG
GACAACTGGGCTCCTTCTCCAACTGCAGCCACAGCAGAAATGACATGTGTCACTCTC
TGGGCCTGACCTGCTTAGAACCCCAGAAGACAACACCTCCAACGACAAGGCCCCCG
CCCACCACAACTCCAGAGCCCACAGCTCCTCCCAGGCTGCAGCTGGTGGCACAGTCT
GGCGGCCAGCACTGTGCCGGCGTGGTGGAGTTCTACAGCGGCAGCCTGGGGGGTAC
CATCAGCTATGAGGCCCAGGACAAGACCCAGGACCTGGAGAACTTCCTCTGCAACA
ACCTCCAGTGTGGCTCCTTCTTGAAGCATCTGCCAGAGACTGAGGCAGGCAGAGCCC
AAGACCCAGGGGAGCCACGGGAACACCAGCCCTTGCCAATCCAATGGAAGATCCAG
AACTCAAGCTGTACCTCCCTGGAGCATTGCTTCAGGAAAATCAAGCCCCAGAAAAG
TGGCCGAGTTCTTGCCCTCCTTTGCTCAGGTTTCCAGCCCAAGGTGCAGAGCCGTCT
GGTGGGGGGCAGCAGCATCTGTGAAGGCACCGTGGAGGTGCGCCAGGGGGCTCAGT
GGGCAGCCCTGTGTGACAGCTCTTCAGCCAGGAGCTCGCTGCGGTGGGAGGAGGTG
TGCCGGGAGCAGCAGTGTGGCAGCGTCAACTCCTATCGAGTGCTGGACGCTGGTGA
CCCAACATCCCGGGGGCTCTTCTGTCCCCATCAGAAGCTGTCCCAGTGCCACGAACT
TTGGGAGAGAAATTCCTACTGCAAGAAGGTGTTTGTCACATGCCAGGATCCAAACCC
CGCAGGCCTGGCCGCAGGCACGGTGGCAAGCATCATCCTGGCCCTGGTGCTCCTGGT

GGTGCTGCTGGTCGTGTGCGGCCCCCTTGCCTACAAGAAGCTAGTGAAGAAATTCCG
CCAGAAGAAGCAGCGCCAGTGGATTGGCCCAACGGGAATGAACCAAAACATGTCTT
TCCATCGCAACCACACGGCAACCGTCCGATCCCATGCTGAGAACCCCACAGCCTCCC
ACGTGGATAACGAATACAGCCAACCTCCCAGGAACTCCCACCTGTCAGCTTATCCAG
CTCTGGAAGGGGCTCTGCATCGCTCCTCCATGCAGCCTGACAACTCCTCCGACAGTG
ACTATGATCTGCATGGGGCTCAGAGGCTGTAAAGAACTGGGATCCATGAGCAAAAA
GCCGAGAGCCAGACCTGTTTGTCCTGAGAAAACTGTCCGCTCTTCACTTGAAATCAT
GTCCCTATTTCTACCCCGGCCAGAACATGGACAGAGGCCAGAAGCCTTCCGGACAG
GCGCTGCTGCCCCGAGTGGCAGGCCAGCTCACACTCTGCTGCACAACAGCTCGGCC
GCCCCTCCACTTGTGGAAGCTGTGGTGGGCAGAGCCCCAAAACAAGCAGCCTTCCA
ACTAGAGACTCGGGGGTGTCTGAAGGGGGCCCCCTTTCCCTGCCCGCTGGGGAGCG
GCGTCTCAGTGAAATCGGCTTTCTCCTCAGACTCTGTCCCTGGTAAGGAGTGACAAG
GAAGCTCACAGCTGGGCGAGTGCATTTTGAATAGTTTTTTGTAAGTAGTGCTTTTCCT
CCTTCCTGACAAATCGAGCGCTTTGGCCTCTTCTGTGCAGCATCCACCCCTGCGGAT
CCCTCTGGGGAGGACAGGAAGGGGACTCCCGGAGACCTCTGCAGCCGTGGTGGTCA
GAGGCTGCTCACCTGAGCACAAAGACAGCTCTGCACATTCACCGCAGCTGCCAGCC
AGGGGTCTGGGTGGGCACCACCCTGACCCACAGCGTCACCCCACTCCCTCTGTCTTA
TGACTCCCCTCCCCAACCCCCTCATCTAAAGACACCTTCCTTTCCACTGGCTGTCAAG
CCCACAGGGCACCAGTGCCACCCAGGGCCCGGCACAAAGGGGCGCCTAGTAAACCT
TAACCAACTTGGTTTTTTGCTTCACCCAGCAATTAAAAGTCCCAAGCTGAGGTAGTT
TCAGTCCATCACAGTTCATCTTCTAACCCAAGAGTCAGAGATGGGGCTGGTCATGTT
CCTTTGGTTTGAATAACTCCCTTGACGAAAACAGACTCCTCTAGTACTTGGAGATCTT
GGACGTACACCTAATCCCATGGGGCCTCGGCTTCCTTAACTGCAAGTGAGAAGAGG
AGGTCTACCCAGGAGCCTCGGGTCTGATCAAGGGAGAGGCCAGGCGCAGCTCACTG
CGGCGGCTCCCTAAGAAGGTGAAGCAACATGGGAACACATCCTAAGACAGGTCCTT
TCTCCACGCCATTTGATGCTGTATCTCCTGGGAGCACAGGCATCAATGGTCCAAGCC
GCATAATAAGTCTGGAAGAGCAAAAGGGAGTTACTAGGATATGGGGTGGGCTGCTC
CCAGAATCTGCTCAGCTTTCTGCCCCCACCAACACCCTCCAACCAGGCCTTGCCTTCT
GAGAGCCCCCGTGGCCAAGCCCAGGTCACAGATCTTCCCCCGACCATGCTGGGAAT
CCAGAAACAGGGACCCCATTTGTCTTCCCATATCTGGTGGAGGTGAGGGGCTCCTC
AAAAGGGAACTGAGAGGCTGCTCTTAGGGAGGGCAAAGGTTCGGGGGCAGCCAGT

GTCTCCCATCAGTGCCTTTTTTAATAAAAGCTCTTTCATCTATAGTTTGGCCACCATA
CAGTGGCCTCAAAGCAACCATGGCCTACTTAAAAACCAAACCAAAAATAAAGAGTT
TAGTTGAGGAGAAAAAAAAAAAAAAAAAAAAAAA

[00358] By "cluster of differentiation 7 (CD7)" is meant a protein having at least about 85% amino acid sequence identity to NCBI Accession No. **NP_006128.1** or fragment thereof and having immunomodulatory activity. An exemplary amino acid sequence is provided below.

>NP_006128.1 T-cell antigen CD7 precursor [Homo sapiens]
MAGPPRLLLLPLLLALARGLPGALAAQEVQQSPHCTTVPVGASVNITCSTSGGLRGIYLR
QLGPQPQDIIYYEDGVVPTTDRRFRGRIDFSGSQDNLTITMHRLQLSDTGTYTCQAITEV
NVYGSGTLVLVTEEQSQGWHRCSDAPPRASALPAPPTGSALPDPQTASALPDPPAASAL
PAALAVISFLLGLGLGVACVLARTQIKKLCSWRDKNSAACVVYEDMSHSRCNTLSSPNQ
YQ

[00359] By "cluster of differentiation 7 (CD7)" is meant a nucleic acid encoding a CD7 polypeptide. An exemplary CD7 nucleic acid sequence is provided below.

>NM_006137.7 Homo sapiens CD7 molecule (CD7), mRNA
CTCTCTGAGCTCTGAGCGCCTGCGGTCTCCTGTGTGCTGCTCTCTGTGGGGTCCTGTA
GACCCAGAGAGGCTCAGCTGCACTCGCCCGGCTGGGAGAGCTGGGTGTGGGGAACA
TGGCCGGGCCTCCGAGGCTCCTGCTGCTGCCCCTGCTTCTGGCGCTGGCTCGCGGCC
TGCCTGGGGCCCTGGCTGCCCAAGAGGTGCAGCAGTCTCCCCACTGCACGACTGTCC
CCGTGGGAGCCTCCGTCAACATCACCTGCTCCACCAGCGGGGGCCTGCGTGGGATCT
ACCTGAGGCAGCTCGGGCCACAGCCCCAAGACATCATTTACTACGAGGACGGGGTG
GTGCCCACTACGGACAGACGGTTCCGGGGCCGCATCGACTTCTCAGGGTCCCAGGA
CAACCTGACTATCACCATGCACCGCCTGCAGCTGTCGGACACTGGCACCTACACCTG
CCAGGCCATCACGGAGGTCAATGTCTACGGCTCCGGCACCCTGGTCCTGGTGACAG
AGGAACAGTCCCAAGGATGGCACAGATGCTCGGACGCCCCACCAAGGGCCTCTGCC
CTCCCTGCCCCACCGACAGGCTCCGCCCTCCCTGACCCGCAGACAGCCTCTGCCCTC
CCTGACCCGCCAGCAGCCTCTGCCCTCCCTGCGGCCCTGGCGGTGATCTCCTTCCTCC
TCGGGCTGGGCCTGGGGGTGGCGTGTGTGCTGGCGAGGACACAGATAAAGAAACTG
TGCTCGTGGCGGGATAAGAATTCGGCGGCATGTGTGGTGTACGAGGACATGTCGCA
CAGCCGCTGCAACACGCTGTCCTCCCCCAACCAGTACCAGTGACCCAGTGGGCCCCT
GCACGTCCCGCCTGTGGTCCCCCCAGCACCTTCCCTGCCCCACCATGCCCCCCACCC

TGCCACACCCCTCACCCTGCTGTCCTCCCACGGCTGCAGCAGAGTTTGAAGGGCCCA
GCCGTGCCCAGCTCCAAGCAGACACACAGGCAGTGGCCAGGCCCCACGGTGCTTCT
CAGTGGACAATGATGCCTCCTCCGGGAAGCCTTCCCTGCCCAGCCCACGCCGCCACC
GGGAGGAAGCCTGACTGTCCTTTGGCTGCATCTCCCGACCATGGCCAAGGAGGGCTT
TTCTGTGGGATGGGCCTGGGCACGCGGCCCTCTCCTGTCAGTGCCGGCCCACCCACC
AGCAGGCCCCCAACCCCCAGGCAGCCCGGCAGAGGACGGGAGGAGACCAGTCCCC
CACCCAGCCGTACCAGAAATAAAGGCTTCTGTGCTTCC

[00360] By "cluster of differentiation 30 (CD30)" is meant a protein having at least about 85% amino acid sequence identity to NCBI Accession No. **NP_001234.3** or fragment thereof and having immunomodulatory activity. An exemplary amino acid sequence is provided below.

>NP_001234.3 tumor necrosis factor receptor superfamily member 8 isoform 1 precursor [Homo sapiens]

MRVLLAALGLLFLGALRAFPQDRPFEDTCHGNPSHYYDKAVRRCCYRCPMGLFPTQQC
PQRPTDCRKQCEPDYYLDEADRCTACVTCSRDDLVEKTPCAWNSSRVCECRPGMFCST
SAVNSCARCFFHSVCPAGMIVKFPGTAQKNTVCEPASPGVSPACASPENCKEPSSGTIPQ
AKPTPVSPATSSASTMPVRGGTRLAQEAASKLTRAPDSPSSVGRPSSDPGLSPTQPCPEGS
GDCRKQCEPDYYLDEAGRCTACVSCSRDDLVEKTPCAWNSSRTCECRPGMICATSATN
SCARCVPYPICAAETVTKPQDMAEKDTTFEAPPLGTQPDCNPTPENGEAPASTSPTQSLL
VDSQASKTLPIPTSAPVALSSTGKPVLDAGPVLFWVILVLVVVVGSSAFLLCHRRACRKR
IRQKLHLCYPVQTSQPKLELVDSRPRRSSTQLRSGASVTEPVAEERGLMSQPLMETCHSV
GAAYLESLPLQDASPAGGPSSPRDLPEPRVSTEHTNNKIEKIYIMKADTVIVGTVKAELPE
GRGLAGPAEPELEEELEADHTPHYPEQETEPPLGSCSDVMLSVEEEGKEDPLPTAASGK

[00361] By "cluster of differentiation 30 (CD30)" is meant a nucleic acid encoding a CD30 polypeptide. An exemplary CD30 nucleic acid sequence is provided below. >NM_001243.5 Homo sapiens TNF receptor superfamily member 8 (TNFRSF8), transcript variant 1, mRNA

CTGAGTCATCTCTGCACGTGTTTGCCCCCTTTTTTCTTCGCTGCTTGTAGCTAAGTGTT
CCTGGAACCAATTTGATACGGGAGAACTAAGGCTGAAACCTCGGAGGAACAACCAC
TTTTGAAGTGACTTCGCGGCGTGCGTTGGGTGCGGACTAGGTGGCCGCGGCGGGAGT
GTGCTGGAGCCTGAAGTCCACGCGCGCGGCTGAGAACCGCCGGGACCGCACGTGGG
CGCCGCGCGCTTCCCCCGCTTCCCAGGTGGGCGCCGGCCGCCAGGCCACCTCACGTC
CGGCCCCGGGGATGCGCGTCCTCCTCGCCGCGCTGGGACTGCTGTTCCTGGGGGCGC

TACGAGCCTTCCCACAGGATCGACCCTTCGAGGACACCTGTCATGGAAACCCCAGCC
ACTACTATGACAAGGCTGTCAGGAGGTGCTGTTACCGCTGCCCCATGGGGCTGTTCC
CGACACAGCAGTGCCCACAGAGGCCTACTGACTGCAGGAAGCAGTGTGAGCCTGAC
TACTACCTGGATGAGGCCGACCGCTGTACAGCCTGCGTGACTTGTTCTCGAGACGAC
CTCGTGGAGAAGACGCCGTGTGCATGGAACTCCTCCCGTGTCTGCGAATGTCGACCC
GGCATGTTCTGTTCCACGTCTGCCGTCAACTCCTGTGCCCGCTGCTTCTTCCATTCTG
TCTGTCCGGCAGGGATGATTGTCAAGTTCCCAGGCACGGCGCAGAAGAACACGGTC
TGTGAGCCGGCTTCCCCAGGGGTCAGCCCTGCCTGTGCCAGCCCAGAGAACTGCAA
GGAACCCTCCAGTGGCACCATCCCCCAGGCCAAGCCCACCCCGGTGTCCCCAGCAA
CCTCCAGTGCCAGCACCATGCCTGTAAGAGGGGGCACCCGCCTCGCCCAGGAAGCT
GCTTCTAAACTGACGAGGGCTCCCGACTCTCCCTCCTCTGTGGGAAGGCCTAGTTCA
GATCCAGGTCTGTCCCCAACACAGCCATGCCCAGAGGGGTCTGGTGATTGCAGAAA
GCAGTGTGAGCCCGACTACTACCTGGACGAGGCCGGCCGCTGCACGGCCTGCGTGA
GCTGTTCTCGAGATGACCTTGTGGAGAAGACGCCATGTGCATGGAACTCCTCCCGCA
CCTGCGAATGTCGACCTGGCATGATCTGTGCCACATCAGCCACCAACTCCTGTGCCC
GCTGTGTCCCCTACCCAATCTGTGCAGCAGAGACGGTCACCAAGCCCCAGGATATG
GCTGAGAAGGACACCACCTTTGAGGCGCCACCCCTGGGGACCCAGCCGGACTGCAA
CCCCACCCCAGAGAATGGCGAGGCGCCTGCCAGCACCAGCCCCACTCAGAGCTTGC
TGGTGGACTCCCAGGCCAGTAAGACGCTGCCCATCCCAACCAGCGCTCCCGTCGCTC
TCTCCTCCACGGGGAAGCCCGTTCTGGATGCAGGGCCAGTGCTCTTCTGGGTGATCC
TGGTGTTGGTTGTGGTGGTCGGCTCCAGCGCCTTCCTCCTGTGCCACCGGAGGGCCT
GCAGGAAGCGAATTCGGCAGAAGCTCCACCTGTGCTACCCGGTCCAGACCTCCCAG
CCCAAGCTAGAGCTTGTGGATTCCAGACCCAGGAGGAGCTCAACGCAGCTGAGGAG
TGGTGCGTCGGTGACAGAACCCGTCGCGGAAGAGCGAGGGTTAATGAGCCAGCCAC
TGATGGAGACCTGCCACAGCGTGGGGGCAGCCTACCTGGAGAGCCTGCCGCTGCAG
GATGCCAGCCCGGCCGGGGGCCCCTCGTCCCCCAGGGACCTTCCTGAGCCCCGGGT
GTCCACGGAGCACACCAATAACAAGATTGAGAAAATCTACATCATGAAGGCTGACA
CCGTGATCGTGGGGACCGTGAAGGCTGAGCTGCCGGAGGGCCGGGGCCTGGCGGGG
CCAGCAGAGCCCGAGTTGGAGGAGGAGCTGGAGGCGGACCATACCCCCCACTACCC
CGAGCAGGAGACAGAACCGCCTCTGGGCAGCTGCAGCGATGTCATGCTCTCAGTGG
AAGAGGAAGGGAAAGAAGACCCCTTGCCCACAGCTGCCTCTGGAAAGTGAGGCCTG

GGCTGGGCTGGGGCTAGGAGGGCAGCAGGGTGGCCTCTGGGAGGCCAGGATGGCAC
TGTTGGCACCGAGGTTGGGGGCAGAGGCCCATCTGGCCTGAACTGAGGCTCCAGCA
TCTAGTGGTGGACCGGCCGGTCACTGCAGGGGTCTGGTGGTCTCTGCTTGCATCCCC
AACTTAGCTGTCCCCTGACCCAGAGCCTAGGGGATCCGGGGCTTGTACAGAAGAGA
CAGTCCAAGGGGACTGGATCCCAGCAGTGATGTTGGTTGAGGCAGCAAACAGATGG
CAGGATGGGCACTGCCGAGAACAGCATTGGTCCCAGAGCCCTGGGCATCAGACCTT
AACCACCAGGCCCACAGCCCAGCGAGGGAGAGGTCGTGAGGCCAGCTCCCGGGGCC
CCTGTAACCCTACTCTCCTCTCTCCCTGGACCTCAGAGGTGACACCCATTGGGCCCTT
CCGGCATGCCCCCAGTTACTGTAAATGTGGCCCCCAGTGGGCATGGAGCCAGTGCCT
GTGGTTGTTTCTCCAGAGTCAAAAGGGAAGTCGAGGGATGGGGCGTCGTCAGCTGG
CACTGTCTCTGCTGCAGCGGCCACACTGTACTCTGCACTGGTGTGAGGGCCCCTGCC
TGGACTGTGGGACCCTCCTGGTGCTGCCCACCTTCCCTGTCCTGTAGCCCCCTCGGTG
GGCCCAGGGCCTAGGGCCCAGGATCAAGTCACTCATCTCAGAATGTCCCCACCAAT
CCCCGCCACAGCAGGCGCCTCGGGTCCCAGATGTCTGCAGCCCTCAGCAGCTGCAG
ACCGCCCCTCACCAACCCAGAGAACCTGCTTTACTTTGCCCAGGGACTTCCTCCCCA
TGTGAACATGGGGAACTTCGGGCCCTGCCTGGAGTCCTTGACCGCTCTCTGTGGGCC
CCACCCACTCTGTCCTGGGAAATGAAGAAGCATCTTCCTTAGGTCTGCCCTGCTTGC
AAATCCACTAGCACCGACCCCACCACCTGGTTCCGGCTCTGCACGCTTTGGGGTGTG
GATGTCGAGAGGCACCACGGCCTCACCCAGGCATCTGCTTTACTCTGGACCATAGGA
AACAAGACCGTTTGGAGGTTTCATCAGGATTTTGGGTTTTTCACATTTCACGCTAAG
GAGTAGTGGCCCTGACTTCCGGTCGGCTGGCCAGCTGACTCCCTAGGGCCTTCAGAC
GTGTATGCAAATGAGTGATGGATAAGGATGAGTCTTGGAGTTGCGGGCAGCCTGGA
GACTCGTGGACTTACCGCCTGGAGGCAGGCCCGGGAAGGCTGCTGTTTACTCATCGG
GCAGCCACGTGCTCTCTGGAGGAAGTGATAGTTTCTGAAACCGCTCAGATGTTTTGG
GGAAAGTTGGAGAAGCCGTGGCCTTGCGAGAGGTGGTTACACCAGAACCTGGACAT
TGGCCAGAAGAAGCTTAAGTGGGCAGACACTGTTTGCCCAGTGTTTGTGCAAGGAT
GGAGTGGGTGTCTCTGCATCACCCACAGCCGCAGCTGTAAGGCACGCTGGAAGGCA
CACGCCTGCCAGGCAGGGCAGTCTGGCGCCCATGATGGGAGGGATTGACATGTTTC
AACAAAATAATGCACTTCCTTACCTAGTGGCCCTTCACACAACTTTTGAATCTCTAA
AAATCCATAAAATCCTTAAAGAACTGTAA

[00362] By "cluster of differentiation 33 (CD33)" is meant a protein having at least about 85% amino acid sequence identity to NCBI Accession No. **NP_001763.3** or fragment thereof and having immunomodulatory activity. An exemplary amino acid sequence is provided below.

>NP_001763.3 myeloid cell surface antigen CD33 isoform 1 precursor [Homo sapiens]

MPLLLLLPLLWAGALAMDPNFWLQVQESVTVQEGLCVLVPCTFFHPIPYYDKNSPVHG
YWFREGAIISRDSPVATNKLDQEVQEETQGRFRLLGDPSRNNCSLSIVDARRRDNGSYFF
RMERGSTKYSYKSPQLSVHVTDLTHRPKILIPGTLEPGHSKNLTCSVSWACEQGTPPIFS
WLSAAPTSLGPRTTHSSVLIITPRPQDHGTNLTCQVKFAGAGVTTERTIQLNVTYVPQNP
TTGIFPGDGSGKQETRAGVVHGAIGGAGVTALLALCLCLIFFIVKTHRRKAARTAVGRN
DTHPTTGSASPKHQKKSKLHGPTETSSCSGAAPTVEMDEELHYASLNFHGMNPSKDTST
EYSEVRTQ

[00363] By "cluster of differentiation 33 (CD33)" is meant a nucleic acid encoding a CD33 polypeptide. An exemplary CD33 nucleic acid sequence is provided below. >NM_001772.4 Homo sapiens CD33 molecule (CD33), transcript variant 1, mRNA

CTGCTCACACAGGAAGCCCTGGAAGCTGCTTCCTCAGACATGCCGCTGCTGCTACTG
CTGCCCCTGCTGTGGGCAGGGGCCCTGGCTATGGATCCAAATTTCTGGCTGCAAGTG
CAGGAGTCAGTGACGGTACAGGAGGGTTTGTGCGTCCTCGTGCCCTGCACTTTCTTC
CATCCCATACCCTACTACGACAAGAACTCCCCAGTTCATGGTTACTGGTTCCGGGAA
GGAGCCATTATATCCAGGGACTCTCCAGTGGCCACAAACAAGCTAGATCAAGAAGT
ACAGGAGGAGACTCAGGGCAGATTCCGCCTCCTTGGGGATCCCAGTAGGAACAACT
GCTCCCTGAGCATCGTAGACGCCAGGAGGAGGGATAATGGTTCATACTTCTTTCGGA
TGGAGAGAGGAAGTACCAAATACAGTTACAAATCTCCCCAGCTCTCTGTGCATGTGA
CAGACTTGACCCACAGGCCCAAAATCCTCATCCCTGGCACTCTAGAACCCGGCCACT
CCAAAAACCTGACCTGCTCTGTGTCCTGGGCCTGTGAGCAGGGAACACCCCCGATCT
TCTCCTGGTTGTCAGCTGCCCCCACCTCCCTGGGCCCCAGGACTACTCACTCCTCGGT
GCTCATAATCACCCCACGGCCCCAGGACCACGGCACCAACCTGACCTGTCAGGTGA
AGTTCGCTGGAGCTGGTGTGACTACGGAGAGAACCATCCAGCTCAACGTCACCTATG
TTCCACAGAACCCAACAACTGGTATCTTTCCAGGAGATGGCTCAGGGAAACAAGAG
ACCAGAGCAGGAGTGGTTCATGGGGCCATTGGAGGAGCTGGTGTTACAGCCCTGCT
CGCTCTTTGTCTCTGCCTCATCTTCTTCATAGTGAAGACCCACAGGAGGAAAGCAGC
CAGGACAGCAGTGGGCAGGAATGACACCCACCCTACCACAGGGTCAGCCTCCCCGA

AACACCAGAAGAAGTCCAAGTTACATGGCCCCACTGAAACCTCAAGCTGTTCAGGT
GCCGCCCCTACTGTGGAGATGGATGAGGAGCTGCATTATGCTTCCCTCAACTTTCAT
GGGATGAATCCTTCCAAGGACACCTCCACCGAATACTCAGAGGTCAGGACCCAGTG
AGGAACCCACAAGAGCATCAGGCTCAGCTAGAAGATCCACATCCTCTACAGGTCGG
GGACCAAAGGCTGATTCTTGGAGATTTAACACCCCACAGGCAATGGGTTTATAGAC
ATTATGTGAGTTTCCTGCTATATTAACATCATCTTAGACTTTGCAAGCAGAGAGTCGT
GGAATCAAATCTGTGCTCTTTCATTTGCTAAGTGTATGATGTCACACAAGCTCCTTAA
CCTTCCATGTCTCCATTTTCTTCTCTGTGAAGTAGGTATAAGAAGTCCTATCTCATAG
GGATGCTGTGAGCATTAAATAAAGGTACACATGGAAAACACCA

[00364] By "cluster of differentiation 52 (CD52)" is meant a protein having at least about 85% amino acid sequence identity to NCBI Accession No. **NP_001794.2** or fragment thereof and having immunomodulatory activity. An exemplary amino acid sequence is provided below. >NP_001794.2 CAMPATH-1 antigen precursor [Homo sapiens]

[00365] MKRFLFLLLTISLLVMVQIQTGLSGQNDTSQTSSPSASSNISGGIFLFFVANAIIH LFCFS

[00366] By "cluster of differentiation 52 (CD52)" is meant a nucleic acid encoding a CD52 polypeptide. An exemplary CD52 nucleic acid sequence is provided below. >NM_001803.3 Homo sapiens CD52 molecule (CD52), mRNA

AGACAGCCCTGAGATCACCTAAAAAGCTGCTACCAAGACAGCCACGAAGATCCTAC
CAAAATGAAGCGCTTCCTCTTCCTCCTACTCACCATCAGCCTCCTGGTTATGGTACA
GATACAAACTGGACTCTCAGGACAAAACGACACCAGCCAAACCAGCAGCCCCTCAG
CATCCAGCAACATAAGCGGAGGCATTTTCCTTTTCTTCGTGGCCAATGCCATAATCC
ACCTCTTCTGCTTCAGTTGAGGTGACACGTCTCAGCCTTAGCCCTGTGCCCCCTGAA
ACAGCTGCCACCATCACTCGCAAGAGAATCCCCTCCATCTTTGGGAGGGGTTGATGC
CAGACATCACCAGGTTGTAGAAGTTGACAGGCAGTGCCATGGGGGCAACAGCCAAA
ATAGGGGGGTAATGATGTAGGGGCCAAGCAGTGCCCAGCTGGGGGTCAATAAAGTT
ACCCTTGTACTTGCA

[00367] By "cluster of differentiation 70 (CD70)" is meant a protein having at least about 85% amino acid sequence identity to NCBI Accession No. **NP_001243.1** or fragment thereof and having immunomodulatory activity. An exemplary amino acid sequence is provided below. >NP_001243.1 CD70 antigen isoform 1 [Homo sapiens]

MPEEGSGCSVRRRPYGCVLRAALVPLVAGLVICLVVCIQRFAQAQQQLPLESLGWDVA
ELQLNHTGPQQDPRLYWQGGPALGRSFLHGPELDKGQLRIHRDGIYMVHIQVTLAICSS
TTASRHHPTTLAVGICSPASRSISLLRLSFHQGCTIASQRLTPLARGDTLCTNLTGTLLPSR
NTDETFFGVQWVRP

[00368] By "cluster of differentiation 70 (CD70)" is meant a nucleic acid encoding a CD70 polypeptide. An exemplary CD70 nucleic acid sequence is provided below. >NM_001252.5 Homo sapiens CD70 molecule (CD70), transcript variant 1, mRNA

AGAGAGGGGCAGGCTGGTCCCCTGACAGGTTGAAGCAAGTAGACGCCCAGGAGCCC
CGGGAGGGGGCTGCAGTTTCCTTCCTTCCTTCTCGGCAGCGCTCCGCGCCCCCATCG
CCCCTCCTGCGCTAGCGGAGGTGATCGCCGCGGCGATGCCGGAGGAGGGTTCGGGC
TGCTCGGTGCGGCGCAGGCCCTATGGGTGCGTCCTGCGGGCTGCTTTGGTCCCATTG
GTCGCGGGCTTGGTGATCTGCCTCGTGGTGTGCATCCAGCGCTTCGCACAGGCTCAG
CAGCAGCTGCCGCTCGAGTCACTTGGGTGGGACGTAGCTGAGCTGCAGCTGAATCA
CACAGGACCTCAGCAGGACCCCAGGCTATACTGGCAGGGGGGCCCAGCACTGGGCC
GCTCCTTCCTGCATGGACCAGAGCTGGACAAGGGGCAGCTACGTATCCATCGTGATG
GCATCTACATGGTACACATCCAGGTGACGCTGGCCATCTGCTCCTCCACGACGGCCT
CCAGGCACCACCCCACCACCCTGGCCGTGGGAATCTGCTCTCCCGCCTCCCGTAGCA
TCAGCCTGCTGCGTCTCAGCTTCCACCAAGGTTGTACCATTGCCTCCCAGCGCCTGA
CGCCCCTGGCCCGAGGGGACACACTCTGCACCAACCTCACTGGGACACTTTTGCCTT
CCCGAAACACTGATGAGACCTTCTTTGGAGTGCAGTGGGTGCGCCCCTGACCACTGC
TGCTGATTAGGGTTTTTTAAATTTTATTTTATTTTATTTAAGTTCAAGAGAAAAAGTG
TACACACAGGGGCCACCCGGGGTTGGGGTGGGAGTGTGGTGGGGGGTAGTGGTGGC
AGGACAAGAGAAGGCATTGAGCTTTTTCTTTCATTTTCCTATTAAAAAATACAAAAA
TCA

[00369] By "class II, major histocompatibility complex, transactivator (CIITA)" is meant a protein having at least about 85% amino acid sequence identity to NCBI Accession No. **NP_001273331.1** or fragment thereof and having immunomodulatory activity. An exemplary amino acid sequence is provided below.

>NP_001273331.1 MHC class II transactivator isoform 1 [Homo sapiens]

MRCLAPRPAGSYLSEPQGSSQCATMELGPLEGGYLELLNSDADPLCLYHFYDQMDLAG
EEEIELYSEPDTDTINCDQFSRLLCDMEGDEETREAYANIAELDQYVFQDSQLEGLSKDIF

IEHIGPDEVIGESMEMPAEVGQKSQKRPFPEELPADLKHWKPAEPPTVVTGSLLVGPVSD
CSTLPCLPLPALFNQEPASGQMRLEKTDQIPMPFSSSSLSCLNLPEGPIQFVPTISTLPHGL
WQISEAGTGVSSIFIYHGEVPQASQVPPPSGFTVHGLPTSPDRPGSTSPFAPSATDLPSMPE
PALTSRANMTEHKTSPTQCPAAGEVSNKLPKWPEPVEQFYRSLQDTYGAEPAGPDGILV
EVDLVQARLERSSSKSLERELATPDWAERQLAQGGLAEVLLAAKEHRRPRETRVIAVLG
KAGQGKSYWAGAVSRAWACGRLPQYDFVFSVPCHCLNRPGDAYGLQDLLFSLGPQPL
VAADEVFSHILKRPDRVLLILDGFEELEAQDGFLHSTCGPAPAEPCSLRGLLAGLFQKKL
LRGCTLLLTARPRGRLVQSLSKADALFELSGFSMEQAQAYVMRYFESSGMTEHQDRAL
TLLRDRPLLLSHSHSPTLCRAVCQLSEALLELGEDAKLPSTLTGLYVGLLGRAALDSPPG
ALAELAKLAWELGRRHQSTLQEDQFPSADVRTWAMAKGLVQHPPRAAESELAFPSFLL
QCFLGALWLALSGEIKDKELPQYLALTPRKKRPYDNWLEGVPRFLAGLIFQPPARCLGA
LLGPSAAASVDRKQKVLARYLKRLQPGTLRARQLLELLHCAHEAEEAGIWQHVVQELP
GRLSFLGTRLTPPDAHVLGKALEAAGQDFSLDLRSTGICPSGLGSLVGLSCVTRFRAALS
DTVALWESLQQHGETKLLQAAEEKFTIEPFKAKSLKDVEDLGKLVQTQRTRSSSEDTAG
ELPAVRDLKKLEFALGPVSGPQAFPKLVRILTAFSSLQHLDLDALSENKIGDEGVSQLSA
TFPQLKSLETLNLSQNNITDLGAYKLAEALPSLAASLLRLSLYNNCICDVGAESLARVLP
DMVSLRVMDVQYNKFTAAGAQQLAASLRRCPHVETLAMWTPTIPFSVQEHLQQQDSRI
SLR

[00370] By "class II, major histocompatibility complex, transactivator (CIITA)" is meant a nucleic acid encoding a CIITA polypeptide. An exemplary CIITA nucleic acid sequence is provided below.

>NM_001286402.1 Homo sapiens class II major histocompatibility complex transactivator (CIITA), transcript variant 1, mRNA

GGTTAGTGATGAGGCTAGTGATGAGGCTGTGTGCTTCTGAGCTGGGCATCCGAAGGC
ATCCTTGGGGAAGCTGAGGGCACGAGGAGGGGCTGCCAGACTCCGGGAGCTGCTGC
CTGGCTGGGATTCCTACACAATGCGTTGCCTGGCTCCACGCCCTGCTGGGTCCTACC
TGTCAGAGCCCCAAGGCAGCTCACAGTGTGCCACCATGGAGTTGGGGCCCCTAGAA
GGTGGCTACCTGGAGCTTCTTAACAGCGATGCTGACCCCCTGTGCCTCTACCACTTC
TATGACCAGATGGACCTGGCTGGAGAAGAAGAGATTGAGCTCTACTCAGAACCCGA
CACAGACACCATCAACTGCGACCAGTTCAGCAGGCTGTTGTGTGACATGGAAGGTG
ATGAAGAGACCAGGGAGGCTTATGCCAATATCGCGGAACTGGACCAGTATGTCTTC

CAGGACTCCCAGCTGGAGGGCCTGAGCAAGGACATTTTCATAGAGCACATAGGACC
AGATGAAGTGATCGGTGAGAGTATGGAGATGCCAGCAGAAGTTGGGCAGAAAAGTC
AGAAAAGACCCTTCCCAGAGGAGCTTCCGGCAGACCTGAAGCACTGGAAGCCAGCT
GAGCCCCCCACTGTGGTGACTGGCAGTCTCCTAGTGGGACCAGTGAGCGACTGCTCC
ACCCTGCCCTGCCTGCCACTGCCTGCGCTGTTCAACCAGGAGCCAGCCTCCGGCCAG
ATGCGCCTGGAGAAAACCGACCAGATTCCCATGCCTTTCTCCAGTTCCTCGTTGAGC
TGCCTGAATCTCCCTGAGGGACCCATCCAGTTTGTCCCCACCATCTCCACTCTGCCCC
ATGGGCTCTGGCAAATCTCTGAGGCTGGAACAGGGGTCTCCAGTATATTCATCTACC
ATGGTGAGGTGCCCCAGGCCAGCCAAGTACCCCCTCCCAGTGGATTCACTGTCCACG
GCCTCCCAACATCTCCAGACCGGCCAGGCTCCACCAGCCCCTTCGCTCCATCAGCCA
CTGACCTGCCCAGCATGCCTGAACCTGCCCTGACCTCCCGAGCAAACATGACAGAG
CACAAGACGTCCCCCACCCAATGCCCGGCAGCTGGAGAGGTCTCCAACAAGCTTCC
AAAATGGCCTGAGCCGGTGGAGCAGTTCTACCGCTCACTGCAGGACACGTATGGTG
CCGAGCCCGCAGGCCCGGATGGCATCCTAGTGGAGGTGGATCTGGTGCAGGCCAGG
CTGGAGAGGAGCAGCAGCAAGAGCCTGGAGCGGGAACTGGCCACCCCGGACTGGG
CAGAACGGCAGCTGGCCCAAGGAGGCCTGGCTGAGGTGCTGTTGGCTGCCAAGGAG
CACCGGCGGCCGCGTGAGACACGAGTGATTGCTGTGCTGGGCAAAGCTGGTCAGGG
CAAGAGCTATTGGGCTGGGGCAGTGAGCCGGGCCTGGGCTTGTGGCCGGCTTCCCC
AGTACGACTTTGTCTTCTCTGTCCCCTGCCATTGCTTGAACCGTCCGGGGGATGCCTA
TGGCCTGCAGGATCTGCTCTTCTCCCTGGGCCCACAGCCACTCGTGGCGGCCGATGA
GGTTTTCAGCCACATCTTGAAGAGACCTGACCGCGTTCTGCTCATCCTAGACGGCTT
CGAGGAGCTGGAAGCGCAAGATGGCTTCCTGCACAGCACGTGCGGACCGGCACCGG
CGGAGCCCTGCTCCCTCCGGGGGCTGCTGGCCGGCCTTTTCCAGAAGAAGCTGCTCC
GAGGTTGCACCCTCCTCCTCACAGCCCGGCCCCGGGGCCGCCTGGTCCAGAGCCTGA
GCAAGGCCGACGCCCTATTTGAGCTGTCCGGCTTCTCCATGGAGCAGGCCCAGGCAT
ACGTGATGCGCTACTTTGAGAGCTCAGGGATGACAGAGCACCAAGACAGAGCCCTG
ACGCTCCTCCGGGACCGGCCACTTCTTCTCAGTCACAGCCACAGCCCTACTTTGTGC
CGGGCAGTGTGCCAGCTCTCAGAGGCCCTGCTGGAGCTTGGGGAGGACGCCAAGCT
GCCCTCCACGCTCACGGGACTCTATGTCGGCCTGCTGGGCCGTGCAGCCCTCGACAG
CCCCCCCGGGGCCCTGGCAGAGCTGGCCAAGCTGGCCTGGGAGCTGGGCCGCAGAC
ATCAAAGTACCCTACAGGAGGACCAGTTCCCATCCGCAGACGTGAGGACCTGGGCG

ATGGCCAAAGGCTTAGTCCAACACCCACCGCGGGCCGCAGAGTCCGAGCTGGCCTT
CCCCAGCTTCCTCCTGCAATGCTTCCTGGGGGCCCTGTGGCTGGCTCTGAGTGGCGA
AATCAAGGACAAGGAGCTCCCGCAGTACCTAGCATTGACCCCAAGGAAGAAGAGGC
CCTATGACAACTGGCTGGAGGGCGTGCCACGCTTTCTGGCTGGGCTGATCTTCCAGC
CTCCCGCCCGCTGCCTGGGAGCCCTACTCGGGCCATCGGCGGCTGCCTCGGTGGACA
GGAAGCAGAAGGTGCTTGCGAGGTACCTGAAGCGGCTGCAGCCGGGGACACTGCGG
GCGCGGCAGCTGCTGGAGCTGCTGCACTGCGCCCACGAGGCCGAGGAGGCTGGAAT
TTGGCAGCACGTGGTACAGGAGCTCCCCGGCCGCCTCTCTTTTCTGGGCACCCGCCT
CACGCCTCCTGATGCACATGTACTGGGCAAGGCCTTGGAGGCGGCGGGCCAAGACT
TCTCCCTGGACCTCCGCAGCACTGGCATTTGCCCCTCTGGATTGGGGAGCCTCGTGG
GACTCAGCTGTGTCACCCGTTTCAGGGCTGCCTTGAGCGACACGGTGGCGCTGTGGG
AGTCCCTGCAGCAGCATGGGGAGACCAAGCTACTTCAGGCAGCAGAGGAGAAGTTC
ACCATCGAGCCTTTCAAAGCCAAGTCCCTGAAGGATGTGGAAGACCTGGGAAAGCT
TGTGCAGACTCAGAGGACGAGAAGTTCCTCGGAAGACACAGCTGGGGAGCTCCCTG
CTGTTCGGGACCTAAAGAAACTGGAGTTTGCGCTGGGCCCTGTCTCAGGCCCCCAGG
CTTTCCCCAAACTGGTGCGGATCCTCACGGCCTTTTCCTCCCTGCAGCATCTGGACCT
GGATGCGCTGAGTGAGAACAAGATCGGGGACGAGGGTGTCTCGCAGCTCTCAGCCA
CCTTCCCCCAGCTGAAGTCCTTGGAAACCCTCAATCTGTCCCAGAACAACATCACTG
ACCTGGGTGCCTACAAACTCGCCGAGGCCCTGCCTTCGCTCGCTGCATCCCTGCTCA
GGCTAAGCTTGTACAATAACTGCATCTGCGACGTGGGAGCCGAGAGCTTGGCTCGTG
TGCTTCCGGACATGGTGTCCCTCCGGGTGATGGACGTCCAGTACAACAAGTTCACGG
CTGCCGGGGCCCAGCAGCTCGCTGCCAGCCTTCGGAGGTGTCCTCATGTGGAGACGC
TGGCGATGTGGACGCCCACCATCCCATTCAGTGTCCAGGAACACCTGCAACAACAG
GATTCACGGATCAGCCTGAGATGATCCCAGCTGTGCTCTGGACAGGCATGTTCTCTG
AGGACACTAACCACGCTGGACCTTGAACTGGGTACTTGTGGACACAGCTCTTCTCCA
GGCTGTATCCCATGAGCCTCAGCATCCTGGCACCCGGCCCCTGCTGGTTCAGGGTTG
GCCCCTGCCCGGCTGCGGAATGAACCACATCTTGCTCTGCTGACAGACACAGGCCCG
GCTCCAGGCTCCTTTAGCGCCCAGTTGGGTGGATGCCTGGTGGCAGCTGCGGTCCAC
CCAGGAGCCCCGAGGCCTTCTCTGAAGGACATTGCGGACAGCCACGGCCAGGCCAG
AGGGAGTGACAGAGGCAGCCCCATTCTGCCTGCCCAGGCCCCTGCCACCCTGGGGA
GAAAGTACTTCTTTTTTTTTATTTTTAGACAGAGTCTCACTGTTGCCCAGGCTGGCGT

GCAGTGGTGCGATCTGGGTTCACTGCAACCTCCGCCTCTTGGGTTCAAGCGATTCTT
CTGCTTCAGCCTCCCGAGTAGCTGGGACTACAGGCACCCACCATCATGTCTGGCTAA
TTTTTCATTTTTAGTAGAGACAGGGTTTTGCCATGTTGGCCAGGCTGGTCTCAAACTC
TTGACCTCAGGTGATCCACCCACCTCAGCCTCCCAAAGTGCTGGGATTACAAGCGTG
AGCCACTGCACCGGGCCACAGAGAAAGTACTTCTCCACCCTGCTCTCCGACCAGAC
ACCTTGACAGGGCACACCGGGCACTCAGAAGACACTGATGGGCAACCCCCAGCCTG
CTAATTCCCCAGATTGCAACAGGCTGGGCTTCAGTGGCAGCTGCTTTTGTCTATGGG
ACTCAATGCACTGACATTGTTGGCCAAAGCCAAAGCTAGGCCTGGCCAGATGCACC
AGCCCTTAGCAGGGAAACAGCTAATGGGACACTAATGGGGCGGTGAGAGGGGAAC
AGACTGGAAGCACAGCTTCATTTCCTGTGTCTTTTTTCACTACATTATAAATGTCTCT
TTAATGTCACAGGCAGGTCCAGGGTTTGAGTTCATACCCTGTTACCATTTTGGGGTA
CCCACTGCTCTGGTTATCTAATATGTAACAAGCCACCCCAAATCATAGTGGCTTAAA
ACAACACTCACATTTA

[00371] By "cytotoxic T-lymphocyte associated protein 4 (CTLA-4) polypeptide" is meant a protein having at least about 85% sequence identity to NCBI Accession No. EAW70354.1 or a fragment thereof. An exemplary amino acid sequence is provided below:

>EAW70354.1 cytotoxic T-lymphocyte-associated protein 4 [Homo sapiens]
MACLGFQRHKAQLNLATRTWPCTLLFFLLFIPVFCKAMHVAQPAVVLASSRGIASFVCE
YASPGKATEVRVTVLRQADSQVTEVCAATYMMGNELTFLDDSICTGTSSGNQVNLTIQ
GLRAMDTGLYICKVELMYPPPYYLGIGNGTQIYVIDPEPCPDSDFLLWILAAVSSGLFFY
SFLLTAVSLSKMLKKRSPLTTGVYVKMPPTEPECEKQFQPYFIPIN

[00372] By "cytotoxic T-lymphocyte associated protein 4 (CTLA-4) polynucleotide" is meant a nucleic acid molecule encoding a CTLA-4 polypeptide. The CTLA-4 gene encodes an immunoglobulin superfamily and encodes a protein which transmits an inhibitory signal to T cells. An exemplary CTLA-4 nucleic acid sequence is provided below.

>BC074842.2 Homo sapiens cytotoxic T-lymphocyte-associated protein 4, mRNA (cDNA clone MGC:104099 IMAGE:30915552), complete cds
GACCTGAACACCGCTCCCATAAAGCCATGGCTTGCCTTGGATTTCAGCGGCACAAGG
CTCAGCTGAACCTGGCTACCAGGACCTGGCCCTGCACTCTCCTGTTTTTTCTTCTCTT
CATCCCTGTCTTCTGCAAAGCAATGCACGTGGCCCAGCCTGCTGTGGTACTGGCCAG
CAGCCGAGGCATCGCCAGCTTTGTGTGTGAGTATGCATCTCCAGGCAAAGCCACTGA

GGTCCGGGTGACAGTGCTTCGGCAGGCTGACAGCCAGGTGACTGAAGTCTGTGCGG
CAACCTACATGATGGGGAATGAGTTGACCTTCCTAGATGATTCCATCTGCACGGGCA
CCTCCAGTGGAAATCAAGTGAACCTCACTATCCAAGGACTGAGGGCCATGGACACG
GGACTCTACATCTGCAAGGTGGAGCTCATGTACCCACCGCCATACTACCTGGGCATA
GGCAACGGAACCCAGATTTATGTAATTGATCCAGAACCGTGCCCAGATTCTGACTTC
CTCCTCTGGATCCTTGCAGCAGTTAGTTCGGGGTTGTTTTTTTATAGCTTTCTCCTCAC
AGCTGTTTCTTTGAGCAAAATGCTAAAGAAAAGAAGCCCTCTTACAACAGGGGTCTA
TGTGAAAATGCCCCCAACAGAGCCAGAATGTGAAAAGCAATTTCAGCCTTATTTTAT
TCCCATCAATTGAGAAACCATTATGAAGAAGAGAGTCCATATTTCAATTTCCAAGAG
CTGAGG

[00373] By "cytidine deaminase" is meant a polypeptide or fragment thereof capable of catalyzing a deamination reaction that converts an amino group to a carbonyl group. In one embodiment, the cytidine deaminase converts cytosine to uracil or 5-methylcytosine to thymine. PmCDA1 derived from Petromyzon marinus (Petromyzon marinus cytosine deaminase 1), or AID (Activation-induced cytidine deaminase; AICDA) derived from mammal (e.g., human, swine, bovine, horse, monkey etc.), and APOBEC are exemplary cytidine deaminases.

[00374] The base sequence and amino acid sequence of PmCDA1 and the base sequence and amino acid sequence of human AID are shown below.

>tr|A5H718|A5H718_PETMA Cytosine deaminase OS=Petromyzon marinus OX=7757 PE=2 SV=1

MTDAEYVRIHEKLDIYTFKKQFFNNKKSVSHRCYVLFELKRRGERRACFWGYAVNKPQ
SGTERGIHAEIFSIRKVEEYLRDNPGQFTINWYSSWSPCADCAEKILEWYNQELRGNGHT
LKIWACKLYYEKNARNQIGLWNLRDNGVGLNVMVSEHYQCCRKIFIQSSHNQLNENR
WLEKTLKRAEKRRSELSIMIQVKILHTTKSPAV

>EF094822.1 Petromyzon marinus isolate PmCDA.21 cytosine deaminase mRNA, complete cds

TGACACGACACAGCCGTGTATATGAGGAAGGGTAGCTGGATGGGGGGGGGGGGAA
TACGTTCAGAGAGGACATTAGCGAGCGTCTTGTTGGTGGCCTTGAGTCTAGACACCT
GCAGACATGACCGACGCTGAGTACGTGAGAATCCATGAGAAGTTGGACATCTACAC
GTTTAAGAAACAGTTTTTCAACAACAAAAAATCCGTGTCGCATAGATGCTACGTTCT
CTTTGAATTAAAACGACGGGGTGAACGTAGAGCGTGTTTTTGGGGCTATGCTGTGAA

TAAACCACAGAGCGGGACAGAACGTGGAATTCACGCCGAAATCTTTAGCATTAGAA
AAGTCGAAGAATACCTGCGCGACAACCCCCGGACAATTCACGATAAATTGGTACTCA
TCCTGGAGTCCTTGTGCAGATTGCGCTGAAAAGATCTTAGAATGGTATAACCAGGAG
CTGCGGGGGAACGGCCACACTTTGAAAATCTGGGCTTGCAAACTCTATTACGAGAA
AAATGCGAGGAATCAAATTGGGCTGTGGAACCTCAGAGATAACGGGGTTGGGTTGA
ATGTAATGGTAAGTGAACACTACCAATGTTGCAGGAAAATATTCATCCAATCGTCGC
ACAATCAATTGAATGAGAATAGATGGCTTGAGAAGACTTTGAAGCGAGCTGAAAAA
CGACGGAGCGAGTTGTCCATTATGATTCAGGTAAAAATACTCCACACCACTAAGAGT
CCTGCTGTTTAAGAGGCTATGCGGATGGTTTTC

>tr|Q6QJ80|Q6QJ80_HUMAN Activation-induced cytidine deaminase OS=Homo sapiens OX=9606 GN=AICDA PE=2 SV=1
MDSLLMNRRKFLYQFKNVRWAKGRRETYLCYVVKRRDSATSFSLDFGYLRNKNGCHV
ELLFLRYISDWDLDPGRCYRVTWFTSWSPCYDCARHVADFLRGNPNLSLRIFTARLYFC
EDRKAEPEGLRRLHRAGVQIAIMTFKAPV

>NG_011588.1:5001-15681 Homo sapiens activation induced cytidine deaminase (AICDA), RefSeqGene (LRG_17) on chromosome 12
AGAGAACCATCATTAATTGAAGTGAGATTTTTCTGGCCTGAGACTTGCAGGGAGGCA
AGAAGACACTCTGGACACCACTATGGACAGGTAAAGAGGCAGTCTTCTCGTGGGTG
ATTGCACTGGCCTTCCTCTCAGAGCAAATCTGAGTAATGAGACTGGTAGCTATCCCT
TTCTCTCATGTAACTGTCTGACTGATAAGATCAGCTTGATCAATATGCATATATATTT
TTTGATCTGTCTCCTTTTCTTCTATTCAGATCTTATACGCTGTCAGCCCAATTCTTTCT
GTTTCAGACTTCTCTTGATTTCCCTCTTTTTCATGTGGCAAAAGAAGTAGTGCGTACA
ATGTACTGATTCGTCCTGAGATTTGTACCATGGTTGAAACTAATTTATGGTAATAAT
ATTAACATAGCAAATCTTTAGAGACTCAAATCATGAAAAGGTAATAGCAGTACTGT
ACTAAAAACGGTAGTGCTAATTTTCGTAATAATTTTGTAAATATTCAACAGTAAAAC
AACTTGAAGACACACTTTCCTAGGGAGGCGTTACTGAAATAATTTAGCTATAGTAAG
AAAATTTGTAATTTTAGAAATGCCAAGCATTCTAAATTAATTGCTTGAAAGTCACTA
TGATTGTGTCCATTATAAGGAGACAAATTCATTCAAGCAAGTTATTTAATGTTAAAG
GCCCAATTGTTAGGCAGTTAATGGCACTTTTACTATTAACTAATCTTTCCATTTGTTC

AGACGTAGCTTAACTTACCTCTTAGGTGTGAATTTGGTTAAGGTCCTCATAATGTCTT
TATGTGCAGTTTTTGATAGGTTATTGTCATAGAACTTATTCTATTCCTACATTTATGA
TTACTATGGATGTATGAGAATAACACCTAATCCTTATACTTTACCTCAATTTAACTCC
TTTATAAAGAACTTACATTACAGAATAAAGATTTTTTAAAAATATATTTTTTTGTAGA
GACAGGGTCTTAGCCCAGCCGAGGCTGGTCTCTAAGTCCTGGCCCAAGCGATCCTCC
TGCCTGGGCCTCCTAAAGTGCTGGAATTATAGACATGAGCCATCACATCCAATATAC
AGAATAAAGATTTTTAATGGAGGATTTAATGTTCTTCAGAAAATTTTCTTGAGGTCA
GACAATGTCAAATGTCTCCTCAGTTTACACTGAGATTTTGAAAACAAGTCTGAGCTA
TAGGTCCTTGTGAAGGGTCCATTGGAAATACTTGTTCAAAGTAAAATGGAAAGCAA
AGGTAAAATCAGCAGTTGAAATTCAGAGAAAGACAGAAAAGGAGAAAAGATGAAA
TTCAACAGGACAGAAGGGAAATATATTATCATTAAGGAGGACAGTATCTGTAGAGC
TCATTAGTGATGGCAAAATGACTTGGTCAGGATTATTTTTAACCCGCTTGTTTCTGGT
TTGCACGGCTGGGGATGCAGCTAGGGTTCTGCCTCAGGGAGCACAGCTGTCCAGAG
CAGCTGTCAGCCTGCAAGCCTGAAACACTCCCTCGGTAAAGTCCTTCCTACTCAGGA
CAGAAATGACGAGAACAGGGAGCTGGAAACAGGCCCCTAACCAGAGAAGGGAAGT
AATGGATCAACAAAGTTAACTAGCAGGTCAGGATCACGCAATTCATTTCACTCTGAC
TGGTAACATGTGACAGAAACAGTGTAGGCTTATTGTATTTTCATGTAGAGTAGGACC
CAAAAATCCACCCAAAGTCCTTTATCTATGCCACATCCTTCTTATCTATACTTCCAGG
ACACTTTTTCTTCCTTATGATAAGGCTCTCTCTCTCTCCACACACACACACACACACA
CACACACACACACACACACACACACACAAACACACACCCCGCCAACCAAGGTGCATGT
AAAAAGATGTAGATTCCTCTGCCTTTCTCATCTACACAGCCCAGGAGGGTAAGTTAA
TATAAGAGGGATTTATTGGTAAGAGATGATGCTTAATCTGTTTAACACTGGGCCTCA
AAGAGAGAATTTCTTTTCTTCTGTACTTATTAAGCACCTATTATGTGTTGAGCTTATA
TATACAAAGGGTTATTATATGCTAATATAGTAATAGTAATGGTGGTTGGTACTATGG
TAATTACCATAAAAATTATTATCCTTTTAAAATAAAGCTAATTATTATTGGATCTTTT
TTAGTATTCATTTTATGTTTTTTATGTTTTTGATTTTTTAAAAGACAATCTCACCCTGT
TACCCAGGCTGGAGTGCAGTGGTGCAATCATAGCTTTCTGCAGTCTTGAACTCCTGG
GCTCAAGCAATCCTCCTGCCTTGGCCTCCCAAAGTGTTGGGATACAGTCATGAGCCA
CTGCATCTGGCCTAGGATCCATTTAGATTAAAATATGCATTTTAAATTTTAAAATAAT
ATGGCTAATTTTTACCTTATGTAATGTGTATACTGGCAATAAATCTAGTTTGCTGCCT
AAAGTTTAAAGTGCTTTCCAGTAAGCTTCATGTACGTGAGGGGAGACATTTAAAGTG

AAACAGACAGCCAGGTGTGGTGGCTCACGCCTGTAATCCCAGCACTCTGGGAGGCT
GAGGTGGGTGGATCGCTTGAGCCCTGGAGTTCAAGACCAGCCTGAGCAACATGGCA
AAACGCTGTTTCTATAACAAAAATTAGCCGGGCATGGTGGCATGTGCCTGTGGTCCC
AGCTACTAGGGGGCTGAGGCAGGAGAATCGTTGGAGCCCAGGAGGTCAAGGCTGCA
CTGAGCAGTGCTTGCGCCACTGCACTCCAGCCTGGGTGACAGGACCAGACCTTGCCT
CAAAAAAATAAGAAGAAAAATTAAAAATAAATGGAAACAACTACAAAGAGCTGTT
GTCCTAGATGAGCTACTTAGTTAGGCTGATATTTTGGTATTTAACTTTTAAAGTCAGG
GTCTGTCACCTGCACTACATTATTAAAATATCAATTCTCAATGTATATCCACACAAA
GACTGGTACGTGAATGTTCATAGTACCTTTATTCACAAAACCCCAAAGTAGAGACTA
TCCAAATATCCATCAACAAGTGAACAAATAAACAAAATGTGCTATATCCATGCAAT
GGAATACCACCCTGCAGTACAAAGAAGCTACTTGGGGATGAATCCCAAAGTCATGA
CGCTAAATGAAAGAGTCAGACATGAAGGAGGAGATAATGTATGCCATACGAAATTC
TAGAAAATGAAAGTAACTTATAGTTACAGAAAGCAAATCAGGGCAGGCATAGAGGC
TCACACCTGTAATCCCAGCACTTTGAGAGGCCACGTGGGAAGATTGCTAGAACTCAG
GAGTTCAAGACCAGCCTGGGCAACACAGTGAAACTCCATTCTCCACAAAAATGGGA
AAAAAAGAAAGCAAATCAGTGGTTGTCCTGTGGGGAGGGGAAGGACTGCAAAGAG
GGAAGAAGCTCTGGTGGGGTGAGGGTGGTGATTCAGGTTCTGTATCCTGACTGTGGT
AGCAGTTTGGGGTGTTTACATCCAAAAATATTCGTAGAATTATGCATCTTAAATGGG
TGGAGTTTACTGTATGTAAATTATACCTCAATGTAAGAAAAAATAATGTGTAAGAAA
ACTTTCAATTCTCTTGCCAGCAAACGTTATTCAAATTCCTGAGCCCTTTACTTCGCAA
ATTCTCTGCACTTCTGCCCCGTACCATTAGGTGACAGCACTAGCTCCACAAATTGGA
TAAATGCATTTCTGGAAAAGACTAGGGACAAAATCCAGGCATCACTTGTGCTTTCAT
ATCAACCATGCTGTACAGCTTGTGTTGCTGTCTGCAGCTGCAATGGGGACTCTTGAT
TTCTTTAAGGAAACTTGGGTTACCAGAGTATTTCCACAAATGCTATTCAAATTAGTG
CTTATGATATGCAAGACACTGTGCTAGGAGCCAGAAAACAAAGAGGAGGAGAAATC
AGTCATTATGTGGGAACAACATAGCAAGATATTTAGATCATTTTGACTAGTTAAAAA
AGCAGCAGAGTACAAAATCACACATGCAATCAGTATAATCCAAATCATGTAAATAT
GTGCCTGTAGAAAGACTAGAGGAATAAACACAAGAATCTTAACAGTCATTGTCATT
AGACACTAAGTCTAATTATTATTATTAGACACTATGATATTTGAGATTTAAAAAATC
TTTAATATTTTAAAATTTAGAGCTCTTCTATTTTTCCATAGTATTCAAGTTTGACAAT
GATCAAGTATTACTCTTTCTTTTTTTTTTTTTTTTTTTTTTTTTGAGATGGAGTTTTGGT

CTTGTTGCCCATGCTGGAGTGGAATGGCATGACCATAGCTCACTGCAACCTCCACCT
CCTGGGTTCAAGCAAAGCTGTCGCCTCAGCCTCCCGGGTAGATGGGATTACAGGCG
CCCACCACCACACTCGGCTAATGTTTGTATTTTTAGTAGAGATGGGGTTTCACCATGT
TGGCCAGGCTGGTCTCAAACTCCTGACCTCAGAGGATCCACCTGCCTCAGCCTCCCA
AAGTGCTGGGATTACAGATGTAGGCCACTGCGCCCGGCCAAGTATTGCTCTTATACA
TTAAAAAACAGGTGTGAGCCACTGCGCCCAGCCAGGTATTGCTCTTATACATTAAAA
AATAGGCCGGTGCAGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAAGCCAAGGC
GGGCAGAACACCCGAGGTCAGGAGTCCAAGGCCAGCCTGGCCAAGATGGTGAAAC
CCCGTCTCTATTAAAAATACAAACATTACCTGGGCATGATGGTGGGCGCCTGTAATC
CCAGCTACTCAGGAGGCTGAGGCAGGAGGATCCGCGGAGCCTGGCAGATCTGCCTG
AGCCTGGGAGGTTGAGGCTACAGTAAGCCAAGATCATGCCAGTATACTTCAGCCTG
GGCGACAAAGTGAGACCGTAACAAAAAAAAAAAAATTTAAAAAAAGAAATTTAGA
TCAAGATCCAACTGTAAAAAGTGGCCTAAACACCACATTAAAGAGTTTGGAGTTTAT
TCTGCAGGCAGAAGAGAACCATCAGGGGGTCTTCAGCATGGGAATGGCATGGTGCA
CCTGGTTTTTGTGAGATCATGGTGGTGACAGTGTGGGGAATGTTATTTTGGAGGGAC
TGGAGGCAGACAGACCGGTTAAAAGGCCAGCACAACAGATAAGGAGGAAGAAGAT
GAGGGCTTGGACCGAAGCAGAGAAGAGCAAACAGGGAAGGTACAAATTCAAGAAA
TATTGGGGGGTTTGAATCAACACATTTAGATGATTAATTAAATATGAGGACTGAGGA
ATAAGAAATGAGTCAAGGATGGTTCCAGGCTGCTAGGCTGCTTACCTGAGGTGGCA
AAGTCGGGAGGAGTGGCAGTTTAGGACAGGGGGCAGTTGAGGAATATTGTTTTGAT
CATTTTGAGTTTGAGGTACAAGTTGGACACTTAGGTAAAGACTGGAGGGGAAATCT
GAATATACAATTATGGGACTGAGGAACAAGTTTATTTTATTTTTTGTTTCGTTTTCTT
GTTGAAGAACAAATTTAATTGTAATCCCAAGTCATCAGCATCTAGAAGACAGTGGC
AGGAGGTGACTGTCTTGTGGGTAAGGGTTTGGGGTCCTTGATGAGTATCTCTCAATT
GGCCTTAAATATAAGCAGGAAAAGGAGTTTATGATGGATTCCAGGCTCAGCAGGGC
TCAGGAGGGCTCAGGCAGCCAGCAGAGGAAGTCAGAGCATCTTCTTTGGTTTAGCC
CAAGTAATGACTTCCTTAAAAAGCTGAAGGAAAATCCAGAGTGACCAGATTATAAA
CTGTACTCTTGCATTTTCTCTCCCTCCTCTCACCCACAGCCTCTTGATGAACCGGAGG
AAGTTTCTTTACCAATTCAAAAATGTCCGCTGGGCTAAGGGTCGGCGTGAGACCTAC
CTGTGCTACGTAGTGAAGAGGCGTGACAGTGCTACATCCTTTTCACTGGACTTTGGT
TATCTTCGCAATAAGGTATCAATTAAAGTCGGCTTTGCAAGCAGTTTAATGGTCAAC

TGTGAGTGCTTTTAGAGCCACCTGCTGATGGTATTACTTCCATCCTTTTTTGGCATTT
GTGTCTCTATCACATTCCTCAAATCCTTTTTTTTATTTCTTTTTCCATGTCCATGCACC
CATATTAGACATGGCCCAAAATATGTGATTTAATTCCTCCCCAGTAATGCTGGGCAC
CCTAATACCACTCCTTCCTTCAGTGCCAAGAACAACTGCTCCCAAACTGTTTACCAG
CTTTCCTCAGCATCTGAATTGCCTTTGAGATTAATTAAGCTAAAAGCATTTTTATATG
GGAGAATATTATCAGCTTGTCCAAGCAAAAATTTTAAATGTGAAAAACAAATTGTGT
CTTAAGCATTTTTGAAAATTAAGGAAGAAGAATTTGGGAAAAAATTAACGGTGGCT
CAATTCTGTCTTCCAAATGATTTCTTTTCCCTCCTACTCACATGGGTCGTAGGCCAGT
GAATACATTCAACATGGTGATCCCCAGAAAACTCAGAGAAGCCTCGGCTGATGATT
AATTAAATTGATCTTTCGGCTACCCGAGAGAATTACATTTCCAAGAGACTTCTTCAC
CAAAATCCAGATGGGTTTACATAAACTTCTGCCCACGGGTATCTCCTCTCTCCTAAC
ACGCTGTGACGTCTGGGCTTGGTGGAATCTCAGGGAAGCATCCGTGGGGTGGAAGG
TCATCGTCTGGCTCGTTGTTTGATGGTTATATTACCATGCAATTTTCTTTGCCTACATT
TGTATTGAATACATCCCAATCTCCTTCCTATTCGGTGACATGACACATTCTATTTCAG
AAGGCTTTGATTTTATCAAGCACTTTCATTTACTTCTCATGGCAGTGCCTATTACTTC
TCTTACAATACCCATCTGTCTGCTTTACCAAAATCTATTTCCCCTTTTCAGATCCTCCC
AAATGGTCCTCATAAACTGTCCTGCCTCCACCTAGTGGTCCAGGTATATTTCCACAA
TGTTACATCAACAGGCACTTCTAGCCATTTTCCTTCTCAAAAGGTGCAAAAAGCAAC
TTCATAAACACAAATTAAATCTTCGGTGAGGTAGTGTGATGCTGCTTCCTCCCAACT
CAGCGCACTTCGTCTTCCTCATTCCACAAAAACCCATAGCCTTCCTTCACTCTGCAGG
ACTAGTGCTGCCAAGGGTTCAGCTCTACCTACTGGTGTGCTCTTTTGAGCAAGTTGCT
TAGCCTCTCTGTAACACAAGGACAATAGCTGCAAGCATCCCCAAAGATCATTGCAG
GAGACAATGACTAAGGCTACCAGAGCCGCAATAAAAGTCAGTGAATTTTAGCGTGG
TCCTCTCTGTCTCTCCAGAACGGCTGCCACGTGGAATTGCTCTTCCTCCGCTACATCT
CGGACTGGGACCTAGACCCTGGCCGCTGCTACCGCGTCACCTGGTTCACCTCCTGGA
GCCCCTGCTACGACTGTGCCCGACATGTGGCCGACTTTCTGCGAGGGAACCCCAACC
TCAGTCTGAGGATCTTCACCGCGCGCCTCTACTTCTGTGAGGACCGCAAGGCTGAGC
CCGAGGGGCTGCGGCGGCTGCACCGCGCCGGGGTGCAAATAGCCATCATGACCTTC
AAAGGTGCGAAAGGGCCTTCCGCGCAGGCGCAGTGCAGCAGCCCGCATTCGGGATT
GCGATGCGGAATGAATGAGTTAGTGGGGAAGCTCGAGGGGAAGAAGTGGGCGGGG
ATTCTGGTTCACCTCTGGAGCCGAAATTAAAGATTAGAAGCAGAGAAAAGAGTGAA

TGGCTCAGAGACAAGGCCCCGAGGAAATGAGAAAATGGGGCCAGGGTTGCTTCTTT
CCCCTCGATTTGGAACCTGAACTGTCTTCTACCCCCATATCCCCGCCTTTTTTTCCTTT
TTTTTTTTTTGAAGATTATTTTTACTGCTGGAATACTTTTGTAGAAAACCACGAAAGA
ACTTTCAAAGCCTGGGAAGGGCTGCATGAAAATTCAGTTCGTCTCTCCAGACAGCTT
CGGCGCATCCTTTTGGTAAGGGGCTTCCTCGCTTTTTAAATTTTCTTTCTTTCTCTACA
GTCTTTTTTGGAGTTTCGTATATTTCTTATATTTTCTTATTGTTCAATCACTCTCAGTT
TTCATCTGATGAAAACTTTATTTCTCCTCCACATCAGCTTTTTCTTCTGCTGTTTCACC
ATTCAGAGCCCTCTGCTAAGGTTCCTTTTCCCTCCCTTTTCTTTCTTTTGTTGTTTCAC
ATCTTTAAATTTCTGTCTCTCCCCAGGGTTGCGTTTCCTTCCTGGTCAGAATTCTTTTC
TCCTTTTTTTTTTTTTTTTTTTTTTTTTTTAAACAAACAAACAAAAAACCCAAAAAAAC
TCTTTCCCAATTTACTTTCTTCCAACATGTTACAAAGCCATCCACTCAGTTTAGAAGA
CTCTCCGGCCCCACCGACCCCCAACCTCGTTTTGAAGCCATTCACTCAATTTGCTTCT
CTCTTTCTCTACAGCCCCTGTATGAGGTTGATGACTTACGAGACGCATTTCGTACTTT
GGGACTTTGATAGCAACTTCCAGGAATGTCACACACGATGAAATATCTCTGCTGAAG
ACAGTGGATAAAAAACAGTCCTTCAAGTCTTCTCTGTTTTTATTCTTCAACTCTCACT
TTCTTAGAGTTTACAGAAAAAATATTTATATACGACTCTTTAAAAAGATCTATGTCTT
GAAAATAGAGAAGGAACACAGGTCTGGCCAGGGACGTGCTGCAATTGGTGCAGTTT
TGAATGCAACATTGTCCCCTACTGGGAATAACAGAACTGCAGGACCTGGGAGCATC
CTAAAGTGTCAACGTTTTTCTATGACTTTTAGGTAGGATGAGAGCAGAAGGTAGATC
CTAAAAAGCATGGTGAGAGGATCAAATGTTTTTATATCAACATCCTTTATTATTTGA
TTCATTTGAGTTAACAGTGGTGTTAGTGATAGATTTTTCTATTCTTTTCCCTTGACGTT
TACTTTCAAGTAACACAAACTCTTCCATCAGGCCATGATCTATAGGACCTCCTAATG
AGAGTATCTGGGTGATTGTGACCCCAAACCATCTCTCCAAAGCATTAATATCCAATC
ATGCGCTGTATGTTTTAATCAGCAGAAGCATGTTTTTATGTTTGTACAAAAGAAGAT
TGTTATGGGTGGGGATGGAGGTATAGACCATGCATGGTCACCTTCAAGCTACTTTAA
TAAAGGATCTTAAAATGGGCAGGAGGACTGTGAACAAGACACCCTAATAATGGGTT
GATGTCTGAAGTAGCAAATCTTCTGGAAACGCAAACTCTTTTAAGGAAGTCCCTAAT
TTAGAAACACCCACAAACTTCACATATCATAATTAGCAAACAATTGGAAGGAAGTT
GCTTGAATGTTGGGGAGAGGAAAATCTATTGGCTCTCGTGGGTCTCTTCATCTCAGA
AATGCCAATCAGGTCAAGGTTTGCTACATTTTGTATGTGTGTGATGCTTCTCCCAAA
GGTATATTAACTATATAAGAGAGTTGTGACAAAACAGAATGATAAAGCTGCGAACC

GTGGCACACGCTCATAGTTCTAGCTGCTTGGGAGGTTGAGGAGGGAGGATGGCTTG
AACACAGGTGTTCAAGGCCAGCCTGGGCAACATAACAAGATCCTGTCTCTCAAAAA
AAAAAAAAAAAAAAAGAAAGAGAGAGGGCCGGGCGTGGTGGCTCACGCCTGTAAT
CCCAGCACTTTGGGAGGCCGAGCCGGGCGGATCACCTGTGGTCAGGAGTTTGAGAC
CAGCCTGGCCAACATGGCAAAACCCCGTCTGTACTCAAAATGCAAAAATTAGCCAG
GCGTGGTAGCAGGCACCTGTAATCCCAGCTACTTGGGAGGCTGAGGCAGGAGAATC
GCTTGAACCCAGGAGGTGGAGGTTGCAGTAAGCTGAGATCGTGCCGTTGCACTCCA
GCCTGGGCGACAAGAGCAAGACTCTGTCTCAGAAAAAAAAAAAAAAAAGAGAGAG
AGAGAGAAAGAGAACAATATTTGGGAGAGAAGGATGGGGAAGCATTGCAAGGAAA
TTGTGCTTTATCCAACAAAATGTAAGGAGCCAATAAGGGATCCCTATTTGTCTCTTTT
GGTGTCTATTTGTCCCTAACAACTGTCTTTGACAGTGAGAAAAATATTCAGAATAAC
CATATCCCTGTGCCGTTATTACCTAGCAACCCTTGCAATGAAGATGAGCAGATCCAC
AGGAAAACTTGAATGCACAACTGTCTTATTTTAATCTTATTGTACATAAGTTTGTAA
AAGAGTTAAAAATTGTTACTTCATGTATTCATTTATATTTTATATTATTTTGCGTCTA
ATGATTTTTTATTAACATGATTTCCTTTTCTGATATATTGAAATGGAGTCTCAAAGCT
TCATAAATTTATAACTTTAGAAATGATTCTAATAACAACGTATGTAATTGTAACATT
GCAGTAATGGTGCTACGAAGCCATTTCTCTTGATTTTTAGTAAACTTTTATGACAGCA
AATTTGCTTCTGGCTCACTTTCAATCAGTTAAATAAATGATAAATAATTTTGGAAGCT
GTGAAGATAAAATACCAAATAAAATAATATAAAAGTGATTTATATGAAGTTAAAAT
AAAAAATCAGTATGATGGAATAAACTTG

[00375] Apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like (APOBEC) is a family of evolutionarily conserved cytidine deaminases. Members of this family are C-to-U editing enzymes. The N-terminal domain of APOBEC like proteins is the catalytic domain, while the C-terminal domain is a pseudocatalytic domain. More specifically, the catalytic domain is a zinc dependent cytidine deaminase domain and is important for cytidine deamination. APOBEC family members include APOBEC1, APOBEC2, APOBEC3A, APOBEC3B, APOBEC3C, APOBEC3D ("APOBEC3E" now refers to this), APOBEC3F, APOBEC3G, APOBEC3H, APOBEC4, and Activation-induced (cytidine) deaminase. Many modified cytidine deaminases are commercially available, including but not limited to SaBE3, SaKKH-BE3, VQR-BE3, EQR-BE3, VRER-BE3, YE1-BE3, EE-BE3, YE2-BE3, and YEE-BE3, which are available from Addgene (plasmids 85169, 85170, 85171, 85172, 85173, 85174, 85175, 85176, 85177).

[00376] Other exemplary deaminases that can be fused to Cas9 according to aspects of this disclosure are provided below. It should be understood that, in some embodiments, the active domain of the respective sequence can be used, e.g., the domain without a localizing signal (nuclear localization sequence, without nuclear export signal, cytoplasmic localizing signal).

[00377] Human AID:
MDSLLMNRRKFLYQFKNVRWAKGRRETYLCYVVKRRDSATSFSLDFGYLRNKNGCHV
ELLFLRYISDWDLDPGRCYRVTWFTSWSPCYDCARHVADFLRGNPNLSLRIFTARLYFC
EDRKAEPEGLRRLHRAGVQIAIMTFKDYFYCWNTFVENHERTFKAWEGLHENSVRLSR
QLRRILLPLYEVDDLRDAFRTLGL (underline: nuclear localization sequence; double underline: nuclear export signal)

[00378] Mouse AID:
MDSLLMKQKKFLYHFKNVRWAKGRHETYLCYVVKRRDSATSCSLDFGHLRNKSGCHV
ELLFLRYISDWDLDPGRCYRVTWFTSWSPCYDCARHVAEFLRWNPNLSLRIFTARLYFC
EDRKAEPEGLRRLHRAGVQIGIMTFKDYFYCWNTFVENRERTFKAWEGLHENSVRLTR
QLRRILLPLYEVDDLRDAFRMLGF (underline: nuclear localization sequence; double underline: nuclear export signal)

[00379] Canine AID:
MDSLLMKQRKFLYHFKNVRWAKGRHETYLCYVVKRRDSATSFSLDFGHLRNKSGCHV
ELLFLRYISDWDLDPGRCYRVTWFTSWSPCYDCARHVADFLRGYPNLSLRIFAARLYFC
EDRKAEPEGLRRLHRAGVQIAIMTFKDYFYCWNTFVENREKTFKAWEGLHENSVRLSR
QLRRILLPLYEVDDLRDAFRTLGL (underline: nuclear localization sequence; double underline: nuclear export signal)

[00380] Bovine AID:
MDSLLKKQRQFLYQFKNVRWAKGRHETYLCYVVKRRDSPTSFSLDFGHLRNKAGCHV
ELLFLRYISDWDLDPGRCYRVTWFTSWSPCYDCARHVADFLRGYPNLSLRIFTARLYFC
DKERKAEPEGLRRLHRAGVQIAIMTFKDYFYCWNTFVENHERTFKAWEGLHENSVRLS
RQLRRILLPLYEVDDLRDAFRTLGL (underline: nuclear localization sequence; double underline: nuclear export signal)

[00381] Rat AID
MAVGSKPKAALVGPHWERERIWCFLCSTGLGTQQTGQTSRWLRPAATQDPVSPPRSLL
MKQRKFLYHFKNVRWAKGRHETYLCYVVKRRDSATSFSLDFGYLRNKSGCHVELLFL

RYISDWDLDPGRCYRVTWFTSWSPCYDCARHVADFLRGNPNLSLRIFTARLTGWGALP AGLMSPARPSDYFYCWNTFVENHERTFKAWEGLHENSVRLSRRLRRILLPLYEVDDLR DAFRTLGL

(underline: nuclear localization sequence; double underline: nuclear export signal)

[00382] Mouse APOBEC-3

MGPFCLGCSHRKCYSPIRNLISQETFKFHFKNLGYAKGRKDTFLCYEVTRKDCDSPVSLH HGVFKNKDNI*HAEICFLYWFHDKVLKVLSPREEFKITWYMSWSPCFEC*AEQIVRFLATHHN LSLDIFSSRLYNVQDPETQQNLCRLVQEGAQVAAMDLYEFKKCWKKFVDNGGRRFRP WKRLLTNFRYQDSKLQEILRPCYIPVPSSSSSTLSNICLTKGLPETRFCVEGRRMDPLSEE EFYSQFYNQRVKHLCYYHRMKPYLCYQLEQFNGQAPLKGCLLSEKGKQ*HAEILFLDKIR SMELSQVTITCYLTWSPCPNC*AWQLAAFKRDRPDLILHIYTSRLYFHWKRPFQKGLCSLW QSGILVDVMDLPQFTDCWTNFVNPKRPFWPWKGLEIISRRTQRRLRRIKESWGLQDLVN DFGNLQLGPPMS (italic: nucleic acid editing domain)

[00383] Rat APOBEC-3 :

MGPFCLGCSHRKCYSPIRNLISQETFKFHFKNRLRYAIDRKDTFLCYEVTRKDCDSPVSL HHGVFKNKDNI*HAEICFLYWFHDKVLKVLSPREEFKITWYMSWSPCFEC*AEQVLRFLATHH NLSLDIFSSRLYNIRDPENQQNLCRLVQEGAQVAAMDLYEFKKCWKKFVDNGGRRFRP WKKLLTNFRYQDSKLQEILRPCYIPVPSSSSSTLSNICLTKGLPETRFCVERRRVHLLSEEE FYSQFYNQRVKHLCYYHGVKPYLCYQLEQFNGQAPLKGCLLSEKGKQ*HAEILFLDKIRS MELSQVIITCYLTWSPCPNC*AWQLAAFKRDRPDLILHIYTSRLYFHWKRPFQKGLCSLWQ SGILVDVMDLPQFTDCWTNFVNPKRPFWPWKGLEIISRRTQRRLHRIKESWGLQDLVND FGNLQLGPPMS (italic: nucleic acid editing domain)

[00384] Rhesus macaque APOBEC-3 G:
MVEPMDPRTFVSNFNNRPILSGLNTVWLCCEVKTKDPSGPPLDAKIFQGKVYSKAKYHP
EMRFLRWFHKWRQLHHDQEYKVTWYVSWSPCTRCANSVATFLAKDPKVTLTIFVARL
YYFWKPDYQQALRILCQKRGGPHATMKIMNYNEFQDCWNKFVDGRGKPFKPRNNLPK
HYTLLQATLGELLRHLMDPGTFTSNFNNKPWVSGQHETYLCYKVERLHNDTWVPLNQ
HRGFLRNQAPNIHGFPKGRHAELCFLDLIPFWKLDGQQYRVTCFTSWSPCFSCAQEMAK
FISNNEHVSLCIFAARIYDDQGRYQEGLRALHRDGAKIAMMNYSEFEYCWDTFVDRQG
RPFQPWDGLDEHSQALSGRLRAI (italic: nucleic acid editing domain; underline: cytoplasmic localization signal)

[00385] Chimpanzee APOBEC-3 G:
MKPHFRNPVERMYQDTFSDNFYNRPILSHRNTVWLCYEVKTKGPSRPPLDAKIFRGQVY
SKLKY*HPEMRFFHWFSKWRKLHRDQEYEVTWYISWSPCTKC*TRDVATFLAEDPKVTLTIF
VARLYYFWDPDYQEALRSLCQKRDGPRATMKIMNYDEFQHCWSKFVYSQRELFEPWN
NLPKYYILLHIMLGEILRHSMDPPTFTSNFNNELWVRGRHETYLCYEVERLHNDTWVLL
NQRRGFLCNQAPHKHGFLEGR*HAELCFLDVIPFWKLDLHQDYRVTCFTSWSPCFSC*AQEM
AKFISNNKHVSLCIFAARIYDDQGRCQEGLRTLAKAGAKISIMTYSEFKHCWDTFVDHQ
GCPFQPWDGLEEHSQALSGRLRAILQNQGN

[00386] (italic: nucleic acid editing domain; underline: cytoplasmic localization signal)

[00387] Green monkey APOBEC-3G:
MNPQIRNMVEQMEPDIFVYYFNNRPILSGRNTVWLCYEVKTKDPSGPPLDANIFQGKLY
PEAKD*HPEMKFLHWFRKWRQLHRDQEYEVTWYVSWSPCTRC*ANSVATFLAEDPKVTLTIF
VARLYYFWKPDYQQALRILCQERGGPHATMKIMNYNEFQHCWNEFVDGQGKPFKPRK
NLPKHYTLLHATLGELLRHVMDPGTFTSNFNNKPWVSGQRETYLCYKVERSHNDTWV
LLNQHRGFLRNQAPDRHGFPKGR*HAELCFLDLIPFWKLDDQQYRVTCFTSWSPCFSC*AQK
MAKFISNNKHVSLCIFAARIYDDQGRCQEGLRTLHRDGAKIAVMNYSEFEYCWDTFVD
RQGRPFQPWDGLDEHSQALSGRLRAI

[00388] (italic: nucleic acid editing domain; underline: cytoplasmic localization signal)

[00389] Human APOBEC-3G:
MKPHFRNTVERMYRDTFSYNFYNRPILSRRNTVWLCYEVKTKGPSRPPLDAKIFRGQVY
SELKY*HPEMRFFHWFSKWRKLHRDQEYEVTWYISWSPCTKC*TRDMATFLAEDPKVTLTIF
VARLYYFWDPDYQEALRSLCQKRDGPRATMKIMNYDEFQHCWSKFVYSQRELFEPWN

NLPKYYILLHIMLGEILRHSMDPPTFTFNFNNEPWVRGRHETYLCYEVERMHNDTWVLL
NQRRGFLCNQAPHKHGFLEGR*HAELCFLDVIPFWKLDLDQDYRVTCFTSWSPCFSC*AQEM
AKFISKNKHVSLCIFTARIYDDQGRCQEGLRTLAEAGAKISIMTYSEFKHCWDTFVDHQ
GCPFQPWDGLDEHSQDLSGRLRAILQNQEN

(italic: nucleic acid editing domain; underline: cytoplasmic localization signal)

[00390] Human APOBEC-3F:

MKPHFRNTVERMYRDTFSYNFYNRPILSRRNTVWLCYEVKTKGPSRPRLDAKIFRGQV
YSQPEH*HAEMCFLSWFCGNQLPAYKCFQITWFVSWTPCPDC*VAKLAEFLAEHPNVTLTIS
AARLYYYWERDYRRALCRLSQAGARVKIMDDEEFAYCWENFVYSEGQPFMPWYKFD
DNYAFLHRTLKEILRNPMEAMYPHIFYFHFKNLRKAYGRNESWLCFTMEVVKHHSPVS
WKRGVFRNQVDPETH*CHAERCFLSWFCDDILSPNTNYEVTWYTSWSPCPEC*AGEVAEFLA
RHSNVNLTIFTARLYYFWDTDYQEGLRSLSQEGASVEIMGYKDFKYCWENFVYNDDEP
FKPWKGLKYNFLFLDSKLQEILE

(italic: nucleic acid editing domain)

[00391] Human APOBEC-3B:

MNPQIRNPMERMYRDTFYDNFENEPILYGRSYTWLCYEVKIKRGRSNLLWDTGVFRGQ
VYFKPQY*HAEMCFLSWFCGNQLPAYKCFQITWFVSWTPCPDC*VAKLAEFLSEHPNVTLTI
SAARLYYYWERDYRRALCRLSQAGARVTIMDYEEFAYCWENFVYNEGQQFMPWYKF
DENYAFLHRTLKEILRYLMDPDTFTFNFNNDPLVLRRRQTYLCYEVERLDNGTWVLMD
QHMGFLCNEAKNLLCGFY*GRHAELRFLDLVPSLQLDPAQIYRVTWFISWSPCFSWGC*AGE
VRAFLQENTHVRLRIFAARIYDYDPLYKEALQMLRDAGAQVSIMTYDEFEYCWDTFVY
RQGCPFQPWDGLEEHSQALSGRLRAILQNQGN

(italic: nucleic acid editing domain)

[00392] Rat APOBEC-3B:

MQPQGLGPNAGMGPVCLGCSHRRPYSPIRNPLKKLYQQTFYFHFKNVRYAWGRKNNF
LCYEVNGMDCALPVPLRQGVFRKQGHIHAELCFIYWFHDKVLRVLSPMEEFKVTWYM
SWSPCSKCAEQVARFLAAHRNLSLAIFSSRLYYYLRNPNYQQKLCRLIQEGVHVAAMD
LPEFKKCWNKFVDNDGQPFRPWMRLRINFSFYDCKLQEIFSRMNLLREDVFYLQFNNSH
RVKPVQNRYYRRKSYLCYQLERANGQEPLKGYLLYKKGEQHVEILFLEKMRSMELSQV
RITCYLTWSPCPNCARQLAAFKKDHPDLILRIYTSRLYFWRKKFQKGLCTLWRSGIHVD
VMDLPQFADCWTNFVNPQRPFRPWNELEKNSWRIQRRLRRIKESWGL

[00393] Bovine APOBEC-3B:

DGWEVAFRSGTVLKAGVLGVSMTEGWAGSGHPGQGACVWTPGTRNTMNLLREVLFK
QQFGNQPRVPAPYYRRKTYLCYQLKQRNDLTLDRGCFRNKKQRHAERFIDKINSLDLNP
SQSYKIICYITWSPCPNCANELVNFITRNNHLKLEIFASRLYFHWIKSFKMGLQDLQNAGI
SVAVMTHTEFEDCWEQFVDNQSRPFQPWDKLEQYSASIRRRLQRILTAPI

[00394] Chimpanzee APOBEC-3B:

MNPQIRNPMEWMYQRTFYYNFENEPILYGRSYTWLCYEVKIRRGHSNLLWDTGVFRGQ
MYSQPEHHAEMCFLSWFCGNQLSAYKCFQITWFVSWTPCPDCVAKLAKFLAEHPNVTL
TISAARLYYYWERDYRRALCRLSQAGARVKIMDDEEFAYCWENFVYNEGQPFMPWYK
FDDNYAFLHRTLKEIIRHLMDPDTFTFNFNNDPLVLRRHQTYLCYEVERLDNGTWVLM
DQHMGFLCNEAKNLLCGFYGRHAELRFLDLVPSLQLDPAQIYRVTWFISWSPCFSWGC
AGQVRAFLQENTHVRLRIFAARIYDYDPLYKEALQMLRDAGAQVSIMTYDEFEYCWDT
FVYRQGCPFQPWDGLEEHSQALSGRLRAILQVRASSLCMVPHRPPPPPQSPGPCLPLCSE
PPLGSLLPTGRPAPSLPFLLTASFSFPPPASLPPLPSLSLSPGHLPVPSFHSLTSCSIQPPCSSR
IRETEGWASVSKEGRDLG

[00395] Human APOBEC-3C:

MNPQIRNPMKAMYPGTFYFQFKNLWEANDRNETWLCFTVEGIKRRSVVSWKTGVFRN
QVDSETH*CHAERCFLSWFCDDILSPNTKYQVTWYTSWSPCPDC*AGEVAEFLARHSNVNLTI
FTARLYYFQYPCYQEGLRSLSQEGVAVEIMDYEDFKYCWENFVYNDNEPFKPWKGLKT
NFRLLKRRLRESLQ

(italic: nucleic acid editing domain)

[00396] Gorilla APOBEC3C

MNPQIRNPMKAMYPGTFYFQFKNLWEANDRNETWLCFTVEGIKRRSVVSWKTGVFRN
QVDSETH*CHAERCFLSWECDDILSPNTNYQVTWYTSWSPCPEC*AGEVAEFLARHSNVNLTI
FTARLYYFQDTDYQEGLRSLSQEGVAVKIMDYKDFKYCWENFVYNDDEPFKPWKGLK
YNFRFLKRRLQEILE

(italic: nucleic acid editing domain)

[00397] Human APOBEC-3A:

MEASPASGPRHLMDPHIFTSNFNNGIGRHKTYLCYEVERLDNGTSVKMDQHRGFLHNQ
AKNLLCGFYGR*HAELRFLDLVPSLQLDPAQIYRVTWFISWSPCFSWGC*AGEVRAFLQENTH
VRLRIFAARIYDYDPLYKEALQMLRDAGAQVSIMTYDEFKHCWDTFVDHQGCPFQPWD
GLDEHSQALSGRLRAILQNQGN

(italic: nucleic acid editing domain)

[00398] Rhesus macaque APOBEC-3A:

MDGSPASRPRHLMDPNTFTFNFNNDLSVRGRHQTYLCYEVERLDNGTWVPMDERRGF
LCNKAKNVPCGDYGC*HVELRFLCEVPSWQLDPAQTYRVTWFISWSPC*FRRGCAGQVRVFL
QENKHVRLRIFAARIYDYDPLYQEALRTLRDAGAQVSIMTYEEFKHCWDTFVDRQGRP
FQPWDGLDEHSQALSGRLRAILQNQGN

(italic: nucleic acid editing domain)

[00399] Bovine APOBEC-3A:

MDEYTFTENFNNQGWPSKTYLCYEMERLDGDATIPLDEYKGFVRNKGLDQPEKPC*HAE*
*LYFLGKIHSWNLDRNQHYRLTCFISWSPC*YDCAQKLTTFLKENHHISLHILASRIYTHNRFG
CHQSGLCELQAAGARITIMTFEDFKHCWETFVDHKGKPFQPWEGLNVKSQALCTELQAI
LKTQQN

(italic: nucleic acid editing domain)

[00400] Human APOBEC-3H:

MALLTAETFRLQFNNKRRLRRPYYPRKALLCYQLTPQNGSTPTRGYFENKKKC*HAEICFI*
*NEIKSMGLDETQCYQVTCYLTWSPCSSC*AWELVDFIKAHDHLNLGIFASRLYYHWCKPQQ
KGLRLLCGSQVPVEVMGFPKFADCWENFVDHEKPLSFNPYKMLEELDKNSRAIKRRLE
RIKIPGVRAQGRYMDILCDAEV

(italic: nucleic acid editing domain)

[00401] Rhesus macaque APOBEC-3H:

MALLTAKTFSLQFNNKRRVNKPYYPRKALLCYQLTPQNGSTPTRGHLKNKKKDHAEIR
FINKIKSMGLDETQCYQVTCYLTWSPCPSCAGELVDFIKAHRHLNLRIFASRLYYHWRP
NYQEGLLLLCGSQVPVEVMGLPEFTDCWENFVDHKEPPSFNPSEKLEELDKNSQAIKRR
LERIKSRSVDVLENGLRSLQLGPVTPSSSIRNSR

[00402] Human APOBEC-3D:
MNPQIRNPMERMYRDTFYDNFENEPILYGRSYTWLCYEVKIKRGRSNLLWDTGVFRGP
VLPKRQSNHRQEVYFRFEN*HAEMCFLSWFCGNRLPANRRFQITWFVSWNPC*LPCVVKVT
KFLAEHPNVTLTISAARLYYYRDRDWRWVLLRLHKAGARVKIMDYEDFAYCWENFVC
NEGQPFMPWYKFDDNYASLHRTLKEILRNPMEAMYPHIFYFHFKNLLKACGRNESWLC
FTMEVTKHHSAVFRKRGVFRNQVDPETHC*HAERCFLSWFCDDILSPNTNYEVTWYTSWSP*
*CPEC*AGEVAEFLARHSNVNLTIFTARLCYFWDTDYQEGLCSLSQEGASVKIMGYKDFVS
CWKNFVYSDDEPFKPWKGLQTNFRLLKRRLREILQ
(italic: nucleic acid editing domain)

[00403] Human APOBEC-1 :
MTSEKGPSTGDPTLRRRIEPWEFDVFYDPRELRKEACLLYEIKWGMSRKIWRSSGKNTT
NHVEVNFIKKFTSERDFHPSMSCSITWFLSWSPCWECSQAIREFLSRHPGVTLVIYVARLF
WHMDQQNRQGLRDLVNSGVTIQIMRASEYYHCWRNFVNYPPGDEAHWPQYPPLWMM
LYALELHCIILSLPPCLKISRRWQNHLTFFRLHLQNCHYQTIPPHILLATGLIHPSVAWR

[00404] Mouse APOBEC-1 :
MSSETGPVAVDPTLRRRIEPHEFEVFFDPRELRKETCLLYEINWGGRHSVWRHTSQNTSN
HVEVNFLEKFTTERYFRPNTRCSITWFLSWSPCGECSRAITEFLSRHPYVTLFIYIARLYH
HTDQRNRQGLRDLISSGVTIQIMTEQEYCYCWRNFVNYPPSNEAYWPRYPHLWVKLYV
LELYCIILGLPPCLKILRRKQPQLTFFTITLQTCHYQRIPPHLLWATGLK

[00405] Rat APOBEC-1 :
MSSETGPVAVDPTLRRRIEPHEFEVFFDPRELRKETCLLYEINWGGRHSIWRHTSQNTNK
HVEVNFIEKFTTERYFCPNTRCSITWFLSWSPCGECSRAITEFLSRYPHVTLFIYIARLYHH
ADPRNRQGLRDLISSGVTIQIMTEQESGYCWRNFVNYSPSNEAHWPRYPHLWVRLYVL
ELYCIILGLPPCLNILRRKQPQLTFFTIALQSCHYQRLPPHILWATGLK

[00406] Human APOBEC-2:
MAQKEEAAVATEAASQNGEDLENLDDPEKLKELIELPPFEIVTGERLPANFFKFQFRNVE
YSSGRNKTFLCYVVEAQGKGGQVQASRGYLEDEHAAAHAEEAFFNTILPAFDPALRYN

VTWYVSSSPCAACADRIIKTLSKTKNLRLLILVGRLFMWEEPEIQAALKKLKEAGCKLRI
MKPQDFEYVWQNFVEQEEGESKAFQPWEDIQENFLYYEEKLADILK

[00407] Mouse APOBEC-2:

MAQKEEAAEAAAPASQNGDDLENLEDPEKLKELIDLPPFEIVTGVRLPVNFFKFQFRNV
EYSSGRNKTFLCYVVEVQSKGGQAQATQGYLEDEHAGAHAEEAFFNTILPAFDPALKY
NVTWYVSSSPCAACADRILKTLSKTKNLRLLILVSRLFMWEEPEVQAALKKLKEAGCKL
RIMKPQDFEYIWQNFVEQEEGESKAFEPWEDIQENFLYYEEKLADILK

[00408] Rat APOBEC-2:

MAQKEEAAEAAAPASQNGDDLENLEDPEKLKELIDLPPFEIVTGVRLPVNFFKFQFRNV
EYSSGRNKTFLCYVVEAQSKGGQVQATQGYLEDEHAGAHAEEAFFNTILPAFDPALKY
NVTWYVSSSPCAACADRILKTLSKTKNLRLLILVSRLFMWEEPEVQAALKKLKEAGCKL
RIMKPQDFEYLWQNFVEQEEGESKAFEPWEDIQENFLYYEEKLADILK

[00409] Bovine APOBEC-2:

MAQKEEAAAAAEPASQNGEEVENLEDPEKLKELIELPPFEIVTGERLPAHYFKFQFRNVE
YSSGRNKTFLCYVVEAQSKGGQVQASRGYLEDEHATNHAEEAFFNSIMPTFDPALRYM
VTWYVSSSPCAACADRIVKTLNKTKNLRLLILVGRLFMWEEPEIQAALRKLKEAGCRLR
IMKPQDFEYIWQNFVEQEEGESKAFEPWEDIQENFLYYEEKLADILK

[00410] Petromyzon marinus CDA1 (pmCDAl)

MTDAEYVRIHEKLDIYTFKKQFFNNKKSVSHRCYVLFELKRRGERRACFWGYAVNKPQ
SGTERGIHAEIFSIRKVEEYLRDNPGQFTINWYSSWSPCADCAEKILEWYNQELRGNGHT
LKIWACKLYYEKNARNQIGLWNLRDNGVGLNVMVSEHYQCCRKIFIQSSHNQLNENR
WLEKTLKRAEKRRSELSFMIQVKILHTTKSPAV

[00411] Human APOBEC3G D316R D317R

MKPHFRNTVERMYRDTFSYNFYNRPILSRRNTVWLCYEVKTKGPSRPPLDAKIFRGQVY
SELKYHPEMRFFHWFSKWRKLHRDQEYEVTWYISWSPCTKCTRDMATFLAEDPKVTLT
IFVARLYYFWDPDYQEALRSLCQKRDGPRATMKFNYDEFQHCWSKFVYSQRELFEPWN
NLPKYYILLHFMLGEILRHSMDPPTFTFNFNNEPWVRGRHETYLCYEVERMHNDTWVL
LNQRRGFLCNQAPHKHGFLEGRHAELCFLDVIPFWKLDLDQDYRVTC
FTSWSPCFSCAQEMAKFISKKHVSLCIFTARIYRRQGRCQEGLRTLAEAGAKISFTYSEFK
HCWDTFVDHQGCPFQPWDGLDEHSQDLSGRLRAILQNQEN

[00412] Human APOBEC3G chain A

MDPPTFTFNFNNEPWWGRHETYLCYEVERMHNDTWVLLNQRRGFLCNQAPHKHGFLE
GRHAELCFLDVIPFWKLDLDQDYRVTCFTSWSPCFSCAQEMAKFISKNKHVSLCIFTARI
YDDQGRCQEGLRTLAEAGAKISF TYSEFKHCWDTFVDHQGCPFQPWDGLD
EHSQDLSGRLRAILQ

[00413] Human APOBEC3G chain A D120R D121R
MDPPTFTFNFNNEPWVRGRHETYLCYEVERMHNDTWVLLNQRRGFLCNQAPHKHGFL
EGRHAELCFLDVIPFWKLDLDQDYRVTCFTSWSPCFSCAQEMAKFISKNKHVSLCIFTAR
IYRRQGRCQEGLRTLAEAGAKISFMTYSEFKHCWDTFVDHQGCPFQPWDGLDEHSQDL
SGRLRAILQ

[00414] The term "deaminase" or "deaminase domain" refers to a protein or fragment thereof that catalyzes a deamination reaction. In some embodiments, the deaminase or deaminase domain is a variant of a naturally-occurring deaminase from an organism. In some embodiments, the deaminase or deaminase domain does not occur in nature. For example, in some embodiments, the deaminase or deaminase domain is at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75% at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5% identical to a naturally-occurring deaminase. In some embodiments, the deaminase is a cytosine deaminase or an adenosine deaminase.

[00415] “Detect” refers to identifying the presence, absence or amount of the analyte to be detected.

[00416] By "detectable label" is meant a composition that when linked to a molecule of interest renders the latter detectable, via spectroscopic, photochemical, biochemical, immunochemical, or chemical means. For example, useful labels include radioactive isotopes, magnetic beads, metallic beads, colloidal particles, fluorescent dyes, electron-dense reagents, enzymes (for example, as commonly used in an ELISA), biotin, digoxigenin, or haptens.

[00417] By “disease” is meant any condition or disorder that damages or interferes with the normal function of a cell, tissue, or organ. In one embodiment, the disease is a neoplasia or cancer (e.g., multiple myeloma).

[00418] The term “effective amount,” as used herein, refers to an amount of a biologically active agent that is sufficient to elicit a desired biological response. In some embodiments, an effective amount of a fusion protein provided herein, e.g., of a cytidine deaminase or an

adenosine deaminase nucleobase editor comprising a nCas9 domain and one or more deaminase domains (e.g., cytidine deaminase, adenosine deaminase) may refer to the amount of the fusion protein that is sufficient to induce editing of a target site specifically bound and edited by the cytidine deaminase or adenosine deaminase nucleobase editors. As will be appreciated by the skilled artisan, the effective amount of an agent, e.g., a fusion protein, may vary depending on various factors as, for example, on the desired biological response, e.g., on the specific allele, genome, or target site to be edited, on the cell or tissue being targeted, and on the agent being used. In the context of a CAR-T cell, "an effective amount refers" to the quantity of cells necessary to administer to a patient to achieve a therapeutic response.

[00419] In some embodiments, an effective amount of a fusion protein provided herein, e.g., of a fusion protein comprising a nCas9 domain and a cytidine deaminase or adenosine deaminase may refer to the amount of the fusion protein that is sufficient to induce editing of a target site specifically bound and edited by the fusion protein. As will be appreciated by the skilled artisan, the effective amount of an agent, e.g., a fusion protein, a nuclease, a cytidine deaminase or adenosine deaminase, a hybrid protein, a protein dimer, a complex of a protein (or protein dimer) and a polynucleotide, or a polynucleotide, may vary depending on various factors as, for example, on the desired biological response, e.g., on the specific allele, genome, or target site to be edited, on the cell or tissue being targeted, and on the agent being used.

[00420] "Epitope," as used herein, means an antigenic determinant. An epitope is the part of an antigen molecule that by its structure determines the specific antibody molecule that will recognize and bind it.

[00421] By "fragment" is meant a portion of a polypeptide or nucleic acid molecule. This portion contains, at least about 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90% of the entire length of the reference nucleic acid molecule or polypeptide. A fragment may contain 10, 20, 30, 40, 50, 60, 70, 80, 90, or 100, 200, 300, 400, 500, 600, 700, 800, 900, or 1000 nucleotides or amino acids.

[00422] "Graft versus host disease" (GVHD) refers to a pathological condition where transplanted cells of a donor generate an immune response against cells of the host.

[00423] "Host versus graft disease" (HVGD) refers to a pathological condition where the immune system of a host generates an immune response against transplanted cells of a donor.

[00424] "Hybridization" means hydrogen bonding, which may be Watson-Crick, Hoogsteen or reversed Hoogsteen hydrogen bonding, between complementary nucleobases. For example, adenine and thymine are complementary nucleobases that pair through the formation of hydrogen bonds.

[00425] By "immune cell" is meant a cell of the immune system capable of generating an immune response.

[00426] By "immune effector cell" is meant a lymphocyte, once activated, capable of effecting an immune response upon a target cell. A T cell is an exemplary immune effector cell.

[00427] By "immune response regulation gene" or "immune response regulator" is meant a gene that encodes a polypeptide that is involved in regulation of a immune response. An immune response regulation gene may regulate immune response in multiple mechanisms or on different levels. For example, an immune response regulation gene may inhibit or facilitate the activation of an immune cell, e.g. a T cell. An immune response regulation gene may increase or decrease the activation threshold of a immune cell. In some embodiments, the immune response regulation gene positively regulates an immune cell signal transduction pathway. In some embodiments, the immune response regulation gene negatively regulates an immune cell signal transduction pathway. In some embodiments, the immune response regulation gene encodes an antigen, an antibody, a cytokine, or a neuroendocrine. In some embodiments, the immune response regulation gene encodes a Cblb protein.

[00428] By "immunogenic gene" is meant a gene that encodes a polypeptide that is able to elicit an immune response. For example, an immunogenic gene may encode an immunogen that elicits an immune response. In some embodiments, an immunogenic gene encodes a cell surface protein. In some embodiments, an immunogenic gene encodes a cell surface antigen or a cell surface marker. In some embodiments, the cell surface marker is a T cell marker or a B cell marker. In some embodiments, an immunogenic gene encodes a CD2, CD3e, CD3 delta, CD3 gamma, TRAC, TRBC1, TRBC2, CD4, CD5, CD7, CD8, CD19, CD23, CD27, CD28, CD30, CD33, CD52, CD70, CD127, CD122, CD130, CD132, CD38, CD69, CD11a, CD58, CD99, CD103, CCR4, CCR5, CCR6, CCR9, CCR10, CXCR3, CXCR4, CLA, CD161, B2M, or CIITA polypeptide.

[00429] The term "inhibitor of base repair" or "IBR" refers to a protein that is capable in inhibiting the activity of a nucleic acid repair enzyme, for example a base excision repair

enzyme. In some embodiments, the IBR is an inhibitor of inosine base excision repair. Exemplary inhibitors of base repair include inhibitors of APE1, Endo III, Endo IV, Endo V, Endo VIII, Fpg, hOGGl, hNEILl, T7 Endol, T4PDG, UDG, hSMUGl, and hAAG. In some embodiments, the IBR is an inhibitor of Endo V or hAAG. In some embodiments, the IBR is a catalytically inactive EndoV or a catalytically inactive hAAG.

[00430] The terms "isolated," "purified," or "biologically pure" refer to material that is free to varying degrees from components which normally accompany it as found in its native state. "Isolate" denotes a degree of separation from original source or surroundings. "Purify" denotes a degree of separation that is higher than isolation. A "purified" or "biologically pure" protein is sufficiently free of other materials such that any impurities do not materially affect the biological properties of the protein or cause other adverse consequences. That is, a nucleic acid or peptide of this invention is purified if it is substantially free of cellular material, viral material, or culture medium when produced by recombinant DNA techniques, or chemical precursors or other chemicals when chemically synthesized. Purity and homogeneity are typically determined using analytical chemistry techniques, for example, polyacrylamide gel electrophoresis or high-performance liquid chromatography. The term "purified" can denote that a nucleic acid or protein gives rise to essentially one band in an electrophoretic gel. For a protein that can be subjected to modifications, for example, phosphorylation or glycosylation, different modifications may give rise to different isolated proteins, which can be separately purified.

[00431] By "isolated polynucleotide" is meant a nucleic acid (e.g., a DNA) that is free of the genes which, in the naturally-occurring genome of the organism from which the nucleic acid molecule of the invention is derived, flank the gene. The term therefore includes, for example, a recombinant DNA that is incorporated into a vector; into an autonomously replicating plasmid or virus; or into the genomic DNA of a prokaryote or eukaryote; or that exists as a separate molecule (for example, a cDNA or a genomic or cDNA fragment produced by PCR or restriction endonuclease digestion) independent of other sequences. In addition, the term includes an RNA molecule that is transcribed from a DNA molecule, as well as a recombinant DNA that is part of a hybrid gene encoding additional polypeptide sequence.

[00432] By an "isolated polypeptide" is meant a polypeptide of the invention that has been separated from components that naturally accompany it. Typically, the polypeptide is isolated when it is at least 60%, by weight, free from the proteins and naturally-occurring organic

molecules with which it is naturally associated. Preferably, the preparation is at least 75%, more preferably at least 90%, and most preferably at least 99%, by weight, a polypeptide of the invention. An isolated polypeptide of the invention may be obtained, for example, by extraction from a natural source, by expression of a recombinant nucleic acid encoding such a polypeptide; or by chemically synthesizing the protein. Purity can be measured by any appropriate method, for example, column chromatography, polyacrylamide gel electrophoresis, or by HPLC analysis.

[00433] The term "linker," as used herein, refers to a bond (e.g., covalent bond), chemical group, or a molecule linking two molecules or moieties, e.g., two domains of a fusion protein, such as, for example, a nuclease-inactive Cas9 domain and a nucleic acid-editing domain (e.g., a cytidine deaminase, adenosine deaminase) or in the context of a chimeric antigen receptor, a linker linking a variable heavy (VH) region to a constant heavy (CH) region. In some embodiments, the linker joins two domains of a fusion protein, such as, for example, a nuclease-inactive Cas9 domain and a nucleic acid-editing domain (*e.g.*, a cytidine deaminase, adenosine deaminase). In some embodiments, a linker joins a gRNA binding domain of an RNA-programmable nuclease, including a Cas9 nuclease domain, and the catalytic domain of a nucleic-acid editing protein. In some embodiments, a linker joins a dCas9 and a nucleic-acid editing protein. Typically, the linker is positioned between, or flanked by, two groups, molecules, or other moieties and connected to each one via a covalent bond, thus connecting the two. In some embodiments, the linker is an amino acid or a plurality of amino acids (e.g., a peptide or protein). In some embodiments, the linker is an organic molecule, group, polymer, or chemical moiety. In some embodiments, the linker is 5-100 amino acids in length, for example, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 35, 45, 50, 55, 60, 60, 65, 70, 70, 75, 80, 85, 90, 90, 95, 100, 101, 102, 103, 104, 105, 110, 120, 130, 140, 150, 160, 175, 180, 190, or 200 amino acids in length. Longer or shorter linkers are also contemplated. In some embodiments, a linker comprises the amino acid sequence SGSETPGTSESATPES, which may also be referred to as the XTEN linker. In some embodiments, a linker comprises the amino acid sequence SGGS. In some embodiments, a linker comprises $(SGGS)_n$, $(GGGS)_n$, $(GGGGS)_n$, $(G)_n$, $(EAAAK)_n$, $(GGS)_n$, SGSETPGTSESATPES, or $(XP)_n$ motif, or a combination of any of these, wherein n is independently an integer between 1 and 30, and wherein X is any amino acid. In some embodiments, n is 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, or 15.

[00434] In some embodiments, the chimeric antigen receptor comprises at least one linker. The at least one linker joins, or links, a variable heavy (VH) region to a constant heavy (CH) region of the extracellular binding domain of the chimeric antigen receptor. Linkers can also link a variable light (VL) region to a variable constant (VC) region of the extracellular binding domain.

[00435] In some embodiments, the domains of the cytidine deaminase or adenosine deaminase nucleobase editor are fused via a linker that comprises the amino acid sequence of SGGSSGSETPGTSESATPESSGGS, SGGSSGGSSGSETPGTSESATPESSGGSSGGS, or GGSGGSPGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGSEPATSGGSGGS. In some embodiments, domains of the cytidine deaminase or adenosine deaminase nucleobase editor are fused via a linker comprising the amino acid sequence SGSETPGTSESATPES, which may also be referred to as the XTEN linker. In some embodiments, the linker is 24 amino acids in length. In some embodiments, the linker comprises the amino acid sequence SGGSSGGSSGSETPGTSESATPES. In some embodiments, the linker is 40 amino acids in length. In some embodiments, the linker comprises the amino acid sequence SGGSSGGSSGSETPGTSESATPESSGGSSGGSSGGSSGGS. In some embodiments, the linker is 64 amino acids in length. In some embodiments, the linker comprises the amino acid sequence SGGSSGGSSGSETPGTSESATPESSGGSSGGSSGGSSGGSSGSETPGTSESATPESSGGSSGGS. In some embodiments, the linker is 92 amino acids in length. In some embodiments, the linker comprises the amino acid sequence PGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGSEPATS.

[00436] By "marker" is meant any protein or polynucleotide having an alteration in expression level or activity that is associated with a disease or disorder.

[00437] The term "mutation," as used herein, refers to a substitution of a residue within a sequence, *e.g.*, a nucleic acid or amino acid sequence, with another residue, or a deletion or insertion of one or more residues within a sequence. Mutations are typically described herein by identifying the original residue followed by the position of the residue within the sequence and by the identity of the newly substituted residue. Various methods for making the amino acid substitutions (mutations) provided herein are well known in the art, and are provided by, for

example, Green and Sambrook, *Molecular Cloning: A Laboratory Manual* (4th ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (2012)).

[00438] "Neoplasia" refers to cells or tissues exhibiting abnormal growth or proliferation. The term neoplasia encompasses cancer and solid tumors.

[00439] By "nuclear factor of activated T cells 1 (NFATc1) polypeptide" is meant a protein having at least about 85% amino acid sequence identity to NCBI Accession No. NM_172390.2 or a fragment thereof and is a component of the activated T cell DNA-binding transcription complex. An exemplary amino acid sequence is provided below.

[00440] >NP_765978.1 nuclear factor of activated T-cells, cytoplasmic 1 isoform A [Homo sapiens]

MPSTSFPVPSKFPLGPAAAVFGRGETLGPAPRAGGTMKSAEEEHYGYASSNVSPALPLPT
AHSTLPAPCHNLQTSTPGIIPPADHPSGYGAALDGGPAGYFLSSGHTRPDGAPALESPRIE
ITSCLGLYHNNNQFFHDVEVEDVLPSSKRSPSTATLSLPSLEAYRDPSCLSPASSLSSRSC
NSEASSYESNYSYPYASPQTSPWQSPCVSPKTTDPEEGFPRGLGACTLLGSPRHSPSTSPR
ASVTEESWLGARSSRPASPCNKRKYSLNGRQPPYSPHHSPTPSPHGSPRVSVTDDSWLG
NTTQYTSSAIVAAINALTTDSSLDLGDGVPVKSRKTTLEQPPSVALKVEPVGEDLGSPPPP
ADFAPEDYSSFQHIRKGGFCDQYLAVPQHPYQWAKPKPLSPTSYMSPTLPALDWQLPSH
SGPYELRIEVQPKSHHRAHYETEGSRGAVKASAGGHPIVQLHGYLENEPLMLQLFIGTA
DDRLLRPHAFYQVHRITGKTVSTTSHEAILSNTKVLEIPLLPENSMRAVIDCAGILKLRNS
DIELRKGETDIGRKNTRVRLVFRVHVPQPSGRTLSLQVASNPIECSQRSAQELPLVEKQST
DSYPVVGGKKMVLSGHNFLQDSKVIFVEKAPDGHHVWEMEAKTDRDLCKPNSLVVEIP
PFRNQRITSPVHVSFYVCNGKRKRSQYQRFTYLPANGNAIFLTVSREHERVGCFF

[00441] By "nuclear factor of activated T cells 1 (NFATc1) polynucleotide" is meant a nucleic acid molecule encoding a NFATc1 polypeptide. The NFATc1 gene encodes a protein that is involved in in the inducible expression of cytokine genes, especially IL-2 and IL-4, in T-cells. An exemplary nucleic acid sequenced is provided below.

[00442] >NM_172390.2 Homo sapiens nuclear factor of activated T cells 1 (NFATC1), transcript variant 1, mRNA

GGCGGGCGCTCGGCGACTCGTCCCCGGGGCCCCGCGCGGGCCCGGGCAGCAGGGGCGTGAT
GTCACGGCA
GGGAGGGGGCGCGGGAGCCGCCGGGCCGGCGGGGAGGCGGGGGAGGTGTTTTCCAGCTTTA
AAAAGGCAG

GAGGCAGAGCGCGGCCCTGCGTCAGAGCGAGACTCAGAGGCTCCGAACTCGCCGGCGGAGT
CGCCGCGCC
AGATCCCAGCAGCAGGGCGCGGGCACCGGGGCGCGGGCAGGGCTCGGAGCCACCGCGCAG
GTCCTAGGGC
CGCGGCCGGGCCCCGCCACGCGCGCACACGCCCCTCGATGACTTTCCTCCGGGGCGCGCGGC
GCTGAGCC
CGGGGCGAGGGCTGTCTTCCCGGAGACCCGACCCCGGCAGCGCGGGGCGGCCGCTTCTCCT
GTGCCTCCG
CCCGCCGCTCCACTCCCCGCCGCCGCCGCGCGGATGCCAAGCACCAGCTTTCCAGTCCCTTC
CAAGTTTC
CACTTGGCCCTGCGGCTGCGGTCTTCGGGAGAGGAGAAACTTTGGGGCCCGCGCCGCGCGCC
GGCGGCAC
CATGAAGTCAGCGGAGGAAGAACACTATGGCTATGCATCCTCCAACGTCAGCCCCGCCCTGC
CGCTCCCC
ACGGCGCACTCCACCCTGCCGGCCCCGTGCCACAACCTTCAGACCTCCACACCGGGCATCAT
CCCGCCGG
CGGATCACCCCTCGGGGTACGGAGCAGCTTTGGACGGTGGGCCCGCGGGCTACTTCCTCTCC
TCCGGCCA
CACCAGGCCTGATGGGGCCCCTGCCCTGGAGAGTCCTCGCATCGAGATAACCTCGTGCTTGG
GCCTGTAC
CACAACAATAACCAGTTTTTCCACGATGTGGAGGTGGAAGACGTCCTCCCTAGCTCCAAACG
GTCCCCCT
CCACGGCCACGCTGAGTCTGCCCAGCCTGGAGGCCTACAGAGACCCCTCGTGCCTGAGCCCG
GCCAGCAG
CCTGTCCTCCCGGAGCTGCAACTCAGAGGCCTCCTCCTACGAGTCCAACTACTCGTACCCGT
ACGCGTCC
CCCCAGACGTCGCCATGGCAGTCTCCCTGCGTGTCTCCCAAGACCACGGACCCCGAGGAGGG
CTTTCCCC
GCGGGCTGGGGGCCTGCACACTGCTGGGTTCCCCGCGGCACTCCCCCTCCACCTCGCCCCGC
GCCAGCGT
CACTGAGGAGAGCTGGCTGGGTGCCCGCTCCTCCAGACCCGCGTCCCCTTGCAACAAGAGG
AAGTACAGC
CTCAACGGCCGGCAGCCGCCCTACTCACCCCACCACTCGCCCACGCCGTCCCCGCACGGCTC
CCCGCGGG

TCAGCGTGACCGACGACTCGTGGTTGGGCAACACCACCCAGTACACCAGCTCGGCCATCGTG
GCCGCCAT
CAACGCGCTGACCACCGACAGCAGCCTGGACCTGGGAGATGGCGTCCCTGTCAAGTCCCGC
AAGACCACC
CTGGAGCAGCCGCCCTCAGTGGCGCTCAAGGTGGAGCCCGTCGGGGAGGACCTGGGCAGCC
CCCCGCCCC
CGGCCGACTTCGCGCCCGAAGACTACTCCTCTTTCCAGCACATCAGGAAGGGCGGCTTCTGC
GACCAGTA
CCTGGCGGTGCCGCAGCACCCCTACCAGTGGGCGAAGCCCAAGCCCCTGTCCCCTACGTCCT
ACATGAGC
CCGACCCTGCCCGCCCTGGACTGGCAGCTGCCGTCCCACTCAGGCCCGTATGAGCTTCGGAT
TGAGGTGC
AGCCCAAGTCCCACCACCGAGCCCACTACGAGACGGAGGGCAGCCGGGGGGCCGTGAAGGC
GTCGGCCGG
AGGACACCCCATCGTGCAGCTGCATGGCTACTTGGAGAATGAGCCGCTGATGCTGCAGCTTT
TCATTGGG
ACGGCGGACGACCGCCTGCTGCGCCCGCACGCCTTCTACCAGGTGCACCGCATCACAGGGA
AGACCGTGT
CCACCACCAGCCACGAGGCCATCCTCTCCAACACCAAAGTCCTGGAGATCCCACTCCTGCCG
GAGAACAG
CATGCGAGCCGTCATTGACTGTGCCGGAATCCTGAAACTCAGAAACTCCGACATTGAACTTC
GGAAAGGA
GAGACGGACATCGGGAGGAAGAACACACGGGTACGGCTGGTGTTCCGCGTTCACGTCCCGC
AACCCAGCG
GCCGCACGCTGTCCCTGCAGGTGGCCTCCAACCCCATCGAATGCTCCCAGCGCTCAGCTCAG
GAGCTGCC
TCTGGTGGAGAAGCAGAGCACGGACAGCTATCCGGTCGTGGGCGGGAAGAAGATGGTCCTG
TCTGGCCAC
AACTTCCTGCAGGACTCCAAGGTCATTTTCGTGGAGAAAGCCCCAGATGGCCACCATGTCTG
GGAGATGG
AAGCGAAAACTGACCGGGACCTGTGCAAGCCGAATTCTCTGGTGGTTGAGATCCCGCCATTT
CGGAATCA
GAGGATAACCAGCCCCGTTCACGTCAGTTTCTACGTCTGCAACGGGAAGAGAAAGCGAAGC
CAGTACCAG

CGTTTCACCTACCTTCCCGCCAACGGTAACGCCATCTTTCTAACCGTAAGCCGTGAACATGA
GCGCGTGG
GGTGCTTTTTCTAAAGACGCAGAAACGACGTCGCCGTAAAGCAGCGTGGCGTGTTGCACATT
TAACTGTG
TGATGTCCCGTTAGTGAGACCGAGCCATCGATGCCCTGAAAAGGAAAGGAAAAGGGAAGCT
TCGGATGCA
TTTTCCTTGATCCCTGTTGGGGGTGGGGGGCGGGGGTTGCATACTCAGATAGTCACGGTTAT
TTTGCTTC
TTGCGAATGTATAACAGCCAAGGGGAAAACATGGCTCTTCTGCTCCAAAAAACTGAGGGGG
TCCTGGTGT
GCATTTGCACCCTAAAGCTGCTTACGGTGAAAAGGCAAATAGGTATAGCTATTTTGCAGGCA
CCTTTAGG AATAAACTTTGCTTTTAAGCCTGTAAAAAAAAAAAAAA

[00443] The term "nuclear localization sequence," "nuclear localization signal," or "NLS" refers to an amino acid sequence that promotes import of a protein into the cell nucleus. Nuclear localization sequences are known in the art and described, for example, in Plank *et al.*, International PCT application, PCT/EP2000/011690, filed November 23, 2000, published as WO/2001/038547 on May 31, 2001, the contents of which are incorporated herein by reference for their disclosure of exemplary nuclear localization sequences. In other embodiments, the NLS is an optimized NLS described, for example, by Koblan et al., Nature Biotech. 2018 doi:10.1038/nbt.4172. Optimized sequences useful in the methods of the invention are shown at FIGS. 8A-8E and 9. In some embodiments, an NLS comprises the amino acid sequence PKKKRKVEGADKRTADGSEFES PKKKRKV, KRTADGSEFESPKKKRKV, KRPAATKKAGQAKKKK, KKTELQTTNAENKTKKL, KRGINDRNFWRGENGRKTR, RKSGKIAAIVVKRPRK, PKKKRKV, or MDSLLMNRRKFLYQFKNVRWAKGRRETYLC.

[00444] The terms "nucleic acid" and "nucleic acid molecule," as used herein, refer to a compound comprising a nucleobase and an acidic moiety, e.g., a nucleoside, a nucleotide, or a polymer of nucleotides. Typically, polymeric nucleic acids, e.g., nucleic acid molecules comprising three or more nucleotides are linear molecules, in which adjacent nucleotides are linked to each other via a phosphodiester linkage. In some embodiments, "nucleic acid" refers to individual nucleic acid residues (*e.g.* nucleotides and/or nucleosides). In some embodiments, "nucleic acid" refers to an oligonucleotide chain comprising three or more individual nucleotide residues. As used herein, the terms "oligonucleotide" and "polynucleotide" can be used

interchangeably to refer to a polymer of nucleotides (e.g., a string of at least three nucleotides). In some embodiments, "nucleic acid" encompasses RNA as well as single and/or double-stranded DNA. Nucleic acids may be naturally occurring, for example, in the context of a genome, a transcript, an mRNA, tRNA, rRNA, siRNA, snRNA, a plasmid, cosmid, chromosome, chromatid, or other naturally occurring nucleic acid molecule. On the other hand, a nucleic acid molecule may be a non-naturally occurring molecule, *e.g.*, a recombinant DNA or RNA, an artificial chromosome, an engineered genome, or fragment thereof, or a synthetic DNA, RNA, DNA/RNA hybrid, or including non-naturally occurring nucleotides or nucleosides. Furthermore, the terms "nucleic acid," "DNA," "RNA," and/or similar terms include nucleic acid analogs, *e.g.,* analogs having other than a phosphodiester backbone. Nucleic acids can be purified from natural sources, produced using recombinant expression systems and optionally purified, chemically synthesized, *etc.* Where appropriate, e.g., in the case of chemically synthesized molecules, nucleic acids can comprise nucleoside analogs such as analogs having chemically modified bases or sugars, and backbone modifications. A nucleic acid sequence is presented in the 5′ to 3′ direction unless otherwise indicated. In some embodiments, a nucleic acid is or comprises natural nucleosides (*e.g.* adenosine, thymidine, guanosine, cytidine, uridine, deoxyadenosine, deoxythymidine, deoxyguanosine, and deoxycytidine); nucleoside analogs (e.g., 2-aminoadenosine, 2-thiothymidine, inosine, pyrrolo-pyrimidine, 3-methyl adenosine, 5-methylcytidine, 2-aminoadenosine, C5-bromouridine, C5-fluorouridine, C5-iodouridine, C5-propynyl-uridine, C5-propynyl-cytidine, C5-methylcytidine, 2-aminoadenosine, 7-deazaadenosine, 7-deazaguanosine, 8-oxoadenosine, 8-oxoguanosine, O(6)-methylguanine, and 2-thiocytidine); chemically modified bases; biologically modified bases (e.g., methylated bases); intercalated bases; modified sugars (., 2′-e.g.,fluororibose, ribose, 2′-deoxyribose, arabinose, and hexose); and/or modified phosphate groups (e.g., phosphorothioates and 5′-*N*-phosphoramidite linkages).

[00445] The term "nucleic acid programmable DNA binding protein" or "napDNAbp" refers to a protein that associates with a nucleic acid (e.g., DNA or RNA), such as a guide nucleic acid, that guides the napDNAbp to a specific nucleic acid sequence. For example, a Cas9 protein can associate with a guide RNA that guides the Cas9 protein to a specific DNA sequence that has complementary to the guide RNA. In some embodiments, the napDNAbp, the napDNAbp is a Cas9 domain, for example a nuclease active Cas9, a Cas9 nickase (nCas9), or a nuclease inactive

Cas9 (dCas9). Examples of nucleic acid programmable DNA binding proteins include, without limitation, Cas9 (e.g., dCas9 and nCas9), CasX, CasY, Cpfl, Cas12b/C2c1, and Cas12c/C2c3. Other nucleic acid programmable DNA binding proteins are also within the scope of this disclosure, though they may not be specifically listed in this disclosure.

[00446] As used herein, "obtaining" as in "obtaining an agent" includes synthesizing, purchasing, or otherwise acquiring the agent.

[00447] By "Programmed cell death 1 (PDCD1 or PD-1) polypeptide" is meant a protein having at least about 85% amino acid sequence identity to NCBI Accession No. AJS10360.1 or a fragment thereof. The PD-1 protein is thought to be involved in T cell function regulation during immune reactions and in tolerance conditions. An exemplary B2M polypeptide sequence is provided below.

[00448] >AJS10360.1 programmed cell death 1 protein [Homo sapiens]
MQIPQAPWPVVWAVLQLGWRPGWFLDSPDRPWNPPTFSPALLVVTEGDNATFTCSFSN
TSESFVLNWYRMSPSNQTDKLAAFPEDRSQPGQDCRFRVTQLPNGRDFHMSVVRARRN
DSGTYLCGAISLAPKAQIKESLRAELRVTERRAEVPTAHPSPSPRPAGQFQTLVVGVVGG
LLGSLVLLVWVLAVICSRAARGTIGARRTGQPLKEDPSAVPVFSVDYGELDFQWREKTP
EPPVPCVPEQTEYATIVFPSGMGTSSPARRGSADGPRSAQPLRPEDGHCSWPL

[00449] By "Programmed cell death 1 (PDCD1 or PD-1) polynucleotide" is meant a nucleic acid molecule encoding a PD-1 polypeptide. The PDCD1 gene encodes an inhibitory cell surface receptor that inhibits T-cell effector functions in an antigen-specific manner. An exemplary PDCD1 nucleic acid sequence is provided below.

[00450] AY238517.1 Homo sapiens programmed cell death 1 (PDCD1) mRNA, complete cds
ATGCAGATCCCACAGGCGCCCTGGCCAGTCGTCTGGGCGGTGCTACAACTGGGCTG
GCGGCCAGGATGGTTCTTAGACTCCCCAGACAGGCCCTGGAACCCCCCCACCTTCTC
CCCAGCCCTGCTCGTGGTGACCGAAGGGGACAACGCCACCTTCACCTGCAGCTTCTC
CAACACATCGGAGAGCTTCGTGCTAAACTGGTACCGCATGAGCCCCAGCAACCAGA
CGGACAAGCTGGCCGCCTTCCCCGAGGACCGCAGCCAGCCCGGCCAGGACTGCCGC
TTCCGTGTCACACAACTGCCCAACGGGCGTGACTTCCACATGAGCGTGGTCAGGGCC
CGGCGCAATGACAGCGGCACCTACCTCTGTGGGGCCATCTCCCTGGCCCCCAAGGC
GCAGATCAAAGAGAGCCTGCGGGCAGAGCTCAGGGTGACAGAGAGAAGGGCAGAA
GTGCCCACAGCCCACCCCAGCCCCTCACCCAGGCCAGCCGGCCAGTTCCAAACCCTG

GTGGTTGGTGTCGTGGGCGGCCTGCTGGGCAGCCTGGTGCTGCTAGTCTGGGTCCTG
GCCGTCATCTGCTCCCGGGCCGCACGAGGGACAATAGGAGCCAGGCGCACCGGCCA
GCCCCTGAAGGAGGACCCCTCAGCCGTGCCTGTGTTCTCTGTGGACTATGGGGAGCT
GGATTTCCAGTGGCGAGAGAAGACCCCGGAGCCCCCCGTGCCCTGTGTCCCTGAGC
AGACGGAGTATGCCACCATTGTCTTTCCTAGCGGAATGGGCACCTCATCCCCCGCCC
GCAGGGGCTCAGCTGACGGCCCTCGGAGTGCCCAGCCACTGAGGCCTGAGGATGGA
CACTGCTCTTGGCCCCTCTGA

[00451] The term “recombinant" as used herein in the context of proteins or nucleic acids refers to proteins or nucleic acids that do not occur in nature, but are the product of human engineering. For example, in some embodiments, a recombinant protein or nucleic acid molecule comprises an amino acid or nucleotide sequence that comprises at least one, at least two, at least three, at least four, at least five, at least six, or at least seven mutations as compared to any naturally occurring sequence.

[00452] By “reduces” or “increases” is meant a negative or positive alteration, respectively, of at least 10%, 25%, 50%, 75%, or 100%.

[00453] By “reference” is meant a standard or control condition.

[00454] A "reference sequence" is a defined sequence used as a basis for sequence comparison. A reference sequence may be a subset of or the entirety of a specified sequence; for example, a segment of a full-length cDNA or gene sequence, or the complete cDNA or gene sequence. For polypeptides, the length of the reference polypeptide sequence will generally be at least about 16 amino acids, at least about 20 amino acids, more at least about 25 amino acids, and even more preferably about 35 amino acids, about 50 amino acids, or about 100 amino acids. For nucleic acids, the length of the reference nucleic acid sequence will generally be at least about 50 nucleotides, at least about 60 nucleotides, at least about 75 nucleotides, and about 100 nucleotides or about 300 nucleotides or any integer thereabout or therebetween.

[00455] The term "RNA-programmable nuclease," and "RNA-guided nuclease" are used with (e.g., binds or associates with) one or more RNA(s) that is not a target for cleavage. In some embodiments, an RNA-programmable nuclease, when in a complex with an RNA, may be referred to as a nuclease:RNA complex. Typically, the bound RNA(s) is referred to as a guide RNA (gRNA). gRNAs can exist as a complex of two or more RNAs, or as a single RNA molecule. gRNAs that exist as a single RNA molecule may be referred to as single-guide RNAs

(sgRNAs), though "gRNA" is used interchangeably to refer to guide RNAs that exist as either single molecules or as a complex of two or more molecules. Typically, gRNAs that exist as single RNA species comprise two domains: (1) a domain that shares homology to a target nucleic acid (e.g., and directs binding of a Cas9 complex to the target); and (2) a domain that binds a Cas9 protein. In some embodiments, domain (2) corresponds to a sequence known as a tracrRNA, and comprises a stem-loop structure. For example, in some embodiments, domain (2) is identical or homologous to a tracrRNA as provided in Jinek et ah, Science 337:816-821(2012), the entire contents of which is incorporated herein by reference. Other examples of gRNAs (e.g., those including domain 2) can be found in U.S. Provisional Patent Application No. 61/874,682, filed September 6, 2013, entitled "Switchable Cas9 Nucleases and Uses Thereof," and U.S. Provisional Patent Application, No. 61/874,746, filed September 6, 2013, entitled "Delivery System For Functional Nucleases," the entire contents of each are hereby incorporated by reference in their entirety. In some embodiments, a gRNA comprises two or more of domains (1) and (2), and may be referred to as an "extended gRNA." For example, an extended gRNA will, e.g., bind two or more Cas9 proteins and bind a target nucleic acid at two or more distinct regions, as described herein. The gRNA comprises a nucleotide sequence that complements a target site, which mediates binding of the nuclease/RNA complex to said target site, providing the sequence specificity of the nuclease:RNA complex. In some embodiments, the RNA-programmable nuclease is the (CRIS PR-associated system) Cas9 endonuclease, for example, Cas9 (Csnl) from Streptococcus pyogenes (see, e.g., "Complete genome sequence of an Ml strain of Streptococcus pyogenes." Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K., Primeaux C, Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P., Qian Y., Jia H.G., Najar F.Z., Ren Q., Zhu H., Song L., White J., Yuan X., Clifton S.W., Roe B.A., McLaughlin R.E., Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001); "CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III." Deltcheva E., Chylinski K., Sharma CM., Gonzales K., Chao Y., Pirzada Z.A., Eckert M.R., Vogel J., Charpentier E., Nature 471:602-607(2011).

[00456] By "specifically binds" is meant a nucleic acid molecule, polypeptide, or complex thereof (e.g., a nucleic acid programmable DNA binding protein, a guide nucleic acid, and a chimeric antigen receptor), but which does not substantially recognize and bind other molecules in a sample, for example, a biological sample. For example, a chimeric antigen receptor

specifically binds to a particular marker expressed on the surface of a cell, but does not bind to other polypeptides, carbohydrates, lipids, or any other compound on the surface of the cell.

[00457] Nucleic acid molecules useful in the methods of the invention include any nucleic acid molecule that encodes a polypeptide of the invention or a fragment thereof. Such nucleic acid molecules need not be 100% identical with an endogenous nucleic acid sequence, but will typically exhibit substantial identity. Polynucleotides having "substantial identity" to an endogenous sequence are typically capable of hybridizing with at least one strand of a double-stranded nucleic acid molecule. Nucleic acid molecules useful in the methods of the invention include any nucleic acid molecule that encodes a polypeptide of the invention or a fragment thereof. Such nucleic acid molecules need not be 100% identical with an endogenous nucleic acid sequence, but will typically exhibit substantial identity. Polynucleotides having "substantial identity" to an endogenous sequence are typically capable of hybridizing with at least one strand of a double-stranded nucleic acid molecule. By "hybridize" is meant pair to form a double-stranded molecule between complementary polynucleotide sequences (e.g., a gene described herein), or portions thereof, under various conditions of stringency. (See, e.g., Wahl, G. M. and S. L. Berger (1987) Methods Enzymol. 152:399; Kimmel, A. R. (1987) Methods Enzymol. 152:507).

[00458] For example, stringent salt concentration will ordinarily be less than about 750 mM NaCl and 75 mM trisodium citrate, preferably less than about 500 mM NaCl and 50 mM trisodium citrate, and more preferably less than about 250 mM NaCl and 25 mM trisodium citrate. Low stringency hybridization can be obtained in the absence of organic solvent, e.g., formamide, while high stringency hybridization can be obtained in the presence of at least about 35% formamide, and more preferably at least about 50% formamide. Stringent temperature conditions will ordinarily include temperatures of at least about 30° C, more preferably of at least about 37° C, and most preferably of at least about 42° C. Varying additional parameters, such as hybridization time, the concentration of detergent, e.g., sodium dodecyl sulfate (SDS), and the inclusion or exclusion of carrier DNA, are well known to those skilled in the art. Various levels of stringency are accomplished by combining these various conditions as needed. In a one: embodiment, hybridization will occur at 30° C in 750 mM NaCl, 75 mM trisodium citrate, and 1% SDS. In another embodiment, hybridization will occur at 37° C in 500 mM NaCl, 50 mM trisodium citrate, 1% SDS, 35% formamide, and 100 μg/ml denatured salmon sperm DNA

(ssDNA). In another embodiment, hybridization will occur at 42° C in 250 mM NaCl, 25 mM trisodium citrate, 1% SDS, 50% formamide, and 200 μg/ml ssDNA. Useful variations on these conditions will be apparent to those skilled in the art.

[00459] For most applications, washing steps that follow hybridization will also vary in stringency. Wash stringency conditions can be defined by salt concentration and by temperature. As above, wash stringency can be increased by decreasing salt concentration or by increasing temperature. For example, stringent salt concentration for the wash steps will preferably be less than about 30 mM NaCl and 3 mM trisodium citrate, and most preferably less than about 15 mM NaCl and 1.5 mM trisodium citrate. Stringent temperature conditions for the wash steps will ordinarily include a temperature of at least about 25° C, more preferably of at least about 42° C, and even more preferably of at least about 68° C. In an embodiment, wash steps will occur at 25° C in 30 mM NaCl, 3 mM trisodium citrate, and 0.1% SDS. In a more preferred embodiment, wash steps will occur at 42 C in 15 mM NaCl, 1.5 mM trisodium citrate, and 0.1% SDS. In a more preferred embodiment, wash steps will occur at 68° C in 15 mM NaCl, 1.5 mM trisodium citrate, and 0.1% SDS. Additional variations on these conditions will be apparent to those skilled in the art. Hybridization techniques are well known to those skilled in the art and are described, for example, in Benton and Davis (Science 196:180, 1977); Grunstein and Hogness (Proc. Natl. Acad. Sci., USA 72:3961, 1975); Ausubel et al. (Current Protocols in Molecular Biology, Wiley Interscience, New York, 2001); Berger and Kimmel (Guide to Molecular Cloning Techniques, 1987, Academic Press, New York); and Sambrook et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, New York.

[00460] By "subject" is meant a mammal, including, but not limited to, a human or non-human mammal, such as a bovine, equine, canine, ovine, or feline. Subjects include livestock, domesticated animals raised to produce labor and to provide commodities, such as food, including without limitation, cattle, goats, chickens, horses, pigs, rabbits, and sheep.

[00461] By "substantially identical" is meant a polypeptide or nucleic acid molecule exhibiting at least 50% identity to a reference amino acid sequence (for example, any one of the amino acid sequences described herein) or nucleic acid sequence (for example, any one of the nucleic acid sequences described herein). In one embodiment, such a sequence is at least 60%, 80% or 85%, 90%, 95% or even 99% identical at the amino acid level or nucleic acid to the sequence used for comparison.

[00462] Sequence identity is typically measured using sequence analysis software (for example, Sequence Analysis Software Package of the Genetics Computer Group, University of Wisconsin Biotechnology Center, 1710 University Avenue, Madison, Wis. 53705, BLAST, BESTFIT, GAP, or PILEUP/PRETTYBOX programs). Such software matches identical or similar sequences by assigning degrees of homology to various substitutions, deletions, and/or other modifications. Conservative substitutions typically include substitutions within the following groups: glycine, alanine; valine, isoleucine, leucine; aspartic acid, glutamic acid, asparagine, glutamine; serine, threonine; lysine, arginine; and phenylalanine, tyrosine. In an exemplary approach to determining the degree of identity, a BLAST program may be used, with a probability score between e^{-3} and e^{-100} indicating a closely related sequence.

[00463] Because RNA-programmable nucleases (e.g., Cas9) use RNA:DNA hybridization to target DNA cleavage sites, these proteins can be targeted, in principle, to any sequence specified by the guide RNA. Methods of using RNA-programmable nucleases, such as Cas9, for site-specific cleavage (e.g., to modify a genome) are known in the art (see e.g., Cong, L. et ah, Multiplex genome engineering using CRISPR/Cas systems. Science 339, 819-823 (2013); Mali, P. et ah, RNA-guided human genome engineering via Cas9. Science 339, 823-826 (2013); Hwang, W.Y. et ah, Efficient genome editing in zebrafish using a CRISPR-Cas system. Nature biotechnology 31, 227-229 (2013); Jinek, M. et ah, RNA-programmed genome editing in human cells. eLife 2, e00471 (2013); Dicarlo, J.E. et ah, Genome engineering in Saccharomyces cerevisiae using CRISPR-Cas systems. Nucleic acids research (2013); Jiang, W. et ah RNA-guided editing of bacterial genomes using CRISPR-Cas systems. Nature biotechnology 31, 233-239 (2013); the entire contents of each of which are incorporated herein by reference).

[00464] By "tet methylcytosine dioxygenase 2 (TET2) polypeptide" is meant a protein having at least about 85% amino acid sequence identity to NCBI Accession No. FM992369.1 or a fragment thereof and having catalytic activity to convert methylcytosine to 5-hydroxymethylcytosine. Defects in the gene have been associated with myeloproliferative disorders, and the enzyme's ability to methylate cytosine contributes to transcriptional regulation. An exemplary TET2 amino acid sequence is provided below.

>CAX30492.1 tet oncogene family member 2 [Homo sapiens]

MEQDRTNHVEGNRLSPFLIPSPPICQTEPLATKLQNGSPLPERAHPEVNGDTKWHSFKSY
YGIPCMKGSQNSRVSPDFTQESRGYSKCLQNGGIKRTVSEPSLSGLLQIKKLKQDQKAN

GERRNFGVSQERNPGESSQPNVSDLSDKKESVSSVAQENAVKDFTSFSTHNCSGPENPEL
QILNEQEGKSANYHDKNIVLLKNKAVLMPNGATVSASSVEHTHGELLEKTLSQYYPDC
VSIAVQKTTSHINAINSQATNELSCEITHPSHTSGQINSAQTSNSELPPKPAAVVSEACDA
DDADNASKLAAMLNTCSFQKPEQLQQQKSVFEICPSPAENNIQGTTKLASGEEFCSGSSS
NLQAPGGSSERYLKQNEMNGAYFKQSSVFTKDSFSATTTPPPPSQLLLSPPPPLPQVPQLP
SEGKSTLNGGVLEEHHHYPNQSNTTLLREVKIEGKPEAPPSQSPNPSTHVCSPSPMLSERP
QNNCVNRNDIQTAGTMTVPLCSEKTRPMSEHLKHNPPIFGSSGELQDNCQQLMRNKEQ
EILKGRDKEQTRDLVPPTQHYLKPGWIELKAPRFHQAESHLKRNEASLPSILQYQPNLSN
QMTSKQYTGNSNMPGGLPRQAYTQKTTQLEHKSQMYQVEMNQGQSQGTVDQHLQFQ
KPSHQVHFSKTDHLPKAHVQSLCGTRFHFQQRADSQTEKLMSPVLKQHLNQQASETEPF
SNSHLLQHKPHKQAAQTQPSQSSHLPQNQQQQQKLQIKNKEEILQTFPHPQSNNDQQRE
GSFFGQTKVEECFHGENQYSKSSEFETHNVQMGLEEVQNINRRNSPYSQTMKSSACKIQ
VSCSNNTHLVSENKEQTTHPELFAGNKTQNLHHMQYFPNNVIPKQDLLHRCFQEQEQK
SQQASVLQGYKNRNQDMSGQQAAQLAQQRYLIHNHANVFPVPDQGGSHTQTPPQKDT
QKHAALRWHLLQKQEQQQTQQPQTESCHSQMHRPIKVEPGCKPHACMHTAPPENKTW
KKVTKQENPPASCDNVQQKSIIETMEQHLKQFHAKSLFDHKALTLKSQKQVKVEMSGP
VTVLTRQTTAAELDSHTPALEQQTTSSEKTPTKRTAASVLNNFIESPSKLLDTPIKNLLDT
PVKTQYDFPSCRCVEQIIEKDEGPFYTHLGAGPNVAAIREIMEERFGQKGKAIRIERVIYT
GKEGKSSQGCPIAKWVVRRSSSEEKLLCLVRERAGHTCEAAVIVILILVWEGIPLSLADK
LYSELTETLRKYGTLTNRRCALNEERTCACQGLDPETCGASFSFGCSWSMYYNGCKFA
RSKIPRKFKLLGDDPKEEEKLESHLQNLSTLMAPTYKKLAPDAYNNQIEYEHRAPECRL
GLKEGRPFSGVTACLDFCAHAHRDLHNMQNGSTLVCTLTREDNREFGGKPEDEQLHVL
PLYKVSDVDEFGSVEAQEEKKRSGAIQVLSSFRRKVRMLAEPVKTCRQRKLEAKKAAA
EKLSSLENSSNKNEKEKSAPSRTKQTENASQAKQLAELLRLSGPVMQQSQQPQPLQKQP
PQPQQQQRPQQQQPHHPQTESVNSYSASGSTNPYMRRPNPVSPYPNSSHTSDIYGSTSPM
NFYSTSSQAAGSYLNSSNPMNPYPGLLNQNTQYPSYQCNGNLSVDNCSPYLGSYSPQSQ
PMDLYRYPSQDPLSKLSLPPIHTLYQPRFGNSQSFTSKYLGYGNQNMQGDGFSSCTIRPN
VHHVGKLPPYPTHEMDGHFMGATSRLPPNLSNPNMDYKNGEHHSPSHIIHNYSAAPGM
FNSSLHALHLQNKENDMLSHTANGLSKMLPALNHDRTACVQGGLHKLSDANGQEKQP
LALVQGVASGAEDNDEVWSDSEQSFLDPDIGGVAVAPTHGSILIECAKRELHATTPLKN

PNRNHPTRISLVFYQHKSMNEPKHGLALWEAKMAEKAREKEEECEKYGPDYVPQKSH
GKKVKREPAEPHETSEPTYLRFIKSLAERTMSVTTDSTVTTSPYAFTRVTGPYNRYI

[00465] By "tet methylcytosine dioxygenase 2 (TET2) polynucleotide" is meant a nucleic acid molecule encoding a TET2 polypeptide. The TETs polypeptide encodes a methylcytosine dioxygenase and has transcription regulatory activity. An exemplary TET2 nucleic acid is presented below.

>FM992369.1 Homo sapiens mRNA for tet oncogene family member 2 (TET2 gene)
CCGTGCCATCCCAACCTCCCACCTCGCCCCCAACCTTCGCGCTTGCTCTGCTTCTTCT
CCCAGGGGTGGAGACCCGCCGAGGTCCCCGGGGTTCCCGAGGGCTGCACCCTTCCC
CGCGCTCGCCAGCCCTGGCCCCTACTCCGCGCTGGTCCGGGCGCACCACTCCCCCCG
CGCCACTGCACGGCGTGAGGGCAGCCCAGGTCTCCACTGCGCGCCCCGCTGTACGG
CCCCAGGTGCCGCCGGCCTTTGTGCTGGACGCCCGGTGCGGGGGGCTAATTCCCTGG
GAGCCGGGGCTGAGGGCCCCAGGGCGGCGGCGCAGGCCGGGGCGGAGCGGGAGGA
GGCCGGGGCGGAGCAGGAGGAGGCCCGGGCGGAGGAGGAGAGCCGGCGGTAGCGG
CAGTGGCAGCGGCGAGAGCTTGGGCGGCCGCCGCCGCCTCCTCGCGAGCGCCGCGC
GCCCGGGTCCCGCTCGCATGCAAGTCACGTCCGCCCCCTCGGCGCGGCCGCCCCGAG
ACGCCGGCCCCGCTGAGTGATGAGAACAGACGTCAAACTGCCTTATGAATATTGAT
GCGGAGGCTAGGCTGCTTTCGTAGAGAAGCAGAAGGAAGCAAGATGGCTGCCCTTT
AGGATTTGTTAGAAAGGAGACCCGACTGCAACTGCTGGATTGCTGCAAGGCTGAGG
GACGAGAACGAGGCTGGCAAACATTCAGCAGCACACCCTCTCAAGATTGTTTACTTG
CCTTTGCTCCTGTTGAGTTACAACGCTTGGAAGCAGGAGATGGGCTCAGCAGCAGCC
AATAGGACATGATCCAGGAAGAGCAAATTCAACTAGAGGGCAGCCTTGTGGATGGC
CCCGAAGCAAGCCTGATGGAACAGGATAGAACCAACCATGTTGAGGGCAACAGACT
AAGTCCATTCCTGATACCATCACCTCCCATTTGCCAGACAGAACCTCTGGCTACAAA
GCTCCAGAATGGAAGCCCACTGCCTGAGAGAGCTCATCCAGAAGTAAATGGAGACA
CCAAGTGGCACTCTTTCAAAAGTTATTATGGAATACCCTGTATGAAGGGAAGCCAGA
ATAGTCGTGTGAGTCCTGACTTTACACAAGAAAGTAGAGGGTATTCCAAGTGTTTGC
AAAATGGAGGAATAAAACGCACAGTTAGTGAACCTTCTCTCTCTGGGCTCCTTCAGA
TCAAGAAATTGAAACAAGACCAAAAGGCTAATGGAGAAAGACGTAACTTCGGGGTA
AGCCAAGAAAGAAATCCAGGTGAAAGCAGTCAACCAAATGTCTCCGATTTGAGTGA
TAAGAAAGAATCTGTGAGTTCTGTAGCCCAAGAAAATGCAGTTAAAGATTTCACCA

GTTTTTCAACACATAACTGCAGTGGGCCTGAAAATCCAGAGCTTCAGATTCTGAATG
AGCAGGAGGGGAAAAGTGCTAATTACCATGACAAGAACATTGTATTACTTAAAAAC
AAGGCAGTGCTAATGCCTAATGGTGCTACAGTTTCTGCCTCTTCCGTGGAACACACA
CATGGTGAACTCCTGGAAAAAACACTGTCTCAATATTATCCAGATTGTGTTTCCATT
GCGGTGCAGAAAACCACATCTCACATAAATGCCATTAACAGTCAGGCTACTAATGA
GTTGTCCTGTGAGATCACTCACCCATCGCATACCTCAGGGCAGATCAATTCCGCACA
GACCTCTAACTCTGAGCTGCCTCCAAAGCCAGCTGCAGTGGTGAGTGAGGCCTGTGA
TGCTGATGATGCTGATAATGCCAGTAAACTAGCTGCAATGCTAAATACCTGTTCCTT
TCAGAAACCAGAACAACTACAACAACAAAAATCAGTTTTTGAGATATGCCCATCTCC
TGCAGAAAATAACATCCAGGGAACCACAAAGCTAGCGTCTGGTGAAGAATTCTGTT
CAGGTTCCAGCAGCAATTTGCAAGCTCCTGGTGGCAGCTCTGAACGGTATTTAAAAC
AAAATGAAATGAATGGTGCTTACTTCAAGCAAAGCTCAGTGTTCACTAAGGATTCCT
TTTCTGCCACTACCACACCACCACCACCATCACAATTGCTTCTTTCTCCCCCTCCTCC
TCTTCCACAGGTTCCTCAGCTTCCTTCAGAAGGAAAAAGCACTCTGAATGGTGGAGT
TTTAGAAGAACACCACCACTACCCCAACCAAAGTAACACAACACTTTTAAGGGAAG
TGAAAATAGAGGGTAAACCTGAGGCACCACCTTCCCAGAGTCCTAATCCATCTACA
CATGTATGCAGCCCTTCTCCGATGCTTTCTGAAAGGCCTCAGAATAATTGTGTGAAC
AGGAATGACATACAGACTGCAGGGACAATGACTGTTCCATTGTGTTCTGAGAAAAC
AAGACCAATGTCAGAACACCTCAAGCATAACCCACCAATTTTTGGTAGCAGTGGAG
AGCTACAGGACAACTGCCAGCAGTTGATGAGAAACAAAGAGCAAGAGATTCTGAAG
GGTCGAGACAAGGAGCAAACACGAGATCTTGTGCCCCCAACACAGCACTATCTGAA
ACCAGGATGGATTGAATTGAAGGCCCCTCGTTTTCACCAAGCGGAATCCCATCTAAA
ACGTAATGAGGCATCACTGCCATCAATTCTTCAGTATCAACCCAATCTCTCCAATCA
AATGACCTCCAAACAATACACTGGAAATTCCAACATGCCTGGGGGGCTCCCAAGGC
AAGCTTACACCCAGAAAACAACACAGCTGGAGCACAAGTCACAAATGTACCAAGTT
GAAATGAATCAAGGGCAGTCCCAAGGTACAGTGGACCAACATCTCCAGTTCCAAAA
ACCCTCACACCAGGTGCACTTCTCCAAAACAGACCATTTACCAAAAGCTCATGTGCA
GTCACTGTGTGGCACTAGATTTCATTTTCAACAAAGAGCAGATTCCCAAACTGAAAA
ACTTATGTCCCCAGTGTTGAAACAGCACTTGAATCAACAGGCTTCAGAGACTGAGCC
ATTTTCAAACTCACACCTTTTGCAACATAAGCCTCATAAACAGGCAGCACAAACACA
ACCATCCCAGAGTTCACATCTCCCTCAAAACCAGCAACAGCAGCAAAAATTACAAA

TAAAGAATAAAGAGGAAATACTCCAGACTTTTCCTCACCCCCAAAGCAACAATGAT
CAGCAAAGAGAAGGATCATTCTTTGGCCAGACTAAAGTGGAAGAATGTTTTCATGG
TGAAAATCAGTATTCAAAATCAAGCGAGTTCGAGACTCATAATGTCCAAATGGGAC
TGGAGGAAGTACAGAATATAAATCGTAGAAATTCCCCTTATAGTCAGACCATGAAA
TCAAGTGCATGCAAAATACAGGTTTCTTGTTCAAACAATACACACCTAGTTTCAGAG
AATAAAGAACAGACTACACATCCTGAACTTTTTGCAGGAAACAAGACCCAAAACTT
GCATCACATGCAATATTTTCCAAATAATGTGATCCCAAAGCAAGATCTTCTTCACAG
GTGCTTTCAAGAACAGGAGCAGAAGTCACAACAAGCTTCAGTTCTACAGGGATATA
AAAATAGAAACCAAGATATGTCTGGTCAACAAGCTGCGCAACTTGCTCAGCAAAGG
TACTTGATACATAACCATGCAAATGTTTTTCCTGTGCCTGACCAGGGAGGAAGTCAC
ACTCAGACCCCTCCCCAGAAGGACACTCAAAAGCATGCTGCTCTAAGGTGGCATCTC
TTACAGAAGCAAGAACAGCAGCAAACACAGCAACCCCAAACTGAGTCTTGCCATAG
TCAGATGCACAGGCCAATTAAGGTGGAACCTGGATGCAAGCCACATGCCTGTATGC
ACACAGCACCACCAGAAAACAAAACATGGAAAAAGGTAACTAAGCAAGAGAATCC
ACCTGCAAGCTGTGATAATGTGCAGCAAAAGAGCATCATTGAGACCATGGAGCAGC
ATCTGAAGCAGTTTCACGCCAAGTCGTTATTTGACCATAAGGCTCTTACTCTCAAAT
CACAGAAGCAAGTAAAAGTTGAAATGTCAGGGCCAGTCACAGTTTTGACTAGACAA
ACCACTGCTGCAGAACTTGATAGCCACACCCCAGCTTTAGAGCAGCAAACAACTTCT
TCAGAAAAGACACCAACCAAAAGAACAGCTGCTTCTGTTCTCAATAATTTTATAGAG
TCACCTTCCAAATTACTAGATACTCCTATAAAAAATTTATTGGATACACCTGTCAAG
ACTCAATATGATTTCCCATCTTGCAGATGTGTAGAGCAAATTATTGAAAAAGATGAA
GGTCCTTTTTATACCCATCTAGGAGCAGGTCCTAATGTGGCAGCTATTAGAGAAATC
ATGGAAGAAAGGTTTGGACAGAAGGGTAAAGCTATTAGGATTGAAAGAGTCATCTA
TACTGGTAAAGAAGGCAAAAGTTCTCAGGGATGTCCTATTGCTAAGTGGGTGGTTCG
CAGAAGCAGCAGTGAAGAGAAGCTACTGTGTTTGGTGCGGGAGCGAGCTGGCCACA
CCTGTGAGGCTGCAGTGATTGTGATTCTCATCCTGGTGTGGGAAGGAATCCCGCTGT
CTCTGGCTGACAAACTCTACTCGGAGCTTACCGAGACGCTGAGGAAATACGGCACG
CTCACCAATCGCCGGTGTGCCTTGAATGAAGAGAGAACTTGCGCCTGTCAGGGGCTG
GATCCAGAAACCTGTGGTGCCTCCTTCTCTTTTGGTTGTTCATGGAGCATGTACTACA
ATGGATGTAAGTTTGCCAGAAGCAAGATCCCAAGGAAGTTTAAGCTGCTTGGGGAT
GACCCAAAAGAGGAAGAGAAACTGGAGTCTCATTTGCAAAACCTGTCCACTCTTAT

GGCACCAACATATAAGAAACTTGCACCTGATGCATATAATAATCAGATTGAATATG
AACACAGAGCACCAGAGTGCCGTCTGGGTCTGAAGGAAGGCCGTCCATTCTCAGGG
GTCACTGCATGTTTGGACTTCTGTGCTCATGCCCACAGAGACTTGCACAACATGCAG
AATGGCAGCACATTGGTATGCACTCTCACTAGAGAAGACAATCGAGAATTTGGAGG
AAAACCTGAGGATGAGCAGCTTCACGTTCTGCCTTTATACAAAGTCTCTGACGTGGA
TGAGTTTGGGAGTGTGGAAGCTCAGGAGGAGAAAAAACGGAGTGGTGCCATTCAGG
TACTGAGTTCTTTTCGGCGAAAAGTCAGGATGTTAGCAGAGCCAGTCAAGACTTGCC
GACAAAGGAAACTAGAAGCCAAGAAAGCTGCAGCTGAAAAGCTTTCCTCCCTGGAG
AACAGCTCAAATAAAAATGAAAAGGAAAAGTCAGCCCCATCACGTACAAAACAAA
CTGAAAACGCAAGCCAGGCTAAACAGTTGGCAGAACTTTTGCGACTTTCAGGACCA
GTCATGCAGCAGTCCCAGCAGCCCCAGCCTCTACAGAAGCAGCCACCACAGCCCCA
GCAGCAGCAGAGACCCCAGCAGCAGCAGCCACATCACCCTCAGACAGAGTCTGTCA
ACTCTTATTCTGCTTCTGGATCCACCAATCCATACATGAGACGGCCCAATCCAGTTA
GTCCTTATCCAAACTCTTCACACACTTCAGATATCTATGGAAGCACCAGCCCTATGA
ACTTCTATTCCACCTCATCTCAAGCTGCAGGTTCATATTTGAATTCTTCTAATCCCAT
GAACCCTTACCCTGGGCTTTTGAATCAGAATACCCAATATCCATCATATCAATGCAA
TGGAAACCTATCAGTGGACAACTGCTCCCCATATCTGGGTTCCTATTCTCCCCAGTCT
CAGCCGATGGATCTGTATAGGTATCCAAGCCAAGACCCTCTGTCTAAGCTCAGTCTA
CCACCCATCCATACACTTTACCAGCCAAGGTTTGGAAATAGCCAGAGTTTTACATCT
AAATACTTAGGTTATGGAAACCAAAATATGCAGGGAGATGGTTTCAGCAGTTGTAC
CATTAGACCAAATGTACATCATGTAGGGAAATTGCCTCCTTATCCCACTCATGAGAT
GGATGGCCACTTCATGGGAGCCACCTCTAGATTACCACCCAATCTGAGCAATCCAAA
CATGGACTATAAAAATGGTGAACATCATTCACCTTCTCACATAATCCATAACTACAG
TGCAGCTCCGGGCATGTTCAACAGCTCTCTTCATGCCCTGCATCTCCAAAACAAGGA
GAATGACATGCTTTCCCACACAGCTAATGGGTTATCAAAGATGCTTCCAGCTCTTAA
CCATGATAGAACTGCTTGTGTCCAAGGAGGCTTACACAAATTAAGTGATGCTAATGG
TCAGGAAAAGCAGCCATTGGCACTAGTCCAGGGTGTGGCTTCTGGTGCAGAGGACA
ACGATGAGGTCTGGTCAGACAGCGAGCAGAGCTTTCTGGATCCTGACATTGGGGGA
GTGGCCGTGGCTCCAACTCATGGGTCAATTCTCATTGAGTGTGCAAAGCGTGAGCTG
CATGCCACAACCCCTTTAAAGAATCCCAATAGGAATCACCCCACCAGGATCTCCCTC
GTCTTTTACCAGCATAAGAGCATGAATGAGCCAAAACATGGCTTGGCTCTTTGGGAA

GCCAAAATGGCTGAAAAAGCCCGTGAGAAAGAGGAAGAGTGTGAAAAGTATGGCC
CAGACTATGTGCCTCAGAAATCCCATGGCAAAAAAGTGAAACGGGAGCCTGCTGAG
CCACATGAAACTTCAGAGCCCACTTACCTGCGTTTCATCAAGTCTCTTGCCGAAAGG
ACCATGTCCGTGACCACAGACTCCACAGTAACTACATCTCCATATGCCTTCACTCGG
GTCACAGGGCCTTACAACAGATATATATGAAGATATATATGATATCACCCCCTTTTG
TTGGTTACCTCACTTGAAAAGACCACAACCAACCTGTCAGTAGTATAGTTCTCATGA
CGTGGGCAGTGGGGAAAGGTCACAGTATTCATGACAAATGTGGTGGGAAAAACCTC
AGCTCACCAGCAACAAAAGAGGTTATCTTACCATAGCACTTAATTTTCACTGGCTCC
CAAGTGGTCACAGATGGCATCTAGGAAAAGACCAAAGCATTCTATGCAAAAAGAAG
GTGGGGAAGAAAGTGTTCCGCAATTTACATTTTTAAACACTGGTTCTATTATTGGAC
GAGATGATATGTAAATGTGATCCCCCCCCCCCGCTTACAACTCTACACATCTGTGAC
CACTTTTAATAATATCAAGTTTGCATAGTCATGGAACACAAATCAAACAAGTACTGT
AGTATTACAGTGACAGGAATCTTAAAATACCATCTGGTGCTGAATATATGATGTACT
GAAATACTGGAATTATGGCTTTTTGAAATGCAGTTTTTACTGTAATCTTAACTTTTAT
TTATCAAAATAGCTACAGGAAACATGAATAGCAGGAAAACACTGAATTTGTTTGGA
TGTTCTAAGAAATGGTGCTAAGAAAATGGTGTCTTTAATAGCTAAAAATTTAATGCC
TTTATATCATCAAGATGCTATCAGTGTACTCCAGTGCCCTTGAATAATAGGGGTACC
TTTTCATTCAAGTTTTTATCATAATTACCTATTCTTACACAAGCTTAGTTTTTAAAATG
TGGACATTTTAAAGGCCTCTGGATTTTGCTCATCCAGTGAAGTCCTTGTAGGACAAT
AAACGTATATATGTACATATATACACAAACATGTATATGTGCACACACATGTATATG
TATAAATATTTTAAATGGTGTTTTAGAAGCACTTTGTCTACCTAAGCTTTGACAACTT
GAACAATGCTAAGGTACTGAGATGTTTAAAAAACAAGTTTACTTTCATTTTAGAATG
CAAAGTTGATTTTTTTAAGGAAACAAAGAAAGCTTTTAAAATATTTTTGCTTTTAGCC
ATGCATCTGCTGATGAGCAATTGTGTCCATTTTTAACACAGCCAGTTAAATCCACCA
TGGGGCTTACTGGATTCAAGGGAATACGTTAGTCCACAAAACATGTTTTCTGGTGCT
CATCTCACATGCTATACTGTAAAACAGTTTTATACAAAATTGTATGACAAGTTCATT
GCTCAAAAATGTACAGTTTTAAGAATTTTCTATTAACTGCAGGTAATAATTAGCTGC
ATGCTGCAGACTCAACAAAGCTAGTTCACTGAAGCCTATGCTATTTTATGGATCATA
GGCTCTTCAGAGAACTGAATGGCAGTCTGCCTTTGTGTTGATAATTATGTACATTGT
GACGTTGTCATTTCTTAGCTTAAGTGTCCTCTTTAACAAGAGGATTGAGCAGACTGA
TGCCTGCATAAGATGAATAAACAGGGTTAGTTCCATGTGAATCTGTCAGTTAAAAAG

AAACAAAAACAGGCAGCTGGTTTGCTGTGGTGGTTTTAAATCATTAATTTGTATAAA
GAAGTGAAAGAGTTGTATAGTAAATTAAATTGTAAACAAAACTTTTTTAATGCAATG
CTTTAGTATTTTAGTACTGTAAAAAAATTAAATATATACATATATATATATATATATA
TATATATATATATGAGTTTGAAGCAGAATTCACATCATGATGGTGCTACTCAGCCTG
CTACAAATATATCATAATGTGAGCTAAGAATTCATTAAATGTTTGAGTGATGTTCCT
ACTTGTCATATACCTCAACACTAGTTTGGCAATAGGATATTGAACTGAGAGTGAAAG
CATTGTGTACCATCATTTTTTTCCAAGTCCTTTTTTTTATTGTTAAAAAAAAAAGCAT
ACCTTTTTTCAATACTTGATTTCTTAGCAAGTATAACTTGAACTTCAACCTTTTTGTTC
TAAAAATTCAGGGATATTTCAGCTCATGCTCTCCCTATGCCAACATGTCACCTGTGTT
TATGTAAAATTGTTGTAGGTTAATAAATATATTCTTTGTCAGGGATTTAACCCTTTTA
TTTTGAATCCCTTCTATTTTACTTGTACATGTGCTGATGTAACTAAAACTAATTTTGT
AAATCTGTTGGCTCTTTTTATTGTAAAGAAAAGCATTTTAAAAGTTTGAGGAATCTTT
TGACTGTTTCAAGCAGGAAAAAAAAATTACATGAAAATAGAATGCACTGAGTTGAT
AAAGGGAAAAATTGTAAGGCAGGAGTTTGGCAAGTGGCTGTTGGCCAGAGACTTAC
TTGTAACTCTCTAAATGAAGTTTTTTTGATCCTGTAATCACTGAAGGTACATACTCCA
TGTGGACTTCCCTTAAACAGGCAAACACCTACAGGTATGGTGTGCAACAGATTGTAC
AATTACATTTTGGCCTAAATACATTTTTGCTTACTAGTATTTAAAATAAATTCTTAAT
CAGAGGAGGCCTTTGGGTTTTATTGGTCAAATCTTTGTAAGCTGGCTTTTGTCTTTTT
AAAAAATTTCTTGAATTTGTGGTTGTGTCCAATTTGCAAACATTTCCAAAAATGTTTG
CTTTGCTTACAAACCACATGATTTTAATGTTTTTTGTATACCATAATATCTAGCCCCA
AACATTTGATTACTACATGTGCATTGGTGATTTTGATCATCCATTCTTAATATTTGAT
TTCTGTGTCACCTACTGTCATTTGTTAAACTGCTGGCCAACAAGAACAGGAAGTATA
GTTTGGGGGGTTGGGGAGAGTTTACATAAGGAAGAGAAGAAATTGAGTGGCATATT
GTAAATATCAGATCTATAATTGTAAATATAAAACCTGCCTCAGTTAGAATGAATGGA
AAGCAGATCTACAATTTGCTAATATAGGAATATCAGGTTGACTATATAGCCATACTT
GAAAATGCTTCTGAGTGGTGTCAACTTTACTTGAATGAATTTTTCATCTTGATTGACG
CACAGTGATGTACAGTTCACTTCTGAAGCTAGTGGTTAACTTGTGTAGGAAACTTTT
GCAGTTTGACACTAAGATAACTTCTGTGTGCATTTTTCTATGCTTTTTTAAAAACTAG
TTTCATTTCATTTTCATGAGATGTTTGGTTTATAAGATCTGAGGATGGTTATAAATAC
TGTAAGTATTGTAATGTTATGAATGCAGGTTATTTGAAAGCTGTTTATTATTATATCA

TTCCTGATAATGCTATGTGAGTGTTTTTAATAAAATTTATATTTATTTAATGCACTCT
AAGTGTTGTCTTCCT

[00466] By "transforming growth factor receptor 2 (TGFBRII) polypeptide" is meant a protein having at least about 85% sequence identity to NCBI Accession No. ABG65632.1 or a fragment thereof and having immunosuppressive activity. An exemplary amino acid sequence is provided below.

>ABG65632.1 transforming growth factor beta receptor II [Homo sapiens]
MGRGLLRGLWPLHIVLWTRIASTIPPHVQKSVNNDMIVTDNNGAVKFPQLCKFCDVRFS
TCDNQKSCMSNCSITSICEKPQEVCVAVWRKNDENITLETVCHDPKLPYHDFILEDAASP
KCIMKEKKKPGETFFMCSCSSDECNDNIIFSEEYNTSNPDLLLVIFQVTGISLLPPLGVAIS
VIIIFYCYRVNRQQKLSSTWETGKTRKLMEFSEHCAIILEDDRSDISSTCANNINHNTELLP
IELDTLVGKGRFAEVYKAKLKQNTSEQFETVAVKIFPYEEYASWKTEKDIFSDINLKHEN
ILQFLTAEERKTELGKQYWLITAFHAKGNLQEYLTRHVISWEDLRKLGSSLARGIAHLHS
DHTPCGRPKMPIVHRDLKSSNILVKNDLTCCLCDFGLSLRLDPTLSVDDLANSGQVGTA
RYMAPEVLESRMNLENVESFKQTDVYSMALVLWEMTSRCNAVGEVKDYEPPFGSKVR
EHPCVESMKDNVLRDRGRPEIPSFWLNHQGIQMVCETLTECWDHDPEARLTAQCVAER
FSELEHLDRLSGRSCSEEKIPEDGSLNTTK

[00467] By "transforming growth factor receptor 2 (TGFBRII) polynucleotide" is meant a nucleic acid that encodes a TGFBRII polypeptide. The TGFBRII gene encodes a transmembrane protein having serine/threonine kinase activity. An exemplary TGFBRII nucleic acid is provided below.

>M85079.1 Human TGF-beta type II receptor mRNA, complete cds
GTTGGCGAGGAGTTTCCTGTTTCCCCCGCAGCGCTGAGTTGAAGTTGAGTGAGTCAC
TCGCGCGCACGGAGCGACGACACCCCCGCGCGTGCACCCGCTCGGGACAGGAGCCG
GACTCCTGTGCAGCTTCCCTCGGCCGCCGGGGGCCTCCCCGCGCCTCGCCGGCCTCC
AGGCCCCTCCTGGCTGGCGAGCGGGCGCCACATCTGGCCCGCACATCTGCGCTGCCG
GCCCGGCGCGGGGTCCGGAGAGGGCGCGGCGCGGAGCGCAGCCAGGGGTCCGGGA
AGGCGCCGTCCGTGCGCTGGGGGCTCGGTCTATGACGAGCAGCGGGGTCTGCCATG
GGTCGGGGGCTGCTCAGGGGCCTGTGGCCGCTGCACATCGTCCTGTGGACGCGTATC
GCCAGCACGATCCCACCGCACGTTCAGAAGTCGGTTAATAACGACATGATAGTCAC
TGACAACAACGGTGCAGTCAAGTTTCCACAACTGTGTAAATTTTGTGATGTGAGATT

TTCCACCTGTGACAACCAGAAATCCTGCATGAGCAACTGCAGCATCACCTCCATCTG
TGAGAAGCCACAGGAAGTCTGTGTGGCTGTATGGAGAAAGAATGACGAGAACATAA
CACTAGAGACAGTTTGCCATGACCCCAAGCTCCCCTACCATGACTTTATTCTGGAAG
ATGCTGCTTCTCCAAAGTGCATTATGAAGGAAAAAAAAAAAGCCTGGTGAGACTTTCT
TCATGTGTTCCTGTAGCTCTGATGAGTGCAATGACAACATCATCTTCTCAGAAGAAT
ATAACACCAGCAATCCTGACTTGTTGCTAGTCATATTTCAAGTGACAGGCATCAGCC
TCCTGCCACCACTGGGAGTTGCCATATCTGTCATCATCATCTTCTACTGCTACCGCGT
TAACCGGCAGCAGAAGCTGAGTTCAACCTGGGAAACCGGCAAGACGCGGAAGCTCA
TGGAGTTCAGCGAGCACTGTGCCATCATCCTGGAAGATGACCGCTCTGACATCAGCT
CCACGTGTGCCAACAACATCAACCACAACACAGAGCTGCTGCCCATTGAGCTGGAC
ACCCTGGTGGGGAAAGGTCGCTTTGCTGAGGTCTATAAGGCCAAGCTGAAGCAGAA
CACTTCAGAGCAGTTTGAGACAGTGGCAGTCAAGATCTTTCCCTATGAGGAGTATGC
CTCTTGGAAGACAGAGAAGGACATCTTCTCAGACATCAATCTGAAGCATGAGAACA
TACTCCAGTTCCTGACGGCTGAGGAGCGGAAGACGGAGTTGGGGAAACAATACTGG
CTGATCACCGCCTTCCACGCCAAGGGCAACCTACAGGAGTACCTGACGCGGCATGT
CATCAGCTGGGAGGACCTGCGCAAGCTGGGCAGCTCCCTCGCCCGGGGGATTGCTC
ACCTCCACAGTGATCACACTCCATGTGGGAGGCCCAAGATGCCCATCGTGCACAGG
GACCTCAAGAGCTCCAATATCCTCGTGAAGAACGACCTAACCTGCTGCCTGTGTGAC
TTTGGGCTTTCCCTGCGTCTGGACCCTACTCTGTCTGTGGATGACCTGGCTAACAGTG
GGCAGGTGGGAACTGCAAGATACATGGCTCCAGAAGTCCTAGAATCCAGGATGAAT
TTGGAGAATGCTGAGTCCTTCAAGCAGACCGATGTCTACTCCATGGCTCTGGTGCTC
TGGGAAATGACATCTCGCTGTAATGCAGTGGGAGAAGTAAAAGATTATGAGCCTCC
ATTTGGTTCCAAGGTGCGGGAGCACCCCTGTGTCGAAAGCATGAAGGACAACGTGT
TGAGAGATCGAGGGCGACCAGAAATTCCCAGCTTCTGGCTCAACCACCAGGGCATC
CAGATGGTGTGTGAGACGTTGACTGAGTGCTGGGACCACGACCCAGAGGCCCGTCT
CACAGCCCAGTGTGTGGCAGAACGCTTCAGTGAGCTGGAGCATCTGGACAGGCTCT
CGGGGAGGAGCTGCTCGGAGGAGAAGATTCCTGAAGACGGCTCCCTAAACACTACC
AAATAGCTCTTATGGGGCAGGCTGGGCATGTCCAAAGAGGCTGCCCCTCTCACCAA
A

[00468] By "T Cell Immunoreceptor with Ig and ITIM Domains (TIGIT) polypeptide" is meant a protein having at least about 85% sequence identity to NCBI Accession No.

ACD74757.1 or a fragment thereof and having immunomodulatory activity. An exemplary TIGIT amino acid sequence is provided below.

>ACD74757.1 T cell immunoreceptor with Ig and ITIM domains [Homo sapiens]

MRWCLLLIWAQGLRQAPLASGMMTGTIETTGNISAEKGGSIILQCHLSSTTAQVTQVNW
EQQDQLLAICNADLGWHISPSFKDRVAPGPGLGLTLQSLTVNDTGEYFCIYHTYPDGTY
TGRIFLEVLESSVAEHGARFQIPLLGAMAATLVVICTAVIVVVALTRKKKALRIHSVEGD
LRRKSAGQEEWSPSAPSPPGSCVQAEAAPAGLCGEQRGEDCAELHDYFNVLSYRSLGN
CSFFTETG

[00469] By "T Cell Immunoreceptor With Ig And ITIM Domains (TIGIT) polynucleotide" is meant a nucleic acid encoding a TIGIT polypeptide. The TIGIT gene encodes an inhibitory immune receptor that is associated with neoplasia and T cell exhaustion. An exemplary nucleic acid sequence is provided below.

>EU675310.1 Homo sapiens T cell immunoreceptor with Ig and ITIM domains (TIGIT) mRNA, complete cds

CGTCCTATCTGCAGTCGGCTACTTTCAGTGGCAGAAGAGGCCACATCTGCTTCCTGT
AGGCCCTCTGGGCAGAAGCATGCGCTGGTGTCTCCTCCTGATCTGGGCCCAGGGGCT
GAGGCAGGCTCCCCTCGCCTCAGGAATGATGACAGGCACAATAGAAACAACGGGGA
ACATTTCTGCAGAGAAAGGTGGCTCTATCATCTTACAATGTCACCTCTCCTCCACCA
CGGCACAAGTGACCCAGGTCAACTGGGAGCAGCAGGACCAGCTTCTGGCCATTTGT
AATGCTGACTTGGGGTGGCACATCTCCCCATCCTTCAAGGATCGAGTGGCCCCAGGT
CCCGGCCTGGGCCTCACCCTCCAGTCGCTGACCGTGAACGATACAGGGGAGTACTTC
TGCATCTATCACACCTACCCTGATGGGACGTACACTGGGAGAATCTTCCTGGAGGTC
CTAGAAAGCTCAGTGGCTGAGCACGGTGCCAGGTTCCAGATTCCATTGCTTGGAGCC
ATGGCCGCGACGCTGGTGGTCATCTGCACAGCAGTCATCGTGGTGGTCGCGTTGACT
AGAAAGAAGAAAGCCCTCAGAATCCATTCTGTGGAAGGTGACCTCAGGAGAAAATC
AGCTGGACAGGAGGAATGGAGCCCCAGTGCTCCCTCACCCCCAGGAAGCTGTGTCC
AGGCAGAAGCTGCACCTGCTGGGCTCTGTGGAGAGCAGCGGGGAGAGGACTGTGCC
GAGCTGCATGACTACTTCAATGTCCTGAGTTACAGAAGCCTGGGTAACTGCAGCTTC
TTCACAGAGACTGGTTAGCAACCAGAGGCATCTTCTGG

[00470] By "T Cell Receptor Alpha Constant (TRAC) polypeptide" is meant a protein having at least about 85% amino acid sequence identity to NCBI Accession No. P01848.2 or fragment

thereof and having immunomodulatory activity. An exemplary amino acid sequence is provided below.

>sp|P01848.2|TRAC_HUMAN RecName: Full=T cell receptor alpha constant

IQNPDPAVYQLRDSKSSDKSVCLFTDFDSQTNVSQSKDSDVYITDKTVLDMRSMDFKSN
SAVAWSNKSDFACANAFNNSIIPEDTFFPSPESSCDVKLVEKSFETDTNLNFQNLSVIGFR
ILLLKVAGFNLLMTLRLWSS

[00471] By "T Cell Receptor Alpha Constant (TRAC) polynucleotide" is meant a nucleic acid encoding a TRAC polypeptide. Exemplary TRAC nucleic acid sequences are provided below.

UCSC human genome database, Gene ENSG00000277734.8 Human T-cell receptor alpha chain (TCR-alpha)

catgctaatcctccggcaaacctctgtttcctcctcaaaaggcaggaggtcggaaagaataaacaatgagagtcacattaaaaacacaaaat
cctacggaaatactgaagaatgagtctcagcactaaggaaaagcctccagcagctcctgctttctgagggtgaaggatagacgctgtggct
ctgcatgactcactagcactctatcacggccatattctggcagggtcagtggctccaactaacatttgtttggtactttacagtttattaaatagat
gtttatatggagaagctctcatttctttctcagaagagcctggctaggaaggtggatgaggcaccatatttcattttgcaggtgaaattcctgaga
tgtaaggagctgctgtgacttgctcaaggccttatatcgagtaaacggtagtgctggggcttagacgcaggtgttctgatttatagttcaaaac
ctctatcaatgagagagcaatctcctggtaatgtgatagatttcccaacttaatgccaacataccataaacctcccattctgctaatgcccagcc
taagttggggagaccactccagattccaagatgtacagtttgctttgctgggccttttttcccatgcctgcctttactctgccagagttatattgctg
gggttttgaagaagatcctattaaataaaagaataagcagtattattaagtagccctgcatttcaggtttccttgagtggcaggccaggcctgg
ccgtgaacgttcactgaaatcatggcctcttggccaagattgatagcttgtgcctgtccctgagtcccagtccatcacgagcagctggtttcta
agatgctatttcccgtataaagcatgagaccgtgacttgccagccccacagagccccgcccttgtccatcactggcatctggactccagcct
gggttggggcaaagagggaaatgagatcatgtcctaaccctgatcctcttgtcccacagATATCCAGAACCCTGACCCT
GCCGTGTACCAGCTGAGAGACTCTAAATCCAGTGACAAGTCTGTCTGCCTATTCACC
GATTTTGATTCTCAAACAAATGTGTCACAAAGTAAGGATTCTGATGTGTATATCACA
GACAAAACTGTGCTAGACATGAGGTCTATGGACTTCAAGAGCAACAGTGCTGTGGC
CTGGAGCAACAAATCTGACTTTGCATGTGCAAACGCCTTCAACAACAGCATTATTCC
AGAAGACACCTTCTTCCCCAGCCCAGgtaagggcagctttggtgccttcgcaggctgtttccttgcttcaggaatgg
ccaggttctgcccagagctctggtcaatgatgtctaaaactcctctgattggtggtctcggccttatccattgccaccaaaaccctcttttttacta
agaaacagtgagccttgttctggcagtccagagaatgacacgggaaaaaagcagatgaagagaaggtggcaggagagggcacgtggc
ccagcctcagtctctccaactgagttcctgcctgcctgcctttgctcagactgtttgccccttactgctcttctaggcctcattctaagccccttct
ccaagttgcctctccttatttctccctgtctgccaaaaaatctttcccagctcactaagtcagtctcacgcagtcactcattaacccaccaatcact

gattgtgccggcacatgaatgcaccaggtgttgaagtggaggaattaaaaagtcagatgaggggtgtgcccagaggaagcaccattctagt
tgggggagcccatctgtcagctgggaaaagtccaaataacttcagattggaatgtgttttaactcagggttgagaaaacagctaccttcagga
caaaagtcagggaagggctctctgaagaaatgctacttgaagataccagccctaccaagggcagggagaggaccctatagaggcctggg
acaggagctcaatgagaaaggagaagagcagcaggcatgagttgaatgaaggaggcagggccgggtcacagggccttctaggccatg
agagggtagacagtattctaaggacgccagaaagctgttgatcggcttcaagcaggggagggacacctaatttgcttttcttttttttttttttttt
ttttttttttgagatggagttttgctcttgttgcccaggctggagtgcaatggtgcatcttggctcactgcaacctccgcctcccaggttcaagtg
attctcctgcctcagcctcccgagtagctgagattacaggcacccgccaccatgcctggctaatttttttgtatttttagtagagacagggtttcac
tatgttggccaggctggtctcgaactcctgacctcaggtgatccacccgcttcagcctcccaaagtgctgggattacaggcgtgagccacca
cacccggcctgcttttcttaaagatcaatctgagtgctgtacggagagtgggttgtaagccaagagtagaagcagaaagggagcagttgca
gcagagagatgatggaggcctgggcagggtggtggcagggaggtaaccaacaccattcaggtttcaaaggtagaaccatgcagggatg
agaaagcaaagaggggatcaaggaaggcagctggattttggcctgagcagctgagtcaatgatagtgccgtttactaagaagaaaccaag
gaaaaaatttggggtgcagggatcaaaacttttttggaacatatgaaagtacgtgtttatactctttatggcccttgtcactatgtatgcctcgctgc
ctccattggactctagaatgaagccaggcaagagcagggtctatgtgtgatggcacatgtggccagggtcatgcaacatgtactttgtacaa
acagtgtatattgagtaaatagaaatggtgtccaggagccgaggtatcggtcctgccagggccaggggctctccctagcaggtgctcatatg
ctgtaagttccctccagatctctccacaaggaggcatggaaaggctgtagttgttcacctgcccaagaactaggaggtctggggtgggaga
gtcagcctgctctggatgctgaaagaatgtctgtttttccttttagAAAGTTCCTGTGATGTCAAGCTGGTCGAGA
AAAGCTTTGAAACAGgtaagacaggggtctagcctgggtttgcacaggattgcggaagtgatgaacccgcaataaccctgc
ctggatgagggagtgggaagaaattagtagatgtgggaatgaatgatgaggaatggaaacagcggttcaagacctgcccagagctgggt
ggggtctctcctgaatccctctcaccatctctgactttccattctaagcactttgaggatgagtttctagcttcaatagaccaaggactctctccta
ggcctctgtattcctttcaacagctccactgtcaagagagccagagagagcttctgggtggcccagctgtgaaatttctgagtcccttagggat
agccctaaacgaaccagatcatcctgaggacagccaagaggttttgccttctttcaagacaagcaacagtactcacataggctgtgggcaat
ggtcctgtctctcaagaatcccctgccactcctcacacccaccctgggcccatattcatttccatttgagttgttcttattgagtcatccttcctgtg
gtagcggaactcactaaggggcccatctggacccgaggtattgtgatgataaattctgagcacctaccccatccccagaagggctcagaaa
taaaataagagccaagtctagtcggtgtttcctgtcttgaaacacaatactgttggccctggaagaatgcacagaatctgtttgtaaggggatat
gcacagaagctgcaagggacaggaggtgcaggagctgcaggcctcccccacccagcctgctctgccttggggaaaaccgtgggtgtgt
cctgcaggccatgcaggcctgggacatgcaagcccataaccgctgtggcctcttggttttacagATACGAACCTAAACTTT
CAAAACCTGTCAGTGATTGGGTTCCGAATCCTCCTCCTGAAAGTGGCCGGGTTTAAT
CTGCTCATGACGCTGCGGCTGTGGTCCAGCTGAGgtgaggggccttgaagctgggagtggggtttaggga
cgcgggtctctgggtgcatcctaagctctgagagcaaacctccctgcagggtcttgcttttaagtccaaagcctgagcccaccaaactctcct
acttcttcctgttacaaattcctcttgtgcaataataatggcctgaaacgctgtaaaatatcctcatttcagccgcctcagttgcacttctcccctat
gaggtaggaagaacagttgtttagaaacgaagaaactgaggccccacagctaatgagtggaggaagagagacacttgtgtacaccacatg

ccttgtgttgtacttctctcaccgtgtaacctcctcatgtcctctctccccagtacggctctcttagctcagtagaaagaagacattacactcatatt
acaccccaatcctggctagagtctccgcaccctcctccccagggtccccagtcgtcttgctgacaactgcatcctgttccatcaccatcaaa
aaaaaactccaggctgggtgcgggggctcacacctgtaatcccagcactttgggaggcagaggcaggaggagcacaggagctggaga
ccagcctgggcaacacagggagaccccgcctctacaaaaagtgaaaaaattaaccaggtgtggtgctgcacacctgtagtcccagctactt
aagaggctgagatgggaggatcgcttgagccctggaatgttgaggctacaatgagctgtgattgcgtcactgcactccagcctggaagaca
aagcaagatcctgtctcaaataataaaaaaaataagaactccagggtacatttgctcctagaactctaccacatagccccaaacagagccatc
accatcacatccctaacagtcctgggtcttcctcagtgtccagcctgacttctgttcttcctcattccagATCTGCAAGATTGTAA
GACAGCCTGTGCTCCCTCGCTCCTTCCTCTGCATTGCCCCTCTTCTCCCTCTCCAAAC
AGAGGGAACTCTCCTACCCCCAAGGAGGTGAAAGCTGCTACCACCTCTGTGCCCCCC
CGGCAATGCCACCAACTGGATCCTACCCGAATTTATGATTAAGATTGCTGAAGAGCT
GCCAAACACTGCTGCCACCCCCTCTGTTCCCTTATTGCTGCTTGTCACTGCCTGACAT
TCACGGCAGAGGCAAGGCTGCTGCAGCCTCCCCTGGCTGTGCACATTCCCTCCTGCT
CCCCAGAGACTGCCTCCGCCATCCCACAGATGATGGATCTTCAGTGGGTTCTCTTGG
GCTCTAGGTCCTGCAGAATGTTGTGAGGGGTTTATTTTTTTTTAATAGTGTTCATAAA
GAAATACATAGTATTCTTCTTCTCAAGACGTGGGGGGAAATTATCTCATTATCGAGG
CCCTGCTATGCTGTGTATCTGGGCGTGTTGTATGTCCTGCTGCCGATGCCTTCATTAA
AATGATTTGGAAGAGCAGA

[00472] Nucleotides in lower cases above are untranslated regions or introns, and nucleotides in upper cases are exons.

[00473] >X02592.1 Human mRNA for T-cell receptor alpha chain (TCR-alpha)

TTTTGAAACCCTTCAAAGGCAGAGACTTGTCCAGCCTAACCTGCCTGCTGCTCCTAG
CTCCTGAGGCTCAGGGCCCTTGGCTTCTGTCCGCTCTGCTCAGGGCCCTCCAGCGTG
GCCACTGCTCAGCCATGCTCCTGCTGCTCGTCCCAGTGCTCGAGGTGATTTTTACCCT
GGGAGGAACCAGAGCCCAGTCGGTGACCCAGCTTGGCAGCCACGTCTCTGTCTCTG
AAGGAGCCCTGGTTCTGCTGAGGTGCAACTACTCATCGTCTGTTCCACCATATCTCTT
CTGGTATGTGCAATACCCCAACCAAGGACTCCAGCTTCTCCTGAAGTACACATCAGC
GGCCACCCTGGTTAAAGGCATCAACGGTTTTGAGGCTGAATTTAAGAAGAGTGAAA
CCTCCTTCCACCTGACGAAACCCTCAGCCCATATGAGCGACGCGGCTGAGTACTTCT
GTGCTGTGAGTGATCTCGAACCGAACAGCAGTGCTTCCAAGATAATCTTTGGATCAG
GGACCAGACTCAGCATCCGGCCAAATATCCAGAACCCTGACCCTGCCGTGTACCAG
CTGAGAGACTCTAAATCCAGTGACAAGTCTGTCTGCCTATTCACCGATTTTGATTCTC

AAACAAATGTGTCACAAAGTAAGGATTCTGATGTGTATATCACAGACAAAACTGTG
CTAGACATGAGGTCTATGGACTTCAAGAGCAACAGTGCTGTGGCCTGGAGCAACAA
ATCTGACTTTGCATGTGCAAACGCCTTCAACAACAGCATTATTCCAGAAGACACCTT
CTTCCCCAGCCCAGAAAGTTCCTGTGATGTCAAGCTGGTCGAGAAAAGCTTTGAAAC
AGATACGAACCTAAACTTTCAAAACCTGTCAGTGATTGGGTTCCGAATCCTCCTCCT
GAAAGTGGCCGGGTTTAATCTGCTCATGACGCTGCGGCTGTGGTCCAGCTGAGATCT
GCAAGATTGTAAGACAGCCTGTGCTCCCTCGCTCCTTCCTCTGCATTGCCCCTCTTCT
CCCTCTCCAAACAGAGGGAACTCTCCTACCCCCAAGGAGGTGAAAGCTGCTACCAC
CTCTGTGCCCCCCCCGGTAATGCCACCAACTGGATCCTACCCGAATTTATGATTAAGA
TTGCTGAAGAGCTGCCAAACACTGCTGCCACCCCCTCTGTTCCCTTATTGCTGCTTGT
CACTGCCTGACATTCACGGCAGAGGCAAGGCTGCTGCAGCCTCCCCTGGCTGTGCAC
ATTCCCTCCTGCTCCCCAGAGACTGCCTCCGCCATCCCACAGATGATGGATCTTCAG
TGGGTTCTCTTGGGCTCTAGGTCCTGGAGAATGTTGTGAGGGGTTTATTTTTTTTTAA
TAGTGTTCATAAAGAAATACATAGTATTCTTCTTCTCAAGACGTGGGGGGAAATTAT
CTCATTATCGAGGCCCTGCTATGCTGTGTGTCTGGGCGTGTTGTATGTCCTGCTGCCG
ATGCCTTCATTAAAATGATTTGGAA

[00474] By "T cell receptor beta constant 1 polypeptide (TRBC1)" is meant a protein having at least about 85% amino acid sequence identity to NCBI Accession No. P01850 or fragment thereof and having immunomodulatory activity. An exemplary amino acid sequence is provided below.

[00475] .>sp|P01850|TRBC1_HUMAN T cell receptor beta constant 1 OS=Homo sapiens OX=9606 GN=TRBC1 PE=1
SV=4DLNKVFPPEVAVFEPSEAEISHTQKATLVCLATGFFPDHVELSWWVNGKEVHSGV
STDPQPLKEQPALNDSRYCLSSRLRVSATFWQNPRNHFRCQVQFYGLSENDEWTQDRA
KPVTQIVSAEAWGRADCGFTSVSYQQGVLSATILYEILLGKATLYAVLVSALVLMAMV
KRKDF

[00476] By "T cell receptor beta constant 1 polynucleotide (**TRBC1**)" is meant a nucleic acid encoding a TRBC1 polypeptide. An exemplary TRBC1 nucleic acid sequence is provided below. >
X00437.1CTGGTCTAGAATATTCCACATCTGCTCTCACTCTGCCATGGACTCCTGGACC
TTCTGCTGTGTGTCCCTTTGCATCCTGGTAGCGAAGCATACAGATGCTGGAGTTATCC

AGTCACCCCGCCATGAGGTGACAGAGATGGGACAAGAAGTGACTCTGAGATGTAAA
CCAATTTCAGGCCACAACTCCCTTTTCTGGTACAGACAGACCATGATGCGGGGACTG
GAGTTGCTCATTTACTTTAACAACAACGTTCCGATAGATGATTCAGGGATGCCCGAG
GATCGATTCTCAGCTAAGATGCCTAATGCATCATTCTCCACTCTGAAGATCCAGCCC
TCAGAACCCAGGGACTCAGCTGTGTACTTCTGTGCCAGCAGTTTCTCGACCTGTTCG
GCTAACTATGGCTACACCTTCGGTTCGGGGACCAGGTTAACCGTTGTAGAGGACCTG
AACAAGGTGTTCCCACCCGAGGTCGCTGTGTTTGAGCCATCAGAAGCAGAGATCTCC
CACACCCAAAAGGCCACACTGGTGTGCCTGGCCACAGGCTTCTTCCCCGACCACGTG
GAGCTGAGCTGGTGGGTGAATGGGAAGGAGGTGCACAGTGGGGTCAGCACAGACCC
GCAGCCCCTCAAGGAGCAGCCCGCCCTCAATGACTCCAGATACTGCCTGAGCAGCC
GCCTGAGGGTCTCGGCCACCTTCTGGCAGAACCCCCGCAACCACTTCCGCTGTCAAG
TCCAGTTCTACGGGCTCTCGGAGAATGACGAGTGGACCCAGGATAGGGCCAAACCC
GTCACCCAGATCGTCAGCGCCGAGGCCTGGGGTAGAGCAGACTGTGGCTTTACCTC
GGTGTCCTACCAGCAAGGGGTCCTGTCTGCCACCATCCTCTATGAGATCCTGCTAGG
GAAGGCCACCCTGTATGCTGTGCTGGTCAGCGCCCTTGTGTTGATGGCCATGGTCAA
GAGAAAGGATTTCTGAAGGCAGCCCTGGAAGTGGAGTTAGGAGCTTCTAACCCGTC
ATGGTTCAATACACATTCTTCTTTTGCCAGCGCTTCTGAAGAGCTGCTCTCACCTCTC
TGCATCCCAATAGATATCCCCCTATGTGCATGCACACCTGCACACTCACGGCTGAAA
TCTCCCTAACCCAGGGGGAC

[00477] By "T cell receptor beta constant 2 polypeptide (TRBC2)" is meant a protein having at least about 85% amino acid sequence identity to NCBI Accession No. A0A5B9 or fragment thereof and having immunomodulatory activity. An exemplary amino acid sequence is provided below.

.>sp|A0A5B9|TRBC2_HUMAN T cell receptor beta constant 2 OS=Homo sapiens OX=9606 GN=TRBC2 PE=1 SV=2DLKNVFPPKVAVFEPSEAEISHTQKATLVCLATGFYPDHVELSWWVNGKEVHSGV STDPQPLKEQPALNDSRYCLSSRLRVSATFWQNPRNHFRCQVQFYGLSENDEWTQDRA KPVTQIVSAEAWGRADCGFTSESYQQGVLSATILYEILLGKATLYAVLVSALVLMAMV KRKDSRG

[00478] By "T cell receptor beta constant 2 polynucleotide (**TRBC2**)" is meant a nucleic acid encoding a TRAC polypeptide. An exemplary TRBC2 nucleic acid sequence is provided below.

[00479] >NG_001333.2:655095-656583 Homo sapiens T cell receptor beta locus (TRB) on chromosome7

AGGACCTGAAAAACGTGTTCCCACCCGAGGTCGCTGTGTTTGAGCCATCAGAAGCA
GAGATCTCCCACACCCAAAAGGCCACACTGGTATGCCTGGCCACAGGCTTCTACCCC
GACCACGTGGAGCTGAGCTGGTGGGTGAATGGGAAGGAGGTGCACAGTGGGGTCAG
CACAGACCCGCAGCCCCTCAAGGAGCAGCCCGCCCTCAATGACTCCAGATACTGCC
TGAGCAGCCGCCTGAGGGTCTCGGCCACCTTCTGGCAGAACCCCCGCAACCACTTCC
GCTGTCAAGTCCAGTTCTACGGGCTCTCGGAGAATGACGAGTGGACCCAGGATAGG
GCCAAACCCGTCACCCAGATCGTCAGCGCCGAGGCCTGGGGTAGAGCAGGTGAGTG
GGGCCTGGGGAGATGCCTGGAGGAGATTAGGTGAGACCAGCTACCAGGGAAAATG
GAAAGATCCAGGTAGCGGACAAGACTAGATCCAGAAGAAAGCCAGAGTGGACAAG
GTGGGATGATCAAGGTTCACAGGGTCAGCAAAGCACGGTGTGCACTTCCCCCACCA
AGAAGCATAGAGGCTGAATGGAGCACCTCAAGCTCATTCTTCCTTCAGATCCTGACA
CCTTAGAGCTAAGCTTTCAAGTCTCCCTGAGGACCAGCCATACAGCTCAGCATCTGA
GTGGTGTGCATCCCATTCTCTTCTGGGGTCCTGGTTTCCTAAGATCATAGTGACCACT
TCGCTGGCACTGGAGCAGCATGAGGGAGACAGAACCAGGGCTATCAAAGGAGGCTG
ACTTTGTACTATCTGATATGCATGTGTTTGTGGCCTGTGAGTCTGTGATGTAAGGCTC
AATGTCCTTACAAAGCAGCATTCTCTCATCCATTTTTCTTCCCCTGTTTTCTTTCAGAC
TGTGGCTTCACCTCCGGTAAGTGAGTCTCTCCTTTTTCTCTCTATCTTTCGCCGTCTCT
GCTCTCGAACCAGGGCATGGAGAATCCACGGACACAGGGGCGTGAGGGAGGCCAG
AGCCACCTGTGCACAGGTGCCTACATGCTCTGTTCTTGTCAACAGAGTCTTACCAGC
AAGGGGTCCTGTCTGCCACCATCCTCTATGAGATCTTGCTAGGGAAGGCCACCTTGT
ATGCCGTGCTGGTCAGTGCCCTCGTGCTGATGGCCATGGTAAGGAGGAGGGTGGGA
TAGGGCAGATGATGGGGGCAGGGGATGGAACATCACACATGGGCATAAAGGAATCT
CAGAGCCAGAGCACAGCCTAATATATCCTATCACCTCAATGAAACCATAATGAAGC
CAGACTGGGGAGAAAATGCAGGGAATATCACAGAATGCATCATGGGAGGATGGAG
ACAACCAGCGAGCCCTACTCAAATTAGGCCTCAGAGCCCGCCTCCCCTGCCCTACTC
CTGCTGTGCCATAGCCCCTGAAACCCTGAAAATGTTCTCTCTTCCACAGGTCAAGAG
AAAGGATTCCAGAGGCTAG

[00480] As used herein “transduction” means to transfer a gene or genetic material to a cell via a viral vector.

[00481] "Transformation," as used herein refers to the process of introducing a genetic change in a cell produced by the introduction of exogenous nucleic acid.

[00482] "Transfection" refers to the transfer of a gene or genetical material to a cell via a chemical or physical means.

[00483] By "translocation" is meant the rearrangement of nucleic acid segments between non-homologous chromosomes.

[00484] As used herein, the terms "treat," treating," "treatment," and the like refer to reducing or ameliorating a disorder and/or a symptom associated therewith. It will be appreciated that, although not precluded, treating a disorder or condition does not require that the disorder, condition or symptoms associated therewith be eliminated.

[00485] The term "uracil glycosylase inhibitor" or "UGI," as used herein, refers to a protein that is capable of inhibiting a uracil-DNA glycosylase base-excision repair enzyme. In some embodiments, the polypeptide further contains one or more (e.g., 1, 2, 3, 4, 5) Uracil glycosylase inhibitors. In some embodiments, a UGI domain comprises a wild-type UGI or a modified version thereof. In some embodiments, the UGI proteins provided herein include fragments of UGI and proteins homologous to a UGI or a UGI fragment. For example, in some embodiments, a UGI domain comprises a fragment of the amino acid sequence set forth herein below. In some embodiments, a UGI fragment comprises an amino acid sequence that comprises at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% of an exemplary UGI sequence provided herein. In some embodiments, a UGI comprises an amino acid sequence homologous to the amino acid sequence set forth herein below, or an amino acid sequence homologous to a fragment of the amino acid sequence set forth herein below. In some embodiments, proteins comprising UGI or fragments of UGI or homologs of UGI or UGI fragments are referred to as "UGI variants." A UGI variant shares homology to UGI, or a fragment thereof. For example, a UGI variant is at least 70% identical, at least 75% identical, at least 80% identical, at least 85% identical, at least 90% identical, at least 95% identical, at least 96% identical, at least 97% identical, at least 98% identical, at least 99% identical, at least 99.5% identical, or at least 99.9% identical to a wild type UGI or a UGI as set forth herein. In some embodiments, the UGI variant comprises a fragment of UGI, such that the fragment is at least 70% identical, at least 80% identical, at least 90% identical, at least 95% identical, at least 96% identical, at least 97%

identical, at least 98% identical, at least 99% identical, at least 99.5% identical, or at least 99.9% to the corresponding fragment of wild-type UGI or a UGI as set forth below. In some embodiments, the UGI comprises the following amino acid sequence:

[00486] >splP14739IUNGI_BPPB2 Uracil-DNA glycosylase inhibitor

[00487] MTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTDENV MLLT S D APE YKPW ALVIQDS NGENKIKML

[00488] The term "vector" refers to a means of introducing a nucleic acid sequence into a cell, resulting in a transformed cell. Vectors include plasmids, transposons, phages, viruses, liposomes, and episome. "Expression vectors" are nucleic acid sequences comprising the nucleotide sequence to be expressed in the recipient cell. Expression vectors may include additional nucleic acid sequences to promote and/or facilitate the expression of the of the introduced sequence such as start, stop, enhancer, promoter, and secretion sequences.

[00489] By "zeta chain of T cell receptor associated protein kinase 70 (ZAP70) polypeptide" is meant a protein having at least about 85% amino acid sequence identity to NCBI Accession No. AAH53878.1 and having kinase activity. An exemplary amino acid sequence is provided below.

[00490] >AAH53878.1 Zeta-chain (TCR) associated protein kinase 70kDa [Homo sapiens]
MPDPAAHLPFFYGSISRAEAEEHLKLAGMADGLFLLRQCLRSLGGYVLSLVHDVRFHHF
PIERQLNGTYAIAGGKAHCGPAELCEFYSRDPDGLPCNLRKPCNRPSGLEPQPGVFDCLR
DAMVRDYVRQTWKLEGEALEQAIISQAPQVEKLIATTAHERMPWYHSSLTREEAERKL
YSGAQTDGKFLLRPRKEQGTYALSLIYGKTVYHYLISQDKAGKYCIPEGTKFDTLWQLV
EYLKLKADGLIYCLKEACPNSSASNASGAAAPTLPAHPSTLTHPQRRIDTLNSDGYTPEP
ARITSPDKPRPMPMDTSVYESPYSDPEELKDKKLFLKRDNLLIADIELGCGNFGSVRQGV
YRMRKKQIDVAIKVLKQGTEKADTEEMMREAQIMHQLDNPYIVRLIGVCQAEALMLV
MEMAGGGPLHKFLVGKREEIPVSNVAELLHQVSMGMKYLEEKNFVHRDLAARNVLLV
NRHYAKISDFGLSKALGADDSYYTARSAGKWPLKWYAPECINFRKFSSRSDVWSYGVT
MWEALSYGQKPYKKMKGPEVMAFIEQGKRMECPPECPPELYALMSDCWIYKWEDRPD
FLTVEQRMRACYYSLASKVEGPPGSTQKAEAACA

[00491] By "zeta chain of T cell receptor associated protein kinase 70 (ZAP70) polynucleotide" is meant a nucleic acid encoding a ZAP70 polypeptide. The ZAP70 gene encodes a tyrosine kinase that is involved in T cell development and lymphocyte activation.

Absence of functional ZAP10 can lead to a severe combined immunodeficiency characterized by the lack of CD8+ T cells. An exemplary ZAP70 nucleic acid sequence is provided below.

[00492] >BC053878.1 Homo sapiens zeta-chain (TCR) associated protein kinase 70kDa, mRNA (cDNA clone MGC:61743 IMAGE:5757161), complete cds

GCTTGCCGGAGCTCAGCAGACACCAGGCCTTCCGGGCAGGCCTGGCCCACCGTGGG
CCTCAGAGCTGCTGCTGGGGCATTCAGAACCGGCTCTCCATTGGCATTGGGACCAGA
GACCCCGCAAGTGGCCTGTTTGCCTGGACATCCACCTGTACGTCCCCAGGTTTCGGG
AGGCCCAGGGGCGATGCCAGACCCCGCGGCGCACCTGCCCTTCTTCTACGGCAGCA
TCTCGCGTGCCGAGGCCGAGGAGCACCTGAAGCTGGCGGGCATGGCGGACGGGCTC
TTCCTGCTGCGCCAGTGCCTGCGCTCGCTGGGCGGCTATGTGCTGTCGCTCGTGCAC
GATGTGCGCTTCCACCACTTTCCCATCGAGCGCCAGCTCAACGGCACCTACGCCATT
GCCGGCGGCAAAGCGCACTGTGGACCGGCAGAGCTCTGCGAGTTCTACTCGCGCGA
CCCCGACGGGCTGCCCTGCAACCTGCGCAAGCCGTGCAACCGGCCGTCGGGCCTCG
AGCCGCAGCCGGGGGTCTTCGACTGCCTGCGAGACGCCATGGTGCGTGACTACGTG
CGCCAGACGTGGAAGCTGGAGGGCGAGGCCCTGGAGCAGGCCATCATCAGCCAGGC
CCCGCAGGTGGAGAAGCTCATTGCTACGACGGCCCACGAGCGGATGCCCTGGTACC
ACAGCAGCCTGACGCGTGAGGAGGCCGAGCGCAAACTTTACTCTGGGGCGCAGACC
GACGGCAAGTTCCTGCTGAGGCCGCGGAAGGAGCAGGGCACATACGCCCTGTCCCT
CATCTATGGGAAGACGGTGTACCACTACCTCATCAGCCAAGACAAGGCGGGCAAGT
ACTGCATTCCCGAGGGCACCAAGTTTGACACGCTCTGGCAGCTGGTGGAGTATCTGA
AGCTGAAGGCGGACGGGCTCATCTACTGCCTGAAGGAGGCCTGCCCCAACAGCAGT
GCCAGCAACGCCTCAGGGGCTGCTGCTCCCACACTCCCAGCCCACCCATCCACGTTG
ACTCATCCTCAGAGACGAATCGACACCCTCAACTCAGATGGATACACCCCTGAGCC
AGCACGCATAACGTCCCCAGACAAACCGCGGCCGATGCCCATGGACACGAGCGTGT
ATGAGAGCCCCTACAGCGACCCAGAGGAGCTCAAGGACAAGAAGCTCTTCCTGAAG
CGCGATAACCTCCTCATAGCTGACATTGAACTTGGCTGCGGCAACTTTGGCTCAGTG
CGCCAGGGCGTGTACCGCATGCGCAAGAAGCAGATCGACGTGGCCATCAAGGTGCT
GAAGCAGGGCACGGAGAAGGCAGACACGGAAGAGATGATGCGCGAGGCGCAGATC
ATGCACCAGCTGGACAACCCCTACATCGTGCGGCTCATTGGCGTCTGCCAGGCCGAG
GCCCTCATGCTGGTCATGGAGATGGCTGGGGGCGGGCCGCTGCACAAGTTCCTGGTC
GGCAAGAGGGAGGAGATCCCTGTGAGCAATGTGGCCGAGCTGCTGCACCAGGTGTC

CATGGGGATGAAGTACCTGGAGGAGAAGAACTTTGTGCACCGTGACCTGGCGGCCC
GCAACGTCCTGCTGGTTAACCGGCACTACGCCAAGATCAGCGACTTTGGCCTCTCCA
AAGCACTGGGTGCCGACGACAGCTACTACACTGCCCGCTCAGCAGGGAAGTGGCCG
CTCAAGTGGTACGCACCCGAATGCATCAACTTCCGCAAGTTCTCCAGCCGCAGCGAT
GTCTGGAGCTATGGGGTCACCATGTGGGAGGCCTTGTCCTACGGCCAGAAGCCCTAC
AAGAAGATGAAAGGGCCGGAGGTCATGGCCTTCATCGAGCAGGGCAAGCGGATGG
AATGCCCACCAGAGTGTCCACCCGAACTGTACGCACTCATGAGTGACTGCTGGATCT
ACAAGTGGGAGGATCGCCCCGACTTCCTGACCGTGGAGCAGCGCATGCGAGCCTGT
TACTACAGCCTGGCCAGCAAGGTGGAAGGGCCCCCAGGCAGCACACAGAAGGCTGA
GGCTGCCTGTGCCTGAGCTCCCGCTGCCCAGGGGAGCCCTCCACACCGGCTCTTCCC
CACCCTCAGCCCCACCCCAGGTCCTGCAGTCTGGCTGAGCCCTGCTTGGTTGTCTCC
ACACACAGCTGGGCTGTGGTAGGGGGTGTCTCAGGCCACACCGGCCTTGCATTGCCT
GCCTGGCCCCCTGTCCTCTCTGGCTGGGGAGCAGGGAGGTCCGGGAGGGTGCGGCT
GTGCAGCCTGTCCTGGGCTGGTGGCTCCCGGAGGGCCCTGAGCTGAGGGCATTGCTT
ACACGGATGCCTTCCCCTGGGCCCTGACATTGGAGCCTGGGCATCCTCAGGTGGTCA
GGCGTAGATCACCAGAATAAACCCAGCTTCCCTCTTGAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

[00493] Unless specifically stated or obvious from context, as used herein, the term "or" is understood to be inclusive. Unless specifically stated or obvious from context, as used herein, the terms "a", "an", and "the" are understood to be singular or plural.

[00494] Unless specifically stated or obvious from context, as used herein, the term “about” is understood as within a range of normal tolerance in the art, for example within 2 standard deviations of the mean. About can be understood as within 10%, 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2%, 1%, 0.5%, 0.1%, 0.05%, or 0.01% of the stated value. Unless otherwise clear from context, all numerical values provided herein are modified by the term about.

[00495] Ranges provided herein are understood to be shorthand for all the values within the range. For example, a range of 1 to 50 is understood to include any number, combination of numbers, or sub-range from the group consisting 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50.

[00496] The recitation of a listing of chemical groups in any definition of a variable herein includes definitions of that variable as any single group or combination of listed groups. The recitation of an embodiment for a variable or aspect herein includes that embodiment as any single embodiment or in combination with any other embodiments or portions thereof.

[00497] Any compositions or methods provided herein can be combined with one or more of any of the other compositions and methods provided herein.

BRIEF DESCRIPTION OF THE DRAWINGS

[00498] FIGs. 1A-1B are illustrations of three proteins that impact T cell function. FIG. 1A is an illustratration of the TRAC protein, which is a key component in graft versus host disease. FIG. 1B is an illustratration of the B2M protein, a component of the MHC class 1 antigen presenting complex present on nucleated cells that can be recognized by a host's CD8+ T cells. FIG. 1C is an illustratration of T cell signaling that leads to expression of the PDCD1 gene, and the resulting PD-1 protein acts to inhibit the T cell signaling.

[00499] FIG. 2 is a graph of the percentage of cells with knocked down expression of target genes after base editing. "EP" denotes electroporation.

[00500] FIG. 3 is a graph of the percentages of the observed types of genetic modification in untransduced cells or in cells transduced with a BE4 base editing system or a Cas9 nuclease.

[00501] FIG. 4 is a graph depicting target nucleotide modification percentage as measured by percentage of cells that are negative for target protein expression as determined by flow cytometry (FC) in cells transduced with BE4 and sgRNAs directing BE4 to splice site acceptors (SA) or donors (SD) or that generate a STOP codon. Control cells were mock electroporated (EP).

[00502] FIG. 5 is a diagram of the BE4 system disrupting splice site acceptors (SA), splice donors (SD), or generate STOP codons.

[00503] FIG. 6 is a chart summarizing off-target binding sites of sgRNAs employed to disrupt target genes.

[00504] FIG. 7 is a graph summarizing flow cytometry (FC) data of the percentage of cells edited with BE4 or Cas9 that exhibit reduced protein expression. Cells were either gated to B2M or CD3, the latter being a proxy for TRAC expression.

[00505] FIG. 8A is a scatter plot of FACS data of unedited control cells. FIG. 8B is a scatter plot of FACS data of cells that have been edited at the B2M, TRAC, and PD1 loci.

[00506] FIG. 9 is a graph illustratrating the effectiveness of the base editing techniques described herein to modify specific genes that can negatively impact CAR-T immunotherapy.

[00507] FIG. 10 is a diagram depicting a droplet digital PCR (ddPCR) protocol to detect and quantify gene modifications and translocations.

[00508] FIG. 11 presents two graphs showing the data generated from next generation sequencing (NGS) analysis or ddPCR of cells edited using either the BE4 system or the Cas9 system.

[00509] FIG. 12 is a schematic diagram that illustrates the role Cbl-b plays in suppressing T cell activation.

[00510] FIG. 13 is a graph depicting the efficiency of Cbl-b knockdown by disruption of splice sites. SA = Splice Acceptor; SD = Splice Donor; STOP – generated STOP codon; 2° Only = secondary antibody only; C373 refers to a loss of function variant (C373R); RL1-A::APC-A = laser; ICS = intracellular staining.

[00511] FIG. 14 is a graph illustrating the rate of Cas12b-mediated indels in the GRIN2B and DNMT1 genes in T cells. EP denotes electroporation.

[00512] FIG. 15 is a graph summarizing fluorescence assisted cell sorting (FACS) data of cells transduced via electroporation (EP) with bvCas12b and guide RNAs specific for TRAC, GRIN2B, and DNMT1 and gated for CD3.

[00513] FIG. 16 is a scatter plot of fluorescence assisted cell sorting data of cells transduced CAR-P2A-mCherry lentivirus demonstrating CAR expression.

[00514] FIG. 17 is a scatter plot of fluorescence assisted cell sorting data demonstrating CAR expression in cells transduced with a poly(1,8-octanediol citrate) (POC) lentiviral vector.

[00515] FIG. 18 is graph showing that BE4 produced efficient, durable gene knockout with high product purity.

[00516] FIG. 19A is a representative FACS analysis showing loss of surface expression of a protein due to gene knockout by BE4 or spCas9. FIG. 19B is a graph show that gene knockout by BE4 or spCas9 produces loss of B2M surface expression.

[00517] FIG. 20 is a schematic depicting the locations of B2M, TRAC, and PD-1 target sites. Translocations can be detected when B2M, TRAC, and PD-1 sequences recombine.

[00518] FIG. 21 is a graph showing that multiplexed base editing does not significantly impair cell expansion.

[00519] FIG. 22 is a graph showing that BE4 generated triple-edited T cells with similar on-target editing efficiency and cellular phenotype as spCas9.

[00520] FIG. 23 depicts flow cytometry analysis showing the generation of triple-edited $CD3^-$, $B2M^-$, $PD1^-$ T cells.

[00521] FIG. 24 depicts flow cytometry analysis showing the CAR expression in BE4 and Cas9 edited cells.

[00522] FIG. 25 is a graph showing CAR-T cell killing or antigen positive cells.

[00523] FIG. 26 are graphs showing that Cas12b and BE4 can be paired for efficient multiplex editing in T cells.

[00524] FIG. 27 is a graph showing that Cas12b can direct insertion of a chimeric antigen receptor (CAR) into a locus by introducing into a cell a double-stranded DNA template encoding the CAR in the presence of a Cas12 nuclease and an sgRNA targeting the locus.

[00525] FIG. 28A and 28B are graphs showing protein knockdown (% Negative) using base editing targeting the genes indicated in the figures as determined by flow cytometry, gated with respect to an unedited control. The figures represent results from replicate experiments. Bars for each set of conditions are presented in the order (from left to right) as listed in the key (top to bottom). The identity of each bar in the grouping of eight bar graphs correspond to, from left to right, CD3, CD7, CD52, PD1, B2M CD2, HLADR (CIITA surrogate), and CD5.

DETAILED DESCRIPTION OF THE INVENTION

[00526] The present invention features genetically modified immune cells having enhanced anti-neoplasia activity, resistance to immune suppression, and decreased risk of eliciting a graft versus host reaction or a host versus graft reaction, or a combination thereof. The present invention also features methods for producing and using these modified immune cells (e.g., immune effector cells, such as T cells).

[00527] In one embodiment, a subject having or having a propensity to develop graft versus host disease (GVHD) is administered a CAR-T cell that lacks or has reduced levels of functional TRAC. In one embodiment, a subject having or having a propensity to develop host versus graft

disease (HVGD) is administered a CAR-T cell that lacks or has reduced levels of functional beta2 microglobulin (B2M).

[00528] The modification of immune effector cells to express chimeric antigen receptors and to knockout or knockdown specific genes to diminish the negative impact that their expression can have on immune cell function is accomplished using a base editor system comprising a cytidine deaminase or adenosine deaminase as described herein.

[00529] Autologous, patient-derived chimeric antigen receptor-T cell (CAR-T) therapies have demonstrated remarkable efficacy in treating some hematologic cancers. While these products have led to significant clinical benefit for patients, the need to generate individualized therapies creates substantial manufacturing challenges and financial burdens. Allogeneic CAR-T therapies were developed as a potential solution to these challenges, having similar clinical efficacy profiles to autologous products while treating many patients with cells derived from a single healthy donor, thereby substantially reducing cost of goods and lot-to-lot variability.

[00530] Most first-generation allogeneic CAR-Ts use nucleases to introduce two or more targeted genomic DNA double strand breaks (DSBs) in a target T cell population, relying on error-prone DNA repair to generate mutations that knock out target genes in a semi-stochastic manner. Such nuclease-based gene knockout strategies aim to reduce the risk of graft-versus-host-disease and host rejection of CAR-Ts. However, the simultaneous induction of multiple DSBs results in a final cell product containing large-scale genomic rearrangements such as balanced and unbalanced translocations, and a relatively high abundance of local rearrangements including inversions and large deletions. Furthermore, as increasing numbers of simultaneous genetic modifications are made by induced DSBs, considerable genotoxicity is observed in the treated cell population. This has the potential to significantly reduce the cell expansion potential from each manufacturing run, thereby decreasing the number of patients that can be treated per healthy donor.

[00531] Base editors (BEs) are a class of emerging gene editing reagents that enable highly efficient, user-defined modification of target genomic DNA without the creation of DSBs. Here, an alternative means of producing allogeneic CAR-T cells is proposed by using base editing technology to reduce or eliminate detectable genomic rearrangements while also improving cell expansion. As shown herein, in contrast to a nuclease-only editing strategy, concurrent modification of multiple gene loci, for example, three, four, five, six, seven, eight, night, ten, or

more genetic loci by base editing produces highly efficient gene knockouts with no detectable translocation events.

[00532] In some embodiments, at least 1, 2, 3, 4, 5, 6, 7, 8, or more genes or regulatory elements thereof are modified in an immune cell with the base editing compositions and methods provided herein. In some embodiments, the at least 1, 2, 3, 4, 5, 6, 7, 8, or more genes or regulatory elements thereof comprise one or more genes selected from CD3e, CD3 delta, CD3 gamma, TRAC, TRBC1, and TRBC2. In some embodiments, the at least 1, 2, 3, 4, 5, 6, 7, 8, or more genes or regulatory elements thereof comprise one or more genes selected from CD3e, CD3 delta, CD3 gamma, TRAC, TRBC1, and TRBC2, CD7, and CD52. In some embodiments, the at least 1, 2, 3, 4, 5, 6, 7, 8, or more genes or regulatory elements thereof comprise one or more genes selected from CD3e, CD3 delta, CD3 gamma, TRAC, TRBC1, TRBC2, CD2, CD5, CD7, and CD52. In some embodiments, the at least 1, 2, 3, 4, 5, 6, 7, 8, or more genes or regulatory elements thereof comprise one or more genes selected from TRAC, CD7, and CD52. In some embodiments, the at least 1, 2, 3, 4, 5, 6, 7, 8, or more genes or regulatory elements thereof comprise one or more genes selected from TRAC, CD2, CD5, CD7, and CD52. In some embodiments, the at least 1, 2, 3, 4, 5, 6, 7, 8, or more genes or regulatory elements thereof comprise one or more genes selected from CD2, CD3 epsilon, CD3 gamma, CD3 delta, CD4, CD5, CD7, CD30, CD33, CD52, CD70, and CIITA. In some embodiments, the at least 1, 2, 3, 4, 5, 6, 7, 8, or more genes or regulatory elements thereof are selected from CD2, CD3 epsilon, CD3 gamma, CD3 delta, CD4, CD5, CD7, CD30, CD33, CD52, CD70, and CIITA. In some embodiments, the at least 1, 2, 3, 4, 5, 6, 7, 8, or more genes or regulatory elements thereof comprise one or more genes selected from CD2, CD3 epsilon, CD3 gamma, CD3 delta, CD4, CD5, CD7, CD30, CD33, CD52, CD70, and CIITA. In some embodiments, the at least 1, 2, 3, 4, 5, 6, 7, 8, or more genes or regulatory elements thereof are selected from ACAT1, ACLY, ADORA2A, AXL, B2M , BATF, BCL2L11, BTLA, CAMK2D, cAMP, CASP8, Cblb, CCR5, CD2, CD3D, CD3E, CD3G, CD4, CD5, CD7, CD8A, CD33, CD38, CD52, CD70, CD82, CD86, CD96, CD123, CD160, CD244, CD276, CDK8, CDKN1B, Chi3l1, CIITA, CISH, CSF2CSK, CTLA-4, CUL3, Cyp11a1, DCK, DGKA, DGKZ, DHX37, ELOB(TCEB2), ENTPD1 (CD39), FADD, FAS, GATA3, IL6, IL6R, IL10, IL10RA, IRF4, IRF8, JUNB, Lag3, , LAIR-1 (CD305), LDHA, LIF, LYN, MAP4K4, MAPK14, MCJ, MEF2D, MGAT5, NR4A1, NR4A2, NR4A3, NT5E (CD73), ODC1, OTULINL (FAM105A), PAG1, PDCD1, PDIA3, PHD1 (EGLN2),

PHD2 (EGLN1), PHD3 (EGLN3), PIK3CD, PIKFYVE, PPARa, PPARd, PRDM1, PRKACA, PTEN, PTPN2, PTPN6, PTPN11, PVRIG (CD112R), RASA2, RFXANK, SELPG/PSGL1, SIGLEC15, SLA, SLAMF7, SOCS1, Spry1, Spry2, STK4, SUV39, H1TET2, TGFbRII, TIGIT, Tim-3, TMEM222, TNFAIP3, TNFRSF8 (CD30), TNFRSF10B, TOX, TOX2, , TRAC, TRBC1, TRBC2, UBASH3A, VHL, VISTA, XBP1, YAP1, and ZC3H12A. In some embodiments, at least 8 genes selected from CD2, CD3 epsilon, CD3 gamma, CD3 delta, CD4, CD5, CD7, CD30, CD33, CD52, CD70, and CIITA or regulatory elements thereof are modified with the base editing compositions and methods provided herein.

[00533] In one aspect, provided herein is a universal CAR-T cell. In some embodiments, the CAR-T cell described herein is an allogeneic cell. In some embodiments, the universal CAR-T cell is an allogeneic T cell that can be used to express a desired CAR, and can be universally applicable, irrespective of the donor and the recipient's immunogenic compatibility. An allogenic immune cell may be derived from one or more donors. In certain embodiments, the allogenic immune cell is derived from a single human donor. For example, the allogenic T cell may be derived from PBMCs of a single healthy human donor. In certain embodiments, the allogenic immune cell is derived from multiple human donors. In some embodiments, an universal CAR-T cell may be generated, as described herein by using gene modification to introduce concurrent edits at multiple gene loci, for example, three, four, five, six, seven, eight, nine, ten or more genetic loci. A modification, or concurrent modifications as described herein may be a genetic editing, such as a base editing, generated by a base editor. The base editor may be a C base editor or A base editor. As is discussed herein, base editing may be used to achieve a gene disruption, such that the gene is not expressed. A modification by base editing may be used to achieve a reduction in gene expression. In some embodiments base editor may be used to introduce a genetic modification such that the edited gene does not generate a structurally or functionally viable protein product. In some embodiments, a modification, such as the concurrent modifications described herein may comprise a genetic editing, such as base editing, such that the expression or functionality of the gene product is altered in any way. For example, the expression of the gene product may be enhanced or upregulated as compared to baseline expression levels. In some embodiments the activity or functionality of the gene product may be upregulated as a result of the base editing, or multiple base editing events acting in concert.

[00534] In some embodiments, generation of universal CAR-T cell may be advantageous over autologous T cell (CAR-T), which may be difficult to generate for an urgent use. Allogeneic approaches are preferred over autologous cell preparation for a number of situations related to uncertainty of engineering autologous T cells to express a CAR and finally achieving the desired cellular products for a transplant at the time of medical emergency. However, for allogeneic T cells, or "off-the-shelf" T cells, it is important to carefully negotiate the host's reactivity to the CAR-T cells (HVGD) as well as the allogeneic T cell's potential hostility towards a host cell (GVHD). Given the scenario, base editing can be successfully used to generate multiple simultaneous gene editing events, such that (a) it is possible to generate a platform cell type that is devoid of or expresses low amounts of an endogenous T cell receptor, for example, a TCR alpha chain (such a via base editing of TRAC), or a TCR beta chain (such a through base editing of TRBC1/TRBC2); (b) it is possible to reduce or down regulate expression of antigens that may be incompatible to a host tissue system and vice versa.

[00535] In some embodiments, the methods described herein can be used to generate an autologous T cell expressing a CAR-T.

[00536] In some embodiment, multiple base editing events can be accomplished in a single electroporation event, thereby reducing electroporation event associated toxicity. Any known methods for incorporation of exogenous genetic material into a cell may be used to replace electroporation, and such methods known in the art are hereby contemplated for use in any of the methods described herein.

[00537] In one experiment, the base editor BE4 demonstrated high efficiency multiplex base editing of three cell surface targets in T cells (TRAC, B2M, and PD-1), knocking out gene expression by 95%, 95% and 88%, respectively, in a single electroporation to generate cell populations with high percentages of cells with reduced protein expression of B2M and CD3. Editing each of these genes may be useful in the creation of CAR-T cell therapies with improved therapeutic properties. Each of the genes was silenced by a single targeted base change (C to T) without the creation of double strand breaks. As a result, the BE4-treated cells also did not show any measurable translocations (large-scale genomic rearrangements), whereas cells receiving the same three edits with a nuclease did show detectable genomic rearrangements.

[00538] Thus, coupling nuclease-based knockout of the *TRAC* gene with simultaneous BE-mediated knockout of two additional genes yields a homogeneous allogeneic T cell population

with minimal genomic rearrangements, . In some embodiments, the simultaneous BE mediated knockout or knockdown, or a combination thereof, may be performed in 2 additional genes, or 3 additional genes, or 4 additional genes, or 5 additional genes, or 6 additional genes, or 7 additional genes, or 8 additional genes, or 9 additional genes, or 10 additional genes, or 11 additional genes, or 12 additional genes, or more, to yield a homogenous allogeneic T cell population with minimal genomic rearrangements, and enabling targeted insertion of a CAR transgene at the TRAC locus. In some embodiments, the disclosure provides three simultaneous gene knockouts or knockdowns, by base editing along with a CAR transgene at the TRAC locus. In some embodiments, the disclosure provides four simultaneous gene knockouts or knockdowns, by base editing along with a CAR transgene at the TRAC locus. In some embodiments, the disclosure provides five simultaneous gene knockouts or knockdowns, by base editing along with a CAR transgene at the TRAC locus. In some embodiments, the disclosure provides six simultaneous gene knockouts or knockdowns, by base editing along with a CAR transgene at the TRAC locus. In some embodiments, the disclosure provides seven simultaneous gene knockouts or knockdowns, by base editing along with a CAR transgene at the TRAC locus. In some embodiments, the disclosure provides eight simultaneous gene knockouts or knockdowns, by base editing along with a CAR transgene at the TRAC locus. In some embodiments, the disclosure provides nine simultaneous gene knockouts or knockdowns, by base editing along with a CAR transgene at the TRAC locus. In some embodiments, the disclosure provides ten simultaneous gene knockouts or knockdowns, by base editing along with a CAR transgene at the TRAC locus. In some embodiments, the disclosure provides eleven simultaneous gene knockouts or knockdowns, by base editing along with a CAR transgene at the TRAC locus. In some embodiments, the disclosure provides twelve simultaneous gene knockouts or knockdowns, by base editing along with a CAR transgene at the TRAC locus. In some embodiments, the disclosure provides thirteen simultaneous gene knockouts or knockdowns, by base editing along with a CAR transgene at the TRAC locus. In some embodiments, the disclosure provides fourteen simultaneous gene knockouts or knockdowns, by base editing along with a CAR transgene at the TRAC locus. In some embodiments, the disclosure provides fifteen simultaneous gene knockouts or knockdowns, by base editing along with a CAR transgene at the TRAC locus. In some embodiments, the disclosure provides sixteen simultaneous gene knockouts or knockdowns, by base editing along with a CAR transgene at the

TRAC locus. In some embodiments, the disclosure provides seventeen simultaneous gene knockouts or knockdowns, by base editing along with a CAR transgene at the TRAC locus. In some embodiments, the disclosure provides eighteen simultaneous gene knockouts or knockdowns, by base editing along with a CAR transgene at the TRAC locus. In some embodiments, the disclosure provides nineteen simultaneous gene knockouts or knockdowns, by base editing along with a CAR transgene at the TRAC locus. In some embodiments, the disclosure provides twenty simultaneous gene knockouts or knockdowns, by base editing along with a CAR transgene at the TRAC locus. Taken together, this demonstrates that base editing alone or in combination with a single nuclease knockout and CAR insertion is a useful strategy for generating allogeneic T cells with minimal genomic rearrangements compared to nuclease-alone approaches. This method addresses known limitations of multiplex-edited T cell products and are a promising development towards the next generation of precision cell based therapies.

Chimeric Antigen Receptor and CAR-T cells

[00539] The invention provides immune cells modified using nucleobase editors described herein that express chimeric antigen receptors. Modification of immune cells to express a chimeric antigen receptor can enhance an immune cell's immunoreactive activity, wherein the chimeric antigen receptor has an affinity for an epitope on an antigen, wherein the antigen is associated with an altered fitness of an organism. For example, the chimeric antigen receptor can have an affinity for an epitope on a protein expressed in a neoplastic cell. Because the CAR-T cells can act independently of major histocompatibility complex (MHC), activated CAR-T cells can kill the neoplastic cell expressing the antigen. The direct action of the CAR-T cell evades neoplastic cell defensive mechanisms that have evolved in response to MHC presentation of antigens to immune cells.

[00540] In some embodiments, the invention provides immune effector cells that express chimeric antigen receptors that target B cells involved in an autoimmune response (e.g., B cells of a subject that express antibodies generated against the subject's own tissues).

[00541] Some embodiments comprise autologous immune cell immunotherapy, wherein immune cells are obtained from a subject having a disease or altered fitness characterized by cancerous or otherwise altered cells expressing a surface marker. The obtained immune cells are genetically modified to express a chimeric antigen receptor and are effectively redirected against

specific antigens. Thus, in some embodiments, immune cells are obtained from a subject in need of CAR-T immunotherapy. In some embodiments, these autologous immune cells are cultured and modified shortly after they are obtained from the subject. In other embodiments, the autologous cells are obtained and then stored for future use. This practice may be advisable for individuals who may be undergoing parallel treatment that will diminish immune cell counts in the future. In allogeneic immune cell immunotherapy, immune cells can be obtained from a donor other than the subject who will be receiving treatment. The immune cells, after modification to express a chimeric antigen receptor, are administered to a subject for treating a neoplasia. In some embodiments, immune cells to be modified to express a chimeric antigen receptor can be obtained from pre-existing stock cultures of immune cells.

[00542] Immune cells and/or immune effector cells can be isolated or purified from a sample collected from a subject or a donor using standard techniques known in the art. For example, immune effector cells can be isolated or purified from a whole blood sample by lysing red blood cells and removing peripheral mononuclear blood cells by centrifugation. The immune effector cells can be further isolated or purified using a selective purification method that isolates the immune effector cells based on cell-specific markers such as CD25, CD3, CD4, CD8, CD28, CD45RA, or CD45RO. In one embodiment, CD25+ is used as a marker to select regulatory T cells. In another embodiment, the invention provides T cells that have targeted gene knockouts at the TCR constant region (TRAC), which is responsible for TCRαβ surface expression. TCR alphabeta-deficient CAR T cells are compatible with allogeneic immunotherapy (Qasim et al., Sci. Transl. Med. 9, eaaj2013 (2017); Valton et al., Mol Ther. 2015 Sep; 23(9): 1507–1518). If desired, residual TCRalphabeta T cells are removed using CliniMACS magnetic bead depletion to minimize the risk of GVHD. In another embodiment, the invention provides donor T cells selected ex vivo to recognize minor histocompatibility antigens expressed on recipient hematopoietic cells, thereby minimizing the risk of graft-versus-host disease (GVHD), which is the main cause of morbidity and mortality after transplantation (Warren et al., Blood 2010;115(19):3869-3878). Another technique for isolating or purifying immune effector cells is flow cytometry. In fluorescence activated cell sorting a fluorescently labelled antibody with affinity for an immune effector cell marker is used to label immune effector cells in a sample. A gating strategy appropriate for the cells expressing the marker is used to segregate the cells. For example, T lymphocytes can be separated from other cells in a sample by using, for example, a

fluorescently labeled antibody specific for an immune effector cell marker (e.g., CD4, CD8, CD28, CD45) and corresponding gating strategy. In one embodiment, a CD45 gating strategy is employed. In some embodiments, a gating strategy for other markers specific to an immune effector cell is employed instead of, or in combination with, the CD45 gating strategy.

[00543] The immune effector cells contemplated in the invention are effector T cells. In some embodiments, the effector T cell is a naïve $CD8^+$ T cell, a cytotoxic T cell, or a regulatory T (Treg) cell. In some embodiments, the effector T cells are thymocytes, immature T lymphocytes, mature T lymphocytes, resting T lymphocytes, or activated T lymphocytes. In some embodiments the immune effector cell is a $CD4^+$ $CD8^+$ T cell or a $CD4^-$ $CD8^-$ T cell. In some embodiments the immune effector cell is a T helper cell. In some embodiments the T helper cell is a T helper 1 (Th1), a T helper 2 (Th2) cell, or a helper T cell expressing CD4 (CD4+ T cell). In some embodiments, the immune effector cell is any other subset of T cells. The modified immune effector cell may express, in addition to the chimeric antigen receptor, an exogenous cytokine, a different chimeric receptor, or any other agent that would enhance immune effector cell signaling or function. For example, coexpression of the chimeric antigen receptor and a cytokine may enhance the CAR-T cell's ability to lyse a target cell.

[00544] Chimeric antigen receptors as contemplated in the present invention comprise an extracellular binding domain, a transmembrane domain, and an intracellular domain. Binding of an antigen to the extracellular binding domain can activate the CAR-T cell and generate an effector response, which includes CAR-T cell proliferation, cytokine production, and other processes that lead to the death of the antigen expressing cell. In some embodiments of the present invention, the chimeric antigen receptor further comprises a linker.

[00545] The extracellular binding domain of a chimeric antigen receptor contemplated herein comprises an amino acid sequence of an antibody, or an antigen binding fragment thereof, that has an affinity for a specific antigen. In various embodiments, the CAR specifically binds 5T4. Exemplary anti-5T4 CARs include, without limitation, CART-5T4 (Oxford BioMedica plc) and UCART-5T4 (Cellectis SA).

[00546] In various embodiments, the CAR specifically binds Alpha-fetoprotein. Exemplary anti-Alpha-fetoprotein CARs include, without limitation, ET-1402 (Eureka Therapeutics Inc). In various embodiments, the CAR specifically binds Axl. Exemplary anti-Axl CARs include, without limitation, CCT-301-38 (F1 Oncology Inc). In various embodiments, the CAR

specifically binds B7H6. Exemplary anti-B7H6 CARs include, without limitation, CYAD-04 (Celyad SA).

[00547] In various embodiments, the CAR specifically binds BCMA. Exemplary anti-BCMA CARs include, without limitation, ACTR-087 + SEA-BCMA (Seattle Genetics Inc), ALLO-715 (Cellectis SA), ARI-0002 (Institut d'Investigacions Biomediques August Pi I Sunyer), bb-2121 (bluebird bio Inc), bb-21217 (bluebird bio Inc), CART-BCMA (University of Pennsylvania), CT-053 (Carsgen Therapeutics Ltd), Descartes-08 (Cartesian Therapeutics), FCARH-143 (Juno Therapeutics Inc), ICTCAR-032 (Innovative Cellular Therapeutics Co Ltd), IM21 CART (Beijing Immunochina Medical Science & Technology Co Ltd), JCARH-125 (Memorial Sloan-Kettering Cancer Center), KITE-585 (Kite Pharma Inc), LCAR-B38M (Nanjing Legend Biotech Co Ltd), LCAR-B4822M (Nanjing Legend Biotech Co Ltd), MCARH-171 (Memorial Sloan-Kettering Cancer Center), P-BCMA-101 (Poseida Therapeutics Inc), P-BCMA-ALLO1 (Poseida Therapeutics Inc), spCART-269 (Shanghai Unicar-Therapy Bio-medicine Technology Co Ltd), and BCMA02/bb2121 (bluebird bio Inc). The polypeptide sequence of the BCMA02/bb2121 CAR is provided below:

MALPVTALLLPLALLLHAARPDIVLTQSPPSLAMSLGKRATISCRASESVTILGSHLIHW
YQQKPGQPPTLLIQLASNVQTGVPARFSGSGSRTDFTLTIDPVEEDDVAVYYCLQSRTIP
RTFGGGTKLEIKGSTSGSGKPGSGEGSTKGQIQLVQSGPELKKPGETVKISCKASGYTFT
DYSINWVKRAPGKGLKWMGWINTETREPAYAYDFRGRFAFSLETSASTAYLQINNLKYED
TATYFCALDYSYAMDYWGQGTSVTVSSAAATTTPAPRPPTPAPTIASQPLSLRPEACRPA
AGGAVHTRGLDFACDIYIWAPLAGTCGVLLLSLVITLYCKRGRKKLLYIFKQPFMRPVQT
TQEEDGCSCRFPEEEEGGCELRVKFSRSADAPAYQQGQNQLYNELNLGRREEYDVLDKRR
GRDPEMGGKPRRKNPQEGLYNELQKDKMAEAYSEIGMKGERRRGKGHDGLYQGLSTATKD
TYDALHMQALPPR

[00548] In various embodiments, the CAR specifically binds CCK2R. Exemplary anti-CCK2R CARs include, without limitation, anti-CCK2R CAR-T adaptor molecule (CAM) + anti-FITC CAR T-cell therapy (cancer), Endocyte/Purdue (Purdue University),

[00549] In various embodiments, the CAR specifically binds a CD antigen. Exemplary anti-CD antigen CARs include, without limitation, VM-802 (ViroMed Co Ltd). In various embodiments, the CAR specifically binds CD123. Exemplary anti-CD123 CARs include, without limitation, MB-102 (Fortress Biotech Inc), RNA CART123 (University of Pennsylvania), SFG–iMC-CD123.zeta (Bellicum Pharmaceuticals Inc), and UCART-123

(Cellectis SA). In various embodiments, the CAR specifically binds CD133. Exemplary anti-CD133 CARs include, without limitation, KD-030 (Nanjing Kaedi Biotech Inc). In various embodiments, the CAR specifically binds CD138. Exemplary anti-CD138 CARs include, without limitation, ATLCAR.CD138 (UNC Lineberger Comprehensive Cancer Center) and CART-138 (Chinese PLA General Hospital). In various embodiments, the CAR specifically binds CD171. Exemplary anti-CD171 CARs include, without limitation, JCAR-023 (Juno Therapeutics Inc). In various embodiments, the CAR specifically binds CD19. Exemplary anti-CD19 CARs include, without limitation, 1928z-41BBL (Memorial Sloan-Kettering Cancer Center), 1928z-E27 (Memorial Sloan-Kettering Cancer Center), 19-28z-T2 (Guangzhou Institutes of Biomedicine and Health), 4G7-CARD (University College London), 4SCAR19 (Shenzhen Geno-Immune Medical Institute), ALLO-501 (Pfizer Inc), ATA-190 (QIMR Berghofer Medical Research Institute), AUTO-1 (University College London), AVA-008 (Avacta Ltd), axicabtagene ciloleucel (Kite Pharma Inc), BG-T19 (Guangzhou Bio-gene Technology Co Ltd), BinD-19 (Shenzhen BinDeBio Ltd.), BPX-401 (Bellicum Pharmaceuticals Inc), CAR19h28TM41BBz (Westmead Institute for Medical Research), C-CAR-011 (Chinese PLA General Hospital), CD19CART (Innovative Cellular Therapeutics Co Ltd), CIK-CAR.CD19 (Formula Pharmaceuticals Inc), CLIC-1901 (Ottawa Hospital Research Institute), CSG-CD19 (Carsgen Therapeutics Ltd), CTL-119 (University of Pennsylvania), CTX-101 (CRISPR Therapeutics AG), DSCAR-01 (Shanghai Hrain Biotechnology), ET-190 (Eureka Therapeutics Inc), FT-819 (Memorial Sloan-Kettering Cancer Center), ICAR-19 (Immune Cell Therapy Inc), IM19 CAR-T (Beijing Immunochina Medical Science & Technology Co Ltd), JCAR-014 (Juno Therapeutics Inc), JWCAR-029 (MingJu Therapeutics (Shanghai) Co., Ltd), KD-C-19 (Nanjing Kaedi Biotech Inc), LinCART19 (iCell Gene Therapeutics), lisocabtagene maraleucel (Juno Therapeutics Inc), MatchCART (Shanghai Hrain Biotechnology), MB-CART19.1 (Shanghai Children's Medical Center), PBCAR-0191 (Precision BioSciences Inc), PCAR-019 (PersonGen Biomedicine (Suzhou) Co Ltd), pCAR-19B (Chongqing Precision Biotech Co Ltd), PZ-01 (Pinze Lifetechnology Co Ltd), RB-1916 (Refuge Biotechnologies Inc), SKLB-083019 (Chengdu Yinhe Biomedical Co Ltd), spCART-19 (Shanghai Unicar-Therapy Bio-medicine Technology Co Ltd), TBI-1501 (Takara Bio Inc), TC-110 (TCR2 Therapeutics Inc), TI-1007 (Timmune Biotech Inc), tisagenlecleucel (Abramson Cancer Center of the University of Pennsylvania), U-CART (Shanghai Bioray Laboratory Inc), UCART-19 (Wugen

Inc), UCART-19 (Cellectis SA), vadacabtagene leraleucel (Memorial Sloan-Kettering Cancer Center), XLCART-001 (Nanjing Medical University), and yinnuokati-19 (Shenzhen Innovation Immunotechnology Co Ltd). In various embodiments, the CAR specifically binds CD2. Exemplary anti-CD2 CARs include, without limitation, UCART-2 (Wugen Inc). In various embodiments, the CAR specifically binds CD20. Exemplary anti-CD20 CARs include, without limitation, ACTR-087 (National University of Singapore), ACTR-707 (Unum Therapeutics Inc), CBM-C20.1 (Chinese PLA General Hospital), MB-106 (Fred Hutchinson Cancer Research Center), and MB-CART20.1 (Miltenyi Biotec GmbH).

[00550] In various embodiments, the CAR specifically binds CD22. Exemplary anti-CD22 CARs include, without limitation, anti-CD22 CAR T-cell therapy (B-cell acute lymphoblastic leukemia), University of Pennsylvania (University of Pennsylvania), CD22-CART (Shanghai Unicar-Therapy Bio-medicine Technology Co Ltd), JCAR-018 (Opus Bio Inc), MendCART (Shanghai Hrain Biotechnology), and UCART-22 (Cellectis SA). In various embodiments, the CAR specifically binds CD30. Exemplary anti-CD30 CARs include, without limitation, ATLCAR.CD30 (UNC Lineberger Comprehensive Cancer Center), CBM-C30.1 (Chinese PLA General Hospital), and Hu30-CD28zeta (National Cancer Institute). In various embodiments, the CAR specifically binds CD33. Exemplary anti-CD33 CARs include, without limitation, anti-CD33 CAR gamma delta T-cell therapy (acute myeloid leukemia), TC BioPharm/University College London (University College London), CAR33VH (Opus Bio Inc), CART-33 (Chinese PLA General Hospital), CIK-CAR.CD33 (Formula Pharmaceuticals Inc), UCART-33 (Cellectis SA), and VOR-33 (Columbia University).

[00551] In various embodiments, the CAR specifically binds CD38. Exemplary anti-CD38 CARs include, without limitation, UCART-38 (Cellectis SA). In various embodiments, the CAR specifically binds CD38 A2. Exemplary anti-CD38 A2 CARs include, without limitation, T-007 (TNK Therapeutics Inc). In various embodiments, the CAR specifically binds CD4. Exemplary anti-CD4 CARs include, without limitation, CD4CAR (iCell Gene Therapeutics). In various embodiments, the CAR specifically binds CD44. Exemplary anti-CD44 CARs include, without limitation, CAR-CD44v6 (Istituto Scientifico H San Raffaele). In various embodiments, the CAR specifically binds CD5. Exemplary anti-CD5 CARs include, without limitation, CD5CAR (iCell Gene Therapeutics). In various embodiments, the CAR specifically binds CD7. Exemplary anti-CD7 CARs include, without limitation, CAR-pNK (PersonGen Biomedicine

(Suzhou) Co Ltd), and CD7.CAR/28zeta CAR T cells (Baylor College of Medicine), UCART7 (Washington University in St Louis).

[00552] In various embodiments, the CAR specifically binds CDH17. Exemplary anti-CDH17 CARs include, without limitation, ARB-001.T (Arbele Ltd). In various embodiments, the CAR specifically binds CEA. Exemplary anti-CEA CARs include, without limitation, HORC-020 (HumOrigin Inc). In various embodiments, the CAR specifically binds Chimeric TGF-beta receptor (CTBR). Exemplary anti-Chimeric TGF-beta receptor (CTBR) CARs include, without limitation, CAR-CTBR T cells (bluebird bio Inc). In various embodiments, the CAR specifically binds Claudin18.2. Exemplary anti-Claudin18.2 CARs include, without limitation, CAR-CLD18 T-cells (Carsgen Therapeutics Ltd) and KD-022 (Nanjing Kaedi Biotech Inc).

[00553] In various embodiments, the CAR specifically binds CLL1. Exemplary anti-CLL1 CARs include, without limitation, KITE-796 (Kite Pharma Inc). In various embodiments, the CAR specifically binds DLL3. Exemplary anti-DLL3 CARs include, without limitation, AMG-119 (Amgen Inc). In various embodiments, the CAR specifically binds Dual BCMA/TACI (APRIL). Exemplary anti-Dual BCMA/TACI (APRIL) CARs include, without limitation, AUTO-2 (Autolus Therapeutics Limited). In various embodiments, the CAR specifically binds Dual CD19/CD22. Exemplary anti-Dual CD19/CD22 CARs include, without limitation, AUTO-3 (Autolus Therapeutics Limited) and LCAR-L10D (Nanjing Legend Biotech Co Ltd). In various embodiments, the CAR specifically binds CD19. In various embodiments, the CAR specifically binds Dual CLL1/CD33. Exemplary anti-Dual CLL1/CD33 CARs include, without limitation, ICG-136 (iCell Gene Therapeutics). In various embodiments, the CAR specifically binds Dual EpCAM/CD3. Exemplary anti-Dual EpCAM/CD3 CARs include, without limitation, IKT-701 (Icell Kealex Therapeutics). In various embodiments, the CAR specifically binds Dual ErbB / 4ab. Exemplary anti-Dual ErbB/4ab CARs include, without limitation, LEU-001 (King's College London). In various embodiments, the CAR specifically binds Dual FAP/CD3. Exemplary anti-Dual FAP/CD3 CARs include, without limitation, IKT-702 (Icell Kealex Therapeutics). In various embodiments, the CAR specifically binds EBV. Exemplary anti-EBV CARs include, without limitation, TT-18 (Tessa Therapeutics Pte Ltd).

[00554] In various embodiments, the CAR specifically binds EGFR. Exemplary anti-EGFR CARs include, without limitation, anti-EGFR CAR T-cell therapy (CBLB MegaTAL, cancer), bluebird bio (bluebird bio Inc), anti-EGFR CAR T-cell therapy expressing CTLA-4 checkpoint

inhibitor + PD-1 checkpoint inhibitor mAbs (EGFR-positive advanced solid tumors), Shanghai Cell Therapy Research Institute (Shanghai Cell Therapy Research Institute), CSG-EGFR (Carsgen Therapeutics Ltd), and EGFR-IL12-CART (Pregene (Shenzhen) Biotechnology Co Ltd).

[00555] In various embodiments, the CAR specifically binds EGFRvIII. Exemplary anti-EGFRvIII CARs include, without limitation, KD-035 (Nanjing Kaedi Biotech Inc) and UCART-EgfrVIII (Cellectis SA). In various embodiments, the CAR specifically binds Flt3. Exemplary anti-Flt3 CARs include, without limitation, ALLO-819 (Pfizer Inc) and AMG-553 (Amgen Inc). In various embodiments, the CAR specifically binds Folate receptor. Exemplary anti-Folate receptor CARs include, without limitation, EC17/CAR T (Endocyte Inc). In various embodiments, the CAR specifically binds G250. Exemplary anti-G250 CARs include, without limitation, autologous T-lymphocyte cell therapy (G250-scFV-transduced, renal cell carcinoma), Erasmus Medical Center (Daniel den Hoed Cancer Center).

[00556] In various embodiments, the CAR specifically binds GD2. Exemplary anti-GD2 CARs include, without limitation, 1RG-CART (University College London), 4SCAR-GD2 (Shenzhen Geno-Immune Medical Institute), C7R-GD2.CART cells (Baylor College of Medicine), CMD-501 (Baylor College of Medicine), CSG-GD2 (Carsgen Therapeutics Ltd), GD2-CART01 (Bambino Gesu Hospital and Research Institute), GINAKIT cells (Baylor College of Medicine), iC9-GD2-CAR-IL-15 T-cells (UNC Lineberger Comprehensive Cancer Center), and IKT-703 (Icell Kealex Therapeutics). In various embodiments, the CAR specifically binds GD2 and MUC1. Exemplary anti-GD2/MUC1 CARs include, without limitation, PSMA CAR-T (University of Pennsylvania).

[00557] In various embodiments, the CAR specifically binds GPC3. Exemplary anti-GPC3 CARs include, without limitation, ARB-002.T (Arbele Ltd), CSG-GPC3 (Carsgen Therapeutics Ltd), GLYCAR (Baylor College of Medicine), and TT-14 (Tessa Therapeutics Pte Ltd). In various embodiments, the CAR specifically binds Her2. Exemplary anti-Her2 CARs include, without limitation, ACTR-087 + trastuzumab (Unum Therapeutics Inc), ACTR-707 + trastuzumab (Unum Therapeutics Inc), CIDeCAR (Bellicum Pharmaceuticals Inc), MB-103 (Mustang Bio Inc), RB-H21 (Refuge Biotechnologies Inc), and TT-16 (Baylor College of Medicine). In various embodiments, the CAR specifically binds IL13R. Exemplary anti-IL13R CARs include, without limitation, MB-101 (City of Hope) and YYB-103 (YooYoung

Pharmaceuticals Co Ltd). In various embodiments, the CAR specifically binds integrin beta-7. Exemplary anti-integrin beta-7 CARs include, without limitation, MMG49 CAR T-cell therapy (Osaka University). In various embodiments, the CAR specifically binds LC antigen. Exemplary anti-LC antigen CARs include, without limitation, VM-803 (ViroMed Co Ltd) and VM-804 (ViroMed Co Ltd).

[00558] In various embodiments, the CAR specifically binds mesothelin. Exemplary anti-mesothelin CARs include, without limitation, CARMA-hMeso (Johns Hopkins University), CSG-MESO (Carsgen Therapeutics Ltd), iCasp9M28z (Memorial Sloan-Kettering Cancer Center), KD-021 (Nanjing Kaedi Biotech Inc), m-28z-T2 (Guangzhou Institutes of Biomedicine and Health), MesoCART (University of Pennsylvania), meso-CAR-T + PD-78 (MirImmune LLC), RB-M1 (Refuge Biotechnologies Inc), and TC-210 (TCR2 Therapeutics Inc).

[00559] In various embodiments, the CAR specifically binds MUC1. Exemplary anti-MUC1 CARs include, without limitation, anti-MUC1 CAR T-cell therapy + PD-1 knockout T cell therapy (esophageal cancer/NSCLC), Guangzhou Anjie Biomedical Technology/University of Technology Sydney (Guangzhou Anjie Biomedical Technology Co LTD), ICTCAR-043 (Innovative Cellular Therapeutics Co Ltd), ICTCAR-046 (Innovative Cellular Therapeutics Co Ltd), P-MUC1C-101 (Poseida Therapeutics Inc), and TAB-28z (OncoTab Inc). In various embodiments, the CAR specifically binds MUC16. Exemplary anti-MUC16 CARs include, without limitation, 4H1128Z-E27 (Eureka Therapeutics Inc) and JCAR-020 (Memorial Sloan-Kettering Cancer Center).

[00560] In various embodiments, the CAR specifically binds nfP2X7. Exemplary anti-nfP2X7 CARs include, without limitation, BIL-022c (Biosceptre International Ltd). In various embodiments, the CAR specifically binds PSCA. Exemplary anti-PSCA CARs include, without limitation, BPX-601 (Bellicum Pharmaceuticals Inc). In various embodiments, the CAR specifically binds PSMA. CIK-CAR.PSMA (Formula Pharmaceuticals Inc), and P-PSMA-101 (Poseida Therapeutics Inc). In various embodiments, the CAR specifically binds ROR1. Exemplary anti-ROR1 CARs include, without limitation, JCAR-024 (Fred Hutchinson Cancer Research Center). In various embodiments, the CAR specifically binds ROR2. Exemplary anti-ROR2 CARs include, without limitation, CCT-301-59 (F1 Oncology Inc). In various embodiments, the CAR specifically binds SLAMF7. Exemplary anti-SLAMF7 CARs include, without limitation, UCART-CS1 (Cellectis SA). In various embodiments, the CAR specifically

binds TRBC1. Exemplary anti-TRBC1 CARs include, without limitation, AUTO-4 (Autolus Therapeutics Limited). In various embodiments, the CAR specifically binds TRBC2. Exemplary anti-TRBC2 CARs include, without limitation, AUTO-5 (Autolus Therapeutics Limited). In various embodiments, the CAR specifically binds TSHR. Exemplary anti-TSHR CARs include, without limitation, ICTCAT-023 (Innovative Cellular Therapeutics Co Ltd). In various embodiments, the CAR specifically binds VEGFR-1. Exemplary anti-VEGFR-1 CARs include, without limitation, SKLB-083017 (Sichuan University).

[00561] In various embodiments, the CAR is AT-101 (AbClon Inc); AU-101, AU-105, and AU-180 (Aurora Biopharma Inc); CARMA-0508 (Carisma Therapeutics); CAR-T (Fate Therapeutics Inc); CAR-T (Cell Design Labs Inc); CM-CX1 (Celdara Medical LLC); CMD-502, CMD-503, and CMD-504 (Baylor College of Medicine); CSG-002 and CSG-005 (Carsgen Therapeutics Ltd); ET-1501, ET-1502 , and ET-1504 (Eureka Therapeutics Inc); FT-61314 (Fate Therapeutics Inc); GB-7001 (Shanghai GeneChem Co Ltd); IMA-201 (Immatics Biotechnologies GmbH); IMM-005 and IMM-039 (Immunome Inc); ImmuniCAR (TC BioPharm Ltd); NT-0004 and NT-0009 (BioNTech Cell and Gene Therapies GmbH), OGD-203 (OGD2 Pharma SAS), PMC-005B (PharmAbcine), and TI-7007 (Timmune Biotech Inc).

[00562] In some embodiments the chimeric antigen receptor comprises an amino acid sequence of an antibody. In some embodiments, the chimeric antigen receptor comprises the amino acid sequence of an antigen binding fragment of an antibody. The antibody (or fragment thereof) portion of the extracellular binding domain recognizes and binds to an epitope of an antigen. In some embodiments, the antibody fragment portion of a chimeric antigen receptor is a single chain variable fragment (scFv). An scFV comprises the light and variable fragments of a monoclonal antibody. In other embodiments, the antibody fragment portion of a chimeric antigen receptor is a multichain variable fragment, which can comprise more than one extracellular binding domains and therefore bind to more than one antigen simultaneously. In a multiple chain variable fragment embodiment, a hinge region may separate the different variable fragments, providing necessary spatial arrangement and flexibility.

[00563] In other embodiments, the antibody portion of a chimeric antigen receptor comprises at least one heavy chain and at least one light chain. In some embodiments, the antibody portion of a chimeric antigen receptor comprises two heavy chains, joined by disulfide bridges and two light chains, wherein the light chains are each joined to one of the heavy chains by disulfide

bridges. In some embodiments, the light chain comprises a constant region and a variable region. Complementarity determining regions residing in the variable region of an antibody are responsible for the antibody's affinity for a particular antigen. Thus, antibodies that recognize different antigens comprise different complementarity determining regions. Complementarity determining regions reside in the variable domains of the extracellular binding domain, and variable domains (i.e., the variable heavy and variable light) can be linked with a linker or, in some embodiments, with disulfide bridges.

[00564] In some embodiments, the antigen recognized and bound by the extracellular domain is a protein or peptide, a nucleic acid, a lipid, or a polysaccharide. Antigens can be heterologous, such as those expressed in a pathogenic bacteria or virus. Antigens can also be synthetic; for example, some individuals have extreme allergies to synthetic latex and exposure to this antigen can result in an extreme immune reaction. In some embodiments, the antigen is autologous, and is expressed on a diseased or otherwise altered cell. For example, in some embodiments, the antigen is expressed in a neoplastic cell. In some embodiments, the neoplastic cell is a solid tumor cell. In other embodiments, the neoplastic cell is a hematological cancer, such as a B cell cancer. In some embodiments, the B cell cancer is a lymphoma (e.g., Hodgkins or non-Hodgkins lymphoma) or a leukemia (e.g., B-cell acute lymphoblastic leukemia). Exemplary B-cell lymphomas include Diffuse large B-cell lymphoma (DLBCL), primary mediastinal B-cell lymphoma, follicular lymphoma, Chronic lymphocytic leukemia (CLL), small lymphocytic lymphoma (SLL), mantle cell lymphomas, Marginal zone lymphoma, Burkitt lymphoma, Burkitt-like lymphoma, Lymphoplasmacytic lymphoma (Waldenstrom macroglobulinemia), and hairy cell leukemia. In some embodiments, the B cell cancer is multiple myeloma.

[00565] Antibody-antigen interactions are noncovalent interactions resulting from hydrogen bonding, electrostatic or hydrophobic interactions, or from van der Waals forces. The affinity of extracellular binding domain of the chimeric antigen receptor for an antigen can be calculated with the following formula:

K_A = [Antibody-Antigen]/[Antibody][Antigen], wherein

[Ab] = molar concentration of unoccupied binding sites on the antibody;

[Ag] = molar concentration of unoccupied binding sites on the antigen; and

[Ab-Ag] = molar concentration of the antibody-antigen complex.

[00566] The antibody-antigen interaction can also be characterized based on the dissociation of the antigen from the antibody. The dissociation constant (K_D) is the ratio of the association rate to the dissociation rate and is inversely proportional to the affinity constant. Thus, $K_D = 1/ K_A$. Those skilled in the art will be familiar with these concepts and will know that traditional methods, such as ELISA assays, can be used to calculate these constants.

[00567] The transmembrane domain of the chimeric antigen receptors described herein spans the CAR-T cells lipid bilayer cellular membrane and separates the extracellular binding domain and the intracellular signaling domain. In some embodiments, this domain is derived from other receptors having a transmembrane domain, while in other embodiments, this domain is synthetic. In some embodiments, the transmembrane domain may be derived from a non-human transmembrane domain and, in some embodiments, humanized. By "humanized" is meant having the sequence of the nucleic acid encoding the transmembrane domain optimized such that it is more reliably or efficiently expressed in a human subject. In some embodiments, the transmembrane domain is derived from another transmembrane protein expressed in a human immune effector cell. Examples of such proteins include, but are not limited to, subunits of the T cell receptor (TCR) complex, PD1, or any of the Cluster of Differentiation proteins, or other proteins, that are expressed in the immune effector cell and that have a transmembrane domain. In some embodiments, the transmembrane domain will be synthetic, and such sequences will comprise many hydrophobic residues.

[00568] The chimeric antigen receptor is designed, in some embodiments, to comprise a spacer between the transmembrane domain and the extracellular domain, the intracellular domain, or both. Such spacers can be 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 amino acids in length. In some embodiments, the spacer can be 20, 30, 40, 50, 60, 70, 80, 90, or 100 amino acids in length. In still other embodiments the spacer can be between 100 and 500 amino acids in length. The spacer can be any polypeptide that links one domain to another and are used to position such linked domains to enhance or optimize chimeric antigen receptor function.

[00569] The intracellular signaling domain of the chimeric antigen receptor contemplated herein comprises a primary signaling domain. In some embodiments, the chimeric antigen receptor comprises the primary signaling domain and a secondary, or co-stimulatory, signaling domain. In some embodiments, the primary signaling domain comprises one or more

immunoreceptor tyrosine-based activation motifs, or ITAMs. In some embodiments, the primary signaling domain comprises more than one ITAM. ITAMs incorporated into the chimeric antigen receptor may be derived from ITAMs from other cellular receptors. In some embodiments, the primary signaling domain comprising an ITAM may be derived from subunits of the TCR complex, such as CD3γ, CD3ε, CD3ζ, or CD3δ (see FIG. 1A). In some embodiments, the primary signaling domain comprising an ITAM may be derived from FcRγ, FcRβ, CD5, CD22, CD79a, CD79b, or CD66d. The secondary signaling domain, in some embodiments, is derived from CD28. In other embodiments, the secondary signaling domain is derived from CD2, CD4, CDS, CD8α, CD83, CD134, CD137, ICOS, or CD154.

[00570] Provided herein are also nucleic acids that encode the chimeric antigen receptors described herein. In some embodiments, the nucleic acid is isolated or purified. Delivery of the nucleic acids *ex vivo* can be accomplished using methods known in the art. For example, immune cells obtained from a subject may be transformed with a nucleic acid vector encoding the chimeric antigen receptor. The vector may then be used to transform recipient immune cells so that these cells will then express the chimeric antigen receptor. Efficient means of transforming immune cells include transfection and transduction. Such methods are well known in the art. For example, applicable methods for delivery the nucleic acid molecule encoding the chimeric antigen receptor (and the nucleic acid(s) encoding the base editor) can be found in International Application No. PCT/US2009/040040 and US Patent Nos. 8,450,112; 9,132,153; and 9,669,058, each of which is incorporated herein in its entirety. Additionally, those methods and vectors described herein for delivering the nucleic acid encoding the base editor are applicable to delivering the nucleic acid encoding the chimeric antigen receptor.

[00571] Some aspects of the present invention provide for immune cells comprising a chimeric antigen and an altered endogenous gene that enhances immune cell function, resistance to immunosuppression or inhibition, or a combination thereof. In some embodiments, the altered endogenous gene may be created by base editing. In some embodiments, the base editing may reduce or attenuate the gene expression. In some embodiments, the base editing may reduce or attenuate the gene activation. In some embodiments, the base editing may reduce or attenuate the functionality of the gene product. In some other embodiments, the base editing may activate or enhance the gene expression. In some embodiments, the base editing may increase the functionality of the gene product. In some embodiments, the altered endogenous gene may be

modified or edited in an exon, an intron, an exon-intron injunction, or a regulatory element thereof. The modification may be edit to a single nucleobase in a gene or a regulatory element thereof. The modification may be in a exon, more than one exons, an intron, or more than one introns, or a combination thereof. The modification may be in an open reading frame of a gene. The modification may be in an untranslated region of the gene, for example, a 3'-UTR or a 5'-UTR. In some embodiments, the modification is in a regulatory element of an endogenous gene. In some embodiments, the modification is in a promoter, an enhancer, an operator, a silencer, an insulator, a terminator, a transcription initiation sequence, a translation initiation sequence (e.g. a Kozak sequence), or any combination thereof.

[00572] Allogeneic immune cells expressing an endogenous immune cell receptor as well as a chimeric antigen receptor may recognize and attack host cells, a circumstance termed graft versus host disease (GVHD). The alpha component of the immune cell receptor complex is encoded by the TRAC gene, and in some embodiments, this gene is edited such that the alpha subunit of the TCR complex is nonfunctional or absent. Because this subunit is necessary for endogenous immune cell signaling, editing this gene can reduce the risk of graft versus host disease caused by allogeneic immune cells.

[00573] Host immune cells can potentially recognize allogeneic CAR-T cells as non-self and elicit an immune response to remove the non-self cells. B2M is expressed in nearly all nucleated cells and is associated with MHC class I complex (FIG. 1B). Circulating host $CD8^+$ T cells can recognize this B2M protein as non-self and kill the allogeneic cells. To overcome this graft rejection, in some embodiments, the B2M gene is edited to either knockout or knockdown expression.

[00574] In some embodiments of the present invention, the PDCD1 gene is edited in the CAR-T cell to knockout or knockdown expression. The PDCD1 gene encodes the cell surface receptor PD-1, an immune system checkpoint expressed in immune cells, and it is involved in reducing autoimmunity by promoting apoptosis of antigen specific immune cells. By knocking out or knocking down expression of the PDCD1 gene, the modified CAR-T cells are less likely to apoptose, are more likely to proliferate, and can escape the programmed cell death immune checkpoint.

[00575] The CBLB gene encodes an E3 ubiquitin ligase that plays a significant role in inhibiting immune effector cell activation. Referring to FIG. 1C, the CBLB protein favors the

signaling pathway resulting in immune effector cell tolerance and actively inhibits signaling that leads to immune effector cell activation. Because immune effector cell activation is necessary for the CAR-T cells to proliferate *in vivo* post-transplant, in some embodiments of the present invention the CBLB is edited to knockout or knockdown expression.

[00576] In some embodiments, editing of genes to enhance the function of the immune cell or to reduce immunosuppression or inhibition can occur in the immune cell before the cell is transformed to express a chimeric antigen receptor. In other aspects, editing of genes to enhance the function of the immune cell or to reduce immunosuppression or inhibition can occur in a CAR-T cell, *i.e.*, after the immune cell has been transformed to express a chimeric antigen receptor.

[00577] In some embodiments, the immune cell may comprise a chimeric antigen receptor (CAR) and one or more edited genes, one or more regulatory elements thereof, or combinations thereof, wherein expression of the edited gene is either knocked out or knocked down. In some embodiments, the CAR-T cells have reduced immunogenicity as compared to a similar CAR-T cell but without further having the one or more edited genes as described herein. In some embodiments, the CAR-T cells have lower activation threshold as compared to a similar CAR-T but without further having the one or more edited genes as described herein. In some embodiments, the CAR-T cells have increased anti-neoplasia activity as compared to a similar CAR-T cell but without further having the one or more edited genes as described herein. The one or more genes may be edited by base editing. In some embodiments the one or more genes, or one or more regulatory elements thereof, or combinations thereof, may be selected from a group consisting of: c-abl oncogene 1 (Abl1); c-abl oncogene 2 (Abl2); a disintegrin and metalloprotease domain 8 (Adam8); a disintegrin and metalloprotease domain 17 (Adam 17); adenosine deaminase (Ada); adenosine kinase (Adk); adenosine A2a receptor (Adora2a); adenosine regulating molecule 1 (Adrm1); advanced glycosylation end product-specific receptor (Ager) allograft inflammatory factor 1 (Aif1); autoimmune regulator (Aire); ankyrin repeat and LEM domain (Ankle1); annecin A1 (Anxa1); adapter related protein complex 3 beta 1 sububit (Ap3b1); adapter related protein complex 3 delta 1 sububit (Ap3d1); amyloid beta (A4) precursor protein-binding family B member 1 interacting protein (Apbb1ip); WNT signaling pathway regulator (Apc); arginase liver (Arg 1); arginase type II (Arg 2); autophagy related 5 (Atg5); AtPase Cu++ transporting, alpha polypeptide (Atp7a); 5-azacytidine induced gene 2

(Azi2); beta 2 microglobulin (B2m); BL2-associated agonist of cell dealth (Bad); basic leucine zipper transcription factor, ATF-like (Batf); BCL2-associated X protein (Bax); B cell leukemia/lymphoma 2 (Bcl2); B cell leukemia/lymphoma 2 related protein A1d (Bcl2a1d); B cell leukemia/lymphoma 3 (Bcl3); B cell leukemia/lymphoma 6 (Bcl6); B cell leukemia/lymphoma 10 (Bcl10); B cell leukemia/lymphoma 11a (Bcl11a); B cell leukemia/lymphoma 11b (Bcl11b); Bloom syndrome, RecQ like helicase (Blm); Bmi1 polycomb ring finger oncogene (Bmi1); Bone morphogenic protein 4 (Bmp4); Braf transforming gene (Braf); B and T lymphocyte associated (Btla); butyrophilin, subfamily 2, member A1 (Btn2a1); butyrophilin, subfamily 2, member A2 (Btn2a2); butyrophilin-like 1 (Btnl1); butyrophilin-like 2 (Btnl2); butyrophilin-like 6 (Btnl6); calcium channel, voltage dependent, beta 4 subunit (Cacnb4); caspase recruitment domain family member 11 (Card11); capping protein regulator and myosin 1 linker 2 (Carmil2); Caspase 3 (Casp3); caveolin 1 (Cav1); core-binding factor beta (Cbfb); Casitas B-lineage lymphoma b (Cblb); coil-coil domain containing 88B (Ccdc88b); chemokine (C-C motif) ligand 2 (Ccl2); chemokine (C-C motif) ligand 5 (Ccl5); chemokine (C-C motif) ligand 19 (Ccl19); chemokine (C-C motif) ligand 20 (Ccl20); cyclin D3 (Ccnd3); chemokine (C-C motif) receptor 2 (Ccr2); chemokine (C-C motif) receptor 6 (Ccr6); chemokine (C-C motif) receptor 7 (Ccr7); chemokine (C-C motif) receptor 9 (Ccr9); CD1d1 antigen (Cd1d1); CD1d2 antigen (CD1d2); CD2 antigen (CD2); CD3 antigen, delta polypeptide (CD3d); CD3 antigen, epsilon polypeptide (CD3d); CD4 antigen (Cd4); CD5 antigen (Cd5); CD6 antigen (Cd6); CD8 antigen (Cd8); CD24a antigen (Cd24a); CD27 antigen (CD27); CD28 antigen (Cd28); CD40 ligand (Cd40lg); CD44 antigen (Cd44); CD46 antigen, complement regulatory protein (Cd46); CD47 antigen (Rh-related antigen, integrin-associated signal transducer) (Cd47); CD48 antigen (Cd48); CD59b antigen (Cd59b); CD74 antigen (Cd74); CD80 antigen (Cd80); CD81 antigen (Cd81); CD83 antigen (Cd83); CD86 antigen (Cd86); CD151 antigen (Cd151); CD160 antigen (Cd160); CD209e antigen (Cd209e); CD244 molecule A (Cd244a); CD274 antigen (Cd274); CD276 antigen (Cd276); CD300A molecule (Cd300a); cadherin-like 26(Cdh26); cyclin-dependent kinase (Cdk6); cyclin dependent kinase inhibitor 2A (Cdkn2a); carcinoembryonic antigen-related cell adhesion molecule (Ceacam1); CCAAT/enhancer binding protein (C/EBP), beta (Cebpb); cyclic GMP-AMP synthase (Cgas); chromodomain helicase DNA binding protein 7 (Chd7); cholinergic receptor, nicotinic, alpha polypeptide 7 (Chrna7); C-type lectin domain family 2, member i (Clec2i); C-type lectin domain

family 4, member a2 (Clec4a2); C-type lectin domain family 4, member d (Clec4d); C-type lectin domain family 4, member e (Clec4e); C-type lectin domain family 4, member f (Clec4f); C-type lectin domain family 4, member g (Clec4g); cleft lip and palate associated transmembrane protein 1 (Clptm1); coronin, actin binding protein 1A (Coro1a); cysteine-rich protein 3 (Crip3); c-src tyrosine kinase (Csk); cytotoxic T lymphocyte-associated protein 2 alpha (Ctla2a); cytotoxic T-lymphocyte-associated protein 4 (Ctla4); catenin (cadherin associated protein), beta 1 (Ctnnb1); cytidine 5'-triphosphate synthase (Ctps); coxsackie virus and adenovirus receptor (Cxadr); chemokine (C-X-C motif) ligand 12 (Cxcl12); chemokine (C-X-C motif) receptor (Cxcr4); CYLD lysine 63 deubiquitinase (Cyld); cytochrome P450, family 26, subfamily b, polypeptide (Cyp26b1); dolichyl-di-phosphooligosaccharide-protein glycotransferase (Ddost); deoxyhypusine synthase (Dhps); dicer 1, ribonuclease type III (Dicer1); discs large MAGUK scaffold protein 1 (Dlg1); discs large MAGUK scaffold protein 5 (Dlg5); delta like canonical Notch ligand 4 (Dll4); DnaJ heat shock protein family (Hsp40) member A3 (Dnaja3); dedicator of cytokinesis 2 (Dock2); dedicator of cytokinesis 8 (Dock8); dipeptidylpeptidase 4 (Dpp4); drosha, ribonuclease type III (Drosha); deltex 1, E3 ubiquitin ligase (Dtx1); dual specificity phosphatase 3 (Dusp3); dual specificity phosphatase 10 (Dusp10); dual specificity phosphatase 22 (Dusp22); double homeobox B-like 1 (Duxbl1); Epstein-Barr virus induced gene 3 (Ebi3); ephrin B1 (Efnb1); ephrin B2 (Efnb2); ephrin B3 (Efnb3); early growth response 1(Egr1); early growth response 3 (Egr3); eukaryotic translation initiation factor 2 alpha kinase 4 (Eif2ak4); E74-like factor 4 (Elf4); eomesodermin (Eomes); Eph receptor B4 (Ephb4); Eph receptor B6 (Ephb6); erythropoietin(Epo); erb-b2 receptor tyrosine kinase (Erbb2); coagulation factor II (thrombin) receptor-like 1 (F2rl1); Fas (TNFRSF6)-associated via death domain (Fadd); family with sequence similarity 49, member B (Fam49b); Fanconi anemia, complementation group A (Fanca); Fanconi anemia, complementation group D2 (Fancd2); Fas (TNF receptor superfamily member 6) (Fas); Fc receptor, IgE, high affinity I, gamma polypeptide (Fcer1g); fibrinogen-like protein 1 (Fgl1); fibrinogen-like protein 2 (Fgl2); FK506 binding protein 1a (Fkbp1a); FK506 binding protein 1b ((Fkbp1b); flotillin 2 (Flot2); FMS-like tyrosine kinase (Flt3); forkhead box J1 (Foxj1); forkhead box N1 (Foxn1); forkhead box P1 (Foxp1); forkhead box P3 (Foxp3); fucosyltransferase 7 (Fut7); Fyn proto-oncogene (Fyn); frizzled class receptor 5 (Fzd5); frizzled class receptor 7 (Fzd7); frizzled class receptor 8 (Fzd8); growth arrest and DNA-damage-inducible 45 gamma (Gadd45g); GATA binding protein 3

(GATA3); GTPase, IMAP family member 1 (Gimap1); gap junction protein, alpha 1 (Gja1); GLI-Kruppel family member GLI3 (Gli3); glycerol-3-phosphate acyltransferase, mitochondrial (Gpam); G protein-coupled receptor 18 (Gpr18); gelsolin (Gsn); histocompatibility 2, class II antigen A, alpha (H2-Aa); histocompatibility 2, class II antigen A, beta 1 (H2-Ab1); histocompatibility 2, class II, locus DMa (H2-DMa); histocompatibility 2, M region locus 3(H3-M3); histocompatibility 2, O region alpha locus (H2-Oa); histocompatibility 2, T region locus 23 (H2-T23); hepatitis A virus cellular receptor 2 (Havcr2); haematopoietic 1(hem1); hes family bHLH transcription factor 1 (Hes1); homeostatic iron regulator (Hfe); H2.0-like homeobox (Hlx); HCLS1 binding protein 3 (Hs1bp3); hematopoietic SH2 domain containing (Hsh2d); heat shock protein 90, alpha (cytosolic), class A member 1 (Hsp90aa1); heat shock protein 1 (chaperonin) (Hspd1); heat shock 105kDa/110kDa protein 1(Hsph1); intercellular adhesion molecule 1 (Icam1); inducible T cell co-stimulator (Icos); icos ligand (Icosl); indoleamine 2,3-dioxygenase 1 (Ido1); interferon alpha 1 (Ifna1); interferon alpha 2 (Ifna2); interferon alpha 4 (Ifna4); interferon alpha 5 (Ifna5); interferon alpha 6 (Ifna6); interferon alpha 7 (Ifna7); interferon alpha 9 (Ifna9); interferon alpha 11 (Ifna11); interferon alpha 12 (Ifna12); interferon alpha 13 (Ifna13); interferon alpha 14 (Ifna14); interferon alpha 15 (Ifna15); interferon alpha 16 (Ifna16); interferon alpha B (Ifnab); interferon (alpha and beta) receptor 1(Ifnar1); interferon beta 1 (Ifnb1); interferon gamma (Ifng); interferon kappa (Ifnk); interferon zeta (Ifnz); insulin-like growth factor 1 (Igf1); insulin-like growth factor 2 (Igf2); insulin-like growth factor binding protein 2 (Igfbp2); Indian hedgehog (Ihh); IKAROS family zinc finger 1 (Ikzf1); interleukin 1 beta (Il1b; interleukin 1 family, member 8 (Il1f8); interleukin 1 receptor-like 2 (Il1rl2); interleukin 2 (Il2); interleukin 2 receptor, alpha chain (Il2ra); interleukin 2 receptor, gamma chain (Il2rg); interleukin 4 (Il4); interleukin 4 receptor, alpha (Il4ra); interleukin 6 (Il6); interleukin 6 signal transducer (Il6st); interleukin 7 (Il7); interleukin 7 receptor (Il7r); interleukin 12a (Il12a); interleukin 12b (Il12b); interleukin 12 receptor, beta1 (Il12rb1); interleukin 15 (Il15); interleukin 18 (Il18); interleukin 18 receptor 1 (Il18r1); interleukin 20 receptor beta (Il20rb); interleukin 21 (Il21); interleukin 23, alpha subunit p19 (Il23a); interleukin 27 (Il27); insulin II (Ins2); interferon regulatory factor 1 (Irf1); interferon regulatory factor 4 (Irf4); itchy, E3 ubiquitin protein ligase (Itch); integrin, alpha D (Itgad); integrin alpha L (Itgal); integrin alpha M (Itgam); integrin alpha V (Itgav); integrin alpha X (Itgax); integrin beta 2 (Itgb2); IL2 inducible T cell kinase (Itk); inositol 1,4,5-trisphosphate 3-kinase B (Itpkb); jagged

2 (Jag2); Janus kinase 3 (Jak3); junction adhesion molecule like 9 (Jam9); jumonji domain containing 6 (Jmjd6); K(lysine) acetyltransferase 2A (Kat2a); KDEL (Lys-Asp-Glu-Leu) endoplasmic reticulum protein retention receptor 1 (Kdelr1); KIT proto-oncogene receptor tyrosine kinase (Kit); lymphocyte-activation gene 3 (Lag3); linker for activation of T cells (Lat); lymphocyte transmembrane adaptor 1 (Lax1); lymphocyte protein tyrosine kinase (Lck); lymphocyte cytosolic protein 1 (Lcp1); lymphoid enhancer binding factor 1 (Lef1); leptin (Lep); leptin receptor (Lepr); LFNG O-fucosylpeptide 3-beta-N-acetylglucosaminyltransferase (Lfng); lectin, galactose binding, soluble 1 (Lgals1); lectin, galactose binding, soluble 3 (Lgals3); lectin, galactose binding, soluble 8 (Lgals8); lectin, galactose binding, soluble 9 (Lgals9); ligase IV, DNA, ATP-dependent (Lig4); leukocyte immunoglobulin-like receptor, subfamily B, member 4A (Lilrb4a); limb region 1 like (Lmbrl); LIM domain only 1 (Lmo1); lysyl oxidase-like 3 (Loxl3); leucine rich repeat containing 32 (Lrrc32); lymphocyte antigen 9 (Ly9); MAD1 mitotic arrest deficient 1-like 1 (Mad1l1); v-maf musculoaponeurotic fibrosarcoma oncogene family, protein B (avian) (Mafb); MALT1 paracaspase (Malt1); mitogen-activated protein kinase 8 interacting protein 1 (Mapk8ip10); membrane associated ring-CH-type finger 7 (Marchf7); midkine (Mdk); methyltransferase like 3 (Mettl3); MHC I like leukocyte 2 (Mill2); myelin protein zero-like 2 (Mpzl2); moesin (Msn); mechanistic target of rapamycin kinase (Mtor); myeloblastosis oncogene (Myb); myosin, heavy polypeptide 9, non-muscle (Myh9); non-SMC condensin II complex, subunit H2 (Ncaph2); non-catalytic region of tyrosine kinase adaptor protein 1 (Nck1); non-catalytic region of tyrosine kinase adaptor protein 2 (Nck2); NCK associated protein 1 like (Nckap1l); nuclear receptor co-repressor 1 (Ncor1); nicastrin (Ncstn); Nedd4 family interacting protein 1 (Ndfip1); neural precursor cell expressed, developmentally down-regulated 4 (Nedd4); nuclear factor of activated T cells, cytoplasmic, calcineurin dependent (Nfatc3); nuclear factor of kappa light polypeptide gene enhancer in B cells inhibitor, delta (Nfkbid); non-homologous end joining factor 1 (Nhej1); NFKB activating protein (Nkap); NK2 homeobox 3 (Nkx2-3); NLR family, CARD domain containing 3 (Nlrc3); NLR family, pyrin domain containing 3 (Nlrp3); Notch-regulated ankyrin repeat protein (Nrarp); OTU domain containing 5 (Otud5); purinergic receptor P2X, ligand-gated ion channel, 7 (P2rx7); phosphoprotein associated with glycosphingolipid microdomains 1 (Pag1); POZ (BTB) and AT hook containing zinc finger 1 (Patz1); PRKC, apoptosis, WT1, regulator (Pawr); paired box 1 (Pax1); programmed cell death 1 ligand 2 (Pdcd1lg2); phosphodiesterase 5A, cGMP-specific

(Pde5a); pellino 1 (Peli1); phosphoinositide-3-kinase regulatory subunit (Pik3r6); phospholipase A2, group IIA (Pla2g2a); phospholipase A2, group IID (Pla2g2d); phospholipase A2, group IIE (Pla2g2e); phospholipase A2, group IIF (Pla2g2f); purine-nucleoside phosphorylase (Pnp); protein phosphatase 3, catalytic subunit, beta isoform (Ppp3cb); PR domain containing 1, with ZNF domain (Prdm1); peroxiredoxin 2 (Prdx2); protein kinase, cAMP dependent regulatory, type I, alpha (Prkar1a); protein kinase C, theta 2 (Prkcq); protein kinase C, zeta (Prkcz); protein kinase, DNA activated, catalytic polypeptide (Prkdc); prosaposin (Psap); presenilin 1 (Psen1); presenilin 2 (Psen2); prostaglandin E receptor 4 (subtype EP4) (Ptger4); protein tyrosine phosphatase, non-receptor type 2 (Ptpn2); protein tyrosine phosphatase, non-receptor type 6 (Ptpn6); protein tyrosine phosphatase, non-receptor type 22 (lymphoid) (Ptpn22); protein tyrosine phosphatase, receptor type, C (Ptprc); PYD and CARD domain containing 7 (Pycard); RAB27A, member RAS oncogene family (Rab27a); RAB29, member RAS oncogene family (Rab29); (Rac family small GTPase 2); recombination activating gene 1 (Rag1); recombination activating gene 2 (Rag2); RAS protein activator like 3 (Rasal3); RAS guanyl releasing protein 1 (Rasgrp1); RING CCCH (C3H) domains 1 (Rc3h1); ring finger and CCCH-type zinc finger domains 2 (Rc3h2); ras homolog family member A (Rhoa); ras homolog family member H (Rhoh); receptor (TNFRSF)-interacting serine-threonine kinase 2 (Ripk2); RHO family interacting cell polarization regulator 2 (Ripor2); RAR-related orphan receptor alpha (Rora); RAR-related orphan receptor gamma (Ror); ribosomal protein L22 (Rpl 22); ribosomal protein S6 (Rps6); radical S-adenosyl methionine domain containing 2 (Rsad2); runt related transcription factor 1 (Runx1); runt related transcription factor 2 (Runx2); runt related transcription factor 3 (Runx3); squamous cell carcinoma antigen recognized by T cells (Sart1); SAM and SH3 domain containing 3 (Sash3); special AT-rich sequence binding protein 1 (Satb1); syndecan 4 (Sdc4); selenoprotein K (Selenok); sema domain, immunoglobulin domain (Ig), transmembrane domain (TM) and short cytoplasmic domain, (semaphorin) 4A (Sema4a); surfactant associated protein D (Sftpd); SH3 domain containing ring finger 1 (Sh3rf1); src homology 2 domain-containing transforming protein B (Shb); sonic hedgehog (Shh); signal-regulatory protein alpha (Sirpa); Signal-regulatory protein beta 1A (Sirpb1a); Signal-regulatory protein beta 1B (Sirpb1b); Signal-regulatory protein beta 1C (Sirpb1c); suppression inducing transmembrane adaptor 1 (Sit1); Src-like-adaptor 2 (Sla2); SLAM family member 6 (Slamf6); solute carrier family 4 (anion exchanger), member 1; (Slc4a1); solute carrier family 11 (proton-

coupled divalent metal ion transporters), member 1 (Slc11a1); solute carrier family 46, member 2 (Slc46a2); schlafen 1; SMAD family member 3 (Smad3); SMAD family member 7 (Smad7); suppressor of cytokine signaling 1 (Socs1); suppressor of cytokine signaling 5 (Socs5); suppressor of cytokine signaling 6 (Socs6); SOS Ras/Rac guanine nucleotide exchange factor 1 (Sos1), SOS Ras/Rac guanine nucleotide exchange factor 2 (Sos2), SRY (sex determining region Y)-box 4 (Sox4); sialophorin (Spn); signal transducer and activator of transcription 3 (Stat3); signal transducer and activator of transcription 5A (Stat5A); signal transducer and activator of transcription 5B (Stat5B); serine/threonine kinase 11 (Stk11); syntaxin 11 (Stx11); spleen tyrosine kinase (Syk); T cell-interacting, activating receptor on myeloid cells 1 (Tarm1); T-box 21 (Tbx21); T cell, immune regulator 1, ATPase, H+ transporting, lysosomal V0 protein A3 (Tcirg1); transforming growth factor, beta 1 (Tgfb1); transforming growth factor, beta receptor II (Tgfbr2); thymocyte selection associated (Themis); thymus cell antigen 1, theta (Thy1); T cell immunoreceptor with Ig and ITIM domains (Tigit); transmembrane protein 98 (Tmem98); transmembrane 131 like (Tmem131l); tumor necrosis factor, alpha-induced protein 8-like 2 (Tnfa1p8l2); tumor necrosis factor receptor superfamily, member 4 (Tnfrsf4); tumor necrosis factor receptor superfamily, member 13c (Tnfrsf13c); tumor necrosis factor (ligand) superfamily, member 4 (Tnfsf4); tumor necrosis factor (ligand) superfamily, member 8 (Tnfsf8); tumor necrosis factor (ligand) superfamily, member 9 (Tnfsf9); tumor necrosis factor (ligand) superfamily, member 11 (Tnfsf11); tumor necrosis factor (ligand) superfamily, member 13b (Tnfsf13b); tumor necrosis factor (ligand) superfamily, member 14 (Tnfsf14); tumor necrosis factor (ligand) superfamily, member 18 (Tnfsf18); TNF receptor-associated factor 6 (Traf6); triggering receptor expressed on myeloid cells-like 2 (Trem12); T cell receptor alpha joining 18 (Traj18); three prime repair exonuclease 1 (Trex1); transformation related protein 53 (Trp53); TSC complex subunit 1 (Tsc1); twisted gastrulation BMP signaling modulator 1 (Twsg1); vascular cell adhesion molecule 1 (Vcam1); vanin 1 (Vnn1); V-set and immunoglobulin domain containing 4 (Vsig4); WD repeat and FYVE domain containing 4 (Wdfy4); wingless-type MMTV integration site family, member 1 (Wnt1); wingless-type MMTV integration site family, member 4 (Wnt4); WW domain containing E3 ubiquitin protein ligase 1 (Wwp1); chemokine (C motif) ligand 1 (Xcl1); zinc finger and BTB domain containing 1 (Zbtb1); zinc finger and BTB domain containing 7B (Zbtb7B); zinc finger CCCH type containing 8 (Zc3h8); zinc finger CCCH type containing 12A (Zc3h12a); zinc finger CCCH type containing 12D (Zc3h12d); zinc

finger E-box binding homeobox 1 (Zeb1); zinc finger protein 36, C3H type (Zfp36); zinc finger protein 36, C3H type-like 1 (Zfp36L1); zinc finger protein 36, C3H type-like 2 (Zfp36L2); and zinc finger protein 683 (Zfp683).

[00578] In some embodiments, an immune cell comprises a chimeric antigen receptor and one or more edited genes, a regulatory element thereof, or combinations thereof. An edited gene may be an immune response regulation gene, an immunogenic gene, a checkpoint inhibitor gene, a gene involved in immune responses, a cell surface marker, e.g. a T cell surface marker, or any combination thereof. In some embodiments, an immune cell comprises a chimeric antigen receptor and an edited gene that is associated with activated T cell proliferation, for example, Fyn, Itgad, Itgal, Itgam, Itgb2, Satb1, or, Ephb6, a regulatory elements thereof, or combinations thereof. In some embodiments, an immune cell comprises a chimeric antigen receptor and an edited gene that is associated with alpha-beta T cell activation, for example, Dock2, Rorc, Lef1, or TCF7, their regulatory elements thereof, or combinations thereof. In some embodiments, an immune cell comprises a chimeric antigen receptor and an edited gene that is associated with gamma-delta T cell activation, for example, Jag2, Sox13, Mill2, or Jaml, their regulatory elements thereof, or combinations thereof. In some embodiments, an immune cell comprises a chimeric antigen receptor and an edited gene that is associated with positive regulation of T cell proliferation, for example, Cd24a, Cd86, Epo, Fadd, Icosl, Igf1, Igf2, Igfbp2, Tnfsf4, Tnfsf9, Gpam, Il2, Il2ra, Il4, Stat5a, Stat5b, Gli3, Ihh, Itpkb, Nkap, Shh, Ada, Cd24a, Cd28, Ceacam1, Socs1, Cd83, Cd81, Cd74, Bad, Gata3, interleukin 2, interleukin 2 receptor alpha chain, interleukin 4, interleukin 7, interleukin 12a or FoxP3 or their regulatory elements thereof, or combinations thereof. In some embodiments, an immune cell comprises a chimeric antigen receptor and an edited gene that is negative regulation of T-helper cell proliferation or differentiation, for example, Xcl1, Jak3, Rc3h1, Rc3h2 , Tbx21, Zbtb7b, Tbx21, Zc3h12a, Smad3, Loxl3, Socs5, Zfp35, or Bcl6 or their regulatory elements thereof, or combinations thereof. In some embodiments, the edited gene may be a checkpoint inhibitor gene, for example, such as a PD1 gene, a PDL1 gene, or a member related to or regulating the pathway of their formation or activation.

[00579] In some embodiments, provided herein is an immune cell with an edited TRAC gene (wherein, the TRAC gene may comprise one, two, three, four, five, six, seven eight, nine, ten or more base edits), such that the immune cell does not express an endogenous functional T cell

receptor alpha chain. In some embodiments, the immune cell is a T cell expressing a chimeric antigen receptor (a CAR-T cell). In some embodiments, provided herein is a CAR-T cell with base edits in TRAC gene, such that the CAR-T cell have reduced or negligible or no expression of endogenous T cell receptor alpha protein.

[00580] In some embodiments, the immune cell comprises anedited TRAC gene, and additionally, at least one edited gene. The at least one edited gene may be selected from the list of genes mentioned in the preceding paragraphs. In one embodiment, the immune cell may comprise an edited TRAC gene, an edited PDCD1 gene, an edited CD52 gene, an edited CD7 gene, an edited B2M gene, an edited CD5 gene, an edited CBLB gene, or any combination thereof. In some embodiments, a single modification event (such as electroporation), may introduce one or more gene edits. In some embodiments at least four, five, six, seven, eight, nine, ten, eleven, twelve, thirteen, fourteen, fifteen, sixteen, seventeen, eighteen, nineteen, twenty or more edits may be introduced in one or more genes simultaneously.

[00581] In some embodiments, the immune cell comprises anedited TRAC gene, and anedited PDCD1, CD52, CD7, B2M, CD5, or CBLB gene, or a combination thereof. In some embodiments, the immune cell comprises one or more of edited genes, selected from TRAC, PDCD1, CD52, CD7, B2M, CD5, B2M, CD5, and CBLB gene.

[00582] In some embodiments, the immune cell may comprise an edited TRAC gene, an edited CD2 gene, an edited CD3 epsilon gene, an edited CD3 gamma gene, an edited CD3 delta gene, an edited CD5 gene, an edited CD7 gene, an edited CD30 gene, an edited CD33 gene, an edited B2M gene, an edited CD52 gene, an edited CD70 gene, an edited CBLB gene, an edited CIITA gene, or any combination thereof.

[00583] In some embodiments, provided herein is an immune cell with an edited TRBC1 or TRBC2 gene, such that the immune cell does not express an endogenous functional T cell receptor beta chain. In some embodiments, provided herein is a CAR-T cell with an edited TRBC1/TRBC2 gene, such that the CAR-T cell exhibits reduced or negligible expression or no expression of endogenous T cell receptor beta chain.

[00584] In some embodiments, the immune cell comprises an edited TRBC1/TRBC2 gene, and additionally, at least edited gene. The at least one edited gene may be selected from the list of genes mentioned in the preceding paragraphs. In some embodiments, the immune cell comprises an edited TRBC1/TRBC2 gene, and an edited PDCD1, CD52 or CD7 gene, or a combination

thereof. In some embodiments, the CAR-T cell comprises one or more of base edited genes, selected from TRBC1/TRBC2 gene, PDCD1, CD52, and CD7 genes. In some embodiments, eachedited gene may comprise a single base edit. In some embodiments, each edited gene may comprise multiple base edits at different regions of the gene.

[00585] In some embodiments, the immune cell comprises an edited TRBC1/TRBC2 genes, and anedited PDCD1, CD52, CD7, B2M, CD5, or CBLB gene, or a combination thereof. In some embodiments, the immune cell may be a CAR-T cell. In some embodiments, the CAR-T cell comprises one or moreedited gene, selected from TRBC1/TRBC2, PDCD1, CD52, CD7, B2M, CD5, B2M, CD5, and CBLB gene.

[00586] In some embodiments, the immune cell may comprise an edited TRBC1/TRBC2 gene, an edited CD2 gene, an edited CD3 epsilon gene, an edited CD3 gamma gene, an edited CD3 delta gene, an edited CD5 gene, an edited CD7 gene, an edited CD30 gene, an edited CD33 gene, an edited B2M gene, an edited CD52 gene, an edited CD70 gene, an edited CBLB gene, an edited CIITA gene, or any combination thereof.

[00587] In some embodiments, an immune cell comprises a chimeric antigen receptor and an edited TRAC, B2M, PDCD1, CBLB gene, or a combination thereof, wherein expression of the edited gene is either knocked out or knocked down. In some embodiments, an immune cell comprises a chimeric antigen receptor and an edited TRAC gene, wherein expression of the edited gene is knocked out or knocked down. In some embodiments, an immune cell comprises a chimeric antigen receptor and edited TRAC and B2M genes, wherein expression of the edited genes is either knocked out or knocked down. In some embodiments, an immune cell comprises a chimeric antigen receptor and edited TRAC and PDCD1 genes, wherein expression of the edited genes is either knocked out or knocked down. In some embodiments, an immune cell comprises a chimeric antigen receptor and edited TRAC and CBLB genes, wherein expression of the edited genes is either knocked out or knocked down. In some embodiments, an immune cell comprises a chimeric antigen receptor and edited TRAC, B2M, and PDCD1 genes, wherein expression of the edited genes is either knocked out or knocked down. In some embodiments, an immune cell comprises a chimeric antigen receptor and edited TRAC, B2M, and CBLB genes, wherein expression of the edited genes is either knocked out or knocked down. In some embodiments, an immune cell or immune effector cell comprises a chimeric antigen receptor and edited TRAC, PDCD1, and CBLB genes, wherein expression of the edited genes is either

knocked out or knocked down. In some embodiments, an immune cell comprises a chimeric antigen and edited TRAC, B2M, PDCD1, and CBLB genes, wherein expression of the edited genes is either knocked out or knocked down. In some embodiments, an immune cell comprises a chimeric antigen receptor and an edited B2M gene, wherein expression of the edited genes is either knocked out or knocked down. In some embodiments, an immune cell comprises a chimeric antigen receptor and edited B2M and PDCD1 genes, wherein expression of the edited genes is either knocked out or knocked down. In some embodiments, an immune cell comprises a chimeric antigen receptor and edited B2M and CBLB genes, wherein expression of the edited genes is either knocked out or knocked down. In some embodiments, an immune cell comprises a chimeric antigen receptor and edited B2M, PDCD1, and CBLB genes, wherein expression of the edited genes is either knocked out or knocked down. In some embodiments, an immune cell comprises a chimeric antigen receptor and an edited PDCD gene, wherein expression of the edited genes is either knocked out or knocked down. In some embodiments, an immune cell comprises a chimeric antigen receptor and edited PDCD1 and CBLB genes, wherein expression of the edited genes is either knocked out or knocked down. In some embodiments, an immune cell comprises a chimeric antigen receptor and an edited CBLB, expression of the edited gene is either knocked out or knocked down.

[00588] In some embodiments, an immune cell comprises a chimeric antigen receptor and an edited TRAC, an edited CD2 gene, an edited CD3 epsilon gene, an edited CD3 gamma gene, an edited CD3 delta gene, an edited CD5 gene, an edited CD7 gene, an edited CD30 gene, an edited CD33 gene, an edited B2M gene, an edited CD52 gene, an edited CD70 gene, an edited CBLB gene, an edited CIITA gene, or any combination thereof, wherein expression of the edited gene is either knocked out or knocked down.

[00589] In some embodiments, an immune cell comprises a chimeric antigen receptor and an edited TRBC1 or TRBC2 gene, an edited CD2 gene, an edited CD3 epsilon gene, an edited CD3 gamma gene, an edited CD3 delta gene, an edited CD5 gene, an edited CD7 gene, an edited CD30 gene, an edited CD33 gene, an edited B2M gene, an edited CD52 gene, an edited CD70 gene, an edited CBLB gene, an edited CIITA gene, or any combination thereof, wherein expression of the edited gene is either knocked out or knocked down.

[00590] In some embodiments, an immune cell, including but not limited to any immune cell comprising an edited gene selected from any of the aforementioned gene edits, can be edited to

generate mutations in other genes that enhance the CAR-T's function or reduce immunosuppression or inhibition of the cell. For example, in some embodiments, an immune cell comprises a chimeric antigen receptor and an edited TGFBR2, ZAP70, NFATc1, TET2 gene, or a combination thereof, wherein expression of the edited gene is either knocked out or knocked down. In some embodiments, an immune cell comprises a chimeric antigen receptor and an edited TGFBR2 gene, wherein expression of the edited gene is knocked out or knocked down. In some embodiments, an immune cell comprises a chimeric antigen receptor and edited TGFBR2 and ZAP70 genes, wherein expression of the edited genes is either knocked out or knocked down. In some embodiments, an immune cell comprises a chimeric antigen receptor and edited TGFBR2 and ZAP70 genes, wherein expression of the edited genes is either knocked out or knocked down. In some embodiments, an immune cell comprises a chimeric antigen receptor and edited TGFBR2 and NFATC1 genes, wherein expression of the edited genes is either knocked out or knocked down. In some embodiments, an immune cell comprises a chimeric antigen receptor and edited TGFBR2 and TET2 genes, wherein expression of the edited genes is either knocked out or knocked down. In some embodiments, an immune cell comprises a chimeric antigen receptor and edited TGFBR2, ZAP70, and NFATC1 genes, wherein expression of the edited genes is either knocked out or knocked down. In some embodiments, an immune cell comprises a chimeric antigen receptor and edited TGFBR2, ZAP70, and TET2 genes, wherein expression of the edited genes is either knocked out or knocked down. In some embodiments, an immune cell comprises a chimeric antigen receptor and edited TGFBR2, NFATC1, and TET2 genes, wherein expression of the edited genes is either knocked out or knocked down. In some embodiments, an immune cell comprises a chimeric antigen and edited TGFBR2, ZAP70, NFATC1, and TET2 genes, wherein expression of the edited genes is either knocked out or knocked down. In some embodiments, an immune cell comprises a chimeric antigen receptor and an edited ZAP70 gene, wherein expression of the edited genes is either knocked out or knocked down. In some embodiments, an immune cell comprises a chimeric antigen receptor and edited ZAP70 and NFATC1 genes, wherein expression of the edited genes is either knocked out or knocked down. In some embodiments, an immune cell comprises a chimeric antigen receptor and edited ZAP70 and TET2 genes, wherein expression of the edited genes is either knocked out or knocked down. In some embodiments, an immune cell comprises a chimeric antigen receptor and edited ZAP70, PDCD1, and TET2 genes, wherein expression of

the edited genes is either knocked out or knocked down. In some embodiments, an immune cell comprises a chimeric antigen receptor and an edited PCDC1 gene, wherein expression of the edited genes is either knocked out or knocked down. In some embodiments, an immune cell comprises a chimeric antigen receptor and edited PCDC1 and TET2 genes, wherein expression of the edited genes is either knocked out or knocked down. And in some embodiments, an immune cell comprises a chimeric antigen receptor and an edited TET2, expression of the edited gene is either knocked out or knocked down.

[00591] *Editing of Target Genes in Immune Cells*

[00592] In some embodiments, provided herein is an immune cell with at least one modification in an endogenous gene or regulatory elements thereof. In some embodiments, the immune cell may comprise at least one modification in each of at least two, at least three, four, five, six, seven, eight, nine, ten, eleven, twelve, thirteen, fourteen, fifteen, sixteen, seventeen, eighteen, nineteen, twenty or more endogenous genes or regulatory elements thereof. In some embodiments, the at least one modification is a single nucleobase modification. In some embodiments, the at least one modification is by base editing. The base editing may be positioned at any suitable position of the gene, or in a regulatory element of the gene. Thus, it may be appreciated that a single base editing at a start codon, for example, can completely abolish the expression of the gene. In some embodiments, the base editing may be performed at a site within an exon. In some embodiments, the base editing may be performed at a site on more than one exons. In some embodiments, the base editing may be performed at any exon of the multiple exons in a gene. In some embodiments, base editing may introduce a premature STOP codon into an exon, resulting in either lack of a translated product or in a truncated that may be misfolded and thereby eliminated by degradation, or may produce an unstable mRNA that is readily degraded. In some embodiments, the immune cell is a T cell. In some embodiments, the immune cell is a CAR-T cell.

[00593] In some embodiments, base editing may be performed, for example on exon 1, or exon 2, or exon 3 or exon 4 of human TRAC gene (UCSC genomic database ENSG00000277734.8). In some embodiments, base editing in human TRAC gene is performed at a site within exon 1. In some embodiments, base editing in human TRAC gene is performed at a site within exon 2. In some embodiments, base editing in human TRAC gene is performed at a site within exon 3. In some embodiments, base editing in human TRAC gene is performed at a site within exon 4.

In some embodiments one or more base editing actions can be performed on human TRAC gene, at exon 1, exon 2, exon 3, exon 4 or any combination thereof.

[00594] For example, base editing may be performed on exon 1, or exon 2, or exon 3 or exon 4, of human B2M gene (Chromosome 15, NC_000015.10, 44711492 – 44718877; exemplary mRNA sequence NM_004048). In some embodiments, base editing in human B2M gene is performed at a site within exon 1. In some embodiments, base editing in human B2M gene is performed at a site within exon 2. In some embodiments, base editing in human B2M gene is performed at a site within exon 3. In some embodiments, base editing in human B2M gene is performed at a site within exon 4. In some embodiments one or more base editing actions can be performed on human B2M gene, at exon 1, exon 2, exon 3, exon 4 or any combination thereof.

[00595] In some embodiments, base editing may be performed on an intron. For example, base editing may be performed on an intron. In some embodiments, the base editing may be performed at a site within an intron. In some embodiments, the base editing may be performed at a site on more than one introns. In some embodiments, the base editing may be performed at any exon of the multiple introns in a gene. In some embodiments, one or more base editing may be performed on an exon, an intron or any combination of exons and introns.

[00596] For example, base editing may be performed, for example on any one or more of the introns in human TRAC gene. In some embodiments, base editing in human TRAC gene is performed at a site within intron 1. In some embodiments, base editing in human TRAC gene is performed at a site within intron 2. In some embodiments, base editing in human TRAC gene is performed at a site within intron 3. In some embodiments one or more base editing actions can be performed on human TRAC gene, at exon 1, exon 2, exon 3, exon 4, intron 1, intron 2, intron 3, or any combination thereof. In some embodiments one or more base edits can be performed on the last noncoding exon of human TRAC gene.

[00597] In some embodiments, the modification or base edit may be within a promoter site. In some embodiments, the base edit may be introduced within an alternative promoter site. In some embodiments, the base edit may be in a 5' regulatory element, such as an enhancer. In some embodiment, base editing may be introduced to disrupt the binding site of a nucleic acid binding protein. Exemplary nucleic acid binding proteins may be a polymerase, nuclease, gyrase, topoisomerase, methylase or methyl transferase, transcription factors, enhancer, PABP, zinc finger proteins, among many others.

[00598] In some embodiments, base editing may generate a splice acceptor-splice donor (SA-SD) site. For example, targeted base editing generating a SA-SD, or at a SA-SD site can result in reduced expression of a gene. For example, exon 1 SD site of TRAC at C5 may be targeted for base editing (GT-AT); TRAC exon 3 SA disruption may be targeted (AG-AA); B2M exon 1 SD at C6 position may be disrupted by base editing (GT-AT); B2M exon 3 SA at C6 can be targeted (AG-AA).

[00599] In some embodiments, provided herein is an immune cell with at least one modification in one or more endogenous genes. In some embodiments, the immune cell may have at least one modification in one, two, three, four, five, six, seven, eight, nine, ten, eleven, twelve, thirteen, fourteen, fifteen, sixteen, seventeen, eighteen, nineteen, twenty or more endogenous genes. In some embodiments, the modification generates a premature stop codon in the endogenous genes. In some embodiments, the modification is a single base modification. In some embodiments, the modification is generated by base editing. The premature stop codon may be generated in an exon, an intron, or an untranslated region. In some embodiments, base editing may be used to introduce more than one STOP codon, in one or more alternative reading frames. For example, a premature STOP codon can be introduced at exon 3 C4 position of TRAC (CAA-TAA) by base editing.

[00600] In some embodiments, modification/base edits may be introduced at a 3'-UTR, for example, in a poly adenylation (poly-A) site. In some embodiments, base editing may be performed on a 5'-UTR region.

[00601] *Chimeric Antigen Receptor Insertion into Immune Cell Genes*

[00602] In some embodiments, a chimeric antigen receptor is inserted into the TRAC gene. This has advantages. First, because TRAC is highly expressed in immune cell, the chimeric antigen receptor will be similarly expressed when its construct is designed to insert the chimeric antigen receptor into the TRAC gene such that expression of the receptor is driven by the TRAC promoter. Second, inserting the chimeric antigen receptor into the TRAC gene will knockout TRAC expression. In some embodiments, the gene editing system described herein can be used to insert the chimeric antigen receptor into the TRAC locus. gRNAs specific for the TRAC locus can guide the gene editing system to the locus and initiate double-stranded DNA cleavage. In particular embodiments, the gRNA is used in conjunction with Cas12b. In various embodiments, the gene editing system is used in conjunction with a nucleic acid having a

sequence encoding a CAR receptor. Exemplary guide RNAs are provided in the following Table 1A.

[00603] Table 1A

gRNA sequence	PAM	napDNAbp	Gene Exon
GUUCUGUCUUUUGGUCAGGACAACCGUC UAGCUAUAAGUGCUGCAGGGUGUGAGAA ACUCCUAUUGCUGGACGAUGUCUCUUAC GAGGCAUUAGCACAGAGUCUCUCAGCUG GUACAC	ATTN	BhCas12b nuclease	TRAC KO gRNA 1 (Exon 1)
GUUCUGUCUUUUGGUCAGGACAACCGUC UAGCUAUAAGUGCUGCAGGGUGUGAGGA AACUCCUAUUGCUGGACGAUGUCUCUUA CGAGGCAUUAGCACACCGAUUUUGAUUC UCAAACA	ATTN	BhCas12b nuclease	TRAC KO gRNA 2 (Exon 1)
GUUCUGUCUUUUGGUCAGGACAACCGUC UAGCUAUAAGUGCUGCAGGGUGUGAGAA ACUCCUAUUGCUGGACGAUGUCUCUUAC GAGGCAUUAGCACUCAAACAAAUGUGCA CAAAG	ATTN	BhCas12b nuclease	TRAC KO gRNA 3 (Exon 1)
GUUCUGUCUUUUGGUCAGGACAACCGUC UAGCUAUAAGUGCUGCAGGGUGUGAGAA ACUCCUAUUGCUGGACGAUGUCUCUUAC GAGGCAUUAGCACUCAAACAAAUGUGUC ACAAAG	ATTN	BhCas12b nuclease	TRAC KO gRNA 4 (Exon 1)
GUUCUGUCUUUUGGUCAGGACAACCGUC UAGCUAUAAGUGCUGCAGGGUGUGAGAA ACUCCUAUUGCUGGACGAUGUCUCUUAC GAGGCAUUAGCACUUUGAGAAUCAAAAU CGGUA	ATTN	BhCas12b nuclease	TRAC KO gRNA 5 (Exon 1)
GUUCUGUCUUUUGGUCAGGACAACCGUC UAGCUAUAAGUGCUGCAGGGUGUGAGAA ACUCCUAUUGCUGGACGAUGUCUCUUAC GAGGCAUUAGCACUGAUGUGUAUAUCAC AGACAA	ATTN	BhCas12b nuclease	TRAC KO gRNA 6 (Exon 1)
GUUCUGUCUUUUGGUCAGGACAACCGUC	ATTN	BhCas12b	TRAC KO

UAGCUAUAAGUGCUGCAGGGUGUGAGAA ACUCCUAUUGCUGGACGAUGUCUCUUAC GAGGCAUUAGCAGUUGCUCCAGGCCACA GCAU		nuclease	gRNA 7 (Exon 1)
GUUCUGUCUUUUGGUCAGGACAACCGUC UAGCUAUAAGUGCUGCAGGGUGUGAGAA ACUCCUAUUGCUGGACGAUGUCUCUUAC GAGGCAUUAGCACUUCCAGAAGACACCU UCUUCC	ATTN	BhCas12b nuclease	TRAC KO gRNA 8 (Exon 1)
GUUCUGUCUUUUGGUCAGGACAACCGUC UAGCUAUAAGUGCUGCAGGGUGUGAGAA ACUCCUAUUGCUGGACGAUGUCUCUUAC GAGGCAUUAGCACCAGAAGACACCUUCU UCCCCA	ATTN	BhCas12b nuclease	TRAC KO gRNA 9 (Exon 1)
GUUCUGUCUUUUGGUCAGGACAACCGUC UAGCUAUAAGUGCUGCAGGGUGUGAGAG AAACUCCUAUUGCUGGACGAUGUCUCUU ACGAGGCAUUAGCACGGUUCCGAAUCCUC CUGA	ATTN	BhCas12b nuclease	TRAC KO gRNA 10 (Exon 3)
GUUCUGUCUUUUGGUCAGGACAACCGUC UAGCUAUAAGUGCUGCAGGGUGUGAGAA ACUCCUAUUGCUGGACGAUGUCUCUUAC GAGGCAUUAGCACGGAACCCAAUCACUG ACAGGU	ATTN	BhCas12b nuclease	TRAC KO gRNA 11 (Exon 3)

[00604] A DNA construct encoding the chimeric antigen receptor and nucleic acid containing extended stretches of TRAC DNA that flank the gRNA targeting sequences. Without being bound by theory, the construct binds to the complementary TRAC sequences, and the chimeric antigen receptor DNA, residing in proximity to the TRAC sequences on the construct is then inserted at the site of the lesion, effectively knocking out the TRAC gene and knocking in the chimeric antigen receptor nucleic acid. Table 1 provides guide RNAs for the TRAC gene that can guide the base editing machinery to the TRAC locus, which enables insertion of the chimeric antigen receptor nucleic acid. The first 11 gRNAS are for BhCas12b nuclease. The second set of 11 are for the BvCas12b nuclease. These are all for inserting the CAR at TRAC by creating a double stranded break, and not for base editing.

[00605] Table 1B: TRAC guide RNAs

Guide RNA Target	Guide RNA Spacer	Gene Exon
GTTCTGTCTTTTGGTCAGG ACAACCGTCTAGCTATAAG TGCTGCAGGGTGTGAGAA ACTCCTATTGCTGGACGAT GTCTCTTACGAGGCATTAG CACAGAGTCTCTCAGCTGG TACA	GUUCUGUCUUUUGGUCAG GACAACCGUCUAGCUAUA AGUGCUGCAGGGUGUGAG AAACUCCUAUUGCUGGAC GAUGUCUCUUACGAGGCA UUAGCACAGAGUCUCUCA GCUGGUACA	TRAC KO gRNA 1
GTTCTGTCTTTTGGTCAGG ACAACCGTCTAGCTATAAG TGCTGCAGGGTGTGAGAA ACTCCTATTGCTGGACGAT GTCTCTTACGAGGCATTAG CACACCGATTTTGATTCTC AAAC	GUUCUGUCUUUUGGUCAG GACAACCGUCUAGCUAUA AGUGCUGCAGGGUGUGAG AAACUCCUAUUGCUGGAC GAUGUCUCUUACGAGGCA UUAGCACACCGAUUUUGA UUCUCAAAC	TRAC KO gRNA 2
GTTCTGTCTTTTGGTCAGG ACAACCGTCTAGCTATAAG TGCTGCAGGGTGTGAGAA	GUUCUGUCUUUUGGUCAG GACAACCGUCUAGCUAUA AGUGCUGCAGGGUGUGAG	TRAC KO gRNA 3

Guide RNA Target	Guide RNA Spacer	Gene Exon
ACTCCTATTGCTGGACGAT GTCTCTTACGAGGCATTAG CACTGATTCTCAAACAAAT GTGT	AAACUCCUAUUGCUGGAC GAUGUCUCUUACGAGGCA UUAGCACUGAUUCUCAAA CAAAUGUGU	
GTTCTGTCTTTTGGTCAGG ACAACCGTCTAGCTATAAG TGCTGCAGGGTGTGAGAA ACTCCTATTGCTGGACGAT GTCTCTTACGAGGCATTAG CACTCAAACAAATGTGTCA CAAA	GUUCUGUCUUUUGGUCAG GACAACCGUCUAGCUAUA AGUGCUGCAGGGUGUGAG AAACUCCUAUUGCUGGAC GAUGUCUCUUACGAGGCA UUAGCACUCAAACAAAUG UGUCACAAA	TRAC KO gRNA 4
GTTCTGTCTTTTGGTCAGG ACAACCGTCTAGCTATAAG TGCTGCAGGGTGTGAGAA ACTCCTATTGCTGGACGAT GTCTCTTACGAGGCATTAG CACGTTTGAGAATCAAAAT CGGT	GUUCUGUCUUUUGGUCAG GACAACCGUCUAGCUAUA AGUGCUGCAGGGUGUGAG AAACUCCUAUUGCUGGAC GAUGUCUCUUACGAGGCA UUAGCACGUUUGAGAAUC AAAAUCGGU	TRAC KO gRNA 5
GTTCTGTCTTTTGGTCAGG ACAACCGTCTAGCTATAAG TGCTGCAGGGTGTGAGAA ACTCCTATTGCTGGACGAT GTCTCTTACGAGGCATTAG CACTGATGTGTATATCACA GACA	GUUCUGUCUUUUGGUCAG GACAACCGUCUAGCUAUA AGUGCUGCAGGGUGUGAG AAACUCCUAUUGCUGGAC GAUGUCUCUUACGAGGCA UUAGCACUGAUGUGUAUA UCACAGACA	TRAC KO gRNA 6
GTTCTGTCTTTTGGTCAGG ACAACCGTCTAGCTATAAG TGCTGCAGGGTGTGAGAA ACTCCTATTGCTGGACGAT GTCTCTTACGAGGCATTAG	GUUCUGUCUUUUGGUCAG GACAACCGUCUAGCUAUA AGUGCUGCAGGGUGUGAG AAACUCCUAUUGCUGGAC GAUGUCUCUUACGAGGCA	TRAC KO gRNA 7

Guide RNA Target	Guide RNA Spacer	Gene Exon
CACGTTGCTCCAGGCCACA GCAC	UUAGCACGUUGCUCCAGG CCACAGCAC	
GTTCTGTCTTTTGGTCAGG ACAACCGTCTAGCTATAAG TGCTGCAGGGTGTGAGAA ACTCCTATTGCTGGACGAT GTCTCTTACGAGGCATTAG CACTTCCAGAAGACACCTT CTTC	GUUCUGUCUUUUGGUCAG GACAACCGUCUAGCUAUA AGUGCUGCAGGGUGUGAG AAACUCCUAUUGCUGGAC GAUGUCUCUUACGAGGCA UUAGCACUUCCAGAAGAC ACCUUCUUC	TRAC KO gRNA 8
GTTCTGTCTTTTGGTCAGG ACAACCGTCTAGCTATAAG TGCTGCAGGGTGTGAGAA ACTCCTATTGCTGGACGAT GTCTCTTACGAGGCATTAG CACCAGAAGACACCTTCTT CCCC	GUUCUGUCUUUUGGUCAG GACAACCGUCUAGCUAUA AGUGCUGCAGGGUGUGAG AAACUCCUAUUGCUGGAC GAUGUCUCUUACGAGGCA UUAGCACCAGAAGACACC UUCUUCCCC	TRAC KO gRNA 9
GTTCTGTCTTTTGGTCAGG ACAACCGTCTAGCTATAAG TGCTGCAGGGTGTGAGAA ACTCCTATTGCTGGACGAT GTCTCTTACGAGGCATTAG CACGGTTCCGAATCCTCCT CCTG	GUUCUGUCUUUUGGUCAG GACAACCGUCUAGCUAUA AGUGCUGCAGGGUGUGAG AAACUCCUAUUGCUGGAC GAUGUCUCUUACGAGGCA UUAGCACGGUUCCGAAUC CUCCUCCUG	TRAC KO gRNA 10
GTTCTGTCTTTTGGTCAGG ACAACCGTCTAGCTATAAG TGCTGCAGGGTGTGAGAA ACTCCTATTGCTGGACGAT GTCTCTTACGAGGCATTAG CACGGAACCCAATCACTGA CAGG	GUUCUGUCUUUUGGUCAG GACAACCGUCUAGCUAUA AGUGCUGCAGGGUGUGAG AAACUCCUAUUGCUGGAC GAUGUCUCUUACGAGGCA UUAGCACGGAACCCAAUC ACUGACAGG	TRAC KO gRNA 11

Guide RNA Target	Guide RNA Spacer	Gene Exon
GACCTATAGGGTCAATGAA TCTGTGCGTGTGCCATAAG TAATTAAAAATTACCCACC ACAGGAGCACCTGAAAAC AGGTGCTTGGCACAGAGTC TCTCAGCTGGTACA	GACCUAUAGGGUCAAUGA AUCUGUGCGUGUGCCAUA AGUAAUUAAAAAUUACCC ACCACAGGAGCACCUGAA AACAGGUGCUUGGCACAG AGUCUCUCAGCUGGUACA	TRAC KO gRNA 1
GACCTATAGGGTCAATGAA TCTGTGCGTGTGCCATAAG TAATTAAAAATTACCCACC ACAGGAGCACCTGAAAAC AGGTGCTTGGCACACCGAT TTTGATTCTCAAAC	GACCUAUAGGGUCAAUGA AUCUGUGCGUGUGCCAUA AGUAAUUAAAAAUUACCC ACCACAGGAGCACCUGAA AACAGGUGCUUGGCACAC CGAUUUUGAUUCUCAAAC	TRAC KO gRNA 2
GACCTATAGGGTCAATGAA TCTGTGCGTGTGCCATAAG TAATTAAAAATTACCCACC ACAGGAGCACCTGAAAAC AGGTGCTTGGCACTGATTC TCAAACAAATGTGT	GACCUAUAGGGUCAAUGA AUCUGUGCGUGUGCCAUA AGUAAUUAAAAAUUACCC ACCACAGGAGCACCUGAA AACAGGUGCUUGGCACUG AUUCUCAAACAAAUGUGU	TRAC KO gRNA 3

[00606] Table 1B: Continued

Guide RNA Target	Guide RNA Spacer	Gene Exon
GACCTATAGGGTCAATGA ATCTGTGCGTGTGCCATA AGTAATTAAAAATTACCC ACCACAGGAGCACCTGAA AACAGGTGCTTGGCACTC AAACAAATGTGTCACAAA	GACCUAUAGGGUCAAUG AAUCUGUGCGUGUGCCA UAAGUAAUUAAAAAUUA CCCACCACAGGAGCACC UGAAAACAGGUGCUUGG CACUCAAACAAAUGUGU CACAAA	TRAC KO gRNA 4
GACCTATAGGGTCAATGA ATCTGTGCGTGTGCCATA AGTAATTAAAAATTACCC ACCACAGGAGCACCTGAA AACAGGTGCTTGGCACGT TTGAGAATCAAAATCGGT	GACCUAUAGGGUCAAUG AAUCUGUGCGUGUGCCA UAAGUAAUUAAAAAUUA CCCACCACAGGAGCACC UGAAAACAGGUGCUUGG CACGUUUGAGAAUCAAA AUCGGU	TRAC KO gRNA 5
GACCTATAGGGTCAATGA ATCTGTGCGTGTGCCATA AGTAATTAAAAATTACCC ACCACAGGAGCACCTGAA AACAGGTGCTTGGCACTG ATGTGTATATCACAGACA	GACCUAUAGGGUCAAUG AAUCUGUGCGUGUGCCA UAAGUAAUUAAAAAUUA CCCACCACAGGAGCACC UGAAAACAGGUGCUUGG CACUGAUGUGUAUAUCA CAGACA	TRAC KO gRNA 6
GACCTATAGGGTCAATGA ATCTGTGCGTGTGCCATA AGTAATTAAAAATTACCC ACCACAGGAGCACCTGAA AACAGGTGCTTGGCACGT TGCTCCAGGCCACAGCAC	GACCUAUAGGGUCAAUG AAUCUGUGCGUGUGCCA UAAGUAAUUAAAAAUUA CCCACCACAGGAGCACC UGAAAACAGGUGCUUGG CACGUUGCUCCAGGCCA CAGCAC	TRAC KO gRNA 7

Guide RNA Target	Guide RNA Spacer	Gene Exon
GACCTATAGGGTCAATGA ATCTGTGCGTGTGCCATA AGTAATTAAAAATTACCC ACCACAGGAGCACCTGAA AACAGGTGCTTGGCACTT CCAGAAGACACCTTCTTC	GACCUAUAGGGUCAAUG AAUCUGUGCGUGUGCCA UAAGUAAUUAAAAAUUA CCCACCACAGGAGCACC UGAAAACAGGUGCUUGG CACUUCCAGAAGACACC UUCUUC	TRAC KO gRNA 8
GACCTATAGGGTCAATGA ATCTGTGCGTGTGCCATA AGTAATTAAAAATTACCC ACCACAGGAGCACCTGAA AACAGGTGCTTGGCACCA GAAGACACCTTCTTCCCC	GACCUAUAGGGUCAAUG AAUCUGUGCGUGUGCCA UAAGUAAUUAAAAAUUA CCCACCACAGGAGCACC UGAAAACAGGUGCUUGG CACCAGAAGACACCUUC UUCCCC	TRAC KO gRNA 9
GACCTATAGGGTCAATGA ATCTGTGCGTGTGCCATA AGTAATTAAAAATTACCC ACCACAGGAGCACCTGAA AACAGGTGCTTGGCACGG TTCCGAATCCTCCTCCTG	GACCUAUAGGGUCAAUG AAUCUGUGCGUGUGCCA UAAGUAAUUAAAAAUUA CCCACCACAGGAGCACC UGAAAACAGGUGCUUGG CACGGUUCCGAAUCCUC CUCCUG	TRAC KO gRNA 10
GACCTATAGGGTCAATGA ATCTGTGCGTGTGCCATA AGTAATTAAAAATTACCC ACCACAGGAGCACCTGAA AACAGGTGCTTGGCACGG AACCCAATCACTGACAGG	GACCUAUAGGGUCAAUG AAUCUGUGCGUGUGCCA UAAGUAAUUAAAAAUUA CCCACCACAGGAGCACC UGAAAACAGGUGCUUGG CACGGAACCCAAUCACU GACAGG	TRAC KO gRNA 11

[00607] First 11 gRNAs are for BhCas12b nuclease. Second set of 11 gRNAs are for the BvCas12b nuclease. Scaffold sequence in bold, in first instance.

[00608] In some embodiments, a nucleic acid encoding a chimeric antigen receptor of the present invention can be targeted to the TRAC locus using the BE4 base editor. In some embodiments, the chimeric antigen receptor is targeted to the TRAC locus using a CRISPR/Cas9 base editing system.

[00609] To produce the gene edits described above, immune cells are collected from a subject and contacted with two or more guide RNAs and a nucleobase editor polypeptide comprising a nucleic acid programmable DNA binding protein (napDNAbp) and a cytidine deaminase or adenosine deaminase. In some embodiments, the collected immune cells are contacted with at least one nucleic acid, wherein the at least one nucleic acid encodes two or more guide RNAs and a nucleobase editor polypeptide comprising a nucleic acid programmable DNA binding protein (napDNAbp) and a cytidine deaminase. In some embodiments, the gRNA comprises nucleotide analogs. These nucleotide analogs can inhibit degradation of the gRNA from cellular processes. Table 2 provides target sequences to be used for gRNAs.

Table 2: Exemplary Target Sequences

Target protein	Target residue	gRNA target	gRNA Spacer	BE	Codon change	Residue function
NFATC1	R118	CTCGATGCGAGGACTCTCCA	CUCGAUGCGAGGACUCUCCA	BE	CGC>CAC	Calcineurin binding
	I119	TCTCGATGCGAGGACTCTCC	UCUCGAUGCGAGGACUCUCC	ABE	ATC>ACC	Calcineurin binding
	E120	CATCGAGATAACCTCGTGCT	CAUCGAGAUAACCUCGUGCU	ABE	GAG>GGG	Calcineurin binding
	S172	TGGCCGGGCTCAGGCACGAG	UGGCCGGGCUCAGGCACGAG	BE	AGC>AAC	PHOSPHORYLATION
	W396	GCCCACTGGTAGGGGTGCTG	GCCCACUGGUAGGGGUGCUG	ABE	TGG>CGG	Calcineurin binding
	R439	TGGGCTCGGTGGTGGGACTT	UGGGCUCGGUGGUGGGACUU	BE	CGA>CAA	DNA BINDING
	H441	CGAGCCCACTACGAGACGGA	CGAGCCCACUACGAGACGGA	ABE	CAC>CGC	DNA BINDING
	Y442	CTCGTAGTGGGCTCG	CUCGUAGUGGGCUC	AB	TAC>C	DNA BINDING

		GTGGT	GGUGGU	E	AC	
	K452	GCCGTGAAGGCGTCG GCCGG	GCCGUGAAGGCGUC GGCCGG	ABE	AAG>GGG	DNA BINDING
	R540	GTTTCTGAGTTTCAG GATTC	GUUUCUGAGUUUCA GGAUUC	BE	AGA>AAA	DNA BINDING
	R555	CATCGGGAGGAAGA ACACAC	CAUCGGGAGGAAGA ACACAC	ABE	AGG>GGG	DNA BINDING
	K556	GGAGGAAGAACACA CGGGTA	GGAGGAAGAACACA CGGGUA	ABE	AAG>GGG	DNA BINDING
	Q589	GAGCGCTGGGCTGCA TCAGA	GAGCGCUGGGCUGC AUCAGA	BE	CAG>CAT	DNA BINDING
	E114	TGATCTCGATCCGAG GGCTC	UGAUCUCGAUCCGA GGGCUC	BE	GAG>AAA	Calcineurin binding
	I115	ACGGAGTGATCTCGA TCCGA	ACGGAGUGAUCUCG AUCCGA	ABE	ATC>ACC	Calcineurin binding
	R253	GCGGAGGCATTCGTG CGCCG	GCGGAGGCAUUCGU GCGCCG	ABE	AGG>GGG	NLS
	S99	GCCGCGCTCAGAAAC TTCTG	GCCGCGCUCAGAAAC UUCUG	BE	AGC>AAC	PHOSPHORYLATION
	S107	GGGCCTCGGGCCTGA GCCCT	GGGCCUCGGGCCUG AGCCCU	BE	TCG>TTG	PHOSPHORYLATION
NFATC2	S148	CCTCGGGCTGGCGGC CACCC	CCUCGGGCUGGCGGC CACCC	BE	AGC>AAC	PHOSPHORYLATION
	S236	CCACTCGCCCGTGCC CCGTC	CCACUCGCCCGUGCC CCGUC	BE	TCG>TTG	PHOSPHORYLATION
	S255	GCATTCGTGCGCCGA GGCCT	GCAUUCGUGCGCCG AGGCCU	BE	TCG>TTG	PHOSPHORYLATION
	S268	GAGCCTCACCCCAGC GCTCC	GAGCCUCACCCCAGC GCUCC	BE	TCA>TTA	PHOSPHORYLATION
	S274	GAGGGGCTCCGGGA GCGCTG	GAGGGGCUCCGGGA GCGCUG	BE	AGC>AAC	PHOSPHORYLATION
	S326	AGGGCTGGTCTTCCA CATCT	AGGGCUGGUCUUCC ACAUCU	BE	AGC>AAC	PHOSPHORYLATION
NFATC	S213	GCGGGGAGCCCAGG	GCGGGGAGCCCAGG	AB	TCC>C	PHOSPHORYL

4		CCAAAG	CCAAAG	E	CC	ATION
AKT1	T305	GCCACCATGAAGACC TTTTG	GCCACCAUGAAGACC UUUUG	BE	ACC>A TT	PHOSPHORYL ATION
	T312	TTGCGGCACACCTGA GTACC	UUGCGGCACACCUG AGUACC	BE	ACA>A TA	PHOSPHORYL ATION
	S473	GTAGGAGAACTGGG GGAAGT	GUAGGAGAACUGGG GGAAGU	AB E	TCC>C CC	PHOSPHORYL ATION
	Y474	CTCCTACTCGGCCAG CGGCA	CUCCUACUCGGCCAG CGGCA	AB E	TAC>T GC	PHOSPHORYL ATION
AKT2	T309	GAAAACCTTCTGTGG GACCC	GAAAACCUUCUGUG GGACCC	BE	ACC>A TT	PHOSPHORYL ATION
	S474	AGTAGGAGAACTGG GGGAAG	AGUAGGAGAACUGG GGGAAG	AB E	TCC>C CC	PHOSPHORYL ATION
BLIMP 1	C608 (ZF2)	GTTGCAAGTCTGACA TTTGA	GUUGCAAGUCUGAC AUUUGA	AB E	TGC>C GC	DNA BINDING
	C608 (ZF2)	GTTGCAAGTCTGACA TTTGA	GUUGCAAGUCUGAC AUUUGA	BE	TGC>T AC	DNA BINDING
	H621 (ZF2)	GAAACACTACCTGGT ACACA	GAAACACUACCUGG UACACA	BE	CAC>T AT	DNA BINDING
	C636 (ZF3)	TGTGGCAGACCTACA GTGTA	UGUGGCAGACCUAC AGUGUA	BE	TGC>T AC	DNA BINDING
	C664 (ZF4)	GGGCACACCTTGCAT TGGTA	GGGCACACCUUGCA UUGGUA	AB E	TGC>C GC	DNA BINDING
	Splic e site 1	CTGCGCACCTGGCAT TCATG	CUGCGCACCUGGCAU UCAUG	BE		

GCN2 kinase (IDO pathway)	Exon 1 SD	CCTACCGGTCCGCAAGCGTC	CCUACCGGUCCGCAAGCGUC	BE		KNOCKOUT
	Exon 2 SD	ACTCACACATCTGGATAGGT	ACUCACACAUCUGGAUAGGU	BE		KNOCKOUT
	Exon 5 SD	GACTTACCTAGACCTTCCTG	GACUUACCUAGACCUUCCUG	BE		KNOCKOUT
CBL-B	C373	AATCTTACAGAGCTGAAAAG	AAUCUUACAGAGCUGAAAAG	BE	TGT>TAT	E3 UBIQUITIN LIGASE
	Y665.1	CATCATATTCTTCACTTCCA	CAUCAUAUUCUUCACUUCCA	ABE	TAT>CAC	
	Y665.2	AAGAATATGATGTTCCTCCC	AAGAAUAUGAUGUUCCUCCC	ABE	TAT>TGT	
	K907	CCCCTAAACCACGACCGCGC	CCCCUAAACCACGACCGCGC	ABE	AAA>GGG	
	R911	TCCTGCGCGGTCGTGGTTTA	UCCUGCGCGGUCGUGGUUUA	BE	CGC>CAC	
SHP1	Y377	CCCTACTCTGTGACCAACTG	CCCUACUCUGUGACCAACUG	ABE	TAC>TGC	
IRF4	R96	CGCAGGCGCGTCTTCCAGGT	CGCAGGCGCGUCUUCCAGGU	BE	CGC>CAC	DNA BINDING
	R98	GCACCGCAGGCGCGTCTTCC	GCACCGCAGGCGCGUCUUCC	BE	CGG>CAG	DNA BINDING
	K103	GAACAAGAGCAATGACTTTG	GAACAAGAGCAAUGACUUUG	ABE	AAG>GGG	DNA BINDING
PD1	Exon 1 STOP	CACCTACCTAAGAACCATCC	CACCUACCUAAGAACCAUCC	BE		KNOCKOUT
	Exon 2 STOP	GGGGTTCCAGGGCCTGTCTG	GGGGUUCCAGGGCCUGUCUG	BE		KNOCKOUT

TET2	H1386	GACTTGCACAACATGCAGAA	GACUUGCACAACAUGCAGAA	BE	CAC>TAC	DNA BINDING
	R1302	TTGCCAGAAGCAAGATCCCA	UUGCCAGAAGCAAGAUCCCA	ABE	AGA>GGG	DNA BINDING
	S1290	CCATGAACAACCAAAAGAGA	CCAUGAACAACCAAAAGAGA	ABE	TCA>CCA	DNA BINDING
SMARCA4	T353	TCACCCCCATCCAGAAGCCG	UCACCCCCAUCCAGAAGCCG	BE	ACC>ATT	PHOSPHORYLATION
	S610	ATCTGGCTGGTCTCGTCCAG	AUCUGGCUGGUCUCGUCCAG	BE	AGC>ATC	PHOSPHORYLATION
	S613	GATGAGCGACCTCCCGGTGA	GAUGAGCGACCUCCCGGUGA	ABE	AGC>GGC	PHOSPHORYLATION
	S695	AGACAGCGATGACGTCTCTG	AGACAGCGAUGACGUCUCUG	ABE	AGC>GGC	PHOSPHORYLATION
	S699	ACGTCTCTGAGGTGGACGCG	ACGUCUCUGAGGUGGACGCG	BE	TCT>TTT	PHOSPHORYLATION
	S1452	TTAGGGGAGAGTTTCTCGGC	UUAGGGGAGAGUUUCUCGGC	ABE	TCC>CCC	PHOSPHORYLATION
	S1575	GGAGAGTGAGGAGGAGGAAG	GGAGAGUGAGGAGGAGGAAG	ABE	AGT>GGT	PHOSPHORYLATION
	S1586	AAGGCTCCGAATCCGAATCT	AAGGCUCCGAAUCCGAAUCU	BE	TCC>TTT	PHOSPHORYLATION
	S1627	ATCGTCACTCACGACCGGCT	AUCGUCACUCACGACCGGCU	BE	AGT>AAT	PHOSPHORYLATION
	S1631	TGACAGTGAGGAGGAACAAG	UGACAGUGAGGAGGAACAAG	ABE	AGT>GGT	PHOSPHORYLATION
CDK4	P173	CACCCGTGGTTGTTACACTC	CACCCGUGGUUGUUACACUC	BE	CCC>CTT	
ZAP70	S144	CATCAGCCAGGCCCCGCAGG	CAUCAGCCAGGCCCCGCAGG	ABE	AGC>TGC	PHOSPHORYLATION
	Y292	GGTGTATCCATCTGAGTTGA	GGUGUAUCCAUCUGAGUUGA	ABE	TAC>CAC	PHOSPHORYLATION

	Y292	GGGTGTATCCATCTGAGTTG	GGGUGUAUCCAUCUGAGUUG	ABE	TAC>CAC	PHOSPHORYLATION
	R360	GCGCAAGAAGCAGATCGACG	GCGCAAGAAGCAGAUCGACG	BE	CGC>TGC	Hypermorphic activity
	Y598	TTACTACAGCCTGGCCAGCA	UUACUACAGCCUGGCCAGCA	ABE	TAC>TGC	PHOSPHORYLATION

[00610] The cytidine and adenosine deaminase nucleobase editors used in this invention can act on DNA, including single stranded DNA. Methods of using them to generate modifications in target nucleobase sequences in immune cells are presented.

[00611] In certain embodiments, the fusion proteins provided herein comprise one or more features that improve the base editing activity of the fusion proteins. For example, any of the fusion proteins provided herein may comprise a Cas9 domain that has reduced nuclease activity. In some embodiments, any of the fusion proteins provided herein may have a Cas9 domain that does not have nuclease activity (dCas9), or a Cas9 domain that cuts one strand of a duplexed DNA molecule, referred to as a Cas9 nickase (nCas9). Without wishing to be bound by any particular theory, the presence of the catalytic residue (*e.g.,* H840) maintains the activity of the Cas9 to cleave the non-edited (*e.g.,* non-methylated) strand opposite the targeted nucleobase. Mutation of the catalytic residue (e.g., D10 to A10) prevents cleavage of the edited strand containing the targeted A residue. Such Cas9 variants can generate a single-strand DNA break (nick) at a specific location based on the gRNA-defined target sequence, leading to repair of the non-edited strand, ultimately resulting in a nucleobase change on the non-edited strand.

Adenosine deaminases

[00612] In some embodiments, the fusion proteins of the invention comprise an adenosine deaminase domain. In some embodiments, the adenosine deaminases provided herein are capable of deaminating adenine. In some embodiments, the adenosine deaminases provided herein are capable of deaminating adenine in a deoxyadenosine residue of DNA. The adenosine deaminase may be derived from any suitable organism (e.g., *E. coli*). In some embodiments, the adenine deaminase is a naturally-occurring adenosine deaminase that includes one or more mutations corresponding to any of the mutations provided herein (e.g., mutations in ecTadA). One of skill in the art will be able to identify the corresponding residue in any homologous protein, e.g., by sequence alignment and determination of homologous residues. Accordingly, one of skill in the art would be able to generate mutations in any naturally-occurring adenosine deaminase (e.g., having homology to ecTadA) that corresponds to any of the mutations described herein, e.g., any of the mutations identified in ecTadA. In some embodiments, the adenosine deaminase is from a prokaryote. In some embodiments, the adenosine deaminase is from a bacterium. In some embodiments, the adenosine deaminase is from *Escherichia coli, Staphylococcus aureus, Salmonella typhi,*

Shewanella putrefaciens, Haemophilus influenzae, Caulobacter crescentus, or *Bacillus subtilis*. In some embodiments, the adenosine deaminase is from E. coli.

[00613] In one embodiment, a fusion protein of the invention comprises a wild-type TadA is linked to TadA7.10, which is linked to Cas9 nickase. In particular embodiments, the fusion proteins comprise a single TadA7.10 domain (e.g., provided as a monomer). In other embodiments, the ABE7.10 editor comprises TadA7.10 and TadA(wt), which are capable of forming heterodimers. The relevant sequences follow:

[00614] TadA (wt):

SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTAHAEIM
ALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGAAGSLMDVL
HHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTD

[00615] TadA7.10:

SEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHAEIMA
LRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLH
YPGMNHRVEITEGILADECAALLCYFFRMPRQVFNAQKKAQSSTD

[00616] In some embodiments, the adenosine deaminase comprises an amino acid sequence that is at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5% identical to any one of the amino acid sequences set forth in any of the adenosine deaminases provided herein. It should be appreciated that adenosine deaminases provided herein may include one or more mutations (e.g., any of the mutations provided herein). The disclosure provides any deaminase domains with a certain percent identify plus any of the mutations or combinations thereof described herein. In some embodiments, the adenosine deaminase comprises an amino acid sequence that has 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 21, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, or more mutations compared to a reference sequence, or any of the adenosine deaminases provided herein. In some embodiments, the adenosine deaminase comprises an amino acid sequence that has at least 5, at least 10, at least 15, at least 20, at least 25, at least 30, at least 35, at least 40, at least 45, at least 50, at least 60, at least 70, at least 80, at least 90, at least 100, at least 110, at least 120, at least 130, at

least 140, at least 150, at least 160, or at least 170 identical contiguous amino acid residues as compared to any one of the amino acid sequences known in the art or described herein.

[00617] In some embodiments, the adenosine deaminase comprises a D108X mutation in the TadA reference sequence, or a corresponding mutation in another adenosine deaminase, where X indicates any amino acid other than the corresponding amino acid in the wild-type adenosine deaminase. In some embodiments, the adenosine deaminase comprises a D108G, D108N, D108V, D108A, or D108Y mutation in TadA reference sequence, or a corresponding mutation in another adenosine deaminase. It should be appreciated, however, that additional deaminases may similarly be aligned to identify homologous amino acid residues that can be mutated as provided herein.

[00618] In some embodiments, the adenosine deaminase comprises an A106X mutation in TadA reference sequence, or a corresponding mutation in another adenosine deaminase, where X indicates any amino acid other than the corresponding amino acid in the wild-type adenosine deaminase. In some embodiments, the adenosine deaminase comprises an A106V mutation in TadA reference sequence, or a corresponding mutation in another adenosine deaminase.

[00619] In some embodiments, the adenosine deaminase comprises a E155X mutation in TadA reference sequence, or a corresponding mutation in another adenosine deaminase, where the presence of X indicates any amino acid other than the corresponding amino acid in the wild-type adenosine deaminase. In some embodiments, the adenosine deaminase comprises a E155D, E155G, or E155V mutation in TadA reference sequence, or a corresponding mutation in another adenosine deaminase.

[00620] In some embodiments, the adenosine deaminase comprises a D147X mutation in TadA reference sequence, or a corresponding mutation in another adenosine deaminase, where the presence of X indicates any amino acid other than the corresponding amino acid in the wild-type adenosine deaminase. In some embodiments, the adenosine deaminase comprises a D147Y, mutation in TadA reference sequence, or a corresponding mutation in another adenosine deaminase.

[00621] It should be appreciated that any of the mutations provided herein (e.g., based on the ecTadA amino acid sequence of TadA reference sequence) may be introduced into other adenosine deaminases, such as *S. aureus* TadA (saTadA), or other adenosine deaminases (e.g., bacterial adenosine deaminases). It would be apparent to the skilled artisan how to are homologous to the mutated residues in ecTadA. Thus, any of the mutations identified in ecTadA may be made in other adenosine deaminases that have homologous amino acid

residues. It should also be appreciated that any of the mutations provided herein may be made individually or in any combination in ecTadA or another adenosine deaminase. For example, an adenosine deaminase may contain a D108N, a A106V, a E155V, and/or a D147Y mutation in TadA reference sequence, or a corresponding mutation in another adenosine deaminase. In some embodiments, an adenosine deaminase comprises the following group of mutations (groups of mutations are separated by a ";") in TadA reference sequence, or corresponding mutations in another adenosine deaminase: D108N and A106V; D108N and E155V; D108N and D147Y; A106V and E155V; A106V and D147Y; E155V and D147Y; D108N, A106V, and E55V; D108N, A106V, and D147Y; D108N, E55V, and D147Y; A106V, E55V, and D 147Y; and D108N, A106V, E55V, and D147Y. It should be appreciated, however, that any combination of corresponding mutations provided herein may be made in an adenosine deaminase (e.g., ecTadA).

[00622] In some embodiments, the adenosine deaminase comprises one or more of a H8X, T17X, L18X, W23X, L34X, W45X, R51X, A56X, E59X, E85X, M94X, I95X, V102X, F104X, A106X, R107X, D108X, Kl lOX, M118X, N127X, A138X, F149X, M151X, R153X, Q154X, I156X, and/or K157X mutation in TadA reference sequence, or one or more corresponding mutations in another adenosine deaminase, where the presence of X indicates any amino acid other than the corresponding amino acid in the wild-type adenosine deaminase. In some embodiments, the adenosine deaminase comprises one or more of H8Y, T17S, L18E, W23L, L34S, W45L, R51H, A56E, or A56S, E59G, E85K, or E85G, M94L, 1951, V102A, F104L, A106V, R107C, or R107H, or R107P, D108G, or D108N, or D108V, or D108A, or D108Y, Kl 101, Ml 18K, N127S, A138V, F149Y, M151V, R153C, Q154L, I156D, and/or K157R mutation in TadA reference sequence, or one or more corresponding mutations in another adenosine deaminase.

[00623] In some embodiments, the adenosine deaminase comprises one or more of H8X, D108X, and/or N127X mutation in TadA reference sequence, or one or more corresponding mutations in another adenosine deaminase, where X indicates the presence of any amino acid. In some embodiments, the adenosine deaminase comprises one or more of a H8Y, D108N, and/or N127S mutation in TadA reference sequence, or one or more corresponding mutations in another adenosine deaminase.

[00624] In some embodiments, the adenosine deaminase comprises one or more of H8X, R26X, M61X, L68X, M70X, A106X, D108X, A109X, N127X, D147X, R152X, Q154X, E155X, K161X, Q163X, and/or T166X mutation in TadA reference sequence, or one or more corresponding mutations in another adenosine deaminase, where X indicates the presence of

any amino acid other than the corresponding amino acid in the wild-type adenosine deaminase. In some embodiments, the adenosine deaminase comprises one or more of H8Y, R26W, M61I, L68Q, M70V, A106T, D108N, A109T, N127S, D147Y, R152C, Q154H or Q154R, E155G or E155V or E155D, K161Q, Q163H, and/or T166P mutation in TadA reference sequence, or one or more corresponding mutations in another adenosine deaminase.

[00625] In some embodiments, the adenosine deaminase comprises one, two, three, four, five, or six mutations selected from the group consisting of H8X, D108X, N127X, D147X, R152X, and Q154X in TadA reference sequence, or a corresponding mutation or mutations in another adenosine deaminase, where X indicates the presence of any amino acid other than the corresponding amino acid in the wild-type adenosine deaminase. In some embodiments, the adenosine deaminase comprises one, two, three, four, five, six, seven, or eight mutations selected from the group consisting of H8X, M61X, M70X, D108X, N127X, Q154X, E155X, and Q163X in TadA reference sequence, or a corresponding mutation or mutations in another adenosine deaminase, where X indicates the presence of any amino acid other than the corresponding amino acid in the wild-type adenosine deaminase. In some embodiments, the adenosine deaminase comprises one, two, three, four, or five, mutations selected from the group consisting of H8X, D108X, N127X, E155X, and T166X in TadA reference sequence, or a corresponding mutation or mutations in another adenosine deaminase, where X indicates the presence of any amino acid other than the corresponding amino acid in the wild-type adenosine deaminase. In some embodiments, the adenosine deaminase comprises one, two, three, four, five, or six mutations selected from the group consisting of H8X, A106X, and D108X, or a corresponding mutation or mutations in another adenosine deaminase, where X indicates the presence of any amino acid other than the corresponding amino acid in the wild-type adenosine deaminase. In some embodiments, the adenosine deaminase comprises one, two, three, four, five, six, or seven mutations selected from the group consisting of H8X, R126X, L68X, D108X, N127X, D147X, and E155X in TadA reference sequence, or a corresponding mutation or mutations in another adenosine deaminase, where X indicates the presence of any amino acid other than the corresponding amino acid in the wild-type adenosine deaminase. In some embodiments, the adenosine deaminase comprises one, two, three, four, or five mutations selected from the group consisting of H8X, D108X, A109X, N127X, and E155X in TadA reference sequence, or a corresponding mutation or mutations in another adenosine deaminase, where X indicates the presence of any amino acid other than the corresponding amino acid in the wild-type adenosine deaminase.

[00626] In some embodiments, the adenosine deaminase comprises one, two, three, four, five, or six mutations selected from the group consisting of H8Y, D108N, N127S, D147Y, R152C, and Q154H in TadA reference sequence, or a corresponding mutation or mutations in another adenosine deaminase. In some embodiments, the adenosine deaminase comprises one, two, three, four, five, six, seven, or eight mutations selected from the group consisting of H8Y, M61I, M70V, D108N, N127S, Q154R, E155G, and Q163H in TadA reference sequence, or a corresponding mutation or mutations in another adenosine deaminase. In some embodiments, the adenosine deaminase comprises one, two, three, four, or five mutations selected from the group consisting of H8Y, D108N, N127S, E155V, and T166P in TadA reference sequence, or a corresponding mutation or mutations in another adenosine deaminase. In some embodiments, the adenosine deaminase comprises one, two, three, four, five, or six mutations selected from the group consisting of H8Y, A106T, D108N, N127S, E155D, and K161Q in TadA reference sequence, or a corresponding mutation or mutations in another adenosine deaminase. In some embodiments, the adenosine deaminase comprises one, two, three, four, five, six, or seven mutations selected from the group consisting of H8Y, R126W, L68Q, D108N, N127S, D147Y, and E155V in TadA reference sequence, or a corresponding mutation or mutations in another adenosine deaminase. In some embodiments, the adenosine deaminase comprises one, two, three, four, or five mutations selected from the group consisting of H8Y, D108N, A109T, N127S, and E155G in TadA reference sequence, or a corresponding mutation or mutations in another adenosine deaminase.

[00627] In some embodiments, the adenosine deaminase comprises one or more of the or one or more corresponding mutations in another adenosine deaminase. In some embodiments, the adenosine deaminase comprises a D108N, D108G, or D108V mutation in TadA reference sequence, or corresponding mutations in another adenosine deaminase. In some embodiments, the adenosine deaminase comprises a A106V and D108N mutation in TadA reference sequence, or corresponding mutations in another adenosine deaminase. In some embodiments, the adenosine deaminase comprises R107C and D108N mutations in TadA reference sequence, or corresponding mutations in another adenosine deaminase. In some embodiments, the adenosine deaminase comprises a H8Y, D108N, N127S, D147Y, and Q154H mutation in TadA reference sequence, or corresponding mutations in another adenosine deaminase. In some embodiments, the adenosine deaminase comprises a H8Y, R24W, D108N, N127S, D147Y, and E155V mutation in TadA reference sequence, or corresponding mutations in another adenosine deaminase. In some embodiments, the

adenosine deaminase comprises a D108N, D147Y, and E155V mutation in TadA reference sequence, or corresponding mutations in another adenosine deaminase. In some embodiments, the adenosine deaminase comprises a H8Y, D108N, and N127S mutation in TadA reference sequence, or corresponding mutations in another adenosine deaminase. In some embodiments, the adenosine deaminase comprises a A106V, D108N, D147Y, and E155V mutation in TadA reference sequence, or corresponding mutations in another adenosine deaminase.

[00628] In some embodiments, the adenosine deaminase comprises one or more of S2X, H8X, I49X, L84X, H123X, N127X, I156X, and/or K160X mutation in TadA reference sequence, or one or more corresponding mutations in another adenosine deaminase, where the presence of X indicates any amino acid other than the corresponding amino acid in the wild-type adenosine deaminase. In some embodiments, the adenosine deaminase comprises one or more of S2A, H8Y, I49F, L84F, H123Y, N127S, I156F, and/or K160S mutation in TadA reference sequence, or one or more corresponding mutations in another adenosine deaminase.

[00629] In some embodiments, the adenosine deaminase comprises an L84X mutation adenosine deaminase, where X indicates any amino acid other than the corresponding amino acid in the wild-type adenosine deaminase. In some embodiments, the adenosine deaminase comprises an L84F mutation in TadA reference sequence, or a corresponding mutation in another adenosine deaminase.

[00630] In some embodiments, the adenosine deaminase comprises an H123X mutation in TadA reference sequence, or a corresponding mutation in another adenosine deaminase, where X indicates any amino acid other than the corresponding amino acid in the wild-type adenosine deaminase. In some embodiments, the adenosine deaminase comprises an H123Y mutation in TadA reference sequence, or a corresponding mutation in another adenosine deaminase.

[00631] In some embodiments, the adenosine deaminase comprises an I157X mutation in TadA reference sequence, or a corresponding mutation in another adenosine deaminase, where X indicates any amino acid other than the corresponding amino acid in the wild-type adenosine deaminase. In some embodiments, the adenosine deaminase comprises an I157F mutation in TadA reference sequence, or a corresponding mutation in another adenosine deaminase.

[00632] In some embodiments, the adenosine deaminase comprises one, two, three, four, five, six, or seven mutations selected from the group consisting of L84X, A106X, D108X,

H123X, D147X, E155X, and I156X in TadA reference sequence, or a corresponding mutation or mutations in another adenosine deaminase, where X indicates the presence of any amino acid other than the corresponding amino acid in the wild-type adenosine deaminase. In some embodiments, the adenosine deaminase comprises one, two, three, four, five, or six mutations selected from the group consisting of S2X, I49X, A106X, D108X, D147X, and E155X in TadA reference sequence, or a corresponding mutation or mutations in another adenosine deaminase, where X indicates the presence of any amino acid other than the corresponding amino acid in the wild-type adenosine deaminase. In some embodiments, the adenosine deaminase comprises one, two, three, four, or five mutations selected from the group consisting of H8X, A106X, D108X, N127X, and K160X in TadA reference sequence, or a corresponding mutation or mutations in another adenosine deaminase, where X indicates the presence of any amino acid other than the corresponding amino acid in the wild-type adenosine deaminase.

[00633] In some embodiments, the adenosine deaminase comprises one, two, three, four, five, six, or seven mutations selected from the group consisting of L84F, A106V, D108N, H123Y, D147Y, E155V, and I156F in TadA reference sequence, or a corresponding mutation or mutations in another adenosine deaminase. In some embodiments, the adenosine deaminase comprises one, two, three, four, five, or six mutations selected from the group consisting of S2A, I49F, A106V, D108N, D147Y, and E155V in TadA reference sequence.

[00634] In some embodiments, the adenosine deaminase comprises one, two, three, four, or five mutations selected from the group consisting of H8Y, A106T, D108N, N127S, and K160S in TadA reference sequence, or a corresponding mutation or mutations in another adenosine deaminase.

[00635] In some embodiments, the adenosine deaminase comprises one or more of a E25X, R26X, R107X, A142X, and/or A143X mutation in TadA reference sequence, or one or more corresponding mutations in another adenosine deaminase, where the presence of X indicates any amino acid other than the corresponding amino acid in the wild-type adenosine deaminase. In some embodiments, the adenosine deaminase comprises one or more of E25M, E25D, E25A, E25R, E25V, E25S, E25Y, R26G, R26N, R26Q, R26C, R26L, R26K, R107P, R07K, R107A, R107N, R107W, R107H, R107S, A142N, A142D, A142G, A143D, A143G, A143E, A143L, A143W, A143M, A143S, A143Q, and/or A143R mutation in TadA reference sequence, or one or more corresponding mutations in another adenosine deaminase. In some embodiments, the adenosine deaminase comprises one or more of the mutations

described herein corresponding to TadA reference sequence, or one or more corresponding mutations in another adenosine deaminase.

[00636] In some embodiments, the adenosine deaminase comprises an E25X mutation in TadA reference sequence, or a corresponding mutation in another adenosine deaminase, where X indicates any amino acid other than the corresponding amino acid in the wild-type adenosine deaminase. In some embodiments, the adenosine deaminase comprises an E25M, E25D, E25A, E25R, E25V, E25S, or E25Y mutation in TadA reference sequence, or a corresponding mutation in another adenosine deaminase.

[00637] In some embodiments, the adenosine deaminase comprises an R26X mutation in TadA reference sequence, or a corresponding mutation in another adenosine deaminase, where X indicates any amino acid other than the corresponding amino acid in the wild-type adenosine deaminase. In some embodiments, the adenosine deaminase comprises R26G, R26N, R26Q, R26C, R26L, or R26K mutation in TadA reference sequence, or a corresponding mutation in another adenosine deaminase.

[00638] In some embodiments, the adenosine deaminase comprises an R107X mutation in TadA reference sequence, or a corresponding mutation in another adenosine deaminase, where X indicates any amino acid other than the corresponding amino acid in the wild-type adenosine deaminase. In some embodiments, the adenosine deaminase comprises an R107P, R07K, R107A, R107N, R107W, R107H, or R107S mutation in TadA reference sequence, or a corresponding mutation in another adenosine deaminase.

[00639] In some embodiments, the adenosine deaminase comprises an A142X mutation in TadA reference sequence, or a corresponding mutation in another adenosine deaminase, where X indicates any amino acid other than the corresponding amino acid in the wild-type adenosine deaminase. In some embodiments, the adenosine deaminase comprises an A142N, A142D, A142G, mutation in TadA reference sequence, or a corresponding mutation in another adenosine deaminase.

[00640] In some embodiments, the adenosine deaminase comprises an A143X mutation in TadA reference sequence, or a corresponding mutation in another adenosine deaminase, where X indicates any amino acid other than the corresponding amino acid in the wild-type adenosine deaminase. In some embodiments, the adenosine deaminase comprises an A143D, A143G, A143E, A143L, A143W, A143M, A143S, A143Q, and/or A143R mutation in TadA reference sequence, or a corresponding mutation in another adenosine deaminase.

[00641] In some embodiments, the adenosine deaminase comprises one or more of a H36X, N37X, P48X, I49X, R51X, M70X, N72X, D77X, E134X, S146X, Q154X, K157X, and/or

K161X mutation in TADA reference sequence, or one or more corresponding mutations in another adenosine deaminase, where the presence of X indicates any amino acid other than the corresponding amino acid in the wild-type adenosine deaminase. In some embodiments, the adenosine deaminase comprises one or more of H36L, N37T, N37S, P48T, P48L, I49V, R51H, R51L, M70L, N72S, D77G, E134G, S146R, S146C, Q154H, K157N, and/or K161T mutation in TadA reference sequence, or one or more corresponding mutations in another adenosine deaminase.

[00642] In some embodiments, the adenosine deaminase comprises an H36X mutation in TadA reference sequence, or a corresponding mutation in another adenosine deaminase, where X indicates any amino acid other than the corresponding amino acid in the wild-type adenosine deaminase. In some embodiments, the adenosine deaminase comprises an H36L mutation in TadA reference sequence, or a corresponding mutation in another adenosine deaminase.

[00643] In some embodiments, the adenosine deaminase comprises an N37X mutation in TadA reference sequence, or a corresponding mutation in another adenosine deaminase, where X indicates any amino acid other than the corresponding amino acid in the wild-type adenosine deaminase. In some embodiments, the adenosine deaminase comprises an N37T or N37S mutation in TadA reference sequence, or a corresponding mutation in another adenosine deaminase.

[00644] In some embodiments, the adenosine deaminase comprises an P48X mutation in TadA reference sequence, or a corresponding mutation in another adenosine deaminase, where X indicates any amino acid other than the corresponding amino acid in the wild-type adenosine deaminase. In some embodiments, the adenosine deaminase comprises an P48T or P48L mutation in TadA reference sequence, or a corresponding mutation in another adenosine deaminase.

[00645] In some embodiments, the adenosine deaminase comprises an R51X mutation in TadA reference sequence, or a corresponding mutation in another adenosine deaminase, where X indicates any amino acid other than the corresponding amino acid in the wild-type adenosine deaminase. In some embodiments, the adenosine deaminase comprises an R51H or R51L mutation in TadA reference sequence, or a corresponding mutation in another adenosine deaminase.

[00646] In some embodiments, the adenosine deaminase comprises an S146X mutation in TadA reference sequence, or a corresponding mutation in another adenosine deaminase, where X indicates any amino acid other than the corresponding amino acid in the wild-type

adenosine deaminase. In some embodiments, the adenosine deaminase comprises an S146R or S146C mutation in TadA reference sequence, or a corresponding mutation in another adenosine deaminase.

[00647] In some embodiments, the adenosine deaminase comprises an K157X mutation in TadA reference sequence, or a corresponding mutation in another adenosine deaminase, where X indicates any amino acid other than the corresponding amino acid in the wild-type adenosine deaminase. In some embodiments, the adenosine deaminase comprises a K157N mutation in TadA reference sequence, or a corresponding mutation in another adenosine deaminase.

[00648] In some embodiments, the adenosine deaminase comprises an P48X mutation in TadA reference sequence, or a corresponding mutation in another adenosine deaminase, where X indicates any amino acid other than the corresponding amino acid in the wild-type adenosine deaminase. In some embodiments, the adenosine deaminase comprises a P48S, P48T, or P48A mutation in TadA reference sequence, or a corresponding mutation in another adenosine deaminase.

[00649] In some embodiments, the adenosine deaminase comprises an A142X mutation in TadA reference sequence, or a corresponding mutation in another adenosine deaminase, where X indicates any amino acid other than the corresponding amino acid in the wild-type adenosine deaminase. In some embodiments, the adenosine deaminase comprises a A142N mutation in TadA reference sequence, or a corresponding mutation in another adenosine deaminase.

[00650] In some embodiments, the adenosine deaminase comprises an W23X mutation in TadA reference sequence, or a corresponding mutation in another adenosine deaminase, where X indicates any amino acid other than the corresponding amino acid in the wild-type adenosine deaminase. In some embodiments, the adenosine deaminase comprises a W23R or W23L mutation in TadA reference sequence, or a corresponding mutation in another adenosine deaminase.

[00651] In some embodiments, the adenosine deaminase comprises an R152X mutation in TadA reference sequence, or a corresponding mutation in another adenosine deaminase, where X indicates any amino acid other than the corresponding amino acid in the wild-type adenosine deaminase. In some embodiments, the adenosine deaminase comprises a R152P or R52H mutation in TadA reference sequence, or a corresponding mutation in another adenosine deaminase.

[00652] In one embodiment, the adenosine deaminase may comprise the mutations H36L, R51L, L84F, A106V, D108N, H123Y, S146C, D147Y, E155V, I156F, and K157N. In some embodiments, the adenosine deaminase comprises the following combination of mutations relative to TadA reference sequence, where each mutation of a combination is separated by a "_" and each combination of mutations is between parentheses: (A106V_D108N), (R107C_D108N), (H8Y_D108N_S127S_D147Y_Q154H), (H8Y_R24W_D108N_N127S_D147Y_E155V), (D108N_D147Y_E155V), (H8Y_D108N_S127S), (H8Y_D108N_N127S_D147Y_Q154H), (A106V_D108N_D147Y_E155V) (D108Q_D147Y_E155V) (D108M_D147Y_E155V), (D108L_D147Y_E155V), (D108K_D147Y_E155V), (D108I_D147Y_E155V), (D108F_D147Y_E155V), (A106V_D108N_D147Y), (A106V_D108M_D147Y_E155V), (E59A_A106V_D108N_D147Y_E155V), (E59A cat dead_A106V_D108N_D147Y_E155V), (L84F_A106V_D108N_H123Y_D147Y_E155V_I156Y), (L84F_A106V_D108N_H123Y_D147Y_E155V_I156F), (D103A_D014N), (G22P_D103A_D104N), (G22P_D103A_D104N_S138A) , (D103A_D104N_S138A), (R26G_L84F_A106V_R107H_D108N_H123Y_A142N_A143D_D147Y_E155V_I156F), (E25G_R26G_L84F_A106V_R107H_D108N_H123Y_A142N_A143D_D147Y_E155V_I156F),(E25D_R26G_L84F_A106V_R107K_D108N_H123Y_A142N_A143G_D147Y_E155V_I15 6F), (R26Q_L84F_A106V_D108N_H123Y_A142N_D147Y_E155V_I156F), (E25M_R26G_L84F_A106V_R107P_D108N_H123Y_A142N_A143D_D147Y_E155V_I156F), (R26C_L84F_A106V_R107H_D108N_H123Y_A142N_D147Y_E155V_I156F), (L84F_A106V_D108N_H123Y_A142N_A143L_D147Y_E155V_I156F), (R26G_L84F_A106V_D108N_H123Y_A142N_D147Y_E155V_I156F), (E25A_R26G_L84F_A106V_R107N_D108N_H123Y_A142N_A143E_D147Y_E155V_I15 6F), (R26G_L84F_A106V_R107H_D108N_H123Y_A142N_A143D_D147Y_E155V_I156F), (A106V_D108N_A142N_D147Y_E155V), (R26G_A106V_D108N_A142N_D147Y_E155V), (E25D_R26G_A106V_R107K_D108N_A142N_A143G_D147Y_E155V), (R26G_A106V_D108N_R107H_A142N_A143D_D147Y_E155V), (E25D_R26G_A106V_D108N_A142N_D147Y_E155V), (A106V_R107K_D108N_A142N_D147Y_E155V), (A106V_D108N_A142N_A143G_D147Y_E155V), (A106V_D108N_A142N_A143L_D147Y_E155V),

(H36L_R51L_L84F_A106V_D108N_H123Y_S 146C_D147Y_E155V_I156F _K157N),
(N37T_P48T_M70L_L84F_A106V_D108N_H123Y_D147Y_I49V_E155V_I156F),
(N37S_L84F_A106V_D108N_H123Y_D147Y_E155V_I156F_K161T),
(H36L_L84F_A106V_D108N_H123Y_D147Y_Q154H_E155V_I156F),
(N72S_L84F_A106V_D108N_H123Y_S 146R_D147Y_E155V_I156F),
(H36L_P48L_L84F_A106V_D108N_H123Y_E134G_D147Y_E155V_I156F_ K157N),
(H36L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F),
(L84F_A106V_D108N_H123Y_S146R_D147Y_E155V_I156F_K161T),
(N37S_R51H_D77G_L84F_A106V_D108N_H123Y_D147Y_E155V_I156F),
(R51L_L84F_A106V_D108N_H123Y_D147Y_E155V_I156F_K157N),
(D24G_Q71R_L84F_H96L_A106V_D108N_H123Y_D147Y_E155V_I156F_K160E),
(H36L_G67V_L84F_A106V_D108N_H123Y_S 146T_D147Y_E155V_I156F),
(Q71L_L84F_A106V_D108N_H123Y_L137M_A143E_D147Y_E155V_I156F),
(E25G_L84F_A106V_D108N_H123Y_D147Y_E155V_I156F_Q159L),
(L84F_A91T_F104I_A106V_D108N_H123Y_D147Y_E155V_I156F),
(N72D_L84F_A106V_D108N_H123Y_G125A_D147Y_E155V_I156F),
(P48S_L84F_S97C_A106V_D108N_H123Y_D147Y_E155V_I156F),
(W23G_L84F_A106V_D108N_H123Y_D147Y_E155V_I156F),
(D24G_P48L_Q71R_L84F_A106V_D108N_H123Y_D147Y_E155V_I156F_Q159L),
(L84F_A106V_D108N_H123Y_A142N_D147Y_E155V_I156F),
(H36L_R51L_L84F_A106V_D108N_H123Y_A142N_S146C_D147Y_E155V_I156F_K157N),
(N37S_L84F_A106V_D108N_H123Y_A142N_D147Y_E155V_I156F_K161T),
(L84F_A106V_D108N_D147Y_E155V_I156F),
(R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F_K157N_K161T),
(L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F_K161T),
(L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F_K157N_K160E_K161T),
(L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F_K157N_K160E),
(R74Q_L84F_A106V_D108N_H123Y_D147Y_E155V_I156F),
(R74A_L84F_A106V_D108N_H123Y_D147Y_E155V_I156F),
(L84F_A106V_D108N_H123Y_D147Y_E155V_I156F),
(R74Q_L84F_A106V_D108N_H123Y_D147Y_E155V_I156F),
(L84F_R98Q_A106V_D108N_H123Y_D147Y_E155V_I156F),
(L84F_A106V_D108N_H123Y_R129Q_D147Y_E155V_I156F),

(P48S_L84F_A106V_D108N_H123Y_A142N_D147Y_E155V_I156F), (P48S_A142N), (P48T_I49V_L84F_A106V_D108N_H123Y_A142N_D147Y_E155V_I156F_L157N), (P48T_I49V_A142N),(H36L_P48S_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F _K157N), (H36L_P48S_R51L_L84F_A106V_D108N_H123Y_S146C_A142N_D147Y_E155V_I156F (H36L_P48T_I49V_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F _K157N),(H36L_P48T_I49V_R51L_L84F_A106V_D108N_H123Y_A142N_S146C_D147Y_E155V_ I156F _K157N), (H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F _K157N), (H36L_P48A_R51L_L84F_A106V_D108N_H123Y_A142N_S146C_D147Y_E155V_I156F _K157N), (H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_A142N_D147Y_E155V_I156F _K157N), (W23L_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F _K157N), (W23R_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F _K157N), (W23L_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146R_D147Y_E155V_I156F _K161T), (H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_R152H_E155V_I156F _K157N), (H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_R152P_E155V_I156F _K157N), (W23L_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_R152P_E155V _I156F _K157N), (W23L_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_A142A_S146C_D147Y_E155 V_I156F _K157N), (W23L_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_A142A_S146C_D147Y_R152 P _E155V_I156F _K157N), (W23L_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146R_D147Y_E155V_I156F _K161T),

(W23R_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_R152P_E155V _I156F _K157N),
(H36L_P48A_R51L_L84F_A106V_D108N_H123Y_A142N_S146C_D147Y_R152P_E155 V_I156F _K157N).

Cytidine deaminase

[00653] In addition to adenosine deaminase, the fusion proteins of the invention comprise one or more cytidine deaminases. In some embodiments, the cytidine deaminases provided herein are capable of deaminating cytosine or 5-methylcytosine to uracil or thymine. In some embodiments, the cytidine deaminases provided herein are capable of deaminating cytosine in DNA. The cytidine deaminase may be derived from any suitable organism. In some embodiments, the cytidine deaminase is a naturally-occurring cytidine deaminase that includes one or more mutations corresponding to any of the mutations provided herein. One of skill in the art will be able to identify the corresponding residue in any homologous protein, e.g., by sequence alignment and determination of homologous residues. Accordingly, one of skill in the art would be able to generate mutations in any naturally-occurring cytidine deaminase that corresponds to any of the mutations described herein. In some embodiments, the cytidine deaminase is from a prokaryote. In some embodiments, the cytidine deaminase is from a bacterium. In some embodiments, the cytidine deaminase is from a mammal (e.g., human).

[00654] In some embodiments, the cytidine deaminase comprises an amino acid sequence that is at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5% identical to any one of the cytidine deaminase amino acid sequences set forth herein. It should be appreciated that cytidine deaminases provided herein may include one or more mutations (e.g., any of the mutations provided herein). Some embodiments provide a polynucleotide molecule encoding the cytidine deaminase nucleobase editor polypeptide of any previous aspect or as delineated herein. In some embodiments, the polynucleotide is codon optimized.

[00655] The disclosure provides any deaminase domains with a certain percent identity plus any of the mutations or combinations thereof described herein. In some embodiments, the cytidine deaminase comprises an amino acid sequence that has 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 21, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35,

36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, or more mutations compared to a reference sequence, or any of the cytidine deaminases provided herein. In some embodiments, the cytidine deaminase comprises an amino acid sequence that has at least 5, at least 10, at least 15, at least 20, at least 25, at least 30, at least 35, at least 40, at least 45, at least 50, at least 60, at least 70, at least 80, at least 90, at least 100, at least 110, at least 120, at least 130, at least 140, at least 150, at least 160, or at least 170 identical contiguous amino acid residues as compared to any one of the amino acid sequences known in the art or described herein.

[00656] A fusion protein of the invention second protein comprises two or more nucleic acid editing domains. In some embodiments, the nucleic acid editing domain can catalyze a C to U base change. In some embodiments, the nucleic acid editing domain is a deaminase domain. In some embodiments, the deaminase is a cytidine deaminase. In some embodiments, the deaminase is an apolipoprotein B mRNA-editing complex (APOBEC) family deaminase. In some embodiments, the deaminase is an APOBECl deaminase. In some embodiments, the deaminase is an APOBEC2 deaminase. In some embodiments, the deaminase is an APOBEC3 deaminase. In some embodiments, the deaminase is an APOBEC3 A deaminase. In some embodiments, the deaminase is an APOBEC3B deaminase. In some embodiments, the deaminase is an APOBEC3C deaminase. In some embodiments, the deaminase is an APOBEC3D deaminase. In some embodiments, the deaminase is an APOBEC3E deaminase. In some embodiments, the deaminase is an APOBEC3F deaminase. In some embodiments, the deaminase is an APOBEC3G deaminase. In some embodiments, the deaminase is an APOBEC3H deaminase. In some embodiments, the deaminase is an APOBEC4 deaminase. In some embodiments, the deaminase is an activation-induced deaminase (AID). In some embodiments, the deaminase is a vertebrate deaminase. In some embodiments, the deaminase is an invertebrate deaminase. In some embodiments, the deaminase is a human, chimpanzee, gorilla, monkey, cow, dog, rat, or mouse deaminase. In some embodiments, the deaminase is a human deaminase. In some embodiments, the deaminase is a rat deaminase, e.g., rAPOBECl . In some embodiments, the deaminase is a Petromyzon marinus cytidine deaminase 1 (pmCDAl). In some embodiments, the deminase is a human APOBEC3G. In some embodiments, the deaminase is a fragment of the human APOBEC3G. In some embodiments, the deaminase is a human APOBEC3G variant comprising a D316R D317R mutation. In some embodiments, the deaminase is a fragment of the human APOBEC3G and comprising mutations corresponding to the D316R D317R mutations. In some embodiments, the nucleic acid editing domain is at least 80%, at least

85%, at least 90%, at least 92%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%), or at least 99.5% identical to the deaminase domain of any deaminase described herein.

[00657] In certain embodiments, the fusion proteins provided herein comprise one or more features that improve the base editing activity of the fusion proteins. For example, any of the fusion proteins provided herein may comprise a Cas9 domain that has reduced nuclease activity. In some embodiments, any of the fusion proteins provided herein may have a Cas9 domain that does not have nuclease activity (dCas9), or a Cas9 domain that cuts one strand of a duplexed DNA molecule, referred to as a Cas9 nickase (nCas9).

Cas9 domains of Nucleobase Editors

[00658] In some aspects, a nucleic acid programmable DNA binding protein (napDNAbp) is selected from the group consisting of Cas9, CasX, CasY, Cpf1, Cas12b/C2c1, and Cas12c/C2c3, or active fragments thereof. In another embodiment, the napDNAbp domain comprises a catalytic domain capable of cleaving the reverse complement strand of the nucleic acid sequence. In another embodiment, the napDNAbp domain does not comprise a catalytic domain capable of cleaving the nucleic acid sequence. In another embodiment, the Cas9 is dCas9 or nCas9. In another embodiment, the napDNAbp comprises a nucleobase editor.

[00659] In some embodiments, a nucleic acid programmable DNA binding protein (napDNAbp) is a Cas9 domain. Non-limiting, exemplary Cas9 domains are provided herein. The Cas9 domain may be a nuclease active Cas9 domain, a nuclease inactive Cas9 domain (a nuclease dead Cas9, or dCas9), or a Cas9 nickase (nCas9). In some embodiments, the Cas9 domain is a nuclease active domain. For example, the Cas9 domain may be a Cas9 domain that cuts both strands of a duplexed nucleic acid (e.g., both strands of a duplexed DNA molecule). In some embodiments, the Cas9 domain comprises any one of the amino acid sequences as set forth herein. In some embodiments the Cas9 domain comprises an amino acid sequence that is at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5% identical to any one of the amino acid sequences set forth herein. In some embodiments, the Cas9 domain comprises an amino acid sequence that has 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 21, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50 or more or more mutations compared to any one of the amino acid sequences set forth herein. In some embodiments, the

Cas9 domain comprises an amino acid sequence that has at least 10, at least 15, at least 20, at least 30, at least 40, at least 50, at least 60, at least 70, at least 80, at least 90, at least 100, at least 150, at least 200, at least 250, at least 300, at least 350, at least 400, at least 500, at least 600, at least 700, at least 800, at least 900, at least 1000, at least 1100, or at least 1200 identical contiguous amino acid residues as compared to any one of the amino acid sequences set forth herein.

[00660] In some embodiments, the Cas9 domain is a nuclease-inactive Cas9 domain (dCas9). For example, the dCas9 domain may bind to a duplexed nucleic acid molecule (*e.g.*, via a gRNA molecule) without cleaving either strand of the duplexed nucleic acid molecule. In some embodiments, the nuclease-inactive dCas9 domain comprises a D10X mutation and a H840X mutation of the amino acid sequence set forth herein, or a corresponding mutation in any of the amino acid sequences provided herein, wherein X is any amino acid change. In some embodiments, the nuclease-inactive dCas9 domain comprises a D10A mutation and a H840A mutation of the amino acid sequence set forth herein, or a corresponding mutation in any of the amino acid sequences provided herein. As one example, a nuclease-inactive Cas9 domain comprises the amino acid sequence set forth in Cloning vector pPlatTET-gRNA2 (Accession No. BAV54124).

[00661] MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGAL
LFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEE
DKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRG
HFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLEN
LIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQI
GDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVR
QQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDL
LRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLAR
GNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLL
YEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFK
KIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDRE
MIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSD
GFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKV
VDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHP
VENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKV
LTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSEL
DKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRK

DFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIA
KSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFAT
VRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTV
AYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIK
LPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNE
QKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHL
FTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD
(see, *e.g.*, Qi *et al.*, "Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression." *Cell.* 2013; 152(5):1173-83, the entire contents of which are incorporated herein by reference).

[00662] Additional suitable nuclease-inactive dCas9 domains will be apparent to those of skill in the art based on this disclosure and knowledge in the field, and are within the scope of this disclosure. Such additional exemplary suitable nuclease-inactive Cas9 domains include, but are not limited to, D10A/H840A, D10A/D839A/H840A, and D10A/D839A/H840A/N863A mutant domains (See, *e.g.*, Prashant *et al.*, CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nature Biotechnology*. 2013; 31(9): 833-838, the entire contents of which are incorporated herein by reference). In some embodiments the dCas9 domain comprises an amino acid sequence that is at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5% identical to any one of the dCas9 domains provided herein. In some embodiments, the Cas9 domain comprises an amino acid sequences that has 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 21, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50 or more or more mutations compared to any one of the amino acid sequences set forth herein. In some embodiments, the Cas9 domain comprises an amino acid sequence that has at least 10, at least 15, at least 20, at least 30, at least 40, at least 50, at least 60, at least 70, at least 80, at least 90, at least 100, at least 150, at least 200, at least 250, at least 300, at least 350, at least 400, at least 500, at least 600, at least 700, at least 800, at least 900, at least 1000, at least 1100, or at least 1200 identical contiguous amino acid residues as compared to any one of the amino acid sequences set forth herein.

[00663] In some embodiments, the Cas9 domain is a Cas9 nickase. The Cas9 nickase may be a Cas9 protein that is capable of cleaving only one strand of a duplexed nucleic acid molecule (*e.g.,* a duplexed DNA molecule). In some embodiments the Cas9 nickase cleaves

the target strand of a duplexed nucleic acid molecule, meaning that the Cas9 nickase cleaves the strand that is base paired to (complementary to) a gRNA (*e.g.*, an sgRNA) that is bound to the Cas9. In some embodiments, a Cas9 nickase comprises a D10A mutation and has a histidine at position 840. In some embodiments the Cas9 nickase cleaves the non-target, non-base-edited strand of a duplexed nucleic acid molecule, meaning that the Cas9 nickase cleaves the strand that is not base paired to a gRNA (*e.g.*, an sgRNA) that is bound to the Cas9. In some embodiments, a Cas9 nickase comprises an H840A mutation and has an aspartic acid residue at position 10, or a corresponding mutation. In some embodiments the Cas9 nickase comprises an amino acid sequence that is at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5% identical to any one of the Cas9 nickases provided herein. Additional suitable Cas9 nickases will be apparent to those of skill in the art based on this disclosure and knowledge in the field, and are within the scope of this disclosure.

Cas9 Domains with Reduced PAM Exclusivity

[00664] Some aspects of the disclosure provide Cas9 domains that have different PAM specificities. In one particular embodiment, the invention features nucleobase editor fusion proteins that comprise an nCas9 domain and a dCas9 domain, where each of the Cas9 domains has a different PAM specificity. Typically, Cas9 proteins, such as Cas9 from *S. pyogenes* (spCas9), require a canonical NGG PAM sequence to bind a particular nucleic acid region, where the "N" in "NGG" is adenosine (A), thymidine (T), or cytosine (C), and the G is guanosine. This may limit the ability to edit desired bases within a genome. In some embodiments, the base editing fusion proteins provided herein may need to be placed at a precise location, for example a region comprising a target base that is upstream of the PAM. See e.g., Komor, A.C., *et al.*, "Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage" *Nature* 533, 420-424 (2016), the entire contents of which are hereby incorporated by reference. Accordingly, in some embodiments, any of the fusion proteins provided herein may contain a Cas9 domain that can bind a nucleotide sequence that does not contain a canonical (*e.g.*, NGG) PAM sequence. Cas9 domains that bind to non-canonical PAM sequences have been described in the art and would be apparent to the skilled artisan. For example, Cas9 domains that bind non-canonical PAM sequences have been described in Kleinstiver, B. P., *et al.*, "Engineered CRISPR-Cas9 nucleases with altered PAM specificities" *Nature* 523, 481-485 (2015); and Kleinstiver, B. P., *et al.*,

"Broadening the targeting range of *Staphylococcus aureus* CRISPR-Cas9 by modifying PAM recognition" *Nature Biotechnology* 33, 1293-1298 (2015); the entire contents of each are hereby incorporated by reference. Several PAM variants are described at Table 3 below:

[00665] Table 3. Cas9 proteins and corresponding PAM sequences

Variant	PAM
spCas9	NGG
spCas9-VRQR	NGA
spCas9-VRER	NGCG
xCas9 (sp)	NGN
saCas9	NNGRRT
saCas9-KKH	NNNRRT
spCas9-MQKSER	NGCG
spCas9-MQKSER	NGCN
spCas9-LRKIQK	NGTN
spCas9-LRVSQK	NGTN
spCas9-LRVSQL	NGTN
Cpf1	5' (TTTV)

[00666] In some embodiments, the Cas9 domain is a Cas9 domain from *Staphylococcus aureus* (SaCas9). In some embodiments, the SaCas9 domain is a nuclease active SaCas9, a nuclease inactive SaCas9 (SaCas9d), or a SaCas9 nickase (SaCas9n). In some embodiments, the SaCas9 comprises a N579A mutation, or a corresponding mutation in any of the amino acid sequences provided herein.

[00667] In some embodiments, the SaCas9 domain, the SaCas9d domain, or the SaCas9n domain can bind to a nucleic acid sequence having a non-canonical PAM. In some

embodiments, the SaCas9 domain, the SaCas9d domain, or the SaCas9n domain can bind to a nucleic acid sequence having a NNGRRT PAM sequence. In some embodiments, the SaCas9 domain comprises one or more of a E781X, a N967X, and a R1014X mutation, or a corresponding mutation in any of the amino acid sequences provided herein, wherein X is any amino acid. In some embodiments, the SaCas9 domain comprises one or more of a E781K, a N967K, and a R1014H mutation, or one or more corresponding mutation in any of the amino acid sequences provided herein. In some embodiments, the SaCas9 domain comprises a E781K, a N967K, or a R1014H mutation, or corresponding mutations in any of the amino acid sequences provided herein.

Exemplary SaCas9 sequence

KRNYILGLDIGITSVGYGIIDYETRDVIDAGVRLFKEANVENNEGRRSKRGARRLKRR
RRHRIQRVKKLLFDYNLLTDHSELSGINPYEARVKGLSQKLSEEEFSAALLHLAKRR
GVHNVNEVEEDTGNELSTKEQISRNSKALEEKYVAELQLERLKKDGEVRGSINRFKT
SDYVKEAKQLLKVQKAYHQLDQSFIDTYIDLLETRRTYYEGPGEGSPFGWKDIKEW
YEMLMGHCTYFPEELRSVKYAYNADLYNALNDLNNLVITRDENEKLEYYEKFQIIEN
VFKQKKKPTLKQIAKEILVNEEDIKGYRVTSTGKPEFTNLKVYHDIKDITARKEIIENA
ELLDQIAKILTIYQSSEDIQEELTNLNSELTQEEIEQISNLKGYTGTHNLSLKAINLILDE
LWHTNDNQIAIFNRLKLVPKKVDLSQQKEIPTTLVDDFILSPVVKRSFIQSIKVINAIIK
KYGLPNDIIIELAREKNSKDAQKMINEMQKRNRQTNERIEEIIRTTGKENAKYLIEKIK
LHDMQEGKCLYSLEAIPLEDLLNNPFNYEVDHIIPRSVSFDNSFNNKVLVKQEE**N**SKK
GNRTPFQYLSSSDSKISYETFKKHILNLAKGKGRISKTKKEYLLEERDINRFSVQKDFI
NRNLVDTRYATRGLMNLLRSYFRVNNLDVKVKSINGGFTSFLRRKWKFKKERNKG
YKHHAEDALIIANADFIFKEWKKLDKAKKVMENQMFEEKQAESMPEIETEQEYKEIF
ITPHQIKHIKDFKDYKYSHRVDKKPNRELINDTLYSTRKDDKGNTLIVNNLNGLYDK
DNDKLKKLINKSPEKLLMYHHDPQTYQKLKLIMEQYGDEKNPLYKYYEETGNYLTK
YSKKDNGPVIKKIKYYGNKLNAHLDITDDYPNSRNKVVKLSLKPYRFDVYLDNGVY
KFVTVKNLDVIKKENYYEVNSKCYEEAKKLKKISNQAEFIASFYNNDLIKINGELYR
VIGVNNDLLNRIEVNMIDITYREYLENMNDKRPPRIIKTIASKTQSIKKYSTDILGNLY
EVKSKKHPQIIKKG

[00668] Residue N579 above, which is underlined and in bold, may be mutated (*e.g.*, to a A579) to yield a SaCas9 nickase.

Exemplary SaCas9n sequence

KRNYILGLDIGITSVGYGIIDYETRDVIDAGVRLFKEANVENNEGRRSKRGARRLKRR
RRHRIQRVKKLLFDYNLLTDHSELSGINPYEARVKGLSQKLSEEEFSAALLHLAKRR
GVHNVNEVEEDTGNELSTKEQISRNSKALEEKYVAELQLERLKKDGEVRGSINRFKT
SDYVKEAKQLLKVQKAYHQLDQSFIDTYIDLLETRRTYYEGPGEGSPFGWKDIKEW
YEMLMGHCTYFPEELRSVKYAYNADLYNALNDLNNLVITRDENEKLEYYEKFQIIEN
VFKQKKKPTLKQIAKEILVNEEDIKGYRVTSTGKPEFTNLKVYHDIKDITARKEIIENA
ELLDQIAKILTIYQSSEDIQEELTNLNSELTQEEIEQISNLKGYTGTHNLSLKAINLILDE
LWHTNDNQIAIFNRLKLVPKKVDLSQQKEIPTTLVDDFILSPVVKRSFIQSIKVINAIIK
KYGLPNDIIIELAREKNSKDAQKMINEMQKRNRQTNERIEEIIRTTGKENAKYLIEKIK
LHDMQEGKCLYSLEAIPLEDLLNNPFNYEVDHIIPRSVSFDNSFNNKVLVKQEE**A**SKK
GNRTPFQYLSSSDSKISYETFKKHILNLAKGKGRISKTKKEYLLEERDINRFSVQKDFI
NRNLVDTRYATRGLMNLLRSYFRVNNLDVKVKSINGGFTSFLRRKWKFKKERNKG
YKHHAEDALIIANADFIFKEWKKLDKAKKVMENQMFEEKQAESMPEIETEQEYKEIF
ITPHQIKHIKDFKDYKYSHRVDKKPNRELINDTLYSTRKDDKGNTLIVNNLNGLYDK
DNDKLKKLINKSPEKLLMYHHDPQTYQKLKLIMEQYGDEKNPLYKYYEETGNYLTK
YSKKDNGPVIKKIKYYGNKLNAHLDITDDYPNSRNKVVKLSLKPYRFDVYLDNGVY
KFVTVKNLDVIKKENYYEVNSKCYEEAKKLKKISNQAEFIASFYNNDLIKINGELYR
VIGVNNDLLNRIEVNMIDITYREYLENMNDKRPPRIIKTIASKTQSIKKYSTDILGNLY
EVKSKKHPQIIKKG

[00669] Residue A579 above, which can be mutated from N579 to yield a SaCas9 nickase, is underlined and in bold.

Exemplary SaKKH Cas9

KRNYILGLDIGITSVGYGIIDYETRDVIDAGVRLFKEANVENNEGRRSKRGARRLKRR
RRHRIQRVKKLLFDYNLLTDHSELSGINPYEARVKGLSQKLSEEEFSAALLHLAKRR
GVHNVNEVEEDTGNELSTKEQISRNSKALEEKYVAELQLERLKKDGEVRGSINRFKT
SDYVKEAKQLLKVQKAYHQLDQSFIDTYIDLLETRRTYYEGPGEGSPFGWKDIKEW
YEMLMGHCTYFPEELRSVKYAYNADLYNALNDLNNLVITRDENEKLEYYEKFQIIEN
VFKQKKKPTLKQIAKEILVNEEDIKGYRVTSTGKPEFTNLKVYHDIKDITARKEIIENA
ELLDQIAKILTIYQSSEDIQEELTNLNSELTQEEIEQISNLKGYTGTHNLSLKAINLILDE
LWHTNDNQIAIFNRLKLVPKKVDLSQQKEIPTTLVDDFILSPVVKRSFIQSIKVINAIIK
KYGLPNDIIIELAREKNSKDAQKMINEMQKRNRQTNERIEEIIRTTGKENAKYLIEKIK
LHDMQEGKCLYSLEAIPLEDLLNNPFNYEVDHIIPRSVSFDNSFNNKVLVKQEE**A**SKK

GNRTPFQYLSSSDSKISYETFKKHILNLAKGKGRISKTKKEYLLEERDINRFSVQKDFI
NRNLVDTRYATRGLMNLLRSYFRVNNLDVKVKSINGGFTSFLRRKWKFKKERNKG
YKHHAEDALIIANADFIFKEWKKLDKAKKVMENQMFEEKQAESMPEIETEQEYKEIF
ITPHQIKHIKDFKDYKYSHRVDKKPNR*K*LINDTLYSTRKDDKGNTLIVNNLNGLYDK
DNDKLKKLINKSPEKLLMYHHDPQTYQKLKLIMEQYGDEKNPLYKYYEETGNYLTK
YSKKDNGPVIKKIKYYGNKLNAHLDITDDYPNSRNKVVKLSLKPYRFDVYLDNGVY
KFVTVKNLDVIKKENYYEVNSKCYEEAKKLKKISNQAEFIASFY*K*NDLIKINGELYRV
IGVNNDLLNRIEVNMIDITYREYLENMNDKRPP*H*IIKTIASKTQSIKKYSTDILGNLYE
VKSKKHPQIIKKG.

[00670] Residue A579 above, which can be mutated from N579 to yield a SaCas9 nickase, is underlined and in bold. Residues K781, K967, and H1014 above, which can be mutated from E781, N967, and R1014 to yield a SaKKH Cas9 are underlined and in italics.

[00671] In some embodiments, the Cas9 domain is a Cas9 domain from *Streptococcus pyogenes* (SpCas9). In some embodiments, the SpCas9 domain is a nuclease active SpCas9, a nuclease inactive SpCas9 (SpCas9d), or a SpCas9 nickase (SpCas9n). In some embodiments, the SpCas9 comprises a D9X mutation, or a corresponding mutation in any of the amino acid sequences provided herein, wherein X is any amino acid except for D. In some embodiments, the SpCas9 comprises a D9A mutation, or a corresponding mutation in any of the amino acid sequences provided herein. In some embodiments, the SpCas9 domain, the SpCas9d domain, or the SpCas9n domain can bind to a nucleic acid sequence having a non-canonical PAM. In some embodiments, the SpCas9 domain, the SpCas9d domain, or the SpCas9n domain can bind to a nucleic acid sequence having an NGG, a NGA, or a NGCG PAM sequence. In some embodiments, the SpCas9 domain comprises one or more of a D1134X, a R1334X, and a T1336X mutation, or a corresponding mutation in any of the amino acid sequences provided herein, wherein X is any amino acid. In some embodiments, the SpCas9 domain comprises one or more of a D1134E, R1334Q, and T1336R mutation, or a corresponding mutation in any of the amino acid sequences provided herein. In some embodiments, the SpCas9 domain comprises a D1134E, a R1334Q, and a T1336R mutation, or corresponding mutations in any of the amino acid sequences provided herein. In some embodiments, the SpCas9 domain comprises one or more of a D1134X, a R1334X, and a T1336X mutation, or a corresponding mutation in any of the amino acid sequences provided herein, wherein X is any amino acid. In some embodiments, the SpCas9 domain comprises one or more of a D1134V, a R1334Q, and a T1336R mutation, or a corresponding mutation in any of the amino acid sequences provided herein. In some

embodiments, the SpCas9 domain comprises a D1134V, a R1334Q, and a T1336R mutation, or corresponding mutations in any of the amino acid sequences provided herein. In some embodiments, the SpCas9 domain comprises one or more of a D1134X, a G1217X, a R1334X, and a T1336X mutation, or a corresponding mutation in any of the amino acid sequences provided herein, wherein X is any amino acid. In some embodiments, the SpCas9 domain comprises one or more of a D1134V, a G1217R, a R1334Q, and a T1336R mutation, or a corresponding mutation in any of the amino acid sequences provided herein. In some embodiments, the SpCas9 domain comprises a D1134V, a G1217R, a R1334Q, and a T1336R mutation, or corresponding mutations in any of the amino acid sequences provided herein.

[00672] In some embodiments, the Cas9 domain of any of the fusion proteins provided herein comprises an amino acid sequence that is at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5% identical to a Cas9 polypeptide described herein. In some embodiments, the Cas9 domain of any of the fusion proteins provided herein comprises the amino acid sequence of any Cas9 polypeptide described herein. In some embodiments, the Cas9 domain of any of the fusion proteins provided herein consists of the amino acid sequence of any Cas9 polypeptide described herein.

Exemplary SpCas9

DKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETA
EATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERH
PIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDL
NPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGE
KKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADL
FLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKY
KEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTF
DNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAW
MTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVY
NELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDS
VEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERL
KTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRN
FMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVK
VMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQL

QNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDK
NRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIK
RQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKV
REINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK
ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSM
PQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVV
AKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFE
LENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQ
HKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGA
PAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD

Exemplary SpCas9n

DKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETA
EATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERH
PIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDL
NPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGE
KKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADL
FLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKY
KEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTF
DNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAW
MTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVY
NELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDS
VEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERL
KTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRN
FMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVK
VMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQL
QNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDK
NRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIK
RQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKV
REINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK
ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSM
PQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVV
AKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFE
LENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQ

HKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGA
PAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD

Exemplary SpEQR Cas9

DKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETA
EATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFVEEDKKHERHPI
FGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNP
DNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKK
NGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFL
AAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKE
IFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDN
GSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMT
RKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNE
LTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEI
SGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTY
AHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQ
LIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMG
RHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEK
LYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGK
SDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLV
ETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINN
YHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAK
YFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVN
IVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGF**E**SPTVAYSVLVVAKVE
KGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENG
RKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKH
YLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAF
KYFDTTIDRK**Q**Y**R**STKEVLDATLIHQSITGLYETRIDLSQLGGD

[00673] Residues E1134, Q1334, and R1336 above, which can be mutated from D1134, R1334, and T1336 to yield a SpEQR Cas9, are underlined and in bold.

Exemplary SpVQR Cas9

DKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETA
EATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERH

PIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDL
NPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGE
KKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADL
FLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKY
KEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTF
DNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAW
MTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVY
NELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDS
VEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERL
KTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRN
FMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVK
VMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQL
QNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDK
NRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIK
RQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKV
REINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK
ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSM
PQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGF**V**SPTVAYSVLVV
AKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFE
LENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQ
HKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGA
PAAFKYFDTTIDRK**Q**Y**R**STKEVLDATLIHQSITGLYETRIDLSQLGGD

[00674] Residues V1134, Q1334, and R1336 above, which can be mutated from D1134, R1334, and T1336 to yield a SpVQR Cas9, are underlined and in bold.Exemplary SpVRER Cas9

DKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETA
EATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERH
PIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDL
NPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGE
KKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADL
FLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKY
KEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTF
DNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAW
MTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVY

NELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDS
VEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERL
KTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRN
FMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVK
VMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQL
QNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDK
NRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIK
RQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKV
REINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK
ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSM
PQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGF**V**SPTVAYSVLVV
AKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFE
LENGRKRMLASA**R**ELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQ
HKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGA
PAAFKYFDTTIDRK**E**Y**R**STKEVLDATLIHQSITGLYETRIDLSQLGGD.

[00675] Residues V1134, R1217, Q1334, and R1336 above, which can be mutated from D1134, G1217, R1334, and T1336 to yield a SpVRER Cas9, are underlined and in bold.

High fidelity Cas9 domains

[00676] Some aspects of the disclosure provide high fidelity Cas9 domains. In some embodiments, high fidelity Cas9 domains are engineered Cas9 domains comprising one or more mutations that decrease electrostatic interactions between the Cas9 domain and a sugar-phosphate backbone of a DNA, as compared to a corresponding wild-type Cas9 domain. Without wishing to be bound by any particular theory, high fidelity Cas9 domains that have decreased electrostatic interactions with a sugar-phosphate backbone of DNA may have less off-target effects. In some embodiments, a Cas9 domain (*e.g.,* a wild type Cas9 domain) comprises one or more mutations that decreases the association between the Cas9 domain and a sugar-phosphate backbone of a DNA. In some embodiments, a Cas9 domain comprises one or more mutations that decreases the association between the Cas9 domain and a sugar-phosphate backbone of a DNA by at least 1%, at least 2%, at least 3%, at least 4%, at least 5%, at least 10%, at least 15%, at least 20%, at least 25%, at least 30%, at least 35%, at least 40%, at least 45%, at least 50%, at least 55%, at least 60%, at least 65%, or at least 70%.

[00677] In some embodiments, any of the Cas9 fusion proteins provided herein comprise one or more of a N497X, a R661X, a Q695X, and/or a Q926X mutation, or a corresponding mutation in any of the amino acid sequences provided herein, wherein X is any amino acid. In some embodiments, any of the Cas9 fusion proteins provided herein comprise one or more of a N497A, a R661A, a Q695A, and/or a Q926A mutation, or a corresponding mutation in any of the amino acid sequences provided herein. In some embodiments, the Cas9 domain comprises a D10A mutation, or a corresponding mutation in any of the amino acid sequences provided herein. Cas9 domains with high fidelity are known in the art and would be apparent to the skilled artisan. For example, Cas9 domains with high fidelity have been described in Kleinstiver, B.P., *et al.* "High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects." *Nature* 529, 490-495 (2016); and Slaymaker, I.M., *et al.* "Rationally engineered Cas9 nucleases with improved specificity." *Science* 351, 84-88 (2015); the entire contents of each are incorporated herein by reference.

[00678] High Fidelity Cas9 domain mutations relative to Cas9 are shown in bold and underlines

DKKYSIGL**A**IGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETA
EATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERH
PIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDL
NPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGE
KKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADL
FLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKY
KEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTF
DNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAW
MTRKSEETITPWNFEEVVDKGASAQSFIERMT**A**FDKNLPNEKVLPKHSLLYEYFTVY
NELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDS
VEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERL
KTYAHLFDDKVMKQLKRRRYTGWG**A**LSRKLINGIRDKQSGKTILDFLKSDGFANRN
FM**A**LIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVK
VMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQL
QNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDK
NRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIK
RQLVETR**A**ITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKV
REINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK
ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSM

PQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVV
AKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFE
LENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQ
HKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGA
PAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD

Nucleic acid programmable DNA binding proteins

[00679] Some aspects of the disclosure provide nucleic acid programmable DNA binding proteins, which may be used to guide a protein, such as a base editor, to a specific nucleic acid (e.g., DNA or RNA) sequence. Nucleic acid programmable DNA binding proteins include, without limitation, Cas9 (e.g., dCas9 and nCas9), CasX, CasY, Cpf1, Cas12b/C2c1, and Cas12c/C2c3. One example of a nucleic acid programmable DNA-binding protein that has different PAM specificity than Cas9 is Clustered Regularly Interspaced Short Palindromic Repeats from *Prevotella* and *Francisella* 1 (Cpf1). Similar to Cas9, Cpf1 is also a class 2 CRISPR effector, it has been shown that Cpf1 mediates robust DNA interference with features distinct from Cas9. Cpf1 is a single RNA-guided endonuclease lacking tracrRNA, and it utilizes a T-rich protospacer-adjacent motif (TTN, TTTN, or YTN). Moreover, Cpf1 cleaves DNA via a staggered DNA double-stranded break. Out of 16 Cpf1-family proteins, two enzymes from Acidaminococcus and Lachnospiraceae are shown to have efficient genome-editing activity in human cells. Cpf1 proteins are known in the art and have been described previously, for example Yamano et al., "Crystal structure of Cpf1 in complex with guide RNA and target DNA." *Cell* (165) 2016, p. 949-962; the entire contents of which is hereby incorporated by reference.

[00680] Also useful in the present compositions and methods are nuclease-inactive Cpf1 (dCpf1) variants that may be used as a guide nucleotide sequence-programmable DNA-binding protein domain. The Cpf1 protein has a RuvC-like endonuclease domain that is similar to the RuvC domain of Cas9 but does not have a HNH endonuclease domain, and the N-terminal of Cpf1 does not have the alfa-helical recognition lobe of Cas9. It was shown in Zetsche et al., *Cell,* 163, 759-771, 2015 (which is incorporated herein by reference) that, the RuvC-like domain of Cpf1 is responsible for cleaving both DNA strands and inactivation of the RuvC-like domain inactivates Cpf1 nuclease activity. For example, mutations corresponding to D917A, E1006A, or D1255A in *Francisella novicida* Cpf1 inactivate Cpf1 nuclease activity. In some embodiments, the dCpf1 of the present disclosure comprises mutations corresponding to D917A, E1006A, D1255A, D917A/E1006A, D917A/D1255A,

E1006A/D1255A, or D917A/E1006A/D1255A. It is to be understood that any mutations, e.g., substitution mutations, deletions, or insertions that inactivate the RuvC domain of Cpf1, may be used in accordance with the present disclosure.

[00681] In some embodiments, the nucleic acid programmable DNA binding protein (napDNAbp) of any of the fusion proteins provided herein may be a Cpf1 protein. In some embodiments, the Cpf1 protein is a Cpf1 nickase (nCpf1). In some embodiments, the Cpf1 protein is a nuclease inactive Cpf1 (dCpf1). In some embodiments, the Cpf1, the nCpf1, or the dCpf1 comprises an amino acid sequence that is at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5% identical to a Cpf1 sequence disclosed herein. In some embodiments, the dCpf1comprises an amino acid sequence that is at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at ease 99.5% identical to a Cpf1 sequence disclosed herein, and comprises mutations corresponding to D917A, E1006A, D1255A, D917A/E1006A, D917A/D1255A, E1006A/D1255A, or D917A/E1006A/D1255A. It should be appreciated that Cpf1 from other bacterial species may also be used in accordance with the present disclosure.

[00682] Wild type *Francisella novicida* Cpf1 (D917, E1006, and D1255 are bolded and underlined)

MSIYQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAKDYKKAKQIIDKYH
QFFIEEILSSVCISEDLLQNYSDVYFKLKKSDDDNLQKDFKSAKDTIKKQISEYIKDSE
KFKNLFNQNLIDAKKGQESDLILWLKQSKDNGIELFKANSDITDIDEALEIIKSFKGWT
TYFKGFHENRKNVYSSNDIPTSIIYRIVDDNLPKFLENKAKYESLKDKAPEAINYEQIK
KDLAEELTFDIDYKTSEVNQRVFSLDEVFEIANFNNYLNQSGITKFNTIIGGKFVNGEN
TKRKGINEYINLYSQQINDKTLKKYKMSVLFKQILSDTESKSFVIDKLEDDSDVVTTM
QSFYEQIAAFKTVEEKSIKETLSLLFDDLKAQKLDLSKIYFKNDKSLTDLSQQVFDDY
SVIGTAVLEYITQQIAPKNLDNPSKKEQELIAKKTEKAKYLSLETIKLALEEFNKHRDI
DKQCRFEEILANFAAIPMIFDEIAQNKDNLAQISIKYQNQGKKDLLQASAEDDVKAIK
DLLDQTNNLLHKLKIFHISQSEDKANILDKDEHFYLVFEECYFELANIVPLYNKIRNYI
TQKPYSDEKFKLNFENSTLANGWDKNKEPDNTAILFIKDDKYYLGVMNKKNNKIFD
DKAIKENKGEGYKKIVYKLLPGANKMLPKVFFSAKSIKFYNPSEDILRIRNHSTHTKN
GSPQKGYEKFEFNIEDCRKFIDFYKQSISKHPEWKDFGFRFSDTQRYNSIDEFYREVE
NQGYKLTFENISESYIDSVVNQGKLYLFQIYNKDFSAYSKGRPNLHTLYWKALFDER

NLQDVVYKLNGEAELFYRKQSIPKKITHPAKEAIANKNKDNPKKESVFEYDLIKDKR
FTEDKFFFHCPITINFKSSGANKFNDEINLLLKEKANDVHILSI**D**RGERHLAYYTLVDG
KGNIIKQDTFNIIGNDRMKTNYHDKLAAIEKDRDSARKDWKKINNIKEMKEGYLSQV
VHEIAKLVIEYNAIVVF**E**DLNFGFKRGRFKVEKQVYQKLEKMLIEKLNYLVFKDNEF
DKTGGVLRAYQLTAPFETFKKMGKQTGIIYYVPAGFTSKICPVTGFVNQLYPKYESV
SKSQEFFSKFDKICYNLDKGYFEFSFDYKNFGDKAAKGKWTIASFGSRLINFRNSDKN
HNWDTREVYPTKELEKLLKDYSIEYGHGECIKAAICGESDKKFFAKLTSVLNTILQM
RNSKTGTELDYLISPVADVNGNFFDSRQAPKNMPQDA**D**ANGAYHIGLKGLMLLGRI
KNNQEGKKLNLVIKNEEYFEFVQNRNN

[00683] *Francisella novicida* Cpf1 D917A (A917, E1006, and D1255 are bolded and underlined)

MSIYQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAKDYKKAKQIIDKYH
QFFIEEILSSVCISEDLLQNYSDVYFKLKKSDDDNLQKDFKSAKDTIKKQISEYIKDSE
KFKNLFNQNLIDAKKGQESDLILWLKQSKDNGIELFKANSDITDIDEALEIIKSFKGWT
TYFKGFHENRKNVYSSNDIPTSIIYRIVDDNLPKFLENKAKYESLKDKAPEAINYEQIK
KDLAEELTFDIDYKTSEVNQRVFSLDEVFEIANFNNYLNQSGITKFNTIIGGKFVNGEN
TKRKGINEYINLYSQQINDKTLKKYKMSVLFKQILSDTESKSFVIDKLEDDSDVVTTM
QSFYEQIAAFKTVEEKSIKETLSLLFDDLKAQKLDLSKIYFKNDKSLTDLSQQVFDDY
SVIGTAVLEYITQQIAPKNLDNPSKKEQELIAKKTEKAKYLSLETIKLALEEFNKHRDI
DKQCRFEEILANFAAIPMIFDEIAQNKDNLAQISIKYQNQGKKDLLQASAEDDVKAIK
DLLDQTNNLLHKLKIFHISQSEDKANILDKDEHFYLVFEECYFELANIVPLYNKIRNYI
TQKPYSDEKFKLNFENSTLANGWDKNKEPDNTAILFIKDDKYYLGVMNKKNNKIFD
DKAIKENKGEGYKKIVYKLLPGANKMLPKVFFSAKSIKFYNPSEDILRIRNHSTHTKN
GSPQKGYEKFEFNIEDCRKFIDFYKQSISKHPEWKDFGFRFSDTQRYNSIDEFYREVE
NQGYKLTFENISESYIDSVVNQGKLYLFQIYNKDFSAYSKGRPNLHTLYWKALFDER
NLQDVVYKLNGEAELFYRKQSIPKKITHPAKEAIANKNKDNPKKESVFEYDLIKDKR
FTEDKFFFHCPITINFKSSGANKFNDEINLLLKEKANDVHILSI**A**RGERHLAYYTLVDG
KGNIIKQDTFNIIGNDRMKTNYHDKLAAIEKDRDSARKDWKKINNIKEMKEGYLSQV
VHEIAKLVIEYNAIVVF**E**DLNFGFKRGRFKVEKQVYQKLEKMLIEKLNYLVFKDNEF
DKTGGVLRAYQLTAPFETFKKMGKQTGIIYYVPAGFTSKICPVTGFVNQLYPKYESV
SKSQEFFSKFDKICYNLDKGYFEFSFDYKNFGDKAAKGKWTIASFGSRLINFRNSDKN
HNWDTREVYPTKELEKLLKDYSIEYGHGECIKAAICGESDKKFFAKLTSVLNTILQM

RNSKTGTELDYLISPVADVNGNFFDSRQAPKNMPQDA**D**ANGAYHIGLKGLMLLGRI
KNNQEGKKLNLVIKNEEYFEFVQNRNN

[00684] *Francisella novicida* Cpf1 E1006A (D917, A1006, and D1255 are bolded and underlined)

MSIYQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAKDYKKAKQIIDKYH
QFFIEEILSSVCISEDLLQNYSDVYFKLKKSDDDNLQKDFKSAKDTIKKQISEYIKDSE
KFKNLFNQNLIDAKKGQESDLILWLKQSKDNGIELFKANSDITDIDEALEIIKSFKGWT
TYFKGFHENRKNVYSSNDIPTSIIYRIVDDNLPKFLENKAKYESLKDKAPEAINYEQIK
KDLAEELTFDIDYKTSEVNQRVFSLDEVFEIANFNNYLNQSGITKFNTIIGGKFVNGEN
TKRKGINEYINLYSQQINDKTLKKYKMSVLFKQILSDTESKSFVIDKLEDDSDVVTTM
QSFYEQIAAFKTVEEKSIKETLSLLFDDLKAQKLDLSKIYFKNDKSLTDLSQQVFDDY
SVIGTAVLEYITQQIAPKNLDNPSKKEQELIAKKTEKAKYLSLETIKLALEEFNKHRDI
DKQCRFEEILANFAAIPMIFDEIAQNKDNLAQISIKYQNQGKKDLLQASAEDDVKAIK
DLLDQTNNLLHKLKIFHISQSEDKANILDKDEHFYLVFEECYFELANIVPLYNKIRNYI
TQKPYSDEKFKLNFENSTLANGWDKNKEPDNTAILFIKDDKYYLGVMNKKNNKIFD
DKAIKENKGEGYKKIVYKLLPGANKMLPKVFFSAKSIKFYNPSEDILRIRNHSTHTKN
GSPQKGYEKFEFNIEDCRKFIDFYKQSISKHPEWKDFGFRFSDTQRYNSIDEFYREVE
NQGYKLTFENISESYIDSVVNQGKLYLFQIYNKDFSAYSKGRPNLHTLYWKALFDER
NLQDVVYKLNGEAELFYRKQSIPKKITHPAKEAIANKNKDNPKKESVFEYDLIKDKR
FTEDKFFFHCPITINFKSSGANKFNDEINLLLKEKANDVHILSI**D**RGERHLAYYTLVDG
KGNIIKQDTFNIIGNDRMKTNYHDKLAAIEKDRDSARKDWKKINNIKEMKEGYLSQV
VHEIAKLVIEYNAIVVF**A**DLNFGFKRGRFKVEKQVYQKLEKMLIEKLNYLVFKDNEF
DKTGGVLRAYQLTAPFETFKKMGKQTGIIYYVPAGFTSKICPVTGFVNQLYPKYESV
SKSQEFFSKFDKICYNLDKGYFEFSFDYKNFGDKAAKGKWTIASFGSRLINFRNSDKN
HNWDTREVYPTKELEKLLKDYSIEYGHGECIKAAICGESDKKFFAKLTSVLNTILQM
RNSKTGTELDYLISPVADVNGNFFDSRQAPKNMPQDA**D**ANGAYHIGLKGLMLLGRI
KNNQEGKKLNLVIKNEEYFEFVQNRNN

[00685] *Francisella novicida* Cpf1 D1255A (D917, E1006, and A1255 are bolded and underlined)

MSIYQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAKDYKKAKQIIDKYH
QFFIEEILSSVCISEDLLQNYSDVYFKLKKSDDDNLQKDFKSAKDTIKKQISEYIKDSE
KFKNLFNQNLIDAKKGQESDLILWLKQSKDNGIELFKANSDITDIDEALEIIKSFKGWT

TYFKGFHENRKNVYSSNDIPTSIIYRIVDDNLPKFLENKAKYESLKDKAPEAINYEQIK
KDLAEELTFDIDYKTSEVNQRVFSLDEVFEIANFNNYLNQSGITKFNTIIGGKFVNGEN
TKRKGINEYINLYSQQINDKTLKKYKMSVLFKQILSDTESKSFVIDKLEDDSDVVTTM
QSFYEQIAAFKTVEEKSIKETLSLLFDDLKAQKLDLSKIYFKNDKSLTDLSQQVFDDY
SVIGTAVLEYITQQIAPKNLDNPSKKEQELIAKKTEKAKYLSLETIKLALEEFNKHRDI
DKQCRFEEILANFAAIPMIFDEIAQNKDNLAQISIKYQNQGKKDLLQASAEDDVKAIK
DLLDQTNNLLHKLKIFHISQSEDKANILDKDEHFYLVFEECYFELANIVPLYNKIRNYI
TQKPYSDEKFKLNFENSTLANGWDKNKEPDNTAILFIKDDKYYLGVMNKKNNKIFD
DKAIKENKGEGYKKIVYKLLPGANKMLPKVFFSAKSIKFYNPSEDILRIRNHSTHTKN
GSPQKGYEKFEFNIEDCRKFIDFYKQSISKHPEWKDFGFRFSDTQRYNSIDEFYREVE
NQGYKLTFENISESYIDSVVNQGKLYLFQIYNKDFSAYSKGRPNLHTLYWKALFDER
NLQDVVYKLNGEAELFYRKQSIPKKITHPAKEAIANKNKDNPKKESVFEYDLIKDKR
FTEDKFFFHCPITINFKSSGANKFNDEINLLLKEKANDVHILSI**D**RGERHLAYYTLVDG
KGNIIKQDTFNIIGNDRMKTNYHDKLAAIEKDRDSARKDWKKINNIKEMKEGYLSQV
VHEIAKLVIEYNAIVVF**E**DLNFGFKRGRFKVEKQVYQKLEKMLIEKLNYLVFKDNEF
DKTGGVLRAYQLTAPFETFKKMGKQTGIIYYVPAGFTSKICPVTGFVNQLYPKYESV
SKSQEFFSKFDKICYNLDKGYFEFSFDYKNFGDKAAKGKWTIASFGSRLINFRNSDKN
HNWDTREVYPTKELEKLLKDYSIEYGHGECIKAAICGESDKKFFAKLTSVLNTILQM
RNSKTGTELDYLISPVADVNGNFFDSRQAPKNMPQDA**A**ANGAYHIGLKGLMLLGRI
KNNQEGKKLNLVIKNEEYFEFVQNRNN

[00686] *Francisella novicida* Cpf1 D917A/E1006A (A917, A1006, and D1255 are bolded and underlined)

MSIYQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAKDYKKAKQIIDKYH
QFFIEEILSSVCISEDLLQNYSDVYFKLKKSDDDNLQKDFKSAKDTIKKQISEYIKDSE
KFKNLFNQNLIDAKKGQESDLILWLKQSKDNGIELFKANSDITDIDEALEIIKSFKGWT
TYFKGFHENRKNVYSSNDIPTSIIYRIVDDNLPKFLENKAKYESLKDKAPEAINYEQIK
KDLAEELTFDIDYKTSEVNQRVFSLDEVFEIANFNNYLNQSGITKFNTIIGGKFVNGEN
TKRKGINEYINLYSQQINDKTLKKYKMSVLFKQILSDTESKSFVIDKLEDDSDVVTTM
QSFYEQIAAFKTVEEKSIKETLSLLFDDLKAQKLDLSKIYFKNDKSLTDLSQQVFDDY
SVIGTAVLEYITQQIAPKNLDNPSKKEQELIAKKTEKAKYLSLETIKLALEEFNKHRDI
DKQCRFEEILANFAAIPMIFDEIAQNKDNLAQISIKYQNQGKKDLLQASAEDDVKAIK
DLLDQTNNLLHKLKIFHISQSEDKANILDKDEHFYLVFEECYFELANIVPLYNKIRNYI
TQKPYSDEKFKLNFENSTLANGWDKNKEPDNTAILFIKDDKYYLGVMNKKNNKIFD

DKAIKENKGEGYKKIVYKLLPGANKMLPKVFFSAKSIKFYNPSEDILRIRNHSTHTKN
GSPQKGYEKFEFNIEDCRKFIDFYKQSISKHPEWKDFGFRFSDTQRYNSIDEFYREVE
NQGYKLTFENISESYIDSVVNQGKLYLFQIYNKDFSAYSKGRPNLHTLYWKALFDER
NLQDVVYKLNGEAELFYRKQSIPKKITHPAKEAIANKNKDNPKKESVFEYDLIKDKR
FTEDKFFFHCPITINFKSSGANKFNDEINLLLKEKANDVHILSI**A**RGERHLAYYTLVDG
KGNIIKQDTFNIIGNDRMKTNYHDKLAAIEKDRDSARKDWKKINNIKEMKEGYLSQV
VHEIAKLVIEYNAIVVF**A**DLNFGFKRGRFKVEKQVYQKLEKMLIEKLNYLVFKDNEF
DKTGGVLRAYQLTAPFETFKKMGKQTGIIYYVPAGFTSKICPVTGFVNQLYPKYESV
SKSQEFFSKFDKICYNLDKGYFEFSFDYKNFGDKAAKGKWTIASFGSRLINFRNSDKN
HNWDTREVYPTKELEKLLKDYSIEYGHGECIKAAICGESDKKFFAKLTSVLNTILQM
RNSKTGTELDYLISPVADVNGNFFDSRQAPKNMPQDA**D**ANGAYHIGLKGLMLLGRI
KNNQEGKKLNLVIKNEEYFEFVQNRNN

[00687] *Francisella novicida* Cpf1 D917A/D1255A (A917, E1006, and A1255 are bolded and underlined)

MSIYQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAKDYKKAKQIIDKYH
QFFIEEILSSVCISEDLLQNYSDVYFKLKKSDDDNLQKDFKSAKDTIKKQISEYIKDSE
KFKNLFNQNLIDAKKGQESDLILWLKQSKDNGIELFKANSDITDIDEALEIIKSFKGWT
TYFKGFHENRKNVYSSNDIPTSIIYRIVDDNLPKFLENKAKYESLKDKAPEAINYEQIK
KDLAEELTFDIDYKTSEVNQRVFSLDEVFEIANFNNYLNQSGITKFNTIIGGKFVNGEN
TKRKGINEYINLYSQQINDKTLKKYKMSVLFKQILSDTESKSFVIDKLEDDSDVVTTM
QSFYEQIAAFKTVEEKSIKETLSLLFDDLKAQKLDLSKIYFKNDKSLTDLSQQVFDDY
SVIGTAVLEYITQQIAPKNLDNPSKKEQELIAKKTEKAKYLSLETIKLALEEFNKHRDI
DKQCRFEEILANFAAIPMIFDEIAQNKDNLAQISIKYQNQGKKDLLQASAEDDVKAIK
DLLDQTNNLLHKLKIFHISQSEDKANILDKDEHFYLVFEECYFELANIVPLYNKIRNYI
TQKPYSDEKFKLNFENSTLANGWDKNKEPDNTAILFIKDDKYYLGVMNKKNNKIFD
DKAIKENKGEGYKKIVYKLLPGANKMLPKVFFSAKSIKFYNPSEDILRIRNHSTHTKN
GSPQKGYEKFEFNIEDCRKFIDFYKQSISKHPEWKDFGFRFSDTQRYNSIDEFYREVE
NQGYKLTFENISESYIDSVVNQGKLYLFQIYNKDFSAYSKGRPNLHTLYWKALFDER
NLQDVVYKLNGEAELFYRKQSIPKKITHPAKEAIANKNKDNPKKESVFEYDLIKDKR
FTEDKFFFHCPITINFKSSGANKFNDEINLLLKEKANDVHILSI**A**RGERHLAYYTLVDG
KGNIIKQDTFNIIGNDRMKTNYHDKLAAIEKDRDSARKDWKKINNIKEMKEGYLSQV
VHEIAKLVIEYNAIVVF**E**DLNFGFKRGRFKVEKQVYQKLEKMLIEKLNYLVFKDNEF
DKTGGVLRAYQLTAPFETFKKMGKQTGIIYYVPAGFTSKICPVTGFVNQLYPKYESV

SKSQEFFSKFDKICYNLDKGYFEFSFDYKNFGDKAAKGKWTIASFGSRLINFRNSDKN
HNWDTREVYPTKELEKLLKDYSIEYGHGECIKAAICGESDKKFFAKLTSVLNTILQM
RNSKTGTELDYLISPVADVNGNFFDSRQAPKNMPQDA**A**ANGAYHIGLKGLMLLGRI
KNNQEGKKLNLVIKNEEYFEFVQNRNN

[00688] *Francisella novicida* Cpf1 E1006A/D1255A (D917, A1006, and A1255 are bolded and underlined)

MSIYQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAKDYKKAKQIIDKYH
QFFIEEILSSVCISEDLLQNYSDVYFKLKKSDDDNLQKDFKSAKDTIKKQISEYIKDSE
KFKNLFNQNLIDAKKGQESDLILWLKQSKDNGIELFKANSDITDIDEALEIIKSFKGWT
TYFKGFHENRKNVYSSNDIPTSIIYRIVDDNLPKFLENKAKYESLKDKAPEAINYEQIK
KDLAEELTFDIDYKTSEVNQRVFSLDEVFEIANFNNYLNQSGITKFNTIIGGKFVNGEN
TKRKGINEYINLYSQQINDKTLKKYKMSVLFKQILSDTESKSFVIDKLEDDSDVVTTM
QSFYEQIAAFKTVEEKSIKETLSLLFDDLKAQKLDLSKIYFKNDKSLTDLSQQVFDDY
SVIGTAVLEYITQQIAPKNLDNPSKKEQELIAKKTEKAKYLSLETIKLALEEFNKHRDI
DKQCRFEEILANFAAIPMIFDEIAQNKDNLAQISIKYQNQGKKDLLQASAEDDVKAIK
DLLDQTNNLLHKLKIFHISQSEDKANILDKDEHFYLVFEECYFELANIVPLYNKIRNYI
TQKPYSDEKFKLNFENSTLANGWDKNKEPDNTAILFIKDDKYYLGVMNKKNNKIFD
DKAIKENKGEGYKKIVYKLLPGANKMLPKVFFSAKSIKFYNPSEDILRIRNHSTHTKN
GSPQKGYEKFEFNIEDCRKFIDFYKQSISKHPEWKDFGFRFSDTQRYNSIDEFYREVE
NQGYKLTFENISESYIDSVVNQGKLYLFQIYNKDFSAYSKGRPNLHTLYWKALFDER
NLQDVVYKLNGEAELFYRKQSIPKKITHPAKEAIANKNKDNPKKESVFEYDLIKDKR
FTEDKFFFHCPITINFKSSGANKFNDEINLLLKEKANDVHILSI**D**RGERHLAYYTLVDG
KGNIIKQDTFNIIGNDRMKTNYHDKLAAIEKDRDSARKDWKKINNIKEMKEGYLSQV
VHEIAKLVIEYNAIVVF**A**DLNFGFKRGRFKVEKQVYQKLEKMLIEKLNYLVFKDNEF
DKTGGVLRAYQLTAPFETFKKMGKQTGIIYYVPAGFTSKICPVTGFVNQLYPKYESV
SKSQEFFSKFDKICYNLDKGYFEFSFDYKNFGDKAAKGKWTIASFGSRLINFRNSDKN
HNWDTREVYPTKELEKLLKDYSIEYGHGECIKAAICGESDKKFFAKLTSVLNTILQM
RNSKTGTELDYLISPVADVNGNFFDSRQAPKNMPQDA**A**ANGAYHIGLKGLMLLGRI
KNNQEGKKLNLVIKNEEYFEFVQNRNN

[00689] *Francisella novicida* Cpf1 D917A/E1006A/D1255A (A917, A1006, and A1255 are bolded and underlined)

MSIYQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAKDYKKAKQIIDKYH
QFFIEEILSSVCISEDLLQNYSDVYFKLKKSDDDNLQKDFKSAKDTIKKQISEYIKDSE
KFKNLFNQNLIDAKKGQESDLILWLKQSKDNGIELFKANSDITDIDEALEIIKSFKGWT
TYFKGFHENRKNVYSSNDIPTSIIYRIVDDNLPKFLENKAKYESLKDKAPEAINYEQIK
KDLAEELTFDIDYKTSEVNQRVFSLDEVFEIANFNNYLNQSGITKFNTIIGGKFVNGEN
TKRKGINEYINLYSQQINDKTLKKYKMSVLFKQILSDTESKSFVIDKLEDDSDVVTTM
QSFYEQIAAFKTVEEKSIKETLSLLFDDLKAQKLDLSKIYFKNDKSLTDLSQQVFDDY
SVIGTAVLEYITQQIAPKNLDNPSKKEQELIAKKTEKAKYLSLETIKLALEEFNKHRDI
DKQCRFEEILANFAAIPMIFDEIAQNKDNLAQISIKYQNQGKKDLLQASAEDDVKAIK
DLLDQTNNLLHKLKIFHISQSEDKANILDKDEHFYLVFEECYFELANIVPLYNKIRNYI
TQKPYSDEKFKLNFENSTLANGWDKNKEPDNTAILFIKDDKYYLGVMNKKNNKIFD
DKAIKENKGEGYKKIVYKLLPGANKMLPKVFFSAKSIKFYNPSEDILRIRNHSTHTKN
GSPQKGYEKFEFNIEDCRKFIDFYKQSISKHPEWKDFGFRFSDTQRYNSIDEFYREVE
NQGYKLTFENISESYIDSVVNQGKLYLFQIYNKDFSAYSKGRPNLHTLYWKALFDER
NLQDVVYKLNGEAELFYRKQSIPKKITHPAKEAIANKNKDNPKKESVFEYDLIKDKR
FTEDKFFFHCPITINFKSSGANKFNDEINLLLKEKANDVHILSIARGERHLAYYTLVDG
KGNIIKQDTFNIIGNDRMKTNYHDKLAAIEKDRDSARKDWKKINNIKEMKEGYLSQV
VHEIAKLVIEYNAIVVFADLNFGFKRGRFKVEKQVYQKLEKMLIEKLNYLVFKDNEF
DKTGGVLRAYQLTAPFETFKKMGKQTGIIYYVPAGFTSKICPVTGFVNQLYPKYESV
SKSQEFFSKFDKICYNLDKGYFEFSFDYKNFGDKAAKGKWTIASFGSRLINFRNSDKN
HNWDTREVYPTKELEKLLKDYSIEYGHGECIKAAICGESDKKFFAKLTSVLNTILQM
RNSKTGTELDYLISPVADVNGNFFDSRQAPKNMPQDAANGAYHIGLKGLMLLGRI
KNNQEGKKLNLVIKNEEYFEFVQNRNN

[00690] The Cas9 nuclease has two functional endonuclease domains: RuvC and HNH. Cas9 undergoes a conformational change upon target binding that positions the nuclease domains to cleave opposite strands of the target DNA. The end result of Cas9-mediated DNA cleavage is a double-strand break (DSB) within the target DNA (~3-4 nucleotides upstream of the PAM sequence). The resulting DSB is then repaired by one of two general repair pathways: (1) the efficient but error-prone non-homologous end joining (NHEJ) pathway; or (2) the less efficient but high-fidelity homology directed repair (HDR) pathway.

[00691] The "efficiency" of non-homologous end joining (NHEJ) and/or homology directed repair (HDR) can be calculated by any convenient method. For example, in some cases, efficiency can be expressed in terms of percentage of successful HDR. For example, a

surveyor nuclease assay can be used to generate cleavage products and the ratio of products to substrate can be used to calculate the percentage. For example, a surveyor nuclease enzyme can be used that directly cleaves DNA containing a newly integrated restriction sequence as the result of successful HDR. More cleaved substrate indicates a greater percent HDR (a greater efficiency of HDR). As an illustrative example, a fraction (percentage) of HDR can be calculated using the following equation [(cleavage products)/(substrate plus cleavage products)] (*e.g.*, (b+c)/(a+b+c), where "a" is the band intensity of DNA substrate and "b" and "c" are the cleavage products).

[00692] In some cases, efficiency can be expressed in terms of percentage of successful NHEJ. For example, a T7 endonuclease I assay can be used to generate cleavage products and the ratio of products to substrate can be used to calculate the percentage NHEJ. T7 endonuclease Icleaves mismatched heteroduplex DNA which arises from hybridization of wild-type and mutant DNA strands (NHEJ generates small random insertions or deletions (indels) at the site of the original break). More cleavage indicates a greater percent NHEJ (a greater efficiency of NHEJ). As an illustrative example, a fraction (percentage) of NHEJ can be calculated using the following equation: $(1-(1-(b+c)/(a+b+c))^{1/2})\times 100$, where "a" is the band intensity of DNA substrate and "b" and "c" are the cleavage products (Ran *et. al.*,2013 Sep. 12; 154(6):1380-9; and Ran *et al.*, Nat Protoc. 2013 Nov.; 8(11): 2281–2308).

[00693] The NHEJ repair pathway is the most active repair mechanism, and it frequently causes small nucleotide insertions or deletions (indels) at the DSB site. The randomness of NHEJ-mediated DSB repair has important practical implications, because a population of cells expressing Cas9 and a gRNA or a guide polynucleotide can result in a diverse array of mutations. In most cases, NHEJ gives rise to small indels in the target DNA that result in amino acid deletions, insertions, or frameshift mutations leading to premature stop codons within the open reading frame (ORF) of the targeted gene. The ideal end result is a loss-of-function mutation within the targeted gene.

[00694] While NHEJ-mediated DSB repair often disrupts the open reading frame of the gene, homology directed repair (HDR) can be used to generate specific nucleotide changes ranging from a single nucleotide change to large insertions like the addition of a fluorophore or tag.

[00695] In order to utilize HDR for gene editing, a DNA repair template containing the desired sequence can be delivered into the cell type of interest with the gRNA(s) and Cas9 or Cas9 nickase. The repair template can contain the desired edit as well as additional homologous sequence immediately upstream and downstream of the target (termed left &

right homology arms). The length of each homology arm can be dependent on the size of the change being introduced, with larger insertions requiring longer homology arms. The repair template can be a single-stranded oligonucleotide, double-stranded oligonucleotide, or a double-stranded DNA plasmid. The efficiency of HDR is generally low (<10% of modified alleles) even in cells that express Cas9, gRNA and an exogenous repair template. The efficiency of HDR can be enhanced by synchronizing the cells, since HDR takes place during the S and G2 phases of the cell cycle. Chemically or genetically inhibiting genes involved in NHEJ can also increase HDR frequency.

[00696] In some embodiments, Cas9 is a modified Cas9. A given gRNA targeting sequence can have additional sites throughout the genome where partial homology exists. These sites are called off-targets and need to be considered when designing a gRNA. In addition to optimizing gRNA design, CRISPR specificity can also be increased through modifications to Cas9. Cas9 generates double-strand breaks (DSBs) through the combined activity of two nuclease domains, RuvC and HNH. Cas9 nickase, a D10A mutant of SpCas9, retains one nuclease domain and generates a DNA nick rather than a DSB. The nickase system can also be combined with HDR-mediated gene editing for specific gene edits.

[00697] In some cases, Cas9 is a variant Cas9 protein. A variant Cas9 polypeptide has an amino acid sequence that is different by one amino acid (*e.g.*, has a deletion, insertion, substitution, fusion) when compared to the amino acid sequence of a wild type Cas9 protein. In some instances, the variant Cas9 polypeptide has an amino acid change (*e.g.*, deletion, insertion, or substitution) that reduces the nuclease activity of the Cas9 polypeptide. For example, in some instances, the variant Cas9 polypeptide has less than 50%, less than 40%, less than 30%, less than 20%, less than 10%, less than 5%, or less than 1% of the nuclease activity of the corresponding wild-type Cas9 protein. In some cases, the variant Cas9 protein has no substantial nuclease activity. When a subject Cas9 protein is a variant Cas9 protein that has no substantial nuclease activity, it can be referred to as "dCas9."

[00698] In some cases, a variant Cas9 protein has reduced nuclease activity. For example, a variant Cas9 protein exhibits less than about 20%, less than about 15%, less than about 10%, less than about 5%, less than about 1%, or less than about 0.1%, of the endonuclease activity of a wild-type Cas9 protein, *e.g.*, a wild-type Cas9 protein.

[00699] In some cases, a variant Cas9 protein can cleave the complementary strand of a guide target sequence but has reduced ability to cleave the non-complementary strand of a double stranded guide target sequence. For example, the variant Cas9 protein can have a mutation (amino acid substitution) that reduces the function of the RuvC domain. As a non-

limiting example, in some embodiments, a variant Cas9 protein has a D10A (aspartate to alanine at amino acid position 10) and can therefore cleave the complementary strand of a double stranded guide target sequence but has reduced ability to cleave the non-complementary strand of a double stranded guide target sequence (thus resulting in a single strand break (SSB) instead of a double strand break (DSB) when the variant Cas9 protein cleaves a double stranded target nucleic acid) (see, for example, Jinek *et al.*, Science. 2012 Aug. 17; 337(6096):816-21).

[00700] In some cases, a variant Cas9 protein can cleave the non-complementary strand of a double stranded guide target sequence but has reduced ability to cleave the complementary strand of the guide target sequence. For example, the variant Cas9 protein can have a mutation (amino acid substitution) that reduces the function of the HNH domain (RuvC/HNH/RuvC domain motifs). As a non-limiting example, in some embodiments, the variant Cas9 protein has an H840A (histidine to alanine at amino acid position 840) mutation and can therefore cleave the non-complementary strand of the guide target sequence but has reduced ability to cleave the complementary strand of the guide target sequence (thus resulting in a SSB instead of a DSB when the variant Cas9 protein cleaves a double stranded guide target sequence). Such a Cas9 protein has a reduced ability to cleave a guide target sequence (*e.g.*, a single stranded guide target sequence) but retains the ability to bind a guide target sequence (*e.g.*, a single stranded guide target sequence).

[00701] In some cases, a variant Cas9 protein has a reduced ability to cleave both the complementary and the non-complementary strands of a double stranded target DNA. As a non-limiting example, in some cases, the variant Cas9 protein harbors both the D10A and the H840A mutations such that the polypeptide has a reduced ability to cleave both the complementary and the non-complementary strands of a double stranded target DNA. Such a Cas9 protein has a reduced ability to cleave a target DNA (*e.g.*, a single stranded target DNA) but retains the ability to bind a target DNA (*e.g.*, a single stranded target DNA).

[00702] As another non-limiting example, in some cases, the variant Cas9 protein harbors W476A and W1126A mutations such that the polypeptide has a reduced ability to cleave a target DNA. Such a Cas9 protein has a reduced ability to cleave a target DNA (*e.g.*, a single stranded target DNA) but retains the ability to bind a target DNA (*e.g.*, a single stranded target DNA).

[00703] As another non-limiting example, in some cases, the variant Cas9 protein harbors P475A, W476A, N477A, D1125A, W1126A, and D1127A mutations such that the polypeptide has a reduced ability to cleave a target DNA. Such a Cas9 protein has a reduced

ability to cleave a target DNA (*e.g.*, a single stranded target DNA) but retains the ability to bind a target DNA (*e.g.*, a single stranded target DNA).

[00704] As another non-limiting example, in some cases, the variant Cas9 protein harbors H840A, W476A, and W1126A, mutations such that the polypeptide has a reduced ability to cleave a target DNA. Such a Cas9 protein has a reduced ability to cleave a target DNA (*e.g.*, a single stranded target DNA) but retains the ability to bind a target DNA (*e.g.*, a single stranded target DNA). As another non-limiting example, in some cases, the variant Cas9 protein harbors H840A, D10A, W476A, and W1126A, mutations such that the polypeptide has a reduced ability to cleave a target DNA. Such a Cas9 protein has a reduced ability to cleave a target DNA (*e.g.*, a single stranded target DNA) but retains the ability to bind a target DNA (*e.g.*, a single stranded target DNA). In some embodiments, the variant Cas9 has restored catalytic His residue at position 840 in the Cas9 HNH domain (A840H).

[00705] As another non-limiting example, in some cases, the variant Cas9 protein harbors, H840A, P475A, W476A, N477A, D1125A, W1126A, and D1127A mutations such that the polypeptide has a reduced ability to cleave a target DNA. Such a Cas9 protein has a reduced ability to cleave a target DNA (*e.g.*, a single stranded target DNA) but retains the ability to bind a target DNA (*e.g.*, a single stranded target DNA). As another non-limiting example, in some cases, the variant Cas9 protein harbors D10A, H840A, P475A, W476A, N477A, D1125A, W1126A, and D1127A mutations such that the polypeptide has a reduced ability to cleave a target DNA. Such a Cas9 protein has a reduced ability to cleave a target DNA (*e.g.*, a single stranded target DNA) but retains the ability to bind a target DNA (*e.g.*, a single stranded target DNA). In some cases, when a variant Cas9 protein harbors W476A and W1126A mutations or when the variant Cas9 protein harbors P475A, W476A, N477A, D1125A, W1126A, and D1127A mutations, the variant Cas9 protein does not bind efficiently to a PAM sequence. Thus, in some such cases, when such a variant Cas9 protein is used in a method of binding, the method does not require a PAM sequence. In other words, in some cases, when such a variant Cas9 protein is used in a method of binding, the method can include a guide RNA, but the method can be performed in the absence of a PAM sequence (and the specificity of binding is therefore provided by the targeting segment of the guide RNA). Other residues can be mutated to achieve the above effects (*i.e.*, inactivate one or the other nuclease portions). As non-limiting examples, residues D10, G12, G17, E762, H840, N854, N863, H982, H983, A984, D986, and/or A987 can be altered (*i.e.*, substituted). Also, mutations other than alanine substitutions are suitable.

[00706] In some embodiments, a variant Cas9 protein that has reduced catalytic activity (*e.g.*, when a Cas9 protein has a D10, G12, G17, E762, H840, N854, N863, H982, H983, A984, D986, and/or a A987 mutation, *e.g.*, D10A, G12A, G17A, E762A, H840A, N854A, N863A, H982A, H983A, A984A, and/or D986A), the variant Cas9 protein can still bind to target DNA in a site-specific manner (because it is still guided to a target DNA sequence by a guide RNA) as long as it retains the ability to interact with the guide RNA.

[00707] In some embodiments, the variant Cas protein can be spCas9, spCas9-VRQR, spCas9-VRER, xCas9 (sp), saCas9, saCas9-KKH, spCas9-MQKSER, spCas9-LRKIQK, or spCas9-LRVSQL.

[00708] Alternatives to *S. pyogenes* Cas9 can include RNA-guided endonucleases from the Cpf1 family that display cleavage activity in mammalian cells. CRISPR from *Prevotella* and *Francisella 1* (CRISPR/Cpf1) is a DNA-editing technology analogous to the CRISPR/Cas9 system. Cpf1 is an RNA-guided endonuclease of a class II CRISPR/Cas system. This acquired immune mechanism is found in *Prevotella* and *Francisella* bacteria. Cpf1 genes are associated with the CRISPR locus, coding for an endonuclease that use a guide RNA to find and cleave viral DNA. Cpf1 is a smaller and simpler endonuclease than Cas9, overcoming some of the CRISPR/Cas9 system limitations. Unlike Cas9 nucleases, the result of Cpf1-mediated DNA cleavage is a double-strand break with a short 3′ overhang. Cpf1's staggered cleavage pattern can open up the possibility of directional gene transfer, analogous to traditional restriction enzyme cloning, which can increase the efficiency of gene editing. Like the Cas9 variants and orthologues described above, Cpf1 can also expand the number of sites that can be targeted by CRISPR to AT-rich regions or AT-rich genomes that lack the NGG PAM sites favored by SpCas9. The Cpf1 locus contains a mixed alpha/beta domain, a RuvC-I followed by a helical region, a RuvC-II and a zinc finger-like domain. The Cpf1 protein has a RuvC-like endonuclease domain that is similar to the RuvC domain of Cas9. Furthermore, Cpf1 does not have a HNH endonuclease domain, and the N-terminal of Cpf1 does not have the alpha-helical recognition lobe of Cas9. Cpf1 CRISPR-Cas domain architecture shows that Cpf1 is functionally unique, being classified as Class 2, type V CRISPR system. The Cpf1 loci encode Cas1, Cas2 and Cas4 proteins more similar to types I and III than from type II systems. Functional Cpf1 doesn't need the trans-activating CRISPR RNA (tracrRNA), therefore, only CRISPR (crRNA) is required. This benefits genome editing because Cpf1 is not only smaller than Cas9, but also it has a smaller sgRNA molecule (proximately half as many nucleotides as Cas9). The Cpf1-crRNA complex cleaves target DNA or RNA by identification of a protospacer adjacent motif 5'-YTN-3' in contrast to the

G-rich PAM targeted by Cas9. After identification of PAM, Cpf1 introduces a sticky-end-like DNA double- stranded break of 4 or 5 nucleotides overhang.

Fusion proteins comprising two napDNAbp, a Deaminase Domain

[00709] Some aspects of the disclosure provide fusion proteins comprising a napDNAbp domain having nickase activity (e.g., nCas domain) and a catalytically inactive napDNAbp (e.g., dCas domain) and a nucleobase editor (e.g., adenosine deaminase domain, cytidine deaminase domain), where at least the napDNAbp domains are joined by a linker. It should be appreciated that the Cas domains may be any of the Cas domains or Cas proteins (*e.g.*, dCas9 and nCas9) provided herein. In some embodiments, any of the Cas domains, DNA binding protein domains, or Cas proteins include, without limitation, Cas9 (*e.g.*, dCas9 and nCas9), Cas12a/Cpf1, Cas12b/C2c1, Cas12c/C2c3, Cas12d/CasY, Cas12e/CasX, Cas12g, Cas12h, and Cas12i. One example of a programmable polynucleotide-binding protein that has different PAM specificity than Cas9 is Clustered Regularly Interspaced Short Palindromic Repeats from *Prevotella* and *Francisella*1 (Cpf1). Similar to Cas9, Cpf1 is also a class 2 CRISPR effector. For example, and without limitation, in some embodiments, the fusion protein comprises the structure, where the deaminase is adenosine deaminase or cytidine deaminase:

NH_2-[deaminase]-[nCas domain]-[dCas domain]-COOH;
NH_2-[deaminase]-[dCas domain]-[nCas domain]-COOH;
NH_2-[nCas domain]-[dCas domain]-[deaminase]-COOH;
NH_2-[dCas domain]-[nCas domain]-[deaminase]-COOH;
NH_2-[nCas domain]-[deaminase]-[dCas domain]-COOH;
NH_2-[dCas domain]-[deaminase]-[nCas domain]-COOH;

[00710] In some embodiments, the "-" used in the general architecture above indicates the presence of an optional linker. In some embodiments, the deaminase and a napDNAbp (e.g., Cas domain) are not joined by a linker sequence, but are directly fused. In some embodiments, a linker is present between the deaminase domain and the napDNAbp. In some embodiments, the deaminase or other nucleobase editor is directly fused to dCas and a linker joins dCas and nCas9. In some embodiments, the deaminase and the napDNAbps are fused via any of the linkers provided herein. For example, in some embodiments the deaminase and the napDNAbp are fused via any of the linkers provided below in the section entitled "Linkers". In some embodiments, the dCas domain and the deaminase are

immediately adjacent and the nCas domain is joined to these domains (either 5' or 3') via a linker.

Protospacer Adjacent Motif

[00711] The term "protospacer adjacent motif (PAM)" or PAM-like motif refers to a 2-6 base pair DNA sequence immediately following the DNA sequence targeted by the Cas9 nuclease in the CRISPR bacterial adaptive immune system. In some embodiments, the PAM can be a 5' PAM (*i.e.*, located upstream of the 5' end of the protospacer). In other embodiments, the PAM can be a 3' PAM (*i.e.*, located downstream of the 5' end of the protospacer).

[00712] The PAM sequence is essential for target binding, but the exact sequence depends on a type of Cas protein.

[00713] A base editor provided herein can comprise a CRISPR protein-derived domain that is capable of binding a nucleotide sequence that contains a canonical or non-canonical protospacer adjacent motif (PAM) sequence. A PAM site is a nucleotide sequence in proximity to a target polynucleotide sequence. Some aspects of the disclosure provide for base editors comprising all or a portion of CRISPR proteins that have different PAM specificities. For example, typically Cas9 proteins, such as Cas9 from *S. pyogenes* (spCas9), require a canonical NGG PAM sequence to bind a particular nucleic acid region, where the "N" in "NGG" is adenine (A), thymine (T), guanine (G), or cytosine (C), and the G is guanine. A PAM can be CRISPR protein-specific and can be different between different base editors comprising different CRISPR protein-derived domains. A PAM can be 5' or 3' of a target sequence. A PAM can be upstream or downstream of a target sequence. A PAM can be 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more nucleotides in length. Often, a PAM is between 2-6 nucleotides in length. Several PAM variants are described in Table 1.

[00714] In some embodiments, the SpCas9 has specificity for PAM nucleic acid sequence 5'-NGC-3' or 5'-NGG-3'. In various embodiments of the above aspects, the SpCas9 is a Cas9 or Cas9 variant listed in Table 1. In various embodiments of the above aspects, the modified SpCas9 is spCas9-MQKFRAER. In some embodiments, the variant Cas protein can be spCas9, spCas9-VRQR, spCas9-VRER, xCas9 (sp), saCas9, saCas9-KKH, SpCas9-MQKFRAER, spCas9-MQKSER, spCas9-LRKIQK, or spCas9-LRVSQL. In one specific embodiment, a modified SpCas9 including amino acid substitutions D1135M, S1136Q,

G1218K, E1219F, A1322R, D1332A, R1335E, and T1337R (SpCas9-MQKFRAER) and having specificity for the altered PAM 5'-NGC-3' is used.

[00715] In some embodiments, the PAM is NGT. In some embodiments, the NGT PAM is a variant. In some embodiments, the NGT PAM variant is created through targeted mutations at one or more residues 1335, 1337, 1135, 1136, 1218, and/or 1219. In some embodiments, the NGT PAM variant is created through targeted mutations at one or more residues 1219, 1335, 1337, 1218. In some embodiments, the NGT PAM variant is created through targeted mutations at one or more residues 1135, 1136, 1218, 1219, and 1335. In some embodiments, the NGT PAM variant is selected from the set of targeted mutations provided in Tables 4 and 5 below.

Table 4: NGT PAM Variant Mutations at residues 1219, 1335, 1337, 1218

Variant	E1219V	R1335Q	T1337	G1218
1	F	V	T	
2	F	V	R	
3	F	V	Q	
4	F	V	L	
5	F	V	T	R
6	F	V	R	R
7	F	V	Q	R
8	F	V	L	R
9	L	L	T	
10	L	L	R	
11	L	L	Q	
12	L	L	L	
13	F	I	T	
14	F	I	R	
15	F	I	Q	
16	F	I	L	
17	F	G	C	
18	H	L	N	
19	F	G	C	A
20	H	L	N	V
21	L	A	W	
22	L	A	F	
23	L	A	Y	
24	I	A	W	
25	I	A	F	
26	I	A	Y	

Table 5: NGT PAM Variant Mutations at residues 1135, 1136, 1218, 1219, and 1335

Variant	D1135L	S1136R	G1218S	E1219V	R1335Q
27	G				
28	V				
29	I				
30		A			
31		W			
32		H			
33		K			
34			K		
35			R		
36			Q		
37			T		
38			N		
39				I	
40				A	
41				N	
42				Q	
43				G	
44				L	
45				S	
46				T	
47					L
48					I
49					V
50					N
51					S
52					T
53					F
54					Y
55	N1286Q	I1331F			

[00716] In some embodiments, the NGT PAM variant is selected from variant 5, 7, 28, 31, or 36 in Tables 2 and 3. In some embodiments, the variants have improved NGT PAM recognition.

[00717] In some embodiments, the NGT PAM variants have mutations at residues 1219, 1335, 1337, and/or 1218. In some embodiments, the NGT PAM variant is selected with mutations for improved recognition from the variants provided in Table 6 below.

Table 6: NGT PAM Variant Mutations at residues 1219, 1335, 1337, and 1218

Variant	E1219V	R1335Q	T1337	G1218
1	F	V	T	
2	F	V	R	
3	F	V	Q	
4	F	V	L	

5	F	V	T	R
6	F	V	R	R
7	F	V	Q	R
8	F	V	L	R

[00718] In some embodiments, the NGT PAM is selected from the variants provided in Table 7 below.

Table 7. NGT PAM variants

	NGTN variant	D1135	S1136	G1218	E1219	A1322R	R1335	T1337
Variant 1	LRKIQK	L	R	K	I	-	Q	K
Variant 2	LRSVQK	L	R	S	V	-	Q	K
Variant 3	LRSVQL	L	R	S	V	-	Q	L
Variant 4	LRKIRQK	L	R	K	I	R	Q	K
Variant 5	LRSVRQK	L	R	S	V	R	Q	K
Variant 6	LRSVRQL	L	R	S	V	R	Q	L

[00719] In some embodiments, the Cas9 domain is a Cas9 domain from *Streptococcus pyogenes* (SpCas9). In some embodiments, the SpCas9 domain is a nuclease active SpCas9, a nuclease inactive SpCas9 (SpCas9d), or a SpCas9 nickase (SpCas9n). In some embodiments, the SpCas9 comprises a D9X mutation, or a corresponding mutation in any of the amino acid sequences provided herein may be fused with any of the cytidine deaminases or adenosine deaminases provided herein

[00720] In some embodiments, the SpCas9 domain comprises one or more of a D1135X, a R1335X, and a T1336X mutation, or a corresponding mutation in any of the amino acid sequences provided herein, wherein X is any amino acid. In some embodiments, the SpCas9 domain comprises one or more of a D1135E, R1335Q, and T1336R mutation, or a corresponding mutation in any of the amino acid sequences provided herein. In some embodiments, the SpCas9 domain comprises a D1135E, a R1335Q, and a T1336R mutation, or corresponding mutations in any of the amino acid sequences provided herein. In some embodiments, the SpCas9 domain comprises one or more of a D1135X, a R1335X, and a T1336X mutation, or a corresponding mutation in any of the amino acid sequences provided herein, wherein X is any amino acid. In some embodiments, the SpCas9 domain comprises one or more of a D1135V, a R1335Q, and a T1336R mutation, or a corresponding mutation in any of the amino acid sequences provided herein. In some embodiments, the SpCas9 domain comprises a D1135V, a R1335Q, and a T1336R mutation, or corresponding mutations in any of the amino acid sequences provided herein. In some embodiments, the SpCas9 domain comprises one or more of a D1135X, a G1217X, a R1335X, and a T1336X mutation, or a corresponding mutation in any of the amino acid sequences provided herein,

wherein X is any amino acid. In some embodiments, the SpCas9 domain comprises one or more of a D1135V, a G1217R, a R1335Q, and a T1336R mutation, or a corresponding mutation in any of the amino acid sequences provided herein. In some embodiments, the SpCas9 domain comprises a D1135V, a G1217R, a R1335Q, and a T1336R mutation, or corresponding mutations in any of the amino acid sequences provided herein.

[00721] In some embodiments, the Cas9 domains of any of the fusion proteins provided herein comprises an amino acid sequence that is at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5% identical to a Cas9 polypeptide described herein. In some embodiments, the Cas9 domains of any of the fusion proteins provided herein comprises the amino acid sequence of any Cas9 polypeptide described herein. In some embodiments, the Cas9 domains of any of the fusion proteins provided herein consists of the amino acid sequence of any Cas9 polypeptide described herein.

[00722] In some examples, a PAM recognized by a CRISPR protein-derived domain of a base editor disclosed herein can be provided to a cell on a separate oligonucleotide to an insert (*e.g.*, an AAV insert) encoding the base editor. In such embodiments, providing PAM on a separate oligonucleotide can allow cleavage of a target sequence that otherwise would not be able to be cleaved, because no adjacent PAM is present on the same polynucleotide as the target sequence.

[00723] In an embodiment, *S. pyogenes* Cas9 (SpCas9) can be used as a CRISPR endonuclease for genome engineering. However, others can be used. In some embodiments, a different endonuclease can be used to target certain genomic targets. In some embodiments, synthetic SpCas9-derived variants with non-NGG PAM sequences can be used. Additionally, other Cas9 orthologues from various species have been identified and these "non-SpCas9s" can bind a variety of PAM sequences that can also be useful for the present disclosure. For example, the relatively large size of SpCas9 (approximately 4 kilobase (kb) coding sequence) can lead to plasmids carrying the SpCas9 cDNA that cannot be efficiently expressed in a cell. Conversely, the coding sequence for *Staphylococcus aureus* Cas9 (SaCas9) is approximately 1 kilobase shorter than SpCas9, possibly allowing it to be efficiently expressed in a cell. Similar to SpCas9, the SaCas9 endonuclease is capable of modifying target genes in mammalian cells *in vitro* and in mice *in vivo*. In some embodiments, a Cas protein can target a different PAM sequence. In some embodiments, a target gene can be adjacent to a Cas9 PAM, 5'-NGG, for example. In other embodiments, other Cas9 orthologs can have different PAM requirements. For example, other PAMs such

as those of *S. thermophilus* (5'-NNAGAA for CRISPR1 and 5'-NGGNG for CRISPR3) and *Neisseria meningiditis* (5'-NNNNGATT) can also be found adjacent to a target gene.

[00724] In some embodiments, for a *S. pyogenes* system, a target gene sequence can precede (*i.e.*, be 5' to) a 5'-NGG PAM, and a 20-nt guide RNA sequence can base pair with an opposite strand to mediate a Cas9 cleavage adjacent to a PAM. In some embodiments, an adjacent cut can be or can be about 3 base pairs upstream of a PAM. In some embodiments, an adjacent cut can be or can be about 10 base pairs upstream of a PAM. In some embodiments, an adjacent cut can be or can be about 0-20 base pairs upstream of a PAM. For example, an adjacent cut can be next to, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, or 30 base pairs upstream of a PAM. An adjacent cut can also be downstream of a PAM by 1 to 30 base pairs. The sequences of exemplary SpCas9 proteins capable of binding a PAM sequence follow:

[00725] The amino acid sequence of an exemplary PAM-binding SpCas9 is as follows:

MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGE
TAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHE
RHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEG
DLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLP
GEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYA
DLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPE
KYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQ
RTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRF
AWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFT
VYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECF
DSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEE
RLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFAN
RNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDEL
VKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENT
QLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRS
DKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAG
FIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFY
KVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEI
GKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL
SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVL
VVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYS

LFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLF
VEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTN
LGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD.

[00726] The amino acid sequence of an exemplary PAM-binding SpCas9n is as follows:

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGE
TAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHE
RHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEG
DLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLP
GEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYA
DLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPE
KYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQ
RTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRF
AWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFT
VYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECF
DSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEE
RLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFAN
RNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDEL
VKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENT
QLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRS
DKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAG
FIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFY
KVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEI
GKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL
SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVL
VVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYS
LFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLF
VEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTN
LGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD.

[00727] The amino acid sequence of an exemplary PAM-binding SpEQR Cas9 is as follows:

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGE
TAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFVEEDKKHER
HPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGD
LNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPG

EKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYAD
LFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEK
YKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRT
FDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFA
WMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTV
YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFD
SVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERL
KTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRN
FMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVK
VMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQL
QNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDK
NRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIK
RQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKV
REINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK
ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSM
PQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGF**E**SPTVAYSVLVV
AKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFE
LENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQ
HKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGA
PAAFKYFDTTIDRK**Q**Y**R**STKEVLDATLIHQSITGLYETRIDLSQLGGD.

[00728] In this sequence, residues E1135, Q1335 and R1337, which can be mutated from D1135, R1335, and T1337 to yield a SpEQR Cas9, are underlined and in bold.

[00729] The amino acid sequence of an exemplary PAM-binding SpVQR Cas9 is as follows:

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGE
TAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHE
RHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEG
DLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLP
GEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYA
DLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPE
KYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQ
RTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRF
AWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFT
VYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECF

DSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEE
RLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFAN
RNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDEL
VKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENT
QLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRS
DKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAG
FIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFY
KVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEI
GKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL
SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGF**V**SPTVAYSVL
VVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYS
LFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLF
VEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTN
LGAPAAFKYFDTTIDRK**Q**Y**R**STKEVLDATLIHQSITGLYETRIDLSQLGGD.

[00730] In this sequence, residues V1135, Q1335, and R1336, which can be mutated from D1135, R1335, and T1336 to yield a SpVQR Cas9, are underlined and in bold.

[00731] The amino acid sequence of an exemplary PAM-binding SpVRER Cas9 is as follows:

[00732] MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGAL
LFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEE
DKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRG
HFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLEN
LIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQI
GDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVR
QQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDL
LRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLAR
GNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLL
YEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFK
KIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDRE
MIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSD
GFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKV
VDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHP
VENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKV
LTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSEL

DKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRK
DFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIA
KSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFAT
VRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGF**V**SPTV
AYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIK
LPKYSLFELENGRKRMLASA**R**ELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNE
QKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHL
FTLTNLGAPAAFKYFDTTIDRK**E**Y**R**STKEVLDATLIHQSITGLYETRIDLSQLGGD.

[00733] In some embodiments, the Cas9 domain is a recombinant Cas9 domain. In some embodiments, the recombinant Cas9 domain is a SpyMacCas9 domain. In some embodiments, the SpyMacCas9 domain is a nuclease active SpyMacCas9, a nuclease inactive SpyMacCas9 (SpyMacCas9d), or a SpyMacCas9 nickase (SpyMacCas9n). In some embodiments, the SaCas9 domain, the SaCas9d domain, or the SaCas9n domain can bind to a nucleic acid sequence having a non-canonical PAM. In some embodiments, the SpyMacCas9 domain, the SpCas9d domain, or the SpCas9n domain can bind to a nucleic acid sequence having a NAA PAM sequence.

Exemplary SpyMacCas9

MDKKYSIGLDIGTNSVGWAVITDDYKVPSKKFKVLGNTDRHSIKKNLIGALLFGSGE
TAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHE
RHPIFGNIVDEVAYHEKYPTIYHLRKKLADSTDKADLRLIYLALAHMIKFRGHFLIEG
DLNPDNSDVDKLFIQLVQIYNQLFEENPINASRVDAKAILSARLSKSRRLENLIAQLPG
EKRNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYAD
LFLAAKNLSDAILLSDILRVNSEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEK
YKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRT
FDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFA
WMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTV
YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFD
SVEISGVEDRFNASLGAYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDRGMIEER
LKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANR
NFMQLIHDDSLTFKEDIQKAQVSGQGHSLHEQIANLAGSPAIKKGILQTVKIVDELVK
VMGHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQ
NEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFIKDDSIDNKVLTRSDKNR
GKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQ
LVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREI

NNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKAT
AKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQ
VNIVKKTEIQTVGQNGGLFDDNPKSPLEVTPSKLVPLKKELNPKKYGGYQKPTTAYP
VLLITDTKQLIPISVMNKKQFEQNPVKFLRDRGYQQVGKNDFIKLPKYTLVDIGDGIK
RLWASSKEIHKGNQLVVSKKSQILLYHAHHLDSDLSNDYLQNHNQQFDVLFNEIISFS
KKCKLGKEHIQKIENVYSNKKNSASIEELAESFIKLLGFTQLGATSPFNFLGVKLNQK
QYKGKKDYILPCTEGTLIRQSITGLYETRVDLSKIGED.

[00734] In some cases, a variant Cas9 protein harbors, H840A, P475A, W476A, N477A, D1125A, W1126A, and D1218A mutations such that the polypeptide has a reduced ability to cleave a target DNA or RNA. Such a Cas9 protein has a reduced ability to cleave a target DNA (*e.g.*, a single stranded target DNA) but retains the ability to bind a target DNA (*e.g.*, a single stranded target DNA). As another non-limiting example, in some cases, the variant Cas9 protein harbors D10A, H840A, P475A, W476A, N477A, D1125A, W1126A, and D1218A mutations such that the polypeptide has a reduced ability to cleave a target DNA. Such a Cas9 protein has a reduced ability to cleave a target DNA (*e.g.*, a single stranded target DNA) but retains the ability to bind a target DNA (*e.g.*, a single stranded target DNA). In some cases, when a variant Cas9 protein harbors W476A and W1126A mutations or when the variant Cas9 protein harbors P475A, W476A, N477A, D1125A, W1126A, and D1218A mutations, the variant Cas9 protein does not bind efficiently to a PAM sequence. Thus, in some such cases, when such a variant Cas9 protein is used in a method of binding, the method does not require a PAM sequence. In other words, in some cases, when such a variant Cas9 protein is used in a method of binding, the method can include a guide RNA, but the method can be performed in the absence of a PAM sequence (and the specificity of binding is therefore provided by the targeting segment of the guide RNA). Other residues can be mutated to achieve the above effects (*i.e.*, inactivate one or the other nuclease portions). As non-limiting examples, residues D10, G12, G17, E762, H840, N854, N863, H982, H983, A984, D986, and/or A987 can be altered (*i.e.*, substituted). Also, mutations other than alanine substitutions are suitable. In some embodiments, a CRISPR protein-derived domain of a base editor can comprise all or a portion of a Cas9 protein with a canonical PAM sequence (NGG). In other embodiments, a Cas9-derived domain of a base editor can employ a non-canonical PAM sequence. Such sequences have been described in the art and would be apparent to the skilled artisan. For example, Cas9 domains that bind non-canonical PAM sequences have been described in Kleinstiver, B. P., *et al.*, "Engineered CRISPR-Cas9 nucleases with altered PAM specificities" Nature 523, 481-485 (2015); and

Kleinstiver, B. P., *et al.*, "Broadening the targeting range of Staphylococcus aureus CRISPR-Cas9 by modifying PAM recognition" Nature Biotechnology 33, 1293-1298 (2015); the entire contents of each are hereby incorporated by reference.

[00735] In some embodiments, the Cas9 domain may be replaced with a guide nucleotide sequence-programmable DNA-binding protein domain that has no requirements for a PAM sequence.

[00736] In some embodiments, the nucleic acid programmable DNA binding protein (napDNAbp) is a single effector of a microbial CRISPR-Cas system. Single effectors of microbial CRISPR-Cas systems include, without limitation, Cas9, Cpf1, Cas12b/C2c1, and Cas12c/C2c3. Typically, microbial CRISPR-Cas systems are divided into Class 1 and Class 2 systems. Class 1 systems have multisubunit effector complexes, while Class 2 systems have a single protein effector. For example, Cas9 and Cpf1 are Class 2 effectors. In addition to Cas9 and Cpf1, three distinct Class 2 CRISPR-Cas systems (Cas12b/C2c1 and Cas12c/C2c3) have been described by Shmakov et al., "Discovery and Functional Characterization of Diverse Class 2 CRISPR Cas Systems", *Mol. Cell,* 2015 Nov. 5; 60(3): 385-397, the entire contents of which is hereby incorporated by reference. Effectors of two of the systems, Cas12b/C2c1 and Cas12c/C2c3, contain RuvC-like endonuclease domains related to Cpf1. A third system, contains an effector with two predicated HEPN RNase domains. Production of mature CRISPR RNA is tracrRNA-independent, unlike production of CRISPR RNA by Cas12b/C2c1. Cas12b/C2c1 depends on both CRISPR RNA and tracrRNA for DNA cleavage.

[00737] The crystal structure of *Alicyclobaccillus acidoterrastris* Cas12b/C2c1 (AacC2c1) has been reported in complex with a chimeric single-molecule guide RNA (sgRNA). See e.g., Liu et al., "C2c1-sgRNA Complex Structure Reveals RNA-Guided DNA Cleavage Mechanism", *Mol. Cell,* 2017 Jan. 19; 65(2):310-322, the entire contents of which are hereby incorporated by reference. The crystal structure has also been reported in *Alicyclobacillus acidoterrestris* Cas12b/C2c1 bound to target DNAs as ternary complexes. See e.g., Yang et al., "PAM-dependent Target DNA Recognition and Cleavage by C2C1 CRISPR-Cas endonuclease", *Cell,* 2016 Dec. 15; 167(7):1814-1828, the entire contents of which are hereby incorporated by reference. Catalytically competent conformations of AacC2c1, both with target and non-target DNA strands, have been captured independently positioned within a single RuvC catalytic pocket, with Cas12b/C2c1-mediated cleavage resulting in a staggered seven-nucleotide break of target DNA. Structural comparisons between Cas12b/C2c1 ternary

complexes and previously identified Cas9 and Cpf1 counterparts demonstrate the diversity of mechanisms used by CRISPR-Cas9 systems.

[00738] In some embodiments, the nucleic acid programmable DNA binding protein (napDNAbp) of any of the fusion proteins provided herein may be a Cas12b/C2c1, or a Cas12c/C2c3 protein. In some embodiments, the napDNAbp is a Cas12b/C2c1 protein. In some embodiments, the napDNAbp is a Cas12c/C2c3 protein. In some embodiments, the napDNAbp comprises an amino acid sequence that is at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at ease 99.5% identical to a naturally-occurring Cas12b/C2c1 or Cas12c/C2c3 protein. In some embodiments, the napDNAbp is a naturally-occurring Cas12b/C2c1 or Cas12c/C2c3 protein. In some embodiments, the napDNAbp comprises an amino acid sequence that is at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at ease 99.5% identical to any one of the napDNAbp sequences provided herein. It should be appreciated that Cas12b/C2c1 or Cas12c/C2c3 from other bacterial species may also be used in accordance with the present disclosure. CRISPR-Cas12b is described, for example, by Teng et al., Cell Discovery (2018) 4:63, which is incorporated therein by reference in its entirety.

[00739] Cas12b/C2c1 (uniprot.org/uniprot/T0D7A2#2)

[00740] sp|T0D7A2|C2C1_ALIAG CRISPR-associated endo- nuclease C2c1 OS = *Alicyclobacillus acido- terrestris* (strain ATCC 49025 / DSM 3922/ CIP 106132 / NCIMB 13137/GD3B) GN=c2c1 PE=1 SV=1

MAVKSIKVKLRLDDMPEIRAGLWKLHKEVNAGVRYYTEWLSLLRQENLYRRSPNG
DGEQECDKTAEECKAELLERLRARQVENGHRGPAGSDDELLQLARQLYELLVPQAI
GAKGDAQQIARKFLSPLADKDAVGGLGIAKAGNKPRWVRMREAGEPGWEEEKEKA
ETRKSADRTADVLRALADFGLKPLMRVYTDSEMSSVEWKPLRKGQAVRTWDRDM
FQQAIERMMSWESWNQRVGQEYAKLVEQKNRFEQKNFVGQEHLVHLVNQLQQDM
KEASPGLESKEQTAHYVTGRALRGSDKVFEKWGKLAPDAPFDLYDAEIKNVQRRNT
RRFGSHDLFAKLAEPEYQALWREDASFLTRYAVYNSILRKLNHAKMFATFTLPDAT
AHPIWTRFDKLGGNLHQYTFLFNEFGERRHAIRFHKLLKVENGVAREVDDVTVPISM
SEQLDNLLPRDPNEPIALYFRDYGAEQHFTGEFGGAKIQCRRDQLAHMHRRRGARD
VYLNVSVRVQSQSEARGERRPPYAAVFRLVGDNHRAFVHFDKLSDYLAEHPDDGKL
GSEGLLSGLRVMSVDLGLRTSASISVFRVARKDELKPNSKGRVPFFFPIKGNDNLVAV
HERSQLLKLPGETESKDLRAIREERQRTLRQLRTQLAYLRLLVRCGSEDVGRRERSW

AKLIEQPVDAANHMTPDWREAFENELQKLKSLHGICSDKEWMDAVYESVRRVWRH
MGKQVRDWRKDVRSGERPKIRGYAKDVVGGNSIEQIEYLERQYKFLKSWSFFGKVS
GQVIRAEKGSRFAITLREHIDHAKEDRLKKLADRIIMEALGYVYALDERGKGKWVA
KYPPCQLILLEELSEYQFNNDRPPSENNQLMQWSHRGVFQELINQAQVHDLLVGTM
YAAFSSRFDARTGAPGIRCRRVPARCTQEHNPEPFPWWLNKFVVEHTLDACPLRAD
DLIPTGEGEIFVSPFSAEEGDFHQIHADLNAAQNLQQRLWSDFDISQIRLRCDWGEVD
GELVLIPRLTGKRTADSYSNKVFYTNTGVTYYERERGKKRRKVFAQEKLSEEEAELL
VEADEAREKSVVLMRDPSGIINRGNWTRQKEFWSMV
NQRIEGYLVKQIRSRVPLQDSACENTGDI

[00741] AacCas12b (Alicyclobacillus acidiphilus) - WP_067623834
MAVKSMKVKLRLDNMPEIRAGLWKLHTEVNAGVRYYTEWLSLLRQENLYRRSPNG
DGEQECYKTAEECKAELLERLRARQVENGHCGPAGSDDELLQLARQLYELLVPQAI
GAKGDAQQIARKFLSPLADKDAVGGLGIAKAGNKPRWVRMREAGEPGWEEEKAK
AEARKSTDRTADVLRALADFGLKPLMRVYTDSDMSSVQWKPLRKGQAVRTWDRD
MFQQAIERMMSWESWNQRVGEAYAKLVEQKSRFEQKNFVGQEHLVQLVNQLQQD
MKEASHGLESKEQTAHYLTGRALRGSDKVFEKWEKLDPDAPFDLYDTEIKNVQRRN
TRRFGSHDLFAKLAEPKYQALWREDASFLTRYAVYNSIVRKLNHAKMFATFTLPDA
TAHPIWTRFDKLGGNLHQYTFLFNEFGEGRHAIRFQKLLTVEDGVAKEVDDVTVPIS
MSAQLDDLLPRDPHELVALYFQDYGAEQHLAGEFGGAKIQYRRDQLNHLHARRGA
RDVYLNLSVRVQSQSEARGERRPPYAAVFRLVGDNHRAFVHFDKLSDYLAEHPDDG
KLGSEGLLSGLRVMSVDLGLRTSASISVFRVARKDELKPNSEGRVPFCFPIEGNENLV
AVHERSQLLKLPGETESKDLRAIREERQRTLRQLRTQLAYLRLLVRCGSEDVGRRER
SWAKLIEQPMDANQMTPDWREAFEDELQKLKSLYGICGDREWTEAVYESVRRVWR
HMGKQVRDWRKDVRSGERPKIRGYQKDVVGGNSIEQIEYLERQYKFLKSWSFFGKV
SGQVIRAEKGSRFAITLREHIDHAKEDRLKKLADRIIMEALGYVYALDDERGKGKWV
AKYPPCQLILLEELSEYQFNNDRPPSENNQLMQWSHRGVFQELLNQAQVHDLLVGT
MYAAFSSRFDARTGAPGIRCRRVPARCAREQNPEPFPWWLNKFVAEHKLDGCPLRA
DDLIPTGEGEFFVSPFSAEEGDFHQIHADLNAAQNLQRRLWSDFDISQIRLRCDWGEV
DGEPVLIPRTTGKRTADSYGNKVFYTKTGVTYYERERGKKRRKVFAQEELSEEEAEL
LVEADEAREKSVVLMRDPSGIINRGDWTRQKEFWSMVNQRIEGYLVKQIRSRVRLQ
ESACENTGDI

[00742] BvCas12b (Bacillus sp. V3-13) NCBI Reference Sequence: WP_101661451.1

MAIRSIKLKMKTNSGTDSIYLRKALWRTHQLINEGIAYYMNLLTLYRQEAIGDKTKE
AYQAELINIIRNQQRNNGSSEEHGSDQEILALLRQLYELIIPSSIGESGDANQLGNKFLY
PLVDPNSQSGKGTSNAGRKPRWKRLKEEGNPDWELEKKKDEERKAKDPTVKIFDNL
NKYGLLPLFPLFTNIQKDIEWLPLGKRQSVRKWDKDMFIQAIERLLSWESWNRRVAD
EYKQLKEKTESYYKEHLTGGEEWIEKIRKFEKERNMELEKNAFAPNDGYFITSRQIR
GWDRVYEKWSKLPESASPEELWKVVAEQQNKMSEGFGDPKVFSFLANRENRDIWR
GHSERIYHIAAYNGLQKKLSRTKEQATFTLPDAIEHPLWIRYESPGGTNLNLFKLEEK
QKKNYYVTLSKIIWPSEEKWIEKENIEIPLAPSIQFNRQIKLKQHVKGKQEISFSDYSSR
ISLDGVLGGSRIQFNRKYIKNHKELLGEGDIGPVFFNLVVDVAPLQETRNGRLQSPIG
KALKVISSDFSKVIDYKPKELMDWMNTGSASNSFGVASLLEGMRVMSIDMGQRTSA
SVSIFEVVKELPKDQEQKLFYSINDTELFAIHKRSFLLNLPGEVVTKNNKQQRQERRK
KRQFVRSQIRMLANVLRLETKKTPDERKKAIHKLMEIVQSYDSWTASQKEVWEKEL
NLLTNMAAFNDEIWKESLVELHHRIEPYVGQIVSKWRKGLSEGRKNLAGISMWNIDE
LEDTRRLLISWSKRSRTPGEANRIETDEPFGSSLLQHIQNVKDDRLKQMANLIIMTAL
GFKYDKEEKDRYKRWKETYPACQIILFENLNRYLFNLDRSRRENSRLMKWAHRSIPR
TVSMQGEMFGLQVGDVRSEYSSRFHAKTGAPGIRCHALTEEDLKAGSNTLKRLIEDG
FINESELAYLKKGDIIPSQGGELFVTLSKRYKKDSDNNELTVIHADINAAQNLQKRFW
QQNSEVYRVPCQLARMGEDKLYIPKSQTETIKKYFGKGSFVKNNTEQEVYKWEKSE
KMKIKTDTTFDLQDLDGFEDISKTIELAQEQQKKYLTMFRDPSGYFFNNETWRPQKE
YWSIVNNIIKSCLKKKILSNKVEL

[00743] BhCas12b (Bacillus hisashii) NCBI Reference Sequence: WP_095142515
MAPKKKRKVGIHGVPAAATRSFILKIEPNEEVKKGLWKTHEVLNHGIAYYMNILKLI
RQEAIYEHHEQDPKNPKKVSKAEIQAELWDFVLKMQKCNSFTHEVDKDEVFNILRE
LYEELVPSSVEKKGEANQLSNKFLYPLVDPNSQSGKGTASSGRKPRWYNLKIAGDPS
WEEEKKKWEEDKKKDPLAKILGKLAEYGLIPLFIPYTDSNEPIVKEIKWMEKSRNQS
VRRLDKDMFIQALERFLSWESWNLKVKEEYEKVEKEYKTLEERIKEDIQALKALEQY
EKERQEQLLRDTLNTNEYRLSKRGLRGWREIIQKWLKMDENEPSEKYLEVFKDYQR
KHPREAGDYSVYEFLSKKENHFIWRNHPEYPYLYATFCEIDKKKKDAKQQATFTLA
DPINHPLWVRFEERSGSNLNKYRILTEQLHTEKLKKKLTVQLDRLIYPTESGGWEEK
GKVDIVLLPSRQFYNQIFLDIEEKGKHAFTYKDESIKFPLKGTLGGARVQFDRDHLRR
YPHKVESGNVGRIYFNMTVNIEPTESPVSKSLKIHRDDFPKVVNFKPKELTEWIKDSK
GKKLKSGIESLEIGLRVMSIDLGQRQAAAASIFEVVDQKPDIEGKLFFPIKGTELYAVH
RASFNIKLPGETLVKSREVLRKAREDNLKLMNQKLNFLRNVLHFQQFEDITEREKRV

TKWISRQENSDVPLVYQDELIQIRELMYKPYKDWVAFLKQLHKRLEVEIGKEVKHW
RKSLSDGRKGLYGISLKNIDEIDRTRKFLLRWSLRPTEPGEVRRLEPGQRFAIDQLNH
LNALKEDRLKKMANTIIMHALGYCYDVRKKKWQAKNPACQIILFEDLSNYNPY**E**ER
SRFENSKLM**K**WSRREIPRQVALQGEIYGLQVGEVGAQFSSRFHAKTGSPGIRCSVVT
KEKLQDNRFFKNLQREGRLTLDKIAVLKEGDLYPDKGGEKFISLSKDRKCVTTHADI
NAAQNLQKRFWTRTHGFYKVYCKAYQVDGQTVYIPESKDQKQKIIEEFGEGYFILK
DGVYEWVNAGKLKIKKGSSKQSSSELVDSDILKDSFDLASELKGEKLMLYRDPSGN
VFPSDKWMAAGVFFGKLERILISKLTNQYSISTIEDDSSKQSMKRPAATKKAGQAKK
KK

including the variant termed BvCas12b V4 (S893R/K846R/E837G changes rel. to wt above)

[00744] BhCas12b (V4) is expressed as follows: 5' mRNA Cap---5'UTR---bhCas12b---STOP sequence --- 3'UTR --- 120polyA tail

[00745] 5'UTR:
GGGAAATAAGAGAGAAAAGAAGAGTAAGAAGAAATATAAGAGCCACC

[00746] 3' UTR (TriLink standard UTR)
GCTGGAGCCTCGGTGGCCATGCTTCTTGCCCCTTGGGCCTCCCCCCAGCCCCTCCT
CCCCTTCCTGCACCCGTACCCCCGTGGTCTTTGAATAAAGTCTGA

[00747] Nucleic acid sequence of bhCas12b (V4)
ATGGCCCCAAAGAAGAAGCGGAAGGTCGGTATCCACGGAGTCCCAGCAGCCGCC
ACCAGATCCTTCATCCTGAAGATCGAGCCCAACGAGGAAGTGAAGAAAGGCCTC
TGGAAAACCCACGAGGTGCTGAACCACGGAATCGCCTACTACATGAATATCCTG
AAGCTGATCCGGCAAGAGGCCATCTACGAGCACCACGAGCAGGACCCCAAGAAT
CCCAAGAAGGTGTCCAAGGCCGAGATCCAGGCCGAGCTGTGGGATTTCGTGCTG
AAGATGCAGAAGTGCAACAGCTTCACACACGAGGTGGACAAGGACGAGGTGTTC
AACATCCTGAGAGAGCTGTACGAGGAACTGGTGCCCAGCAGCGTGGAAAAGAA
GGGCGAAGCCAACCAGCTGAGCAACAAGTTTCTGTACCCTCTGGTGGACCCCAA
CAGCCAGTCTGGAAAGGGAACAGCCAGCAGCGGCAGAAAGCCCAGATGGTACA
ACCTGAAGATTGCCGGCGATCCCTCCTGGGAAGAAGAGAAGAAGAAGTGGGAA
GAAGATAAGAAAAAGGACCCGCTGGCCAAGATCCTGGGCAAGCTGGCTGAGTAC
GGACTGATCCCTCTGTTCATCCCCTACACCGACAGCAACGAGCCCATCGTGAAAG
AAATCAAGTGGATGGAAAAGTCCCGGAACCAGAGCGTGCGGCGGCTGGATAAG

GACATGTTCATTCAGGCCCTGGAACGGTTCCTGAGCTGGGAGAGCTGGAACCTG
AAAGTGAAAGAGGAATACGAGAAGGTCGAGAAAGAGTACAAGACCCTGGAAGA
GAGGATCAAAGAGGACATCCAGGCTCTGAAGGCTCTGGAACAGTATGAGAAAG
AGCGGCAAGAACAGCTGCTGCGGGACACCCTGAACACCAACGAGTACCGGCTGA
GCAAGAGAGGCCTTAGAGGCTGGCGGGAAATCATCCAGAAATGGCTGAAAATG
GACGAGAACGAGCCCTCCGAGAAGTACCTGGAAGTGTTCAAGGACTACCAGCGG
AAGCACCCTAGAGAGGCCGGCGATTACAGCGTGTACGAGTTCCTGTCCAAGAAA
GAGAACCACTTCATCTGGCGGAATCACCCTGAGTACCCCTACCTGTACGCCACCT
TCTGCGAGATCGACAAGAAAAAGAAGGACGCCAAGCAGCAGGCCACCTTCACAC
TGGCCGATCCTATCAATCACCCTCTGTGGGTCCGATTCGAGGAAAGAAGCGGCA
GCAACCTGAACAAGTACAGAATCCTGACCGAGCAGCTGCACACCGAGAAGCTGA
AGAAAAAGCTGACAGTGCAGCTGGACCGGCTGATCTACCCTACAGAATCTGGCG
GCTGGGAAGAGAAGGGCAAAGTGGACATTGTGCTGCTGCCCAGCCGGCAGTTCT
ACAACCAGATCTTCCTGGACATCGAGGAAAAGGGCAAGCACGCCTTCACCTACA
AGGATGAGAGCATCAAGTTCCCTCTGAAGGGCACACTCGGCGGAGCCAGAGTGC
AGTTCGACAGAGATCACCTGAGAAGATACCCTCACAAGGTGGAAAGCGGCAACG
TGGGCAGAATCTACTTCAACATGACCGTGAACATCGAGCCTACAGAGTCCCCAG
TGTCCAAGTCTCTGAAGATCCACCGGGACGACTTCCCCAAGGTGGTCAACTTCAA
GCCCAAAGAACTGACCGAGTGGATCAAGGACAGCAAGGGCAAGAAACTGAAGT
CCGGCATCGAGTCCCTGGAAATCGGCCTGAGAGTGATGAGCATCGACCTGGGAC
AGAGACAGGCCGCTGCCGCCTCTATTTTCGAGGTGGTGGATCAGAAGCCCGACA
TCGAAGGCAAGCTGTTTTTCCCAATCAAGGGCACCGAGCTGTATGCCGTGCACAG
AGCCAGCTTCAACATCAAGCTGCCCGGCGAGACACTGGTCAAGAGCAGAGAAGT
GCTGCGGAAGGCCAGAGAGGACAATCTGAAACTGATGAACCAGAAGCTCAACTT
CCTGCGGAACGTGCTGCACTTCCAGCAGTTCGAGGACATCACCGAGAGAGAGAA
GCGGGTCACCAAGTGGATCAGCAGACAAGAGAACAGCGACGTGCCCCTGGTGTA
CCAGGATGAGCTGATCCAGATCCGCGAGCTGATGTACAAGCCTTACAAGGACTG
GGTCGCCTTCCTGAAGCAGCTCCACAAGAGACTGGAAGTCGAGATCGGCAAAGA
AGTGAAGCACTGGCGGAAGTCCCTGAGCGACGGAAGAAAGGGCCTGTACGGCAT
CTCCCTGAAGAACATCGACGAGATCGATCGGACCCGGAAGTTCCTGCTGAGATG
GTCCCTGAGGCCTACCGAACCTGGCGAAGTGCGTAGACTGGAACCCGGCCAGAG
ATTCGCCATCGACCAGCTGAATCACCTGAACGCCCTGAAAGAAGATCGGCTGAA
GAAGATGGCCAACACCATCATCATGCACGCCCTGGGCTACTGCTACGACGTGCG
GAAGAAGAAATGGCAGGCTAAGAACCCCGCCTGCCAGATCATCCTGTTCGAGGA

TCTGAGCAACTACAACCCCTACGAGGAAAGGTCCCGCTTCGAGAACAGCAAGCT
CATGAAGTGGTCCAGACGCGAGATCCCCAGACAGGTTGCACTGCAGGGCGAGAT
CTATGGCCTGCAAGTGGGAGAAGTGGGCGCTCAGTTCAGCAGCAGATTCCACGC
CAAGACAGGCAGCCCTGGCATCAGATGTAGCGTCGTGACCAAAGAGAAGCTGCA
GGACAATCGGTTCTTCAAGAATCTGCAGAGAGAGGGCAGACTGACCCTGGACAA
AATCGCCGTGCTGAAAGAGGGCGATCTGTACCCAGACAAAGGCGGCGAGAAGTT
CATCAGCCTGAGCAAGGATCGGAAGTGCGTGACCACACACGCCGACATCAACGC
CGCTCAGAACCTGCAGAAGCGGTTCTGGACAAGAACCCACGGCTTCTACAAGGT
GTACTGCAAGGCCTACCAGGTGGACGGCCAGACCGTGTACATCCCTGAGAGCAA
GGACCAGAAGCAGAAGATCATCGAAGAGTTCGGCGAGGGCTACTTCATTCTGAA
GGACGGGGTGTACGAATGGGTCAACGCCGGCAAGCTGAAAATCAAGAAGGGCA
GCTCCAAGCAGAGCAGCAGCGAGCTGGTGGATAGCGACATCCTGAAAGACAGCT
TCGACCTGGCCTCCGAGCTGAAAGGCGAAAAGCTGATGCTGTACAGGGACCCCA
GCGGCAATGTGTTCCCCAGCGACAAATGGATGGCCGCTGGCGTGTTCTTCGGAA
AGCTGGAACGCATCCTGATCAGCAAGCTGACCAACCAGTACTCCATCAGCACCA
TCGAGGACGACAGCAGCAAGCAGTCTATGAAAAGGCCGGCGGCCACGAAAAAG
GCCGGCCAGGCAAAAAAGAAAAAG

Fusion proteins comprising a Cas9 domain and a Cytidine Deaminase or Adenosine Deaminase

[00748] Some aspects of the disclosure provide fusion proteins comprising a Cas9 domain or other nucleic acid programmable DNA binding protein and one or more cytidine deaminase or adenosine deaminase domains. It should be appreciated that the Cas9 domain may be any of the Cas9 domains or Cas9 proteins (*e.g.*, dCas9 or nCas9) provided herein. In some embodiments, any of the Cas9 domains or Cas9 proteins (*e.g.*, dCas9 or nCas9) provided herein may be fused with any of the cytidine deaminases provided herein. For example, and without limitation, in some embodiments, the fusion protein comprises the structure:

[00749] NH_2-[cytidine deaminase]-[Cas9 domain]-COOH; or

[00750] NH_2-[Cas9 domain]-[cytidine deaminase]-COOH.

[00751] In some embodiments, the fusion proteins comprising a cytidine deaminase or adenosine deaminase and a napDNAbp (e.g., Cas9 domain) do not include a linker sequence. In some embodiments, a linker is present between the cytidine or adenosine deaminase and the napDNAbp. In some embodiments, the "-" used in the general architecture above

indicates the presence of an optional linker. In some embodiments, cytidine or adenosine deaminase and the napDNAbp are fused via any of the linkers provided herein. For example, in some embodiments the cytidine or adenosine deaminase and the napDNAbp are fused via any of the linkers in the section entitled "Linkers".

Fusion proteins comprising a nuclear localization sequence (NLS)

[00752] In some embodiments, the fusion proteins provided herein further comprise one or more (e.g., 2, 3, 4, 5) nuclear targeting sequences, for example a nuclear localization sequence (NLS). In one embodiment, a bipartite NLS is used. In some embodiments, a NLS comprises an amino acid sequence that facilitates the importation of a protein, that comprises an NLS, into the cell nucleus (*e.g.,* by nuclear transport). In some embodiments, any of the fusion proteins provided herein further comprise a nuclear localization sequence (NLS). In some embodiments, the NLS is fused to the N-terminus of the fusion protein. In some embodiments, the NLS is fused to the C-terminus of the fusion protein. In some embodiments, the NLS is fused to the N-terminus of the Cas9 domain. In some embodiments, the NLS is fused to the C-terminus of the Cas9 domain. In some embodiments, the NLS is fused to the N-terminus of the cytidine or adenosine deaminase. In some embodiments, the NLS is fused to the C-terminus of the cytidine or adenosine deaminase. In some embodiments, the NLS is fused to the fusion protein via one or more linkers. In some embodiments, the NLS is fused to the fusion protein without a linker. In some embodiments, the NLS comprises an amino acid sequence of any one of the NLS sequences provided or referenced herein. Additional nuclear localization sequences are known in the art and would be apparent to the skilled artisan. For example, NLS sequences are described in Plank *et al.*, PCT/EP2000/011690, the contents of which are incorporated herein by reference for their disclosure of exemplary nuclear localization sequences. In some embodiments, an NLS comprises the amino acid sequence KRTADGSEFESPKKKRKV, KRPAATKKAGQAKKKK, KKTELQTTNAENKTKKL, KRGINDRNFWRGENGRKTR, RKSGKIAAIVVKRPRKPKKKRKV, or MDSLLMNRRKFLYQFKNVRWAKGRRETYLC.

[00753] In some embodiments, the general architecture of exemplary Cas9 fusion proteins with a cytidine or adenosine deaminase and a Cas9 domain comprises any one of the following structures, where NLS is a nuclear localization sequence (*e.g.,* any NLS provided herein), NH_2 is the N-terminus of the fusion protein, and COOH is the C-terminus of the fusion protein:

[00754] NH_2-NLS-[cytidine deaminase]-[Cas9 domain]-COOH;

[00755] NH_2-NLS [Cas9 domain]-[cytidine deaminase]-COOH;

[00756] NH_2-[cytidine deaminase]-[Cas9 domain]-NLS-COOH; or

[00757] NH_2-[Cas9 domain]-[cytidine deaminase]-NLS-COOH.

[00758] NH_2-NLS-[adenosine deaminase]-[Cas9 domain]-COOH;

[00759] NH_2-NLS [Cas9 domain]-[adenosine deaminase]-COOH;

[00760] NH_2-[adenosine deaminase]-[Cas9 domain]-NLS-COOH; or

[00761] NH_2-[Cas9 domain]-[adenosine deaminase]-NLS-COOH.

[00762] In some embodiments, the NLS is present in a linker or the NLS is flanked by linkers, for example described herein. A bipartite NLS comprises two basic amino acid clusters, which are separated by a relatively short spacer sequence (hence bipartite - 2 parts, while monopartite NLSs are not). The NLS of nucleoplasmin, KR[PAATKKAGQA]KKKK, is the prototype of the ubiquitous bipartite signal: two clusters of basic amino acids, separated by a spacer of about 10 amino acids.

[00763] The sequence of an exemplary bipartite NLS follows: PKKKRKVEGADKRTADGSEFES PKKKRKV

[00764] In some embodiments, the fusion proteins comprising a cytidine or adenosine deaminase, a Cas9 domain, and an NLS do not comprise a linker sequence. In some embodiments, linker sequences between one or more of the domains or proteins (*e.g.,* cytidine or adenosine deaminase, Cas9 domain or NLS) are present.

[00765] It should be appreciated that the fusion proteins of the present disclosure may comprise one or more additional features. For example, in some embodiments, the fusion protein may comprise inhibitors, cytoplasmic localization sequences, export sequences, such as nuclear export sequences, or other localization sequences, as well as sequence tags that are useful for solubilization, purification, or detection of the fusion proteins. Suitable protein tags provided herein include, but are not limited to, biotin carboxylase carrier protein (BCCP) tags, myc-tags, calmodulin-tags, FLAG-tags, hemagglutinin (HA)-tags, polyhistidine tags, also referred to as histidine tags or His-tags, maltose binding protein (MBP)-tags, nus-tags, glutathione-S-transferase (GST)-tags, green fluorescent protein (GFP)-tags, thioredoxin-tags, S-tags, Softags (e.g., Softag 1, Softag 3), strep-tags , biotin ligase tags, FlAsH tags, V5 tags, and SBP-tags. Additional suitable sequences will be apparent to those of skill in the art. In some embodiments, the fusion protein comprises one or more His tags.

Linkers

[00766] In certain embodiments, linkers may be used to link any of the peptides or peptide domains of the invention. The linker may be as simple as a covalent bond, or it may be a polymeric linker many atoms in length. In certain embodiments, the linker is a polypeptide or based on amino acids. In other embodiments, the linker is not peptide-like. In certain embodiments, the linker is a covalent bond (*e.g.*, a carbon-carbon bond, disulfide bond, carbon-heteroatom bond, *etc.*). In certain embodiments, the linker is a carbon-nitrogen bond of an amide linkage. In certain embodiments, the linker is a cyclic or acyclic, substituted or unsubstituted, branched or unbranched aliphatic or heteroaliphatic linker. In certain embodiments, the linker is polymeric (e.g., polyethylene, polyethylene glycol, polyamide, polyester, *etc.*). In certain embodiments, the linker comprises a monomer, dimer, or polymer of aminoalkanoic acid. In certain embodiments, the linker comprises an aminoalkanoic acid (*e.g.*, glycine, ethanoic acid, alanine, beta-alanine, 3-aminopropanoic acid, 4-aminobutanoic acid, 5-pentanoic acid, *etc.*). In certain embodiments, the linker comprises a monomer, dimer, or polymer of aminohexanoic acid (Ahx). In certain embodiments, the linker is based on a carbocyclic moiety (e.g., cyclopentane, cyclohexane). In other embodiments, the linker comprises a polyethylene glycol moiety (PEG). In other embodiments, the linker comprises amino acids. In certain embodiments, the linker comprises a peptide. In certain embodiments, the linker comprises an aryl or heteroaryl moiety. In certain embodiments, the linker is based on a phenyl ring. The linker may include functionalized moieties to facilitate attachment of a nucleophile (e.g., thiol, amino) from the peptide to the linker. Any electrophile may be used as part of the linker. Exemplary electrophiles include, but are not limited to, activated esters, activated amides, Michael acceptors, alkyl halides, aryl halides, acyl halides, and isothiocyanates.

[00767] In some embodiments, the linker is an amino acid or a plurality of amino acids (e.g., a peptide or protein). In some embodiments, the linker is a bond (e.g., a covalent bond), an organic molecule, group, polymer, or chemical moiety. In some embodiments, the cytidine or adenosine deaminase and the napDNAbp are fused via a linker that comprises 4, 16, 32, or 104 amino acids in length. In some embodiments, the linker is about 3 to about 104 amino acids in length. In some embodiments, any of the fusion proteins provided herein, comprise a cytidine or adenosine deaminase and a Cas9 domain that are fused to each other via a linker. e.g., Various linker lengths and flexibilities between the cytidine or adenosine deaminase and the Cas9 domain can be employed (*e.g.,* ranging from very flexible linkers of the form $(GGGS)_n$, $(GGGGS)_n$, and $(G)_n$ to more rigid linkers of the form $(EAAAK)_n$,

$(SGGS)_n$, SGSETPGTSESATPES (see, e.g., Guilinger JP, Thompson DB, Liu DR. Fusion of catalytically inactive Cas9 to FokI nuclease improves the specificity of genome modification. Nat. Biotechnol. 2014; 32(6): 577-82; the entire contents are incorporated herein by reference) and $(XP)_n$) in order to achieve the optimal length for activity for the cytidine or adenosine deaminase nucleobase editor. In some embodiments, n is 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, or 15. In some embodiments, the linker comprises a $(GGS)_n$ motif, wherein n is 1, 3, or 7. In some embodiments, cytidine deaminase or adenosine deaminase and the Cas9 domain of any of the fusion proteins provided herein are fused via a linker comprising the amino acid sequence SGSETPGTSESATPES.

Cas9 complexes with guide RNAs

[00768] Some aspects of this disclosure provide complexes comprising any of the fusion proteins provided herein, and a guide RNA bound to a Cas9 domain (*e.g.,* a dCas9, a nuclease active Cas9, or a Cas9 nickase) of fusion protein. These complexes are also termed ribonucleoproteins (RNPs). In some embodiments, the guide nucleic acid (e.g., guide RNA) is from 15-100 nucleotides long and comprises a sequence of at least 10 contiguous nucleotides that is complementary to a target sequence. In some embodiments, the guide RNA is 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50 nucleotides long. In some embodiments, the guide RNA comprises a sequence of 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, or 40 contiguous nucleotides that is complementary to a target sequence. In some embodiments, the target sequence is a DNA sequence. In some embodiments, the target sequence is an RNA sequence. In some embodiments, the target sequence is a sequence in the genome of a mammal. In some embodiments, the target sequence is a sequence in the genome of a human. In some embodiments, the 3' end of the target sequence is immediately adjacent to a canonical PAM sequence (NGG). In some embodiments, the guide nucleic acid (e.g., guide RNA) is complementary to a sequence associated with a disease or disorder.

[00769] In some embodiments, the guide RNA is designed to disrupt a splice site (i.e., a splice acceptor (SA) or a splice donor (SD). In some embodiments, the guide RNA is designed such that the base editing results in a premature STOP codon. Tables 8A Table 8B and Table 8C provide a nonexhaustive list of gRNA target sequences designed to disrupt a splice site or to result in a premature STOP codon.

[00770] Provided herein are compositions and methods for base editing in host cells, e.g. immune cells. Further provided herein are compositions comprising a guide polynucleic acid sequence, e.g. a guide RNA sequence, or a combination of 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, or more guide RNAs as provided herein. In some embodiments, a composition for base editing as provided herein further comprises a polynucleotide that encodes a base editor, e.g. a C-base editor or an A-base editor. For example, a composition for base editing may comprise a mRNA sequence encoding a BE, a BE4, an ABE, and a combination of one or more guide RNAs as provided. A composition for base editing may comprise a base editor polypeptide and a compbination of one or more of any guide RNAs provided herein. Such a composition may be used to effect base editing in an immune cell through different delivery approaches, for example, electroporation, nucleofection, viral transduction or transfection. In some embodiments, the composition for base editing comprises an mRNA sequence that encodes a base editor and a combination of one or more guide RNA sequences provided herein for electroporation.

Table 8A: gRNAs: Splice Site and STOP Codons

Gene	Description	Targeting sequence	gRNA Spacer Sequence
VISTA	Exon 1 SD (pos6)	CCTTACCTAGGGACGCAGCC	CCUUACCUAGGGACGCAGCC
	Exon 1 STOP (pos7)	GGATCCCCAGCGCCAGCTGC	GGAUCCCCAGCGCCAGCUGC
	Exon 1 STOP (pos5)	AGCGCCAGCTGCCGGCCTCC	AGCGCCAGCUGCCGGCCUCC
	Exon 1 STOP (pos4)	GCGCCAGCTGCCGGCCTCCA	GCGCCAGCUGCCGGCCUCCA
	Exon 2 STOP (pos8)	CCTGGCTCAGCGCCACGGGC	CCUGGCUCAGCGCCACGGGC
	Exon 2 STOP (pos5)	GCTGCAGGTGCAGACAGGTG	GCUGCAGGUGCAGACAGGUG
	Exon 2 STOP (pos7)	GCGGTACCACGTCTTGTAGA	GCGGUACCACGUCUUGUAGA
	Exon 3 SA (pos4)	TGCCTGTGGGAACAAACAGA	UGCCUGUGGGAACAAACAGA
	Exon 3 SD (pos5)	CTTACTTTCACTATCCTGGG	CUUACUUUCACUAUCCUGGG
	Exon 3 SD (pos8)	TCCCTTACTTTCACTATCCT	UCCCUUACUUUCACUAUCCU
	Exon 3 STOP (pos5)	CTCCCAGGATAGTGAAAGTA	CUCCCAGGAUAGUGAAAGUA
	Exon 4 SA (pos7)	TGATGTCTGAAAGGGCAGAG	UGAUGUCUGAAAGGGCAGAG
	Exon 5 STOP (pos5)	TGCCCAGGAGCTGGTGCGGA	UGCCCAGGAGCUGGUGCGGA

Gene	Description	Targeting sequence	gRNA Spacer Sequence
	Exon 6 SA (pos4)	TTGCTGCCACAGAACCAGAA	UUGCUGCCACAGAACCAGAA
	Exon 6 STOP (pos4)	ATTCAAGGGATTGAAAACCC	AUUCAAGGGAUUGAAAACCC
	Exon 6 STOP (pos8)	ACCTGCCCAGGGGATACCCG	ACCUGCCCAGGGGAUACCCG
	Exon 6 STOP (pos7)	CAGCGGCAGCCTTCTGAGTC	CAGCGGCAGCCUUCUGAGUC
TRAC	Exon 1 STOP 1 (pos5)	GCTACAAACAAGCTCATCTT	GCUACAAACAAGCUCAUCUU
	Exon 1 STOP 2 (pos6)	CCAGCCAAGTACGTAAGTAG	CCAGCCAAGUACGUAAGUAG
	Exon 1 SA (pos9)	CTGGATATCTGTGGGACAAG	CUGGAUAUCUGUGGGACAAG
	Exon 1 SD	CTTACCTGGGCTGGGGAAGA	CUUACCUGGGCUGGGGAAGA
	Exon 3SA	TTCGTATCTGTAAAACCAAG	UUCGUAUCUGUAAAACCAAG
	Exon 3 STOP	TTTCAAAACCTGTCAGTGAT	UUUCAAAACCUGUCAGUGAU
	Exon 3 STOP	TTCAAAACCTGTCAGTGATT	UUCAAAACCUGUCAGUGAUU
Tim-3	Exon 2 SA (pos6)	GGACCCTGCATAGAGAGAGA	GGACCCUGCAUAGAGAGAGA
	Exon 2 STOP (pos5)	TGCCCCAGCAGACGGGCACG	UGCCCCAGCAGACGGGCACG
	Exon 3 SD (pos5)	GTTACCTGGGCCATGTCCCC	GUUACCUGGGCCAUGUCCCC
	Exon 4 SD (pos5)	CTTACTGTTAGATTTATATC	CUUACUGUUAGAUUUAUAUC
	Exon 4 SD (pos4)	TTACTGTTAGATTTATATCA	UUACUGUUAGAUUUAUAUCA
	Exon 5 SA (pos5)	TTTGCTATGGAAACACAAAC	UUUGCUAUGGAAACACAAAC
	Exon 5 STOP (pos8)	TCCATAGCAAATATCCACAT	UCCAUAGCAAAUAUCCACAU
	Exon 7 STOP (pos5)	GCAGCAACCCTCACAACCTT	GCAGCAACCCUCACAACCUU
	Exon 7 STOP (pos 4)	CAGCAACCCTCACAACCTTT	CAGCAACCCUCACAACCUUU
TIGIT	Exon 1 STOP (pos4)	AGGCAGGCTCCCCTCGCCTC	AGGCAGGCUCCCCUCGCCUC
	Exon 2 STOP (5&8)	GGAGCAGCAGGACCAGCTTC	GGAGCAGCAGGACCAGCUUC
	Exon 2 SD (pos9)	CAGGAATACCTGAGCTTTCT	CAGGAAUACCUGAGCUUUCU
	Exon 3 STOP (pos7)	AGGTTCCAGATTCCATTGCT	AGGUUCCAGAUUCCAUUGCU
	Exon 1 STOP	CTGGGCCCAGGGGCTGAGGC	CUGGGCCCAGGGGCUGAGGC
	Exon 2 STOP	GATCGAGTGGCCCCAGGTCC	GAUCGAGUGGCCCCAGGUCC

Gene	Description	Targeting sequence	gRNA Spacer Sequence
TGFbRII	Exon 1 SD (JMG79)	TCACCCGACTTCTGAACGTG	UCACCCGACUUCUGAACGUG
	Exon 3 SD (JMG83)	TTACCTGCCCACTGTTAGCC	UUACCUGCCCACUGUUAGCC
	Exon 2 STOP (JMG80)	GAAGCCACAGGAAGTCTGTG	GAAGCCACAGGAAGUCUGUG
	Exon 3 STOP (JMG81)	ACTCCAGTTCCTGACGGCTG	ACUCCAGUUCCUGACGGCUG
	Exon 3 STOP (JMG82)	ACCTACAGGAGTACCTGACG	ACCUACAGGAGUACCUGACG
	Exon 4 STOP (JMG84)	TTCCCAGAGCACCAGAGCCA	UUCCCAGAGCACCAGAGCCA
	Exon 1 STOP (JMG85)	ACGTTCAGAAGTCGGGTGAG	ACGUUCAGAAGUCGGGUGAG
	Exon 3 STOP (pos8)	TTCAGAGCAGTTTGAGACAG	UUCAGAGCAGUUUGAGACAG
	Exon 1 SD	TCACCCGACTTCTGAACGTG	UCACCCGACUUCUGAACGUG
	Exon 1 Stop	ACGTTCAGAAGTCGGGTGAG	ACGUUCAGAAGUCGGGUGAG
	Exon 2 SD 1	TTTACTATGTCTCAGTGGAT	UUUACUAUGUCUCAGUGGAU
	Exon 2 SD2	CTTTACTATGTCTCAGTGGA	CUUUACUAUGUCUCAGUGGA
	Exon 3 STOP	GAAGCCACAGGAAGTCTGTG	GAAGCCACAGGAAGUCUGUG
	Exon 6 SD	TTACCTGCCCACTGTTAGCC	UUACCUGCCCACUGUUAGCC
	Exon 6 STOP 1	TTCAGAGCAGTTTGAGACAG	UUCAGAGCAGUUUGAGACAG
	Exon 6 STOP 2	ACTCCAGTTCCTGACGGCTG	ACUCCAGUUCCUGACGGCUG
	Exon 6 STOP	ACCTACAGGAGTACCTGACG	ACCUACAGGAGUACCUGACG
	Exon 7 STOP	TTCCCAGAGCACCAGAGCCA	UUCCCAGAGCACCAGAGCCA
	Exon 8 STOP	AGCCAGAAGCTGGGAATTTC	AGCCAGAAGCUGGGAAUUUC
	Isoform ATG	TATCATGTCGTTATTAACTG	UAUCAUGUCGUUAUUAACUG
RFXANK	Exon 2 SA (JMG8)	CCTGCTGGGAAACAGACAAC	CCUGCUGGGAAACAGACAAC
	Exon 2 SD (JMG9)	CACTCACAGTCTAGGGTGGC	CACUCACAGUCUAGGGUGGC
	Exon 2 STOP (pos8)	CAACCGGCAGCGAGGGAACG	CAACCGGCAGCGAGGGAACG
	Exon 3 SA (pos7)	ACAGGGCTGGGGCAGGACAG	ACAGGGCUGGGGCAGGACAG
	Exon 3 STOP (pos8)	CATCCACCAGCTCGCAGCAC	CAUCCACCAGCUCGCAGCAC
	Exon 3 STOP (pos7)	ATCCACCAGCTCGCAGCACA	AUCCACCAGCUCGCAGCACA

Gene	Description	Targeting sequence	gRNA Spacer Sequence
	Exon 3 STOP (pos6)	TCCACCAGCTCGCAGC ACAG	UCCACCAGCUCGCAGC ACAG
	Exon 3 STOP (pos5)	CCACCAGCTCGCAGCA CAGG	CCACCAGCUCGCAGCA CAGG
	Exon 4 SA (JMG10)	TGTCACCTGGCAGGAG GAGG	UGUCACCUGGCAGGAG GAGG
	Exon 4 SA (pos6)	GTCACCTGGCAGGAGG AGGC	GUCACCUGGCAGGAGG AGGC
	Exon 5 SA (pos7)	GGCACCCTGCAGGGAG AAGA	GGCACCCUGCAGGGAG AAGA
	Exon 5 SA (JMG11)	GCACCCTGCAGGGAGA AGAA	GCACCCUGCAGGGAGA AGAA
	Exon 6 SA (pos4)	ATTCTGTCGTGGGTAG GGGC	AUUCUGUCGUGGGUAG GGGC
	Exon 6 SA (JMG12)	CTCCATTCTGTCGTGG GTAG	CUCCAUUCUGUCGUGG GUAG
	Exon 7 SA (pos8)	CCTCGGGCTGCAAAGG AGAG	CCUCGGGCUGCAAAGG AGAG
	Exon 7 SA (pos5)	CGGGCTGCAAAGGAG AGGGG	CGGGCUGCAAAGGAGA GGGG
	Exon 7 SD (pos6)	GCTGACCTTTCCGGTA TCCC	GCUGACCUUUCCGGUA UCCC
	Exon 7 SD (pos5)	CTGACCTTTCCGGTAT CCCA	CUGACCUUUCCGGUAU CCCA
	Exon 8 SA (pos8)	TGTTGCACTGAGATGG GGCA	UGUUGCACUGAGAUGG GGCA
	Exon 8 SA (pos9)	CTGTTGCACTGAGATG GGGC	CUGUUGCACUGAGAUG GGGC
PVRIG (CD112 R)	Exon 1 STOP (pos7)	GCCCTGCAGCCCCCAG AACC	GCCCUGCAGCCCCCAG AACC
	Exon 1 SD (pos5)	CTCACCCGCAGTGACA CACA	CUCACCCGCAGUGACA CACA
	Exon 1 STOP (pos8)	GCAGCACCCAGGGCA GGACC	GCAGCACCCAGGGCAG GACC
	Exon 1 STOP (pos7)	CAGCACCCAGGGCAG GACCA	CAGCACCCAGGGCAGG ACCA
	Exon 2 SA (pos5)	GTCCCTGTGGAACAGC AGCA	GUCCCUGUGGAACAGC AGCA
	Exon 2 STOP (pos8)	GTGGGTTCAAGTTCGG ATGG	GUGGGUUCAAGUUCGG AUGG
	Exon 2 SD (pos 7)	GCCCCACCTGGGTCTG AGCT	GCCCCACCUGGGUCUG AGCU
	Exon 2 SD (pos8)	GGCCCCACCTGGGTCT GAGC	GGCCCCACCUGGGUCU GAGC
	Exon 2 SD (pos4)	CCACCTGGGTCTGAGC TGGG	CCACCUGGGUCUGAGC UGGG
	Exon 2 STOP (pos8)	AGGCCTCCCAGGAGCC CTCA	AGGCCUCCCAGGAGCC CUCA
	Exon 2 STOP (pos4)	CTCCCAGGAGCCCTCA GGGA	CUCCCAGGAGCCCUCA GGGA
	Exon 2 STOP (pos4)	CCCCCAGCTCACAGTC ACCA	CCCCCAGCUCACAGUC ACCA

Gene	Description	Targeting sequence	gRNA Spacer Sequence
	Exon 3 SD (pos8)	GGTCTCACCGGTGCTTATGT	GGUCUCACCGGUGCUUAUGU
	Exon 3 STOP (pos9)	TGCTGCGCCGACATAAGCAC	UGCUGCGCCGACAUAAGCAC
	Exon 4 SA (pos8)	GGCAGGGCTGGGAGAGAGCA	GGCAGGGCUGGGAGAGAGCA
	Exon 4 STOP (pos9)	CGAGAGCACGAGCATGGGTG	CGAGAGCACGAGCAUGGGUG
	Exon 4 STOP (pos6)	GAGCACGAGCATGGGTGAGG	GAGCACGAGCAUGGGUGAGG
	Exon 4 STOP (pos5)	AGCACGAGCATGGGTGAGGA	AGCACGAGCAUGGGUGAGGA
	Exon 4 STOP (pos4)	GCACGAGCATGGGTGAGGAG	GCACGAGCAUGGGUGAGGAG
	Exon 4 SD (pos5)	CTCACCCATGCTCGTGCTCT	CUCACCCAUGCUCGUGCUCU
	Exon 5 SA (pos6)	GGTGCCTGCGCGGGGGAAGG	GGUGCCUGCGCGGGGGAAGG
	Exon 5 SA (pos5)	GTGCCTGCGCGGGGGAAGGA	GUGCCUGCGCGGGGGAAGGA
	Exon 5 SA (pos9)	CTTGGTGCCTGCGCGGGGGA	CUUGGUGCCUGCGCGGGGGA
	Exon 5 STOP (pos6)	GGCCCCAGGGCCCTGCCGCC	GGCCCCAGGGCCCUGCCGCC
	Exon 5 STOP (pos9)	TCTACGCTCAGGCAGGGGAG	UCUACGCUCAGGCAGGGGAG
	Exon 5 STOP (pos4)	CCACCAGGACGGCCCCCCAT	CCACCAGGACGGCCCCCCAU
	Exon 5 STOP (pos5)	AGGCCCAGGCGGCAGGGCCC	AGGCCCAGGCGGCAGGGCCC
	Exon 5 STOP (pos4)	GGCCCAGGCGGCAGGGCCCT	GGCCCAGGCGGCAGGGCCCU
PDCD1	Exon 1 STOP 2 (pos9)	ACGACTGGCCAGGGCGCCTG	ACGACUGGCCAGGGCGCCUG
	Exon 1 STOP 4 (pos7)	CACCGCCCAGACGACTGGCC	CACCGCCCAGACGACUGGCC
	Exon 1 STOP (pos4)	CTACAACTGGGCTGGCGGCC	CUACAACUGGGCUGGCGGCC
	Exon 1 SD	CACCTACCTAAGAACCATCC]	CACCUACCUAAGAACCAUCC]
	Exon 2 SA	GGAGTCTGAGAGATGGAGAG	GGAGUCUGAGAGAUGGAGAG
	Exon 2 STOP 1 (pos8)	CAGCAACCAGACGGACAAGC	CAGCAACCAGACGGACAAGC
	Exon 2 STOP 2 (pos9)	GTGTCACACAACTGCCCAAC	GUGUCACACAACUGCCCAAC
	Exon 3 STOP 1 (pos8)	AGCCGGCCAGTTCCAAACCC	AGCCGGCCAGUUCCAAACCC
	Exon 3 STOP (pos7)	CAGTTCCAAACCCTGGTGGT	CAGUUCCAAACCCUGGUGGU
	Exon 3 STOP 2 (pos5)	CGGCCAGTTCCAAACCCTGG	CGGCCAGUUCCAAACCCUGG

Gene	Description	Targeting sequence	gRNA Spacer Sequence
	Exon 3 STOP (pos5)	GGACCCAGACTAGCAG CACC]	GGACCCAGACUAGCAG CACC]
	Exon 3 SD	GACGTTACCTCGTGCG GCCC	GACGUUACCUCGUGCG GCCC
	Exon 4 SA	TCCCTGCAGAGAAACA CACT	UCCCUGCAGAGAAACA CACU
	Exon 4 SD	GAGACTCACCAGGGGC TGGC	GAGACUCACCAGGGGC UGGC
	Exon 5 SA	CCTCCTTCTTTGAGGA GAAA	CCUCCUUCUUUGAGGA GAAA
	Exon 2 STOP (pos 7)	GGGGTTCCAGGGCCTG TCTG	GGGGUUCCAGGGCCUG UCUG
	Exon 3 SA	TTCTCTCTGGAAGGGC ACAA	UUCUCUCUGGAAGGGC ACAA
	Exon 5 STOP 1 (pos 8)	CCAGTGGCGAGAGAA GACCC	CCAGUGGCGAGAGAAG ACCC
	Exon 5 STOP 2 (pos 5)	TGCCCAGCCACTGAGG CCTG	UGCCCAGCCACUGAGG CCUG
	Exon 1 STOP 1 (pos8)	CGACTGGCCAGGGCGC CTGT	CGACUGGCCAGGGCGC CUGU
	Exon 1 STOP 3 (pos6)	ACCGCCCAGACGACTG GCCA	ACCGCCCAGACGACUG GCCA
Lag3	Exon 1 STOP (pos8)	GTTTCTGCAGCCGCTT TGGG	GUUUCUGCAGCCGCUU UGGG
	Exon 1 SD (pos4)	TTACCTGGAGCCACCC AAAG	UUACCUGGAGCCACCC AAAG
	Exon 2 SA (pos4)	TCACTAGGTGAGCAAA AGAG	UCACUAGGUGAGCAAA AGAG
	Exon 2 STOP (pos8)	GCCTCTCCAGCCAGGG GCTG	GCCUCUCCAGCCAGGG GCUG
	Exon 2 STOP (pos 6)	CTTGGCAGCATCAGCC AGAC	CUUGGCAGCAUCAGCC AGAC
	Exon 3 SA (pos4)	CCACTGGGCGGGAAA GAGAA	CCACUGGGCGGGAAAG AGAA
	Exon 3 SD (pos6)	ACATACTCGAGGCCTG GCCC	ACAUACUCGAGGCCUG GCCC
	Exon 3 STOP (pos5)	CCTGCAGCCCCGCGTC CAGC	CCUGCAGCCCCGCGUC CAGC
	Exon 3 STOP (pos7)	CGCGTCCAGCTGGATG AGCG	CGCGUCCAGCUGGAUG AGCG
	Exon 3 STOP (pos6)	TGGGCCAGGCCTCGAG TATG	UGGGCCAGGCCUCGAG UAUG
	Exon 4 SD (pos4)	GGGAGTTACCCAGAAC AGTG	GGGAGUUACCCAGAAC AGUG
	Exon 4 STOP (pos8)	CCTGCCCCAAGTCAGC CCCA	CCUGCCCCAAGUCAGC CCCA
	Exon 4 STOP (pos9)	GCCAGGGCCGAGTCCC TGTC	GCCAGGGCCGAGUCCC UGUC
	Exon 4 STOP (pos8)	CCAGGGCCGAGTCCCT GTCC	CCAGGGCCGAGUCCCU GUCC
	Exon 4 STOP (pos4)	GCCCCAGGGCCCAGAG TCCA	GCCCCAGGGCCCAGAG UCCA

Gene	Description	Targeting sequence	gRNA Spacer Sequence
	Exon 5 STOP (pos9)	ATGTGAGCCAGGCCCAGGCT	AUGUGAGCCAGGCCCAGGCU
	Exon 5 STOP (pos 8)	GAGGAGTCCACTTGGCAGTG	GAGGAGUCCACUUGGCAGUG
	Exon 6 SA (pos7)	GAGTCACTGAAAAGAGTAGA	GAGUCACUGAAAAGAGUAGA
	Exon 6 STOP (pos6)	CTGGACAAGAACGCTTTGTG	CUGGACAAGAACGCUUUGUG
	Exon 6 STOP (pos7)	CCATCCCAGAGGAGTTTCTC	CCAUCCCAGAGGAGUUUCUC
	Exon 6 STOP (pos4)	TGGCAATGCCAGCTGTACCA	UGGCAAUGCCAGCUGUACCA
	Exon 6 STOP (pos4)	TACCAGGGGGAGAGGCTTCT	UACCAGGGGGAGAGGCUUCU
	Exon 6 STOP (pos8)	GGCATTGCCAAGGCTGGGAA	GGCAUUGCCAAGGCUGGGAA
	Exon 7 SA (pos6)	GGCACCTATGGAGAAAGTAC	GGCACCUAUGGAGAAAGUAC
	Exon 7 STOP (pos4)	AGACAGGTGAGCCAGGGACA	AGACAGGUGAGCCAGGGACA
	Exon 7 SD (pos7)	GGCTCACCTGTCTTCTCCAA	GGCUCACCUGUCUUCUCCAA
	Exon 8 SA (pos8)	GTCGCCACTGTGAGAAGAGA	GUCGCCACUGUGAGAAGAGA
	Exon 8 STOP (pos8)	GCAGGCTCAGAGCAAGATAG	GCAGGCUCAGAGCAAGAUAG
	Exon 8 STOP (pos8)	GCTGGAGCAAGAACCGGAGC	GCUGGAGCAAGAACCGGAGC
CTLA-4	Exon 1 SD (pos 6)	ACTCACCTTTGCAGAAGACA	ACUCACCUUUGCAGAAGACA
	Exon 1 SD	CACTCACCTTTGCAGAAGAC	CACUCACCUUUGCAGAAGAC
	Exon 1 STOP (pos5)	AGGGCCAGGTCCTGGTAGCC	AGGGCCAGGUCCUGGUAGCC
	Exon 2 STOP	GGCCCAGCCTGCTGTGGTAC	GGCCCAGCCUGCUGUGGUAC
	Exon 2 STOP (pos 8)	GCTTCGGCAGGCTGACAGCC	GCUUCGGCAGGCUGACAGCC
	Exon 2 STOP	TATCCAAGGACTGAGGGCCA	UAUCCAAGGACUGAGGGCCA
	Exon 2 STOP	GGAACCCAGATTTATGTAAT	GGAACCCAGAUUUAUGUAAU
	Exon 2 SD	GCTCACCAATTACATAAATC	GCUCACCAAUUACAUAAAUC
	Exon 2 SD	CTCACCAATTACATAAATCT	CUCACCAAUUACAUAAAUCU
	Exon 1 STOP	CTCAGCTGAACCTGGCTACC	CUCAGCUGAACCUGGCUACC
Chi3l1	Exon 1 STOP (pos8)	GGCGTCTCAAACAGGTATCT	GGCGUCUCAAACAGGUAUCU
	Exon 1 SA (pos7)	CAAAGCCTGAAGAGAAATCC	CAAAGCCUGAAGAGAAAUCC

Gene	Description	Targeting sequence	gRNA Spacer Sequence
	Exon 3 SA (pos6)	AGAGCCTGAAGGAGA AGTCT	AGAGCCUGAAGGAGAA GUCU
	Exon 3 STOP (pos4)	TCCCAGTACCGGGAAG GCGA	UCCCAGUACCGGGAAG GCGA
	Exon 4 SA (pos6)	GGTTCCTGTGGAGCAC AGGG	GGUUCCUGUGGAGCAC AGGG
	Exon 4 SA (pos9)	TGGGGTTCCTGTGGAG CACA	UGGGGUUCCUGUGGAG CACA
	Exon 6 SA (pos8)	TCATTTCCTAGATGGG AGAC	UCAUUUCCUAGAUGGG AGAC
	Exon 6 SA (pos4)	TTCCTAGATGGGAGAC AGGC	UUCCUAGAUGGGAGAC AGGC
	Exon 8 SA (pos9)	CCAGGTGTCTGAGGAG GAAG	CCAGGUGUCUGAGGAG GAAG
	Exon 8 SA (pos5)	GTGTCTGAGGAGGAAG GGGA	GUGUCUGAGGAGGAAG GGGA
	Exon 9 SA (pos6)	TAGTCCTGGGTGGGGT AGGG	UAGUCCUGGGUGGGGU AGGG
	Exon 9 SA (pos5)	AGTCCTGGGTGGGGTA GGGT	AGUCCUGGGUGGGGUA GGGU
	Exon 9 SD (pos6)	CATTACCTCATAGTAG GCAA	CAUUACCUCAUAGUAG GCAA
	Exon 9 SD (pos7)	CCATTACCTCATAGTA GGCA	CCAUUACCUCAUAGUA GGCA
	Exon 10 SA (pos7)	ACAGATCTGAGCAGAT AACA	ACAGAUCUGAGCAGAU AACA
	Exon 10 STOP (pos 7)	TCCTACCCACTGGTTG CCCT	UCCUACCCACUGGUUG CCCU
	Exon 11 STOP (pos7)	AGGTGCAGTACCTGAA GGAC	AGGUGCAGUACCUGAA GGAC
	Exon 11 STOP (pos5)	CAGGCAGCTGGCGGGC GCCA	CAGGCAGCUGGCGGGC GCCA
	Exon 11 STOP (pos7)	GACTTCCAGGGCTCCT TCTG	GACUUCCAGGGCUCCU UCUG
CD96	Exon 1 STOP (pos5)	CATCCAGATACATTTT GTCA	CAUCCAGAUACAUUUU GUCA
	Exon 2 STOP (pos5)	ACCTGCCAAACACAGA CAGT	ACCUGCCAAACACAGA CAGU
	Exon 2 STOP (pos7)	CGTGCAGATGCAATGG TCCA	CGUGCAGAUGCAAUGG UCCA
	Exon 3 SA (pos6)	TGTAACTGTAACAAAA CATA	UGUAACUGUAACAAAA CAUA
	Exon 3 SD (pos6)	ACTTACCACCGACCAT GCAT	ACUUACCACCGACCAU GCAU
	Exon 5 SD (pos5)	CTTACCAAAAACCTTG ACTG	CUUACCAAAAACCUUG ACUG
	Exon 5 STOP (pos6)	CCAGTCCAAATCTTCG ATGA	CCAGUCCAAAUCUUCG AUGA
	Exon 5 STOP (pos7)	CAGTCCAAATCTTCGA TGAT	CAGUCCAAAUCUUCGA UGAU
	Exon 7 STOP (pos4)	AAACCATGTGATATTT GCTT	AAACCAUGUGAUAUUU GCUU

Gene	Description	Targeting sequence	gRNA Spacer Sequence
	Exon 8 STOP (pos6)	ATGTTCCACACTTTATTTCC	AUGUUCCACACUUUAUUUCC
	Exon 10 SD (pos4)	TCACGTTGAGGAGTGGTGTT	UCACGUUGAGGAGUGGUGUU
	Exon 13 SA (pos7)	CATTGTCTAGGGATATAAAG	CAUUGUCUAGGGAUAUAAAG
	Exon 13 SA (pos8)	ACATTGTCTAGGGATATAAA	ACAUUGUCUAGGGAUAUAAA
	Exon 13 SA (pos9)	GACATTGTCTAGGGATATAA	GACAUUGUCUAGGGAUAUAA
	Exon 14 STOP (pos4)	TGGCCAGGACATTCCATCTT	UGGCCAGGACAUUCCAUCUU
	Exon 15 SA (pos6)	CCATTCTAGGAACAAAATAT	CCAUUCUAGGAACAAAAUAU
Cblb	Exon 1 STOP	GAGCTTCCAAGTCTTCTCCA	GAGCUUCCAAGUCUUCUCCA
	Exon 1 STOP (JMG44)	TCCCCGAAAAGGTCGAATTT	UCCCCGAAAAGGUCGAAUUU
	Exon 2 STOP	ATGAAGAACAGTCACAGGAC	AUGAAGAACAGUCACAGGAC
	Exon 3 SA	GATTTCGTCTGTAGGCACAA	GAUUUCGUCUGUAGGCACAA
	Exon 4 SD	TAAACTTACCTGAAACAGCC	UAAACUUACCUGAAACAGCC
	Exon 4 STOP	ATTCAGACAGTGCCTTCATG	AUUCAGACAGUGCCUUCAUG
	Exon 6 STOP	GTTGCACTCGATTGGGACAG	GUUGCACUCGAUUGGGACAG
	Exon 6 STOP	TTATTTCAAGCCCTGATTGA	UUAUUUCAAGCCCUGAUUGA
	Exon 7 SD	TTACCTGTGTAACTTTTATA	UUACCUGUGUAACUUUUAUA
	Exon 8 SA (pos8)	ATTGTTCCTGGAATTTGGGG	AUUGUUCCUGGAAUUUGGGG
	Exon 8 SD (JMG48)	ATTATACCTGCCATGCCGTA	AUUAUACCUGCCAUGCCGUA
	Exon 8 SA (pos 5) (JMG46)	GTTCCTGGAATTTGGGGAGG	GUUCCUGGAAUUUGGGGAGG
	Exon 8 STOP (JMG47)	CTGCCATGCCGTAAGGCAAG	CUGCCAUGCCGUAAGGCAAG
	Exon 10 SD (JMG49)	TCTACCTTTGGTGAACCCGT	UCUACCUUUGGUGAACCCGU
	Exon 11 SD (JMG50)	CTTACCTTAGCTCCTTCTAA	CUUACCUUAGCUCCUUCUAA
	Exon 11 STOP	GGGATGTCGACTCCTAGGGG	GGGAUGUCGACUCCUAGGGG
	Exon 11 STOP	CGAGGGCACCATGCTTCAAG	CGAGGGCACCAUGCUUCAAG
	Exon 12 SD	AAACTCACTTTATGCTAGGG	AAACUCACUUUAUGCUAGGG
	Exon 12 SD (JMG51)	CTCACTTTATGCTAGGGAGG	CUCACUUUAUGCUAGGGAGG

Gene	Description	Targeting sequence	gRNA Spacer Sequence
	Exon 16 SA (JMG52)	CTTCACCTGCATTTAAAGAA	CUUCACCUGCAUUUAAAGAA
	Exon 4 STOP (JMG45)	CCACCAGATTAGCTCTGGCC	CCACCAGAUUAGCUCUGGCC
	Exon 10 SD (pos4)	CTACCTTTGGTGAACCCGTT	CUACCUUUGGUGAACCCGUU
BTLA	Exon 1 STOP (pos6)	ATGTTCCAGATGTCCAGATA	AUGUUCCAGAUGUCCAGAUA
	Exon 1 STOP (pos5)	TGTTCCAGATGTCCAGATAT	UGUUCCAGAUGUCCAGAUAU
	Exon 2 STOP (pos8)	AGATAGACAAACAAGTTGGA	AGAUAGACAAACAAGUUGGA
	Exon 2 STOP (pos9)	AGCTTGCACCAAGTCACATG	AGCUUGCACCAAGUCACAUG
	Exon 3 SD (pos6)	ACCCACCTTGGTGCCTTCTC	ACCCACCUUGGUGCCUUCUC
B2M (BE)	Exon 1 SD	ACTCACGCTGGATAGCCTCC	ACUCACGCUGGAUAGCCUCC
	Exon 2 SA (pos9)	TGGAGTACCTGAGGAATATC	UGGAGUACCUGAGGAAUAUC
	Exon 2 STOP (pos6)	TTACCCCACTTAACTATCTT	UUACCCCACUUAACUAUCUU
	Exon 3 SA	TCGATCTATGAAAAAGACAG	UCGAUCUAUGAAAAAGACAG
	Exon 2 STOP	TACCCCACTTAACTATCT	UACCCCACUUAACUAUCU
B2M (ABE)	Exon 1 SD 1 (pos 5)	ACTCACGCTGGATAGCCTCC	ACUCACGCUGGAUAGCCUCC
	Exon 2 SA (pos 4)	CTCAGGTACTCCAAAGATTC	CUCAGGUACUCCAAAGAUUC
	Exon 2 SD (pos 4)	CTTACCCCACTTAACTATCT	CUUACCCCACUUAACUAUCU
TET2	Exon 1 STOP 1 (pos 8)	CATTTGCCAGACAGAACCTC	CAUUUGCCAGACAGAACCUC
	Exon 1 STOP 2 (pos 4)	AAACAAGACCAAAAGGCTAA	AAACAAGACCAAAAGGCUAA
	Exon 1 STOP 3 (pos 7)	GTAAGCCAAGAAAGAAATCC	GUAAGCCAAGAAAGAAAUCC
	Exon 1 STOP 4 (pos 5)	GCTTCAGATTCTGAATGAGC	GCUUCAGAUUCUGAAUGAGC
	Exon 1 STOP 5 (pos 7)	TTAAAACAAAATGAAATGAA	UUAAAACAAAAUGAAAUGAA
	Exon 1 STOP 6 (pos 7)	GTTCCTCAGCTTCCTTCAGA	GUUCCUCAGCUUCCUUCAGA
	Exon 1 STOP 7 (pos 8)	CAAAGAGCAAGAGATTCTGA	CAAAGAGCAAGAGAUUCUGA
	Exon 1 STOP 8 (pos 7)	AAAGAGCAAGAGATTCTGAA	AAAGAGCAAGAGAUUCUGAA
	Exon 1 STOP 9 (pos 4)	ACACAGCACTATCTGAAACC	ACACAGCACUAUCUGAAACC
	Exon 1 STOP 10 (pos 5)	CACCCAGAAAACAACACAGC	CACCCAGAAAACAACACAGC

Gene	Description	Targeting sequence	gRNA Spacer Sequence
	Exon 1 STOP 11 (pos 4)	TACCAAGTTGAAATGAATCA	UACCAAGUUGAAAUGAAUCA
	Exon 1 STOP 12 (pos 7)	ATGAATCAAGGGCAGTCCCA	AUGAAUCAAGGGCAGUCCCA
	Exon 1 STOP 13 (pos 5)	AGGGCAGTCCCAAGGTACAG	AGGGCAGUCCCAAGGUACAG
	Exon 1 STOP 14 (pos 5)	GTTCCAAAAACCCTCACACC	GUUCCAAAAACCCUCACACC
	Exon 1 STOP 15 (pos 5)	GAAACAGCACTTGAATCAAC	GAAACAGCACUUGAAUCAAC
	Exon 1 STOP 16 (pos 5)	ATTACAAATAAAGAATAAAG	AUUACAAAUAAAGAAUAAAG
	Exon 1 STOP 17 (pos 8)	TAATGTCCAAATGGGACTGG	UAAUGUCCAAAUGGGACUGG
	Exon 1 STOP 18 (pos 6)	CAAAGCAAGATCTTCTTCAC	CAAAGCAAGAUCUUCUUCAC
	Exon 1 STOP 19 (pos 5)	ACAACAAGCTTCAGTTCTAC	ACAACAAGCUUCAGUUCUAC
	Exon 1 STOP 20 (pos 6)	CTGCGCAACTTGCTCAGCAA	CUGCGCAACUUGCUCAGCAA
	Exon 1 STOP 21 (pos 5)	CACTCAGACCCCTCCCCAGA	CACUCAGACCCCUCCCCAGA
	Exon 1 STOP 22 (pos 6)	TTTTTCCATGTTTTGTTTTC	UUUUUCCAUGUUUUGUUUUC
	Exon 1 SD (pos 4)	TTACCTACACATCTGCAAGA	UUACCUACACAUCUGCAAGA
	Exon 3 SD (pos 8)	ACACTTACCCACTTAGCAAT	ACACUUACCCACUUAGCAAU
	Exon 7 STOP (pos 5)	CATGCAGAATGGCAGCACAT	CAUGCAGAAUGGCAGCACAU
	Exon 8 STOP 1 (pos 6)	AAGCTCAGGAGGAGAAAAAA	AAGCUCAGGAGGAGAAAAAA
	Exon 8 STOP 2 (pos 8)	CGCAAGCCAGGCTAAACAGT	CGCAAGCCAGGCUAAACAGU
	Exon 9 STOP 1 (pos 8)	TTCTCCCCAGTCTCAGCCGA	UUCUCCCCAGUCUCAGCCGA
	Exon 9 STOP 2 (pos 5)	TGGTCAGGAAAAGCAGCCAT	UGGUCAGGAAAAGCAGCCAU
	Exon 9 STOP 3 (pos 7)	CTAGTCCAGGGTGTGGCTTC	CUAGUCCAGGGUGUGGCUUC
Spry1	Exon 1 STOP 1	CCCCAAAATCAACATGGCAG	CCCCAAAAUCAACAUGGCAG
	Exon 1 STOP 2	TGTGATCCAGCAGCCTTCTT	UGUGAUCCAGCAGCCUUCUU
	Exon 1 STOP 3	GACCAGATCAAGGCCATAAG	GACCAGAUCAAGGCCAUAAG
	Exon 1 STOP 4	CAAGACAAGAAAAGCATGAA	CAAGACAAGAAAAGCAUGAA
	Exon 1 STOP 5	CTGAACAGGGACTGTTAGGA	CUGAACAGGGACUGUUAGGA

Gene	Description	Targeting sequence	gRNA Spacer Sequence
Spry2	Exon 1 STOP 1	CCAGAGCTCAGAGTGG CAAC	CCAGAGCUCAGAGUGG CAAC
	Exon 1 STOP 2	TTGCTGCAGACGCCCC GTGA	UUGCUGCAGACGCCCC GUGA
	Exon 1 STOP 3	CTGCAGACGCCCCGTG ACGG	CUGCAGACGCCCCGUG ACGG
	Exon 1 STOP 4	CGACAAGCAGTGCCTT TGCT	CGACAAGCAGUGCCUU UGCU
	Exon 1 STOP 5	GCCCAGAACGTGATTG ACTA	GCCCAGAACGUGAUUG ACUA
	Exon 1 STOP 6	TGTGCCAGGGGTGTTA TGAC	UGUGCCAGGGGUGUUA UGAC
	Exon 1 STOP 7	CAGATCCAGTCTGATG GCAG	CAGAUCCAGUCUGAUG GCAG
	Exon 1 STOP 8	TGTACACGATGGTCAG CCAT	UGUACACGAUGGUCAG CCAU
CIITA	Exon 1 SD (pos 6)	TTTTACCTTGGGGCTCT GAC	UUUUACCUUGGGGCUC UGAC
	Exon 1 STOP 1 (pos 6)	AGCCCCAAGGTAAAA AGGCC	AGCCCCAAGGUAAAAA GGCC
	Exon 1 STOP 2 (pos 7)	GAGCCCCAAGGTAAA AAGGC	GAGCCCCAAGGUAAAA AGGC
	Exon 2 STOP 1 (pos 8)	CAGCTCACAGTGTGCC ACCA	CAGCUCACAGUGUGCC ACCA
	Exon 2 STOP 2 (pos 7)	TATGACCAGATGGACC TGGC	UAUGACCAGAUGGACC UGGC
	Exon 4 STOP 1 (pos 8)	ACTGGACCAGTATGTC TTCC	ACUGGACCAGUAUGUC UUCC
	Exon 4 STOP 2 (pos 8)	TGTCTTCCAGGACTCC CAGC	UGUCUUCCAGGACUCC CAGC
	Exon 7 STOP 1 (pos 7)	TTCAACCAGGAGCCAG CCTC	UUCAACCAGGAGCCAG CCUC
	Exon 7 STOP 2 (pos 4)	GACCAGATTCCCAGTA TGTT	GACCAGAUUCCCAGUA UGUU
	Exon 7 SD (pos 8)	TAACATACTGGGAATC TGGT	UAACAUACUGGGAAUC UGGU
	Exon 8 SA (pos 8)	AAAGGCACTGCAAGA GACAA	AAAGGCACUGCAAGAG ACAA
	Exon 8 STOP (pos 8)	CTCTGGCAAATCTCTG AGGC	CUCUGGCAAAUCUCUG AGGC
	Exon 9 STOP 1 (pos 4)	AGCCAAGTACCCCCTC CCAG	AGCCAAGUACCCCCUC CCAG
	Exon 9 STOP 2 (pos 7)	ACCTCCCGAGCAAACA TGAC	ACCUCCCGAGCAAACA UGAC
	Exon 9 SD (pos 6)	CCTTACCTGTCATGTTT GCT	CCUUACCUGUCAUGUU UGCU
	Exon 10 SA (pos 5)	TGCTCTGGAGATGGAG AAGC	UGCUCUGGAGAUGGAG AAGC
	Exon 10 STOP 1 (pos 7)	CCCACCCAATGCCCGG CAGC	CCCACCCAAUGCCCGG CAGC
	Exon 10 STOP 2 (pos 4)	AGGCCATTTTGGAAGC TTGT	AGGCCAUUUUGGAAGC UUGU

Gene	Description	Targeting sequence	gRNA Spacer Sequence
	Exon 11 SA (pos 8)	ACCGGCTCTGCAAAGGCCAG	ACCGGCUCUGCAAAGGCCAG
	Exon 11 STOP 1 (pos 6)	TGGTGCAGGCCAGGCTGGAG	UGGUGCAGGCCAGGCUGGAG
	Exon 11 STOP 3 (pos 7)	GAACGGCAGCTGGCCCAAGG	GAACGGCAGCUGGCCCAAGG
	Exon 11 STOP 4 (pos 5)	GGCCCAAGGAGGCCTGGCTG	GGCCCAAGGAGGCCUGGCUG
	Exon 11 STOP 5 (pos 5)	GACACGAGTGATTGCTGTGC	GACACGAGUGAUUGCUGUGC
	Exon 11 STOP 5 (pos 6)	CTGGTCAGGGCAAGAGCTAT	CUGGUCAGGGCAAGAGCUAU
	Exon 11 STOP 5 (pos 8)	GGGCCCACAGCCACTCGTGG	GGGCCCACAGCCACUCGUGG
	Exon 11 STOP 6 (pos 4)	TTCCAGAAGAAGCTGCTCCG	UUCCAGAAGAAGCUGCUCCG
	Exon 11 STOP 7 (pos 8)	CCTGGTCCAGAGCCTGAGCA	CCUGGUCCAGAGCCUGAGCA
	Exon 11 STOP 8 (pos 8)	CAGACATCAAAGTACCCTAC	CAGACAUCAAAGUACCCUAC
	Exon 11 STOP 9 (pos 5)	ACATCAAAGTACCCTACAGG	ACAUCAAAGUACCCUACAGG
	Exon 11 STOP 10 (pos 4)	CGCCCAGGTCCTCACGTCTG	CGCCCAGGUCCUCACGUCUG
	Exon 11 STOP 11 (pos 8)	CTTAGTCCAACACCCACCGC	CUUAGUCCAACACCCACCGC
	Exon 11 STOP 12 (pos 8)	CCTCCTGCAATGCTTCCTGG	CCUCCUGCAAUGCUUCCUGG
	Exon 11 STOP 13 (pos 8)	GAGCCAGCCACAGGGCCCCC	GAGCCAGCCACAGGGCCCCC
	Exon 11 STOP 14 (pos 6)	GGAAGCAGAAGGTGCTTGCG	GGAAGCAGAAGGUGCUUGCG
	Exon 11 STOP 15 (pos 6)	GGCTGCAGCCGGGGACACTG	GGCUGCAGCCGGGGACACUG
	Exon 11 STOP 16 (pos 4)	CTGCCAAATTCCAGCCTCCT	CUGCCAAAUUCCAGCCUCCU
	Exon 11 STOP 17 (pos 8)	GGCGGGCCAAGACTTCTCCC	GGCGGGCCAAGACUUCUCCC
	Exon 12 STOP 1 (pos 6)	AGACTCAGAGGTGAGAGGAG	AGACUCAGAGGUGAGAGGAG
	Exon 14 SA (pos 4)	AGCCTAGGAGGCAAAGAGCA	AGCCUAGGAGGCAAAGAGCA
	Exon 14 STOP 1 (pos 5)	CCCCCAGGCTTTCCCCAAAC	CCCCCAGGCUUUCCCCAAAC
	Exon 14 SD (pos 4)	TCACTCCAGATGCTGCAGGG	UCACUCCAGAUGCUGCAGGG
	Exon 15 SA (pos 4)	AGGCTGCAGGTGGAATCAGA	AGGCUGCAGGUGGAAUCAGA
	Exon 15 STOP 1 (pos 8)	CTTCCCCCAGCTGAAGTCCT	CUUCCCCCAGCUGAAGUCCU
	Exon 15 SD (pos 7)	CACTCACTTGAGGGTTTCCA	CACUCACUUGAGGGUUUCCA

Gene	Description	Targeting sequence	gRNA Spacer Sequence
	Exon 16 SA (pos 5)	CAGACTGCGGGGACAC AGTG	CAGACUGCGGGGACAC AGUG
	Exon 16 SD 1 (pos 8)	CCACTCACCTTAGCCT GAGC	CCACUCACCUUAGCCU GAGC
	Exon 16 SD 2 (pos 7)	CACTCACCTTAGCCTG AGCA	CACUCACCUUAGCCUG AGCA
	Exon 17 SA (pos 8)	GTACAAGCTGTCGGAA ACAG	GUACAAGCUGUCGGAA ACAG
	Exon 17 SD 1 (pos 8)	ACACTCACTCCATCAC CCGG	ACACUCACUCCAUCAC CCGG
	Exon 17 SD 2 (pos 7)	CACTCACTCCATCACC CGGA	CACUCACUCCAUCACC CGGA
	Exon 18 STOP (pos 5)	CGTCCAGTACAACAAG TTCA	CGUCCAGUACAACAAG UUCA
	Exon 19 SA 1 (pos 8)	CCACATCCTGCAAGGG GGGA	CCACAUCCUGCAAGGG GGGA
	Exon 19 SA 2 (pos 7)	CACATCCTGCAAGGGG GGAT	CACAUCCUGCAAGGGG GGAU
	Exon 19 STOP 1 (pos 8)	TGGGCGTCCACATCCT GCAA	UGGGCGUCCACAUCCU GCAA
	Exon 19 STOP 2 (pos 7)	GGGCGTCCACATCCTG CAAG	GGGCGUCCACAUCCUG CAAG
	Exon 19 STOP 3 (pos 6)	GGCGTCCACATCCTGC AAGG	GGCGUCCACAUCCUGC AAGG
	Exon 19 STOP 4 (pos 5)	GCGTCCACATCCTGCA AGGG	GCGUCCACAUCCUGCA AGGG
CD7	Exon 1 STOP (pos 4)	GCCCAAGGTAAGAGCT TCCC	GCCCAAGGUAAGAGCU UCCC
	Exon 1 SD 1 (pos 8)	GCTCTTACCTTGGGCA GCCA	GCUCUUACCUUGGGCA GCCA
	Exon 1 SD 2 (pos 9)	AGCTCTTACCTTGGGC AGCC	AGCUCUUACCUUGGGC AGCC
	Exon 2 SA 1 (pos 8)	TGCACCTCTGGGGAGG ACCT	UGCACCUCUGGGGAGG ACCU
	Exon 2 SA 2 (pos 9)	CTGCACCTCTGGGGAG GACC	CUGCACCUCUGGGGAG GACC
	Exon 2 STOP 1 (pos 7)	CGCCTGCAGCTGTCGG ACAC	CGCCUGCAGCUGUCGG ACAC
	Exon 2 STOP 2 (pos 8)	CACCTGCCAGGCCATC ACGG	CACCUGCCAGGCCAUC ACGG
	Exon 2 SD 1 (pos 6)	CCCTACCTGTCACCAG GACC	CCCUACCUGUCACCAG GACC
	Exon 2 SD 2 (pos 5)	CCTACCTGTCACCAGG ACCA	CCUACCUGUCACCAGG ACCA
	Exon 3 SA (pos 4)	CCTCTGAGAAGGAAAA AAGA	CCUCUGAGAAGGAAAA AAGA
	Exon 3 STOP 1 (pos9)	CAGAGGAACAGTCCCA AGGA	CAGAGGAACAGUCCCA AGGA
CD33	Exon 1 SD 1 (pos 7)	CACTCACCTGCCCACA GCAG	CACUCACCUGCCCACA GCAG
	Exon 1 SD 2 (pos 8)	CCACTCACCTGCCCAC AGCA	CCACUCACCUGCCCAC AGCA

Gene	Description	Targeting sequence	gRNA Spacer Sequence
	Exon 1 SD (pos 9)	GCCACTCACCTGCCCA CAGC	GCCACUCACCUGCCCA CAGC
	Exon 2 SA 1 (pos 8)	AGGGCCCCTGTGGGGA AACG	AGGGCCCCUGUGGGGA AACG
	Exon 2 SA 2 (pos 7)	GGGCCCCTGTGGGGAA ACGA	GGGCCCCUGUGGGGAA ACGA
	Exon 2 STOP 1 (pos 8)	GCAAGTGCAGGAGTCA GTGA	GCAAGUGCAGGAGUCA GUGA
	Exon 2 STOP 2 (pos 6)	CGGAACCAGTAACCAT GAAC	CGGAACCAGUAACCAU GAAC
	Exon 2 STOP 3 (pos 5)	GGAACCAGTAACCATG AACT	GGAACCAGUAACCAUG AACU
	Exon 2 STOP 4 (pos 4)	GAACCAGTAACCATGA ACTG	GAACCAGUAACCAUGA ACUG
	Exon 2 STOP 5 (pos 8)	GCTAGATCAAGAAGTA CAGG	GCUAGAUCAAGAAGUA CAGG
	Exon 2 STOP 6 (pos 8)	AGAAGTACAGGAGGA GACTC	AGAAGUACAGGAGGAG ACUC
	Exon 3 SA 1 (pos 6)	CAAGTCTAGTGAGGAG AAAG	CAAGUCUAGUGAGGAG AAAG
	Exon 3 SA 2 (pos 5)	AAGTCTAGTGAGGAGA AAGA	AAGUCUAGUGAGGAGA AAGA
	Exon 3 SA 3 (pos 4)	AGTCTAGTGAGGAGAA AGAG	AGUCUAGUGAGGAGAA AGAG
	Exon 3 STOP 1 (pos 7)	ACAGGCCCAGGACAC AGAGC	ACAGGCCCAGGACACA GAGC
	Exon 3 STOP 2 (pos 7)	ACCTGTCAGGTGAAGT TCGC	ACCUGUCAGGUGAAGU UCGC
	Exon 3 SD 1 (pos 6)	ACTTACAGGTGACGTT GAGC	ACUUACAGGUGACGUU GAGC
	Exon 4 SA 1 (pos 6)	AACATCTAGGAGAGG AAGAG	AACAUCUAGGAGAGGA AGAG
	Exon 4 STOP 1 (pos 7)	GTTCCACAGAACCCAA CAAC	GUUCCACAGAACCCAA CAAC
	Exon 4 SD 1 (pos 7)	TTCCTACCTGAGCCAT CTCC	UUCCUACCUGAGCCAU CUCC
	Exon 5 SD (pos 8)	ATGCTCACATGAAGAA GATG	AUGCUCACAUGAAGAA GAUG
	Exon 5 STOP 1 (pos 7)	GGGAAACAAGAGACC AGAGC	GGGAAACAAGAGACCA GAGC
	Exon 6 SA 1 (pos 6)	TCACTCTGATGGGAGA CACC	UCACUCUGAUGGGAGA CACC
	Exon 6 SA 2 (pos 5)	CACTCTGATGGGAGAC ACCA	CACUCUGAUGGGAGAC ACCA
	Exon 6 SA 1 (pos 4)	TTTCTTATGGAGAGGA AAGA	UUUCUUAUGGAGAGGA AAGA
CD52	Exon 1 STOP (pos 4)	GTACAGGTAAGAGCA ACGCC	GUACAGGUAAGAGCAA CGCC
	Exon 1 SD (pos7)	CTCTTACCTGTACCAT AACC	CUCUUACCUGUACCAU AACC
	Exon 1 SD (pos 4)	TTACCTGTACCATAAC CAGG	UUACCUGUACCAUAAC CAGG

Gene	Description	Targeting sequence	gRNA Spacer Sequence
	Exon 2 SA (pos 6)	TGTATCTGTAGGAGGA GAAG	UGUAUCUGUAGGAGGA GAAG
	Exon 2 SA (pos 5)	GTATCTGTAGGAGGAG AAGT	GUAUCUGUAGGAGGAG AAGU
	Exon 2 STOP (pos 7)	CAGATACAAACTGGAC TCTC	CAGAUACAAACUGGAC UCUC
CD123	Exon 1 SD (pos 6)	TCTTACCTTCCTTCGTT TGC	UCUUACCUUCCUUCGU UUGC
	Exon 2 SA 1 (pos 8)	TTTGGATCTAAAACGG TGAC	UUUGGAUCUAAAACGG UGAC
	Exon 2 SA 2 (pos 4)	GATCTAAAACGGTGAC AGGT	GAUCUAAAACGGUGAC AGGU
	Exon 2 STOP 1 (pos 8)	AAAGGCTCAGCAGTTG ACCT	AAAGGCUCAGCAGUUG ACCU
	Exon 2 SD (pos 6)	ATTTACCGGCATAGAA TAGT	AUUUACCGGCAUAGAA UAGU
	Exon 3 SA (pos 8)	TCACTGCCTAAGAGAG ACAT	UCACUGCCUAAGAGAG ACAU
	Exon 3 STOP 1 (pos 6)	AGGATCCACGTGGAGA ATGG	AGGAUCCACGUGGAGA AUGG
	Exon 3 STOP 2 (pos 5)	GGATCCACGTGGAGAA TGGT	GGAUCCACGUGGAGAA UGGU
	Exon 3 SD (pos 6)	TCTCACTGTTCTCAGG GAAG	UCUCACUGUUCUCAGG GAAG
	Exon 4 STOP 1 (pos 6)	CCTGCCCAAGGCTTCC CACC	CCUGCCCAAGGCUUCC CACC
	Exon 4 STOP 2 (pos 5)	CTGCCCAAGGCTTCCC ACCT	CUGCCCAAGGCUUCCC ACCU
	Exon 5 SA 1 (pos 6)	GCCTGCTGCGGTAAGC GGTA	GCCUGCUGCGGUAAGC GGUA
	Exon 5 STOP 1 (pos 7)	GATGCTCAGGGAACAC GTAT	GAUGCUCAGGGAACAC GUAU
	Exon 5 STOP 2 (pos 5)	TTCTCAAAGTTCCCAC ATCC	UUCUCAAAGUUCCCAC AUCC
	Exon 5 STOP 3 (pos 4)	TCACAGATTGGTGAGT AGCC	UCACAGAUUGGUGAGU AGCC
	Exon 7 SD (pos 5)	CTCACCTGTTCTGTGA TTAC	CUCACCUGUUCUGUGA UUAC
	Exon 8 STOP 1 (pos 7)	TCCTTCCAGCTACTCA ATCC	UCCUUCCAGCUACUCA AUCC
	Exon 8 STOP 2 (pos 8)	CACAGTACAAATAAGA GCCC	CACAGUACAAAUAAGA GCCC
	Exon 8 STOP 3 (pos 6)	CCCCCCAGCGCTTCGG TGAG	CCCCCCAGCGCUUCGG UGAG
	Exon 8 STOP 4 (pos 5)	CCCCCAGCGCTTCGGT GAGT	CCCCCAGCGCUUCGGU GAGU
	Exon 8 SD (pos 8)	CCACTCACCGAAGCGC TGGG	CCACUCACCGAAGCGC UGGG
	Exon 10 SA (pos 4)	TACCTCGGAGGAAAGA GAAA	UACCUCGGAGGAAAGA GAAA
	Exon 10 STOP (pos 8)	CAGCTTCCAAAACGAC AAGC	CAGCUUCCAAAACGAC AAGC

Gene	Description	Targeting sequence	gRNA Spacer Sequence
	Exon 10 SD (pos 7)	AACATACCAGCTTGTC GTTT	AACAUACCAGCUUGUC GUUU
	Exon 11 SA 1 (pos 8)	AGACCACCTGCAGAGA CGAG	AGACCACCUGCAGAGA CGAG
	Exon 11 SA 2 (pos 5)	CCACCTGCAGAGACGA GAGG	CCACCUGCAGAGACGA GAGG
TRBC1	Exon 1 STOP 1 (pos 8)	CCACACCCAAAAGGCC ACAC	CCACACCCAAAAGGCC ACAC
	Exon 1 STOP 2 (pos 5)	CCCACCAGCTCAGCTC CACG	CCCACCAGCUCAGCUC CACG
	Exon 1 STOP 3 (pos 7)	CGCTGTCAAGTCCAGT TCTA	CGCUGUCAAGUCCAGU UCUA
	Exon 1 STOP 4 (pos 6)	GCTGTCAAGTCCAGTT CTAC	GCUGUCAAGUCCAGUU CUAC
	Exon 1 STOP 5 (pos 5)	CACCCAGATCGTCAGC GCCG	CACCCAGAUCGUCAGC GCCG
	Exon 1 SD (pos 8)	CCACTCACCTGCTCTA CCCC	CCACUCACCUGCUCUA CCCC
	Exon 2 SA (pos 8)	CCACAGTCTGAAAGAA AGCA	CCACAGUCUGAAAGAA AGCA
	Exon 3 SA (pos 5)	GACACTGTTGGCACGG AGGA	GACACUGUUGGCACGG AGGA
	Exon 3 SD (pos 4)	TTACCATGGCCATCAA CACA	UUACCAUGGCCAUCAA CACA
TRBC2	Exon 1 STOP 1 (pos 8)	CCACACCCAAAAGGCC ACAC	CCACACCCAAAAGGCC ACAC
	Exon 1 STOP 2 (pos 5)	CCCACCAGCTCAGCTC CACG	CCCACCAGCUCAGCUC CACG
	Exon 1 STOP 3 (pos 7)	CGCTGTCAAGTCCAGT TCTA	CGCUGUCAAGUCCAGU UCUA
	Exon 1 STOP 4 (pos 6)	GCTGTCAAGTCCAGTT CTAC	GCUGUCAAGUCCAGUU CUAC
	Exon 1 STOP 5 (pos 5)	CACCCAGATCGTCAGC GCCG	CACCCAGAUCGUCAGC GCCG
	Exon 2 SA (pos 8)	CCACAGTCTGAAAGAA AACA	CCACAGUCUGAAAGAA AACA
	Exon 2 SA (pos 7)	CACAGTCTGAAAGAAA ACAG	CACAGUCUGAAAGAAA ACAG
	Exon 3 SD (pos 4)	TTACCATGGCCATCAG CACG	UUACCAUGGCCAUCAG CACG
	Exon 1 SD (pos 8)	CCACTCACCTGCTCTA CCCC	CCACUCACCUGCUCUA CCCC
CISH	Exon 1 STOP	TCTGCGTTCAGGGGTA AGCG	UCUGCGUUCAGGGGUA AGCG
	Exon 1 SD	GCGCTTACCCCTGAAC GCAG	GCGCUUACCCCUGAAC GCAG
	Exon 2 STOP 2	GACTGGGCAGCGGCCC CTGT	GACUGGGCAGCGGCCC CUGU
	Exon 2 STOP 1	GGACTGGGCAGCGGCC CCTG	GGACUGGGCAGCGGCC CCUG
	Exon 2 STOP 3	GTCATGCAGCCCTTGC CTGC	GUCAUGCAGCCCUUGC CUGC

Gene	Description	Targeting sequence	gRNA Spacer Sequence
	Exon 2 STOP 4	TCATGCAGCCCTTGCCTGCT	UCAUGCAGCCCUUGCCUGCU
	Exon 2 STOP 5	CATGCAGCCCTTGCCTGCTG	CAUGCAGCCCUUGCCUGCUG
	Exon 2 SD 1	CTCACCAGATTCCCGAAGGT	CUCACCAGAUUCCCGAAGGU
	Exon 2 SD 2	CAGACTCACCAGATTCCCGA	CAGACUCACCAGAUUCCCGA
	Exon 3 SA 1 (pos 4)	AGCCTAGGCAAGTGCAGAGG	AGCCUAGGCAAGUGCAGAGG
	Exon 3 SA 2 (pos 5)	CAGCCTAGGCAAGTGCAGAG	CAGCCUAGGCAAGUGCAGAG
	Exon 3 SA 3 (pos 7)	ACCAGCCTAGGCAAGTGCAG	ACCAGCCUAGGCAAGUGCAG
	Exon 3 STOP 1 (pos 8)	TGGAACCCCAATACCAGCCT	UGGAACCCCAAUACCAGCCU
	Exon 3 STOP 2 (pos 7)	CACCTGCAGAAGATGCCAGA	CACCUGCAGAAGAUGCCAGA
ACAT1	Exon 1 SD 1 (pos 7)	CGCTCACCTGCACCAGCCTC	CGCUCACCUGCACCAGCCUC
	Exon 3 SA (pos 5)	CTTCCTGGCAAGACACAAGA	CUUCCUGGCAAGACACAAGA
	Exon 3 STOP (pos 5)	AATTCAGGGAGCCATTGAAA	AAUUCAGGGAGCCAUUGAAA
	Exon 3 SD (pos 8)	CTACTGACCTGCCTTTTCAA	CUACUGACCUGCCUUUUCAA
	Exon 5 STOP (pos 7)	GCCTCTCAAAGTCTTATGTG	GCCUCUCAAAGUCUUAUGUG
	Exon 7 STOP (pos 4)	TTCCCATGCTGCTTTACTTC	UUCCCAUGCUGCUUUACUUC
	Exon 8 STOP (pos 8)	TTTAGGTCAACCAGATGTAG	UUUAGGUCAACCAGAUGUAG
	Exon 9 SA (pos 9)	TGTGCCTGAAAGCAAAAATG	UGUGCCUGAAAGCAAAAAUG
	Exon 9 SD (pos 4)	TTACCTACTATTCTTGCCAG	UUACCUACUAUUCUUGCCAG
	Exon 10 SA (pos 6)	AAATGCTGTTTAAAAAAAGG	AAAUGCUGUUUAAAAAAAGG
	Exon 11 STOP (pos 4)	CCCCAAAAAGTGAATATCAA	CCCCAAAAAGUGAAUAUCAA
Cyp11a1	Exon 1 STOP 1 (pos 4)	GTCCAGAATTTCCAGAAGTA	GUCCAGAAUUUCCAGAAGUA
	Exon 2 SA 1 (pos 4)	TCCCTGGAGGGGTGGGGGAG	UCCCUGGAGGGGUGGGGGAG
	Exon 2 SD 1 (pos 4)	TCACTTCAACAGGACTCCTA	UCACUUCAACAGGACUCCUA
	Exon 3 SD 1 (pos 6)	CCTTACACTCAAAGGCAAAG	CCUUACACUCAAAGGCAAAG
	Exon 4 SA (pos 5)	ATGGCTGCAGGGAGAGGAAG	AUGGCUGCAGGGAGAGGAAG
	Exon 4 STOP 1 (pos 8)	GGAGCGCCAGGGGAT	GGAGCGCCAGGGGAUG

Gene	Description	Targeting sequence	gRNA Spacer Sequence
		GCTGG	CUGG
	Exon 4 STOP 2 (pos 8)	TCACGTCCCATGCAGC CACA	UCACGUCCCAUGCAGC CACA
	Exon 6 SA (pos 8)	TGGACGTCTGGTGGGG AGTA	UGGACGUCUGGUGGGG AGUA
	Exon 8 STOP 1(pos 6)	ACTCACATTGATGAGG AAGA	ACUCACAUUGAUGAGG AAGA
	Exon 9 SA (pos 7)	CAGCATCTGAGAAAGG CAGA	CAGCAUCUGAGAAAGG CAGA
	Exon 9 STOP 1 (pos 5)	AATCCAACACCTCAGC GATG	AAUCCAACACCUCAGC GAUG
	Exon 9 STOP 2 (pos 4)	ATCCAACACCTCAGCG ATGT	AUCCAACACCUCAGCG AUGU
GATA3	Exon 1 STOP 1 (pos 8)	CGCGGCGCAGTACCCG CTGC	CGCGGCGCAGUACCCG CUGC
	Exon 1 SD 1 (pos 7)	CACTCACCGTGGTGGG TCGG	CACUCACCGUGGUGGG UCGG
	Exon 1 SD 2 (pos 6)	ACTCACCGTGGTGGGT CGGA	ACUCACCGUGGUGGGU CGGA
	Exon 2 SA 1 (pos 8)	TGGCTCCCTGTGGGGC AACG	UGGCUCCCUGUGGGGC AACG
	Exon 2 STOP 2 (pos 5)	GATTCCAGGGGGAGGC GGTG	GAUUCCAGGGGGAGGC GGUG
	Exon 2 SD 1 (pos 8)	GCTCCTACCTGTGCTG GACC	GCUCCUACCUGUGCUG GACC
	Exon 3 STOP 1 (pos 7)	TCGCCGCCACAGTGGG GTCG	UCGCCGCCACAGUGGG GUCG
	Exon 4 SA (pos 5)	CAGACTGAGAGTGGG GAGAG	CAGACUGAGAGUGGGG AGAG
	Exon 4 STOP 1 (pos 7)	CCTCCTCCAGAGTGTG GTTG	CCUCCUCCAGAGUGUG GUUG
NR4A1	Exon 1 STOP 1 (pos 8)	AGCCATCCCAGGGAGA GAGC	AGCCAUCCCAGGGAGA GAGC
	Exon 1 STOP 2 (pos 7)	GCCATCCCAGGGAGAG AGCT	GCCAUCCCAGGGAGAG AGCU
	Exon 1 STOP 3 (pos 6)	CCATCCCAGGGAGAGA GCTG	CCAUCCCAGGGAGAGA GCUG
	Exon 1 STOP 4 (pos 5)	CTCACAGGCCACCCAC CAGC	CUCACAGGCCACCCAC CAGC
	Exon 2 STOP 1 (pos 8)	CCGCTTCCAGAAGTGC CTGG	CCGCUUCCAGAAGUGC CUGG
	Exon 2 STOP 2 (pos 5)	CTTCCAGAAGTGCCTG GCGG	CUUCCAGAAGUGCCUG GCGG
	Exon 3 SA 1 (pos 5)	ACAACTGCAAAGGAAT GGGT	ACAACUGCAAAGGAAU GGGU
	Exon 3 SA 2 (pos 4)	CAACTGCAAAGGAATG GGTA	CAACUGCAAAGGAAUG GGUA
	Exon 4 SA (pos 4)	GAACTAGGAAGACGG TCCAG	GAACUAGGAAGACGGU CCAG
	Exon 4 STOP 1 (pos 8)	GGCTGACCAGGACCTG TTGC	GGCUGACCAGGACCUG UUGC

Gene	Description	Targeting sequence	gRNA Spacer Sequence
	Exon 4 SD 1 (pos 5)	CTCACCTGTACGCCAGGCGG	CUCACCUGUACGCCAGGCGG
	Exon 4 SD 2 (pos 8)	GCTCTCACCTGTACGCCAGG	GCUCUCACCUGUACGCCAGG
	Exon 5 SA (pos 8)	CTTAGACCTGGCAGGCAGAT	CUUAGACCUGGCAGGCAGAU
	Exon 5 STOP 1 (pos 5)	CAATCCAGTCCCCGAAGCCA	CAAUCCAGUCCCCGAAGCCA
	Exon 5 STOP 2 (pos 4)	AATCCAGTCCCCGAAGCCAC	AAUCCAGUCCCCGAAGCCAC
	Exon 5 SD 1 (pos 6)	ACTCACCGGTGATGAGGACA	ACUCACCGGUGAUGAGGACA
	Exon 5 SD 2 (pos 5)	CTCACCGGTGATGAGGACAA	CUCACCGGUGAUGAGGACAA
	Exon 6 SA (pos 6)	CCGGTCTGCGGGAAGGGTAC	CCGGUCUGCGGGAAGGGUAC
	Exon 6 STOP 1 (pos 8)	TGGGCTGCAGGAGCCGCGGC	UGGGCUGCAGGAGCCGCGGC
NR4A2	Exon 1 STOP 1 (pos 7)	TTGTACCAAATGCCCCTGTC	UUGUACCAAAUGCCCCUGUC
	Exon 1 STOP 2 (pos 8)	CGGACAGCAGTCCTCCATTA	CGGACAGCAGUCCUCCAUUA
	Exon 1 STOP 3 (pos 6)	AGGTGCAGCACAGCCCCATG	AGGUGCAGCACAGCCCCAUG
	Exon 1 STOP 4 (pos 5)	GGTGCAGCACAGCCCCATGT	GGUGCAGCACAGCCCCAUGU
	Exon 1 STOP 5 (pos 7)	AGTTGCCAGATGCGCTTCGA	AGUUGCCAGAUGCGCUUCGA
	Exon 1 STOP 6 (pos 6)	GTTGCCAGATGCGCTTCGAC	GUUGCCAGAUGCGCUUCGAC
	Exon 1 STOP 7 (pos 5)	GTCTCAGCTGCTCGACACGC	GUCUCAGCUGCUCGACACGC
	Exon 3 SD (pos 7)	TTCTTACCCTGGAATAGTCC	UUCUUACCCUGGAAUAGUCC
	Exon 4 SD (pos 5)	ATTACCTGTATGCTAATCGA	AUUACCUGUAUGCUAAUCGA
	Exon 5 STOP 1 (pos 4)	TTGCAATGCGTTCGTGGCTT	UUGCAAUGCGUUCGUGGCUU
	Exon 5 SD (pos 6)	ACTGACCTGTGACCATAGCC	ACUGACCUGUGACCAUAGCC
NR4A3	Exon 2 SA (pos 4)	TATCTGCAGGGACAGAGAAA	UAUCUGCAGGGACAGAGAAA
	Exon 2 STOP 1 (pos 8)	TGCGGCGCAGACATACAGCT	UGCGGCGCAGACAUACAGCU
	Exon 2 STOP 2 (pos 6)	CCCCGCAGGCGGGGGCGTTA	CCCCGCAGGCGGGGGCGUUA
	Exon 3 STOP 1 (pos 4)	TTTCAGAAGTGTCTCAGTGT	UUUCAGAAGUGUCUCAGUGU
	Exon 5 SD (pos 5)	ATTACCTGATGGAAAGTCTG	AUUACCUGAUGGAAAGUCUG
	Exon 6 STOP 1 (pos 4)	CTTCAGTGCCTTCGTG	CUUCAGUGCCUUCGUG

Gene	Description	Targeting sequence	gRNA Spacer Sequence
		GATT	GAUU
	Exon 7 SA (pos 4)	TTTCTGCAGAGGGATAGAGA	UUUCUGCAGAGGGAUAGAGA
	Exon 7 STOP 1 (pos 8)	AGACCACCAGAGTAAGGGAC	AGACCACCAGAGUAAGGGAC
MCJ	Exon 1 STOP (pos 6)	ACTTGCAGCCCTCGGCCAAA	ACUUGCAGCCCUCGGCCAAA
FAS	Exon 1 SD (pos 9)	AGGGCTCACCAGAGGTAGGA	AGGGCUCACCAGAGGUAGGA
	Exon 3 SA (pos 6)	TTCACCTGCCCAAGGAAAAA	UUCACCUGCCCAAGGAAAAA
	Exon 4 SA (pos 7)	CTAAGCCTAGAAAATCAGTT	CUAAGCCUAGAAAAUCAGUU
	Exon 5 SA (pos 5)	ACATCTAGAAAAAAAAATAC	ACAUCUAGAAAAAAAAAUAC
	Exon 5 SD (pos 5)	ATTACCTTCCTCTTTGCACT	AUUACCUUCCUCUUUGCACU
	Exon 6 SA (pos 5)	GATCCTGTAGGTTGGAACAT	GAUCCUGUAGGUUGGAACAU
	Exon 6 STOP 1 (pos 4)	AAGCCACCCCAAGTTAGATC	AAGCCACCCCAAGUUAGAUC
	Exon 6 SD (pos 7)	AACTTACCCCAAACAATTAG	AACUUACCCCAAACAAUUAG
	Exon 7 SD (pos 8)	ATACCTACAGGATTTAAAGT	AUACCUACAGGAUUUAAAGU
	Exon 8 SA (pos 8)	GTTTCCTAGAAAGCAAAAAA	GUUUCCUAGAAAGCAAAAAA
	Exon 9 STOP 1 (pos 6)	AAGTTCAACTGCTTCGTAAT	AAGUUCAACUGCUUCGUAAU
	Exon 9 STOP (pos 5)	AATTCAGACTATCATCCTCA	AAUUCAGACUAUCAUCCUCA
SELPG/PSGL1	Exon1 STOP 1 (pos 6)	GCTTGCAGCTGTGGGACACC	GCUUGCAGCUGUGGGACACC
	Exon1 STOP 2 (pos 8)	GACCACTCAACCAGTGCCCA	GACCACUCAACCAGUGCCCA
	Exon1 STOP 3 (pos 8)	GGAGGCACAGACCACTCCAC	GGAGGCACAGACCACUCCAC
	Exon1 STOP 4 (pos 5)	GGCACAGACAACTCGACTGA	GGCACAGACAACUCGACUGA
	Exon1 STOP 5 (pos 8)	GGAGGCACAGACCACTCCAC	GGAGGCACAGACCACUCCAC
	Exon1 STOP 6 (pos 4)	GCACAGACCACTCAACCCAC	GCACAGACCACUCAACCCAC
	Exon1 STOP 7 (pos 8)	GACCACTCAACCCACAGGCC	GACCACUCAACCCACAGGCC
	Exon1 STOP 8 (pos 8)	GACCACTCAAACCACAGCCA	GACCACUCAAACCACAGCCA
	Exon1 STOP 9 (pos 8)	GACCACTCAACCCACAGCCA	GACCACUCAACCCACAGCCA
	Exon1 STOP 10 (pos 8)	GGAGGCACAGACCACTCCAC	GGAGGCACAGACCACUCCAC

Gene	Description	Targeting sequence	gRNA Spacer Sequence
	Exon1 STOP 11 (pos 8)	GACCACTCAACCAGCAGCCA	GACCACUCAACCAGCAGCCA
CD3		TTCGTATCTGTAAAACCAAG	UUCGUAUCUGUAAAACCAAG
CD7		CCTACCTGTCACCAGGACCA	CCUACCUGUCACCAGGACCA
CD52		CTCTTACCTGTACCATAACC	CUCUUACCUGUACCAUAACC
PD1		CACCTACCTAAGAACCATCC	CACCUACCUAAGAACCAUCC
B2M		ACTCACGCTGGATAGCCTCC	ACUCACGCUGGAUAGCCUCC
CD5		ACTCACCCAGCATCCCCAGC	ACUCACCCAGCAUCCCCAGC
CIITA		CACTCACCTTAGCCTGAGCA	CACUCACCUUAGCCUGAGCA
CD2		CACGCACCTGGACAGCTGAC	CACGCACCUGGACAGCUGAC

Table 8B

Gene	gRNA Name	gRNA target	Orientation	Target base(s)	Predicted Outcome
PDCD1	Ex 1 SD	CACCTACCTAAGAACCATCC	Antisense	C7	Splice donor disruption: GT → AT
PDCD1	Ex 2 SA	GGAGTGTGAGAGATGGAGAG	Antisense	C6	Splice acceptor disruption: AG → AA
PDCD1	Ex 3 SA	TTCTCTGTGGAAGGGCACAA	Antisense	C7	Splice acceptor disruption: AG → AA
PDCD1	Ex 3 SD	GACGTTAGCTCGTGCGGCCC	Antisense	C8	Splice donor disruption: GT → AT
PDCD1	Ex 4 SA	CCTGCAGAGAAACACACTTG	Antisense	C2	Splice acceptor disruption: AG → AA
PDCD1	Ex 2 pmSTOP	GGGGTTCCAGGGCCTGTCTG	Antisense	C7, C8	pmSTOP induction: TGG (Trp) → TAG, TGA, TAA
PDCD1	Ex 3 pmSTOP_1	CAGTTCCAAACCCTGGTGGT	Sense	C7	pmSTOP induction: CAA (Gln) → TAA
PDCD1	Ex 3 pmSTOP_2	GGACCCAGACTAGCAGCACC	Antisense	C5, C6	pmSTOP induction: TGG (Trp) → TAG, TGA, TAA
TRAC	Ex 1 SD	CTTACCTGGGCTGGGGAAGA	Antisense	C5	Splice donor disruption: GT → AT
TRAC	Ex 3 SA	TTCGTATCTGTAAAACCAAG	Antisense	C6	Splice acceptor disruption: AG → AA
TRAC	Ex 3 pmSTOP_1	TTTCAAAACCTGTCAGTGAT	Sense	C4	pmSTOP induction: CAA (Gln) → TAA
TRAC	Ex 3 pmSTOP_2	TTCAAAACCTGTCAGTGATT	Sense	C3	pmSTOP induction: CAA (Gln) → TAA
B2M	Ex 1 SD	ACTCACGCTGGATAGCCTCC	Antisense	C6	Splice donor disruption: GT → AT
B2M	Ex 3 SA	TCGATCTATGAAAAAGACAG	Antisense	C5	Splice acceptor disruption: AG → AA
B2M	Ex 2 pmSTOP	CTTACCCCACTTAACTATCT	Antisense	C7, C8	pmSTOP induction: TGG (Trp) → TAG, TGA, TAA

Table 8C

Gene	Description	gRNA Target	gRNA spacer
ACLY	Exon 1 SA	CCATCGGCTCGCGGCGAGAA	CCAUCGGCUCGCGGCGAGAA
	Exon 2 SA	CCTGTCTGGGAGAGAGAAGC	CCUGUCUGGGAGAGAGAAGC
	Exon 2 SD 1	GCTCACCTGGCTGAGCAGCC	GCUCACCUGGCUGAGCAGCC
	Exon 2 SD 2	CTCACCTGGCTGAGCAGCCA	CUCACCUGGCUGAGCAGCCA
	Exon 3 SA	ACCAAGTTCTGGAACAAAAG	ACCAAGUUCUGGAACAAAAG
	Exon 4 SA 1	GCCAACCTACAGAAAAATTG	GCCAACCUACAGAAAAAUUG
	Exon 4 SA 2	CCAACCTACAGAAAAATTGA	CCAACCUACAGAAAAAUUGA
	Exon 5 SA	AGCCTTGCAGGTGAAGAGAC	AGCCUUGCAGGUGAAGAGAC
	Exon 5 SD 1	CTCAACTCTTTCTTGTCTTC	CUCAACUCUUUCUUGUCUUC
	Exon 5 SD 2	TCAACTCTTTCTTGTCTTCA	UCAACUCUUUCUUGUCUUCA
	Exon 7 SA	CACTACTTCAAGGGGAGCAG	CACUACUUCAAGGGGAGCAG
	Exon 12 SD	ACCTACCGATGTGCTCCCGC	ACCUACCGAUGUGCUCCCGC
	Exon 13 SA	CTGGCGTCTGGGGTGAGATA	CUGGCGUCUGGGGUGAGAUA
	Exon 13 SD	GAGTTACCTTGTGGCATGGC	GAGUUACCUUGUGGCAUGGC
	Exon 14 SD	ATCCTACCTTGCAGGGATCT	AUCCUACCUUGCAGGGAUCU
	Exon 15 SD	TCACGTGAAAGGGTAGACCA	UCACGUGAAAGGGUAGACCA
	Exon 16 SD	ATCTACCTGGGCATAGTTCA	AUCUACCUGGGCAUAGUUCA
	Exon 18 SD	TGATTACCTGTCCCCACCAA	UGAUUACCUGUCCCCACCAA
	Exon 20 SA	CCCCAATCTGCCAAGGAATG	CCCCAAUCUGCCAAGGAAUG
	Exon 20 SD	CCATACCTCAGAGGAGAACA	CCAUACCUCAGAGGAGAACA
	Exon 23 SA	CAAGCTCCTGGGCAGAGATG	CAAGCUCCUGGGCAGAGAUG
	Exon 26 SA	TTATCTAGAAATGAACCCAA	UUAUCUAGAAAUGAACCCAA
ADORA2A	Exon 1 ATG	TGGGCATGGCCACAGACGAC	UGGGCAUGGCCACAGACGAC
	Exon 1 SD	CTGCTCACCGGAGCGGGATG	CUGCUCACCGGAGCGGGAUG
	Exon 2 Stop 1	CAGTTGTTCCAACCTAGCAT	CAGUUGUUCCAACCUAGCAU
	Exon 2 STOP 2	CACTCCCAGGGCTGCGGGGA	CACUCCCAGGGCUGCGGGGA
	Exon 2 STOP 3	CCACTCCCAGGGCTGCGGGG	CCACUCCCAGGGCUGCGGGG
	Exon 2 STOP 4	GCGACGACAGCTGAAGCAGA	GCGACGACAGCUGAAGCAGA
	Exon 2 STOP 5	GGAGAGCCAGCCTCTGCCGG	GGAGAGCCAGCCUCUGCCGG
	Exon 2 STOP 6	ACATGAGCCAGAGAGGGGCG	ACAUGAGCCAGAGAGGGGCG
	Exon 2 STOP 7	GAGGCAGCAAGAACCTTTCA	GAGGCAGCAAGAACCUUUCA
	Exon 2 STOP 8	TGGCCCACACTCCTGGCGGG	UGGCCCACACUCCUGGCGGG
	Exon 2 STOP 9	CGTTGGCCCACACTCCTGGC	CGUUGGCCCACACUCCUGGC
	Exon 2 STOP 10	CTGGGACTCTTGGGCACTCC	CUGGGACUCUUGGGCACUCC
AXL	Exon 2 SA 1	TGCGTGCCTGGAGGGGAGAT	UGCGUGCCUGGAGGGGAGAU
	Exon 2 SA 2	CTGCGTGCCTGGAGGGGAGA	CUGCGUGCCUGGAGGGGAGA
	Exon 3 SA	GGTGATTCTGACAGGGCAAG	GGUGAUUCUGACAGGGCAAG
	Exon 4 SA	AAGCCTAGCGGGGTGGGCAG	AAGCCUAGCGGGGUGGGCAG
	Exon 4 SD	CGGACTCACCTGGAACATGC	CGGACUCACCUGGAACAUGC
	Exon 5 SA 1	AGCCCTAGGGAGTCATATGA	AGCCCUAGGGAGUCAUAUGA

	Exon 5 SA 2	CAGCCCTAGGGAGTCATATG	CAGCCCUAGGGAGUCAUAUG
	Exon 6 SD 1	TCTCACCTGCAGGGTGCAGT	UCUCACCUGCAGGGUGCAGU
	Exon 6 SD 2	GTCTCACCTGCAGGGTGCAG	GUCUCACCUGCAGGGUGCAG
	Exon 7 SA 1	CACAGCCTGAGGAGAGGCAA	CACAGCCUGAGGAGAGGCAA
	Exon 7 SA 2	GCACAGCCTGAGGAGAGGCA	GCACAGCCUGAGGAGAGGCA
	Exon 8 SA 1	GCACTGGAGGACAGGGAAGA	GCACUGGAGGACAGGGAAGA
	Exon 8 SA 2	GGCACTGGAGGACAGGGAAG	GGCACUGGAGGACAGGGAAG
	Exon 8 SD	CACCCACCTCTGGGGTGTCC	CACCCACCUCUGGGGUGUCC
	Exon 9 SA	GCACCTAGGAGGTCCAGAAG	GCACCUAGGAGGUCCAGAAG
	Exon 10 SD	CCCTTACCCAGCTGGTGGAC	CCCUUACCCAGCUGGUGGAC
	Exon 11 SA	CTTCACTATCAGGGGGTATG	CUUCACUAUCAGGGGGUAUG
	Exon 12 SA	TCACTTACAGGTAGCTTCAG	UCACUUACAGGUAGCUUCAG
	Exon 13 SA 1	GTTCACTGCATGCAAGGTTG	GUUCACUGCAUGCAAGGUUG
	Exon 13 SA 2	TGTTCACTGCATGCAAGGTT	UGUUCACUGCAUGCAAGGUU
	Exon 13 SA 3	CTGTTCACTGCATGCAAGGT	CUGUUCACUGCAUGCAAGGU
	Exon 14 SA 1	CTCTCCTGTGGGGGGCCAGA	CUCUCCUGUGGGGGGCCAGA
	Exon 14 SA 2	ACTCTCCTGTGGGGGGCCAG	ACUCUCCUGUGGGGGGCCAG
	Exon 15 SA	GCAACTTGAGGGAGAGAGAA	GCAACUUGAGGGAGAGAGAA
	Exon 17 SA	AGGTACTGGGGAGCCAAGGC	AGGUACUGGGGAGCCAAGGC
	Exon 18 SD	ACCTACCACATCGCTCTTGC	ACCUACCACAUCGCUCUUGC
	Exon 19 SA 1	GGACCACTGTGAGGGGCAGA	GGACCACUGUGAGGGGCAGA
	Exon 19 SA 2	AGGACCACTGTGAGGGGCAG	AGGACCACUGUGAGGGGCAG
	Exon 20 SA	ATACCTAGGGCAGCAAAATG	AUACCUAGGGCAGCAAAAUG
BATF	Exon 1 ATG	GAGGCATGGCTGAAATCTTC	GAGGCAUGGCUGAAAUCUUC
	Exon 1 SD 1	TCTACCTGTTTGCCAGGGGG	UCUACCUGUUUGCCAGGGGG
	Exon 1 SD 2	GACTCTACCTGTTTGCCAGG	GACUCUACCUGUUUGCCAGG
	Exon 2 SA 1	AGTCCTGGGAAGCAGAGACG	AGUCCUGGGAAGCAGAGACG
	Exon 2 SA 2	GAGTCCTGGGAAGCAGAGAC	GAGUCCUGGGAAGCAGAGAC
	Exon 2 SA 3	TGAGTCCTGGGAAGCAGAGA	UGAGUCCUGGGAAGCAGAGA
	Exon 2 SD	ACTTACCAGGTGCAGGGTGT	ACUUACCAGGUGCAGGGUGU
BCL2L11	Exon 1 STOP 1	GGTAGACAATTGCAGCCTG	GGUAGACAAUUGCAGCCUG
	Exon 1 STOP 2	GCCTCCCCAGCTCAGACCTG	GCCUCCCCAGCUCAGACCUG
	Exon 1 STOP 3	TCCCTACAGACAGAGCCACA	UCCCUACAGACAGAGCCACA
	Exon 1 STOP 4	GAGCCACAAGGTAATCCTGA	GAGCCACAAGGUAAUCCUGA
	Exon 3 STOP 1	GCCCAAGAGTTGCGGCGTAT	GCCCAAGAGUUGCGGCGUAU
	Exon 4 SA	AAAATACCTGAAACAACAAA	AAAAUACCUGAAACAACAAA
CAMK2D	Exon 6 SA 1	CAATGACTGCAAAGATACAA	CAAUGACUGCAAAGAUACAA
	Exon 6 SA 2	ACAATGACTGCAAAGATACA	ACAAUGACUGCAAAGAUACA
	Exon 7 SA 3	AATGACTGCATGCAAACACC	AAUGACUGCAUGCAAACACC
	Exon 7 SA 2	ATGACTGCATGCAAACACCA	AUGACUGCAUGCAAACACCA
	Exon 7 SA 1	TGACTGCATGCAAACACCAG	UGACUGCAUGCAAACACCAG
	Exon 7 SD	TACTCACCTTCAGGTCCCGA	UACUCACCUUCAGGUCCCGA

	Exon 8 SD	ACTCACCAAACCACGCCTGC	ACUCACCAAACCACGCCUGC
	Exon 14 SD	TAACTTACCTTTACTCCATC	UAACUUACCUUUACUCCAUC
	Exon 16 SD	AGGTATACCAGCGCTGGGGT	AGGUAUACCAGCGCUGGGGU
	Exon 17 SD 1	TGAATACCTTGTTTCCATCA	UGAAUACCUUGUUUCCAUCA
	Exon 17 SD 2	CTGAATACCTTGTTTCCATC	CUGAAUACCUUGUUUCCAUC
	Exon 19 SA	TCGTGCTAAAGGCAAAAATA	UCGUGCUAAAGGCAAAAAUA
cAMP	Exon 1 SD 1	AGCTCACCATCGTGGGCCTG	AGCUCACCAUCGUGGGCCUG
	Exon 1 SD 2	AAGCTCACCATCGTGGGCCT	AAGCUCACCAUCGUGGGCCU
	Exon 1 SD 3	AAAGCTCACCATCGTGGGCC	AAAGCUCACCAUCGUGGGCC
	Exon 2 SA	ATCCTAGTCAGAGGAGGAAA	AUCCUAGUCAGAGGAGGAAA
	Exon 3 SA	GTTATCCTGGGGTTGTGTAC	GUUAUCCUGGGGUUGUGUAC
CASP8	Exon "3" SA	GAACCTTCAAAGGACCAAGA	GAACCUUCAAAGGACCAAGA
	Exon 1 SD	TCACCCGCTCCACCCTTTCC	UCACCCGCUCCACCCUUUCC
	Exon 2 SA	ATAATCTAAGTCAAAATAAA	AUAAUCUAAGUCAAAAUAAA
	Exon 2.5 SA	AGTCCATCTTTTTAAAAGGC	AGUCCAUCUUUUUAAAAGGC
	Exon 3 SA	CATGACCCTGTGGTGGGAAA	CAUGACCCUGUGGUGGGAAA
	Exon 5 SD	TTACCATTTGAAAATTCATC	UUACCAUUUGAAAAUUCAUC
CCR5	Exon 1 STOP	CATACAGTCAGTATCAATTC	CAUACAGUCAGUAUCAAUUC
	Exon 1 STOP 2	GGTGTCGAAATGAGAAGAAG	GGUGUCGAAAUGAGAAGAAG
	Exon 1 STOP 3	ATGCAGGTGACAGAGACTCT	AUGCAGGUGACAGAGACUCU
	Exon 1 STOP 4	TGGGGAGCAGGAAATATCTG	UGGGGAGCAGGAAAUAUCUG
CD2	Ex3 SD (pos 8)	CACGCACCTGGACAGCTGAC	CACGCACCUGGACAGCUGAC
	Ex3 STOP1 (Pos 4)	TCTCAAAACCAAAGATCTCC	UCUCAAAACCAAAGAUCUCC
	Ex3 STOP2 (Pos 6)	CAACACAACCCTGACCTGTG	CAACACAACCCUGACCUGUG
	Ex4 STOP (pos 4)	AAACAGAGGAGTCGGAGAAA	AAACAGAGGAGUCGGAGAAA
	Ex4 STOP2 (Pos 5)	TCACCAAAAGGAAAAAACAG	UCACCAAAAGGAAAAAACAG
	Ex5 STOP (pos 4)	ACACAAGTTCACCAGCAGAA	ACACAAGUUCACCAGCAGAA
	Ex5 STOP (pos 4)	GTTCAGCCAAAACCTCCCCA	GUUCAGCCAAAACCUCCCCA
	Exon 2 STOP (pos 8)	CTTGGGTCAGGACATCAACT	CUUGGGUCAGGACAUCAACU
	Exon 2 STOP (pos 8)	CGATGATCAGGATATCTACA	CGAUGAUCAGGAUAUCUACA
CD3D	Exon 1 SD 1	AGCCTTACCTTGCGAGAGAA	AGCCUUACCUUGCGAGAGAA
	Exon 1 SD 2	TAGCCTTACCTTGCGAGAGA	UAGCCUUACCUUGCGAGAGA
	Exon 1 STOP	TCGCAAGGTAAGGCTACTCC	UCGCAAGGUAAGGCUACUCC
	Exon 3 SA	GGCACACTGTGGGGGAAGGG	GGCACACUGUGGGGGAAGGG
	Exon 3 STOP	GTGCCAGAGCTGTGTGGAGC	GUGCCAGAGCUGUGUGGAGC
	Exon 4 STOP 1	CCGACACACAAGCTCTGTTG	CCGACACACAAGCUCUGUUG
	Exon 4 STOP 2	GGTCTATCAGGTGAGCGTTG	GGUCUAUCAGGUGAGCGUUG
	Exon 5 STOP	GATGCTCAGTACAGCCACCT	GAUGCUCAGUACAGCCACCU
CD3E	Exon 1 ATG	CCGACTGCATCTTTGTTTCA	CCGACUGCAUCUUUGUUUCA
	Exon 1 SD	ACTCACCTGATAAGAGGCAG	ACUCACCUGAUAAGAGGCAG
	Exon 4 SA	TACCACCTGAAAATGAAAAA	UACCACCUGAAAAUGAAAAA
	Exon 4 STOP	ACACAGACACGTGAGTTTAT	ACACAGACACGUGAGUUUAU

	Exon 5 SA 1	TATATGCTGGGGAGAAAGAA	UAUAUGCUGGGGAGAAAGAA
	Exon 5 SA 2	TTATATGCTGGGGAGAAAGA	UUAUAUGCUGGGGAGAAAGA
	Exon 5 SD	CTGGATTACCTCTTGCCCTC	CUGGAUUACCUCUUGCCCUC
	Exon 6 SA 1	ACACTGTGGGGGGTGGGGTG	ACACUGUGGGGGGUGGGGUG
	Exon 6 SA 2	CACACTGTGGGGGGTGGGGT	CACACUGUGGGGGGUGGGGU
	Exon 6 SA 3	ACACACTGTGGGGGGTGGGG	ACACACUGUGGGGGGUGGGG
	Exon 7 SA 1	TTGTCCTGCGGAGGAAGGAG	UUGUCCUGCGGAGGAAGGAG
	Exon 7 SA 2	TTTGTCCTGCGGAGGAAGGA	UUUGUCCUGCGGAGGAAGGA
	Exon 7 SA 3	TTTTGTCCTGCGGAGGAAGG	UUUUGUCCUGCGGAGGAAGG
	Exon 7 SD 1	GTTACCTCATAGTCTGGGTT	GUUACCUCAUAGUCUGGGUU
	Exon 7 SD 2	CGTTACCTCATAGTCTGGGT	CGUUACCUCAUAGUCUGGGU
CD3G	Exon 1 STOP 1	CATGGAACAGGGGAAGGGCC	CAUGGAACAGGGGAAGGGCC
	Exon 1 STOP 2	CTTCAAGGTAAGGGCCTACT	CUUCAAGGUAAGGGCCUACU
	Exon 2 SD 1	TCTCCTACCTTTGATTGACT	UCUCCUACCUUUGAUUGACU
	Exon 2 SD 2	TTCTCCTACCTTTGATTGAC	UUCUCCUACCUUUGAUUGAC
	Exon 2 STOP	TGGCCCAGTCAATCAAAGGT	UGGCCCAGUCAAUCAAAGGU
	Exon 3 SD	ACATACTTCTGTAATACACT	ACAUACUUCUGUAAUACACU
	Exon 3 STOP 1	TGACTATCAAGAAGATGGTT	UGACUAUCAAGAAGAUGGUU
	Exon 3 STOP 2	TTTAAACCATGTGATATTTT	UUUAAACCAUGUGAUAUUUU
	Exon 4 STOP	CTCTTCCATTGGGTACATAA	CUCUUCCAUUGGGUACAUAA
	Exon 5 STOP	TGACCAGCTCTACCAGGTAA	UGACCAGCUCUACCAGGUAA
	Exon 7 STOP 1	GACCAGTACAGCCACCTTCA	GACCAGUACAGCCACCUUCA
	Exon 7 STOP 2	ACCTTCAAGGAAACCAGTTG	ACCUUCAAGGAAACCAGUUG
CD4	Exon 1 ATG	GGTTCATTGTGGCCTTGCCG	GGUUCAUUGUGGCCUUGCCG
	Exon 2 SA	GAGCGCTAAGTGGAAAAGAA	GAGCGCUAAGUGGAAAAGAA
	Exon 2 SD	AACCCTACCTTTAGTTAAGA	AACCCUACCUUUAGUUAAGA
	Exon 5 SA	GGCAGTCACTGTGGAGGGAA	GGCAGUCACUGUGGAGGGAA
	Exon 6 SA	TGGAAAGCTGGAGGTGGGAA	UGGAAAGCUGGAGGUGGGAA
	Exon 6 SD	CCTCACCTCTCATCACCACC	CCUCACCUCUCAUCACCACC
	Exon 7 SA	AGTGGCTGCAGAGGAACGAG	AGUGGCUGCAGAGGAACGAG
	Exon 10 SA	GCGCTGTCCAGGGACAAGAA	GCGCUGUCCAGGGACAAGAA
	Exon 10 SD	TCCTTACTGAGGACACTGGC	UCCUUACUGAGGACACUGGC
	short alt exon 2 SA	CCATCTGGAGCTTAGGGTCC	CCAUCUGGAGCUUAGGGUCC
	Short CD4 ATG	GGTTGGCATGTGGAGGCAGC	GGUUGGCAUGUGGAGGCAGC
CD5	Ex2 STOP 2 (pos 6)	GGGTCATACCAGCTGAGCCG	GGGUCAUACCAGCUGAGCCG
	Ex3 SA (pos 8)	TGGAAATCTGGGGGTCAGAA	UGGAAAUCUGGGGGUCAGAA
	Ex3 SD (pos 9)	GTTACCCACCTAAGCAGGTC	GUUACCCACCUAAGCAGGUC
	Ex3 STOP (pos 6)	TCTGCCAGCGGCTGAACTGT	UCUGCCAGCGGCUGAACUGU
	Ex3 STOP (pos 5)	CTGCCAGCGGCTGAACTGTG	CUGCCAGCGGCUGAACUGUG
	Ex3 STOP (pos 5/6)	CCTCCCACTGCTTGGAGCTC	CCUCCCACUGCUUGGAGCUC
	Ex3 STOP (pos 8)	GAAGTGCCAGGGCCAGCTGG	GAAGUGCCAGGGCCAGCUGG
	Ex3 STOP (pos 8/9)	CCATGTGCCATCCGTCCTTG	CCAUGUGCCAUCCGUCCUUG

	Ex3 STOP (pos 9)	TTTGCAGCCAGAGCTGGGGC	UUUGCAGCCAGAGCUGGGGC
	Ex4 SA (pos 5)	GGTTCTGCAATGAGACACTC	GGUUCUGCAAUGAGACACUC
	Ex4 STOP (pos 4)	CTCCAGAGCCCACAGGTAAG	CUCCAGAGCCCACAGGUAAG
	Ex4 STOP2 (Pos 5)	ACCACAACTCCAGAGCCCAC	ACCACAACUCCAGAGCCCAC
	Ex5 SA (pos 4)	GAGCTAGGAGAGGAGAGAGC	GAGCUAGGAGAGGAGAGAGC
	Ex5 SD (pos 9)	CTCACTTACCTGAGCAAAGG	CUCACUUACCUGAGCAAAGG
	Ex5 STOP (pos 5)	CTGCAGCTGGTGGCACAGTC	CUGCAGCUGGUGGCACAGUC
	Ex5 STOP (pos 7)	GATCTTCCATTGGATTGGCA	GAUCUUCCAUUGGAUUGGCA
	Ex5 STOP (pos 8)	TGAGGCCCAGGACAAGACCC	UGAGGCCCAGGACAAGACCC
	Ex6 SA (pos 5)	AAACCTGAGAGGGGAAGCAA	AAACCUGAGAGGGGAAGCAA
	Ex6 STOP (pos 4/5)	CTCCCACCGCAGCGAGCTCC	CUCCCACCGCAGCGAGCUCC
	Ex6 STOP (pos 5)	TTTCCAGCCCAAGGTGCAGA	UUUCCAGCCCAAGGUGCAGA
	Ex6 STOP (pos 5)	GGTGCAGAGCCGTCTGGTGG	GGUGCAGAGCCGUCUGGUGG
	Ex6 STOP (pos 6)	AGGTGCAGAGCCGTCTGGTG	AGGUGCAGAGCCGUCUGGUG
	Ex6 STOP (pos 7)	TCCTATCGAGTGCTGGACGC	UCCUAUCGAGUGCUGGACGC
	Ex6 STOP (pos 7)	AAGGTGCAGAGCCGTCTGGT	AAGGUGCAGAGCCGUCUGGU
	Ex6 STOP (pos 8)	CAAGGTGCAGAGCCGTCTGG	CAAGGUGCAGAGCCGUCUGG
	Ex6 STOP (pos 8/9)	GGGCTGCCCACTGAGCCCCC	GGGCUGCCCACUGAGCCCCC
	Ex6 STOP (pos 9)	AGGTGCGCCAGGGGGCTCAG	AGGUGCGCCAGGGGGCUCAG
	Ex7 STOP (pos 4)	GGCCAGGATCCAAACCCCGC	GGCCAGGAUCCAAACCCCGC
	Ex8 STOP (pos 4)	CGCCAGTGGATTGGCCCAAC	CGCCAGUGGAUUGGCCCAAC
	Ex8 STOP (pos 5)	GCGCCAGTGGATTGGCCCAA	GCGCCAGUGGAUUGGCCCAA
	Ex8 STOP (pos 7)	AAGAAGCAGCGCCAGTGGAT	AAGAAGCAGCGCCAGUGGAU
	Ex9 SD (pos 6)	GCTTACCTGGATAAGCTGAC	GCUUACCUGGAUAAGCUGAC
	Ex9 SD1 (Pos 8)	AAAGACACTGGGCAGATGGT	AAAGACACUGGGCAGAUGGU
	Ex10 SA (pos 9)	TTCCAGAGCTGGGGAAAGAA	UUCCAGAGCUGGGGAAAGAA
	Exon 1 SD (pos 6)	ACTCACCCAGCATCCCCAGC	ACUCACCCAGCAUCCCCAGC
	Exon 2 SA (pos 6)	AGCGACTGCAGAAAGAAGAG	AGCGACUGCAGAAAGAAGAG
	Exon 2 STOP (pos 5/6)	CATACCAGCTGAGCCGTCCG	CAUACCAGCUGAGCCGUCCG
CD8A	Exon 1 ATG	AAGGCCATGACGCGCTCCCC	AAGGCCAUGACGCGCUCCCC
	Exon 1 SD	TCACGGAGCAGCAAGGCCAG	UCACGGAGCAGCAAGGCCAG
	Exon 2 SD	CGCGGACCTGGCAGGAAGAC	CGCGGACCUGGCAGGAAGAC
	Exon 3 SD	TCACCTGCGCCCCCCGCCGC	UCACCUGCGCCCCCCGCCGC
	Exon 4 SD 1	CTTACTGTGGTTGCAGTAAA	CUUACUGUGGUUGCAGUAAA
	Exon 4 SD 2	ACTTACTGTGGTTGCAGTAA	ACUUACUGUGGUUGCAGUAA
CD38	Exon 1 ATG 1	TTGGCCATAGGGCTCCAGGC	UUGGCCAUAGGGCUCCAGGC
	Exon 1 ATG 2	GTTGGCCATAGGGCTCCAGG	GUUGGCCAUAGGGCUCCAGG
	Exon 1 STOP	GCGCCAGCAGTGGAGCGGTC	GCGCCAGCAGUGGAGCGGUC
	Exon 2 SD	AATTACCTTGTTGCAAGGTA	AAUUACCUUGUUGCAAGGUA
	Exon 2 STOP	CTATCAGCCACTAATGAAGT	CUAUCAGCCACUAAUGAAGU
	Exon 3 STOP 1	CTGCTCCAAAGAAGAATCTA	CUGCUCCAAAGAAGAAUCUA
	Exon 4 STOP 1	ACTATCAATCTTGCCCAGAC	ACUAUCAAUCUUGCCCAGAC

	Exon 4 STOP 2	TTTTCCAGAATACTGAAACA	UUUUCCAGAAUACUGAAACA
	Exon 4 STOP 3	GTTTTCCAGAATACTGAAAC	GUUUUCCAGAAUACUGAAAC
	Exon 7 SD	TTACCTGTAGATATTCTTGC	UUACCUGUAGAUAUUCUUGC
CD70	Ex1 SD (pos 6)	CTCACCCCAAGTGACTCGAG	CUCACCCCAAGUGACUCGAG
	Ex1 STOP (pos 8)	GTGCATCCAGCGCTTCGCAC	GUGCAUCCAGCGCUUCGCAC
	Ex2 STOP (pos 7)	GAGCTGCAGCTGAATCACAC	GAGCUGCAGCUGAAUCACAC
	Ex3 STOP (pos 5)	CTGGCAGGGGGGCCCAGCAC	CUGGCAGGGGGGCCCAGCAC
	Ex3 STOP (pos 5)	CTCCCAGCGCCTGACGCCCC	CUCCCAGCGCCUGACGCCCC
	Ex3 STOP (pos 8)	CCCCCTGCCAGTATAGCCTG	CCCCCUGCCAGUAUAGCCUG
	Ex3 STOP (pos 9)	CCCCCCTGCCAGTATAGCCT	CCCCCCUGCCAGUAUAGCCU
CD82	Exon 1 ATG	TGAGCCCATCCCGCCAGTCC	UGAGCCCAUCCCGCCAGUCC
	Exon 3 SD	TCACCAGCCCCAGCAGGCAG	UCACCAGCCCCAGCAGGCAG
	Exon 4 SA	AAGTACTGGGGACACAGAGC	AAGUACUGGGGACACAGAGC
	Exon 6 SA 1	ACTTCACCTGGGCAAGGCAG	ACUUCACCUGGGCAAGGCAG
	Exon 6 SA 2	CACTTCACCTGGGCAAGGCA	CACUUCACCUGGGCAAGGCA
	Exon 6 SD	CCGCACACCTCCTGGTACAC	CCGCACACCUCCUGGUACAC
	Exon 7 SA	AGCCCTGCAAGGGCAGAATG	AGCCCUGCAAGGGCAGAAUG
	Exon 8 SA	CCAGGAGCTGTGGGGAGAGG	CCAGGAGCUGUGGGGAGAGG
CD86	Ex2 SD (pos 8)	GTTCTTACCAGAGAGCAGGA	GUUCUUACCAGAGAGCAGGA
	Ex3 SA (pos 5)	GCACCTAAAAAAGAAGGTTA	GCACCUAAAAAAGAAGGUUA
	Ex3 STOP (pos 5)	TTGGCAGGACCAGGAAAACT	UUGGCAGGACCAGGAAAACU
	Ex3 STOP (pos 8)	CAATCTTCAGATCAAGGACA	CAAUCUUCAGAUCAAGGACA
	Ex5 STOP (pos 6/7)	GTAATCCAAGGAATGTGGTC	GUAAUCCAAGGAAUGUGGUC
	Ex6 STOP (pos 9)	AGAGTGAACAGACCAAGAAA	AGAGUGAACAGACCAAGAAA
CD160	Exon 1 STOP	GGACATCCAGTCTGGTGGTG	GGACAUCCAGUCUGGUGGUG
	Exon 2 SA	AATGCATCCTGGAATGGAAA	AAUGCAUCCUGGAAUGGAAA
	Exon 2 SD	GCACTCACCTGTGAATAGAA	GCACUCACCUGUGAAUAGAA
	Exon 2 STOP	TAAAACAGCTGAGACTTAAA	UAAAACAGCUGAGACUUAAA
	Exon 3 STOP 1	GCTTCCTACAAGAAAAGGTC	GCUUCCUACAAGAAAAGGUC
	Exon 3 STOP 2	TTACCCAGACCTTTTCTTGT	UUACCCAGACCUUUUCUUGU
CD244	Exon 1 ATG 1	CAGCATTTCCACAGGACAGA	CAGCAUUUCCACAGGACAGA
	Exon 1 ATG 2	CCAGCATTTCCACAGGACAG	CCAGCAUUUCCACAGGACAG
	Exon 2 SD	ACTTACCAAATACAAAAACC	ACUUACCAAAUACAAAAACC
	Exon 3 SA	GATTCTGATCAGAAAGGCAT	GAUUCUGAUCAGAAAGGCAU
	Exon 4 SA	TGAATTCTGAGGAATACAGA	UGAAUUCUGAGGAAUACAGA
	Exon 5 SA	ATGACATACGTGATTTCTCC	AUGACAUACGUGAUUUCUCC
	Exon 6 SA	TCCTGCTCCTGCACAAGAAA	UCCUGCUCCUGCACAAGAAA
	Exon 8 SA	TCACCCTAGGAGCAAAACAA	UCACCCUAGGAGCAAAACAA
CD276	Exon 1 ATG	CGCAGCATCTTCCTGTGAGG	CGCAGCAUCUUCCUGUGAGG
	Exon 2 SA	GCTCCTGGGGGTAGGGGGAG	GCUCCUGGGGGUAGGGGGAG
	Exon 2 SD	GGTGCTCACCGGCCACCTGC	GGUGCUCACCGGCCACCUGC
	Exon 3 SA 1	AGGGAGCTGGAGGTGACAGA	AGGGAGCUGGAGGUGACAGA

	Exon 3 SA 2	TAGGGAGCTGGAGGTGACAG	UAGGGAGCUGGAGGUGACAG
	Exon 3 SD 1	GCAACCTGTGGGGCTTCTCT	GCAACCUGUGGGGCUUCUCU
	Exon 3 SD 2	AGCAACCTGTGGGGCTTCTC	AGCAACCUGUGGGGCUUCUC
	Exon 4 SA 1	CTCCTGGGGGCGGGGTCAGA	CUCCUGGGGGCGGGGUCAGA
	Exon 4 SA 2	GCTCCTGGGGGCGGGGTCAG	GCUCCUGGGGGCGGGGUCAG
	Exon 4 SD	GGTGCTCACCGGCCACCTGC	GGUGCUCACCGGCCACCUGC
	Exon 5 SA 1	AGGGAGCTGGAGGTGACAGA	AGGGAGCUGGAGGUGACAGA
	Exon 5 SA 2	TAGGGAGCTGGAGGTGACAG	UAGGGAGCUGGAGGUGACAG
	Exon 8 SA	GCAGGGCTGTAAAAAAAGGA	GCAGGGCUGUAAAAAAAGGA
	Exon 9 SA 1	CATCATCTTCATTTCATGAT	CAUCAUCUUCAUUUCAUGAU
	Exon 9 SA 2	CCATCATCTTCATTTCATGA	CCAUCAUCUUCAUUUCAUGA
CDK8	Exon 1 ATG	GTCCATTGTCACAGCCTCTG	GUCCAUUGUCACAGCCUCUG
	Exon 1 SD	CACTCACCCATCTTTCCTCT	CACUCACCCAUCUUUCCUCU
	Exon 10 SD 2	ACTTACTCTGATGTAGGAAG	ACUUACUCUGAUGUAGGAAG
	Exon 10 SD 1	CTTACTCTGATGTAGGAAGT	CUUACUCUGAUGUAGGAAGU
	Exon 12 SD	TTGGAATACCTGATAGTCTG	UUGGAAUACCUGAUAGUCUG
	Exon 13 SA 2	GGAACGCTGGAAAGGAGATG	GGAACGCUGGAAAGGAGAUG
	Exon 13 SA 1	GAACGCTGGAAAGGAGATGA	GAACGCUGGAAAGGAGAUGA
CDKN1B	Exon 1 ATG	ACGTTTGACATCTTTCTCCC	ACGUUUGACAUCUUUCUCCC
	Exon 1 STOP 1	CAAACGTGCGAGTGTCTAAC	CAAACGUGCGAGUGUCUAAC
	Exon 1 STOP 2	CGAGTGGCAAGAGGTGGAGA	CGAGUGGCAAGAGGUGGAGA
	Exon 1 STOP 3	GAGTGGCAAGAGGTGGAGAA	GAGUGGCAAGAGGUGGAGAA
	Exon 1 STOP 4	AGGAGAGCCAGGATGTCAGC	AGGAGAGCCAGGAUGUCAGC
	Exon 1 STOP 5	GGACAGCCAGACGGGGTTAG	GGACAGCCAGACGGGGUUAG
	Exon 1 STOP 6	CGGAGCAATGCGCAGGAATA	CGGAGCAAUGCGCAGGAAUA
	Exon 1 STOP 7	AGGAAGCGACCTGCAACCGA	AGGAAGCGACCUGCAACCGA
	Exon 2 STOP	GAGCAGACGCCCAAGAAGCC	GAGCAGACGCCCAAGAAGCC
CSF2	Exon 1 STOP 1	GCTGCAGAGCCTGCTGCTCT	GCUGCAGAGCCUGCUGCUCU
	Exon 1 STOP 2	GCTCCCAGGGCTGCGTGCTG	GCUCCCAGGGCUGCGUGCUG
	Exon 1 STOP 3	TGCTCCCAGGGCTGCGTGCT	UGCUCCCAGGGCUGCGUGCU
	Exon 1 STOP 4	ATGCTCCCAGGGCTGCGTGC	AUGCUCCCAGGGCUGCGUGC
	Exon 3 SD (pos 10)	AGGCACTCACCGGGGTTGGA	AGGCACUCACCGGGGUUGGA
	Exon 4 STOP 1	CTGGCTCCCAGCAGTCAAAG	CUGGCUCCCAGCAGUCAAAG
CSK	Exon 1 ATG	TGACATCTTCTCAGGAGCTC	UGACAUCUUCUCAGGAGCUC
	Exon 3 SD	TCACGGCATGAGGCTGAGTT	UCACGGCAUGAGGCUGAGUU
	Exon 4 SA 1	GAACCAACTGGGGAGCAGCA	GAACCAACUGGGGAGCAGCA
	Exon 4 SA 2	GGAACCAACTGGGGAGCAGC	GGAACCAACUGGGGAGCAGC
	Exon 4 SD	TCACCTCCACCAGCTGCATG	UCACCUCCACCAGCUGCAUG
	Exon 5 SA 1	GTAGTGCTGCAGGGTGTGGG	GUAGUGCUGCAGGGUGUGGG
	Exon 5 SA 2	TGTAGTGCTGCAGGGTGTGG	UGUAGUGCUGCAGGGUGUGG
	Exon 7 SA 1	CACGTCTGGGGGCAGAGAGG	CACGUCUGGGGGCAGAGAGG
	Exon 7 SA 2	CATCACGTCTGGGGGCAGAG	CAUCACGUCUGGGGGCAGAG

	Exon 9 SA	AGGCTCCCCTGGGGGCAGGA	AGGCUCCCCUGGGGGCAGGA
	Exon 9 SD	TCACTCACAGCGAGAACTTG	UCACUCACAGCGAGAACUUG
	Exon 10 SD	CAGCCCCACCTTCTCTCTCA	CAGCCCCACCUUCUCUCUCA
	Exon 11 SA	GAGAATTTCTGCCATGTGGA	GAGAAUUUCUGCCAUGUGGA
	Exon 11 SD	ATACTCACAATTCTTGGATA	AUACUCACAAUUCUUGGAUA
	Exon 12 SA	CAGGGGCTGTGGCCAGGGGG	CAGGGGCUGUGGCCAGGGGG
CTLA-4	Exon 1 SD (pos 6)	ACTCACCTTTGCAGAAGACA	ACUCACCUUUGCAGAAGACA
	Exon 1 SD	CACTCACCTTTGCAGAAGAC	CACUCACCUUUGCAGAAGAC
	Exon 1 STOP (pos5)	AGGGCCAGGTCCTGGTAGCC	AGGGCCAGGUCCUGGUAGCC
	Exon 2 STOP	GGCCCAGCCTGCTGTGGTAC	GGCCCAGCCUGCUGUGGUAC
	Exon 2 STOP (pos 8)	GCTTCGGCAGGCTGACAGCC	GCUUCGGCAGGCUGACAGCC
	Exon 2 STOP	[[TATCCAAGGACTGAGGGCCA	[[UAUCCAAGGACUGAGGGCCA
	Exon 2 STOP	GGAACCCAGATTTATGTAAT	GGAACCCAGAUUUAUGUAAU
	Exon 2 SD	GCTCACCAATTACATAAATC	GCUCACCAAUUACAUAAAUC
	Exon 2 SD	CTCACCAATTACATAAATCT	CUCACCAAUUACAUAAAUCU
	Exon 1 STOP	CTCAGCTGAACCTGGCTACC	CUCAGCUGAACCUGGCUACC
CUL3	Exon 1/2 SA ATG	TTAACATCTACTACATACAA	UUAACAUCUACUACAUACAA
	Exon 6 SD	CTTACCTGGATATAGTCAAC	CUUACCUGGAUAUAGUCAAC
	Exon 2 STOP	ATCCAGCGTAAGAATAACAG	AUCCAGCGUAAGAAUAACAG
	Exon 4 STOP 1	GTATGTACAACAAAATAATG	GUAUGUACAACAAAAUAAUG
	Exon 4 STOP 2	CGAGATCAAGTTGTACGTTA	CGAGAUCAAGUUGUACGUUA
	Exon 5 STOP	TGCCAGATGTTAATGATTTT	UGCCAGAUGUUAAUGAUUUU
	Exon 9 STOP	ATGTCAGTTCACGTCAAAAC	AUGUCAGUUCACGUCAAAAC
	Exon 14 STOP	GCCCTACAGTCCCTCGCCTG	GCCCUACAGUCCCUCGCCUG
Cyp11a1	Exon 1 STOP 1 (pos 4)	GTCCAGAATTTCCAGAAGTA	GUCCAGAAUUUCCAGAAGUA
	Exon 2 SA 1 (pos 4)	TCCCTGGAGGGGTGGGGGAG	UCCCUGGAGGGGUGGGGGAG
	Exon 2 SD 1 (pos 4)	TCACTTCAACAGGACTCCTA	UCACUUCAACAGGACUCCUA
	Exon 3 SD 1 (pos 6)	CCTTACACTCAAAGGCAAAG	CCUUACACUCAAAGGCAAAG
	Exon 4 SA (pos 5)	ATGGCTGCAGGGAGAGGAAG	AUGGCUGCAGGGAGAGGAAG
	Exon 4 STOP 1 (pos 8)	GGAGCGCCAGGGGATGCTGG	GGAGCGCCAGGGGAUGCUGG
	Exon 4 STOP 2 (pos 8)	TCACGTCCCATGCAGCCACA	UCACGUCCCAUGCAGCCACA
	Exon 6 SA (pos 8)	TGGACGTCTGGTGGGGAGTA	UGGACGUCUGGUGGGGAGUA
	Exon 8 STOP 1(pos 6)	ACTCACATTGATGAGGAAGA	ACUCACAUUGAUGAGGAAGA
	Exon 9 SA (pos 7)	CAGCATCTGAGAAAGGCAGA	CAGCAUCUGAGAAAGGCAGA
	Exon 9 STOP 1 (pos 5)	AATCCAACACCTCAGCGATG	AAUCCAACACCUCAGCGAUG
	Exon 9 STOP 2 (pos 4)	ATCCAACACCTCAGCGATGT	AUCCAACACCUCAGCGAUGU
DCK	Exon 1 ATG	GTGGCCATTCCTTAGTCTTG	GUGGCCAUUCCUUAGUCUUG
	Exon 1 SD	CTTACCGATGTTCCCTTCGA	CUUACCGAUGUUCCCUUCGA
	Exon 2 STOP 1	TTAAACAATTGTGTGAAGAT	UUAAACAAUUGUGUGAAGAU
	Exon 2 STOP 2	TAAACAATTGTGTGAAGATT	UAAACAAUUGUGUGAAGAUU
	Exon 2 STOP 3	CACCATCTGGCAACAGGTTC	CACCAUCUGGCAACAGGUUC
	Exon 2 STOP 4	AAGTACTCAAGATGAATTTG	AAGUACUCAAGAUGAAUUUG

	Exon 3 STOP 1	CAATGTCTCAGAAAAATGGT	CAAUGUCUCAGAAAAAUGGU
	Exon 3 STOP 2	GCTCAGCTTGCCTCTCTGAA	GCUCAGCUUGCCUCUCUGAA
	Exon 4 STOP 1	TTTATCAAGACTGGCATGAC	UUUAUCAAGACUGGCAUGAC
	Exon 4 STOP 2	ATTTGGCCAAAGCCTTGAAT	AUUUGGCCAAAGCCUUGAAU
	Exon 4 STOP 3	TTATCTTCAAGCCACTCCAG	UUAUCUUCAAGCCACUCCAG
	Exon 5 SD	CTTACTTCAGTGTCCTATGC	CUUACUUCAGUGUCCUAUGC
DGKA	Exon 1 ATG	ACCCCATTTTGTTCCGCCTC	ACCCCAUUUUGUUCCGCCUC
	Exon 5 SD	TCACATTCTAACTTGTCTTC	UCACAUUCUAACUUGUCUUC
	Exon 6 SA 1	TGACTGTGGGGTGTTTTAGG	UGACUGUGGGGUGUUUUAGG
	Exon 6 SA 2	GTGACTGTGGGGTGTTTTAG	GUGACUGUGGGGUGUUUUAG
	Exon 6 SA 3	GGTGACTGTGGGGTGTTTTA	GGUGACUGUGGGGUGUUUUA
	Exon 6 SA 4	AGGTGACTGTGGGGTGTTTT	AGGUGACUGUGGGGUGUUUU
	Exon 7 SA 1	AATCTGAGCACAGAGTGGAA	AAUCUGAGCACAGAGUGGAA
	Exon 7 SA 2	TGAAGAATCTGAGCACAGAG	UGAAGAAUCUGAGCACAGAG
	Exon 10 SA	TGATTGGACCTTGGGGAGAA	UGAUUGGACCUUGGGGAGAA
	Exon 11 SA 1	GTGGATCTGAAAGACGAGGT	GUGGAUCUGAAAGACGAGGU
	Exon 11 SA 2	CGTGGATCTGAAAGACGAGG	CGUGGAUCUGAAAGACGAGG
	Exon 12 SD	CTGTACCCGCAGAGCCTCAG	CUGUACCCGCAGAGCCUCAG
	Exon 15 SA	AATCGGAGCCTGAGACAAAG	AAUCGGAGCCUGAGACAAAG
	Exon 16 SD	ACTTACCTCCTCCCCATCTT	ACUUACCUCCUCCCCAUCUU
	Exon 17 SA	AACCTAGGAGTGGAGAAGAC	AACCUAGGAGUGGAGAAGAC
	Exon 18 SA	AGAGGCATCCTGGAGAGTTC	AGAGGCAUCCUGGAGAGUUC
	Exon 20 SA	CCCACAGATCTGAGAGGAGG	CCCACAGAUCUGAGAGGAGG
	Exon 21 SA 1	TTAGGTCTGGGGACGAAGTA	UUAGGUCUGGGGACGAAGUA
	Exon 21 SA 2	CTTAGGTCTGGGGACGAAGT	CUUAGGUCUGGGGACGAAGU
	Exon 22 SA	TGTGGTGCTATAGGAGGCCA	UGUGGUGCUAUAGGAGGCCA
	Iso SA	GACCCTGGAAGAGTTGGGGC	GACCCUGGAAGAGUUGGGGC
DGKZ	Exon 3 SA	CTGACTCCTAGCCACGGGAT	CUGACUCCUAGCCACGGGAU
	Exon 4 SA	CTGCTGCAGGACAGGAAGAG	CUGCUGCAGGACAGGAAGAG
	Exon 5 SD	ACTTACCTCGCGGACATTCC	ACUUACCUCGCGGACAUUCC
	Exon 6 SA 1	AAGGTTGGCTGGGGGAGAAG	AAGGUUGGCUGGGGGAGAAG
	Exon 6 SA 2	AAAGGTTGGCTGGGGGAGAA	AAAGGUUGGCUGGGGGAGAA
	Exon 7 SA 1	ATCCCTGAGGGTAGACAGGA	AUCCCUGAGGGUAGACAGGA
	Exon 7 SA 2	TGGAATCCCTGAGGGTAGAC	UGGAAUCCCUGAGGGUAGAC
	Exon 8 SA 1	GTGGTACTGAGGAGAGCGAG	GUGGUACUGAGGAGAGCGAG
	Exon 8 SA 2	TGTGGTACTGAGGAGAGCGA	UGUGGUACUGAGGAGAGCGA
	Exon 8 SA 3	CTGTGGTACTGAGGAGAGCG	CUGUGGUACUGAGGAGAGCG
	Exon 8 SD 1	AGTACTCACCTGGGGCCTCC	AGUACUCACCUGGGGCCUCC
	Exon 9 SA 1	GTATTCTGCAAGGGAAGCAG	GUAUUCUGCAAGGGAAGCAG
	Exon 9 SA 2	AGTATTCTGCAAGGGAAGCA	AGUAUUCUGCAAGGGAAGCA
	Exon 9 SA 3	GAGTATTCTGCAAGGGAAGC	GAGUAUUCUGCAAGGGAAGC
	Exon 10 SA	CTCCTGGGAGACAAGGTGAG	CUCCUGGGAGACAAGGUGAG

	Exon 11 SA	TTTGCACCCTGGATGGGATG	UUUGCACCCUGGAUGGGAUG
	Exon 12 SA	AGCCTGGGCTCATGGGAAGA	AGCCUGGGCUCAUGGGAAGA
	Exon 13 SD 1	GCTTACCCCACCCCAGTTGA	GCUUACCCCACCCCAGUUGA
	Exon 13 SD 2	TGCTTACCCCACCCCAGTTG	UGCUUACCCCACCCCAGUUG
	Exon 14 SA 1	TAGCCCTGGGGGAGCAGGCA	UAGCCCUGGGGGAGCAGGCA
	Exon 14 SA 2	GTAGCCCTGGGGGAGCAGGC	GUAGCCCUGGGGGAGCAGGC
	Exon 15 SA 1	GGCAACTGGAAAAAGGTCAA	GGCAACUGGAAAAAGGUCAA
	Exon 15 SA 2	GGGCAACTGGAAAAAGGTCA	GGGCAACUGGAAAAAGGUCA
	Exon 15 SD	GCTGCCAACCTCGAGACTCG	GCUGCCAACCUCGAGACUCG
	Exon 17 SA 1	AGCTGTCTGTGGGAGACAGA	AGCUGUCUGUGGGAGACAGA
	Exon 17 SA 2	AAGCTGTCTGTGGGAGACAG	AAGCUGUCUGUGGGAGACAG
	Exon 17 SD	CTCCTCACCTGGGGATGTTC	CUCCUCACCUGGGGAUGUUC
	Exon 18 SD	CCCACTCACCAACGACGTCA	CCCACUCACCAACGACGUCA
	Exon 19 SD 1	GTACTCGCTGTGCAGGGGGG	GUACUCGCUGUGCAGGGGGG
	Exon 19 SD 2	GACGTACTCGCTGTGCAGGG	GACGUACUCGCUGUGCAGGG
	Exon 19 SD 3	GGACGTACTCGCTGTGCAGG	GGACGUACUCGCUGUGCAGG
	Exon 19 SD 4	GGGACGTACTCGCTGTGCAG	GGGACGUACUCGCUGUGCAG
	Exon 20 SA	TGCTGGCTGTCGGGCAGGGC	UGCUGGCUGUCGGGCAGGGC
	Exon 22 SA	GCTCCTGTTGAGAGAAGCAA	GCUCCUGUUGAGAGAAGCAA
	Exon 23 SA	AGTGGTGGCTGTGGGAGAGA	AGUGGUGGCUGUGGGAGAGA
	Exon 24 SA 1	GCTCCTGTGGGGGGAGGTCA	GCUCCUGUGGGGGGAGGUCA
	Exon 24 SA 2	TGCTCCTGTGGGGGGAGGTC	UGCUCCUGUGGGGGGAGGUC
	Exon 24 SD	TCACCGGGGCGTGGGTGAGC	UCACCGGGGCGUGGGUGAGC
	Exon 27 SA	GGAGCTGAGGGCAGGAGGCA	GGAGCUGAGGGCAGGAGGCA
	Exon 27 SD	CCGGCTCACCGTGGTCCAGC	CCGGCUCACCGUGGUCCAGC
	Exon 28 SA 1	GGGGGCTGGAGAAGGGGAGG	GGGGGCUGGAGAAGGGGAGG
	Exon 28 SA 2	TGGGGGGGCTGGAGAAGGGG	UGGGGGGGCUGGAGAAGGGG
	Exon 29 SA 1	CCCGCTGGGGGAGCAGGGGC	CCCGCUGGGGGAGCAGGGGC
	Exon 29 SA 2	TCTCCCCGCTGGGGGAGCAG	UCUCCCCGCUGGGGGAGCAG
	iso exon 18 SA 1	ACACTGGAGGCGGGCGGGGA	ACACUGGAGGCGGGCGGGGA
	iso exon 18 SA 2	CACACTGGAGGCGGGCGGGG	CACACUGGAGGCGGGCGGGG
	iso exon 18 SA 3	CATCACACTGGAGGCGGGCG	CAUCACACUGGAGGCGGGCG
	Isoform exon 1/2 SD	CTACGCGAGCAGGGCGCTCC	CUACGCGAGCAGGGCGCUCC
	Isoform exon 2 SA	TGGCTGTTGGGCACAGAGAC	UGGCUGUUGGGCACAGAGAC
	Isoform exon 4 SA	ACTGACTTCTGCTGCAGGAC	ACUGACUUCUGCUGCAGGAC
DHX37	Exon 1 ATG 1	TCCCCATGGCGACTAGGCCA	UCCCCAUGGCGACUAGGCCA
	Exon 1 ATG 2	TTCCCCATGGCGACTAGGCC	UUCCCCAUGGCGACUAGGCC
	Exon 4 SD 1	GAACCTGCATTTCCGGGGAG	GAACCUGCAUUUCCGGGGAG
	Exon 4 SD 2	GTGGCGAACCTGCATTTCCG	GUGGCGAACCUGCAUUUCCG
	Exon 6 SA	TCACTGGGGGAGGAAGAACA	UCACUGGGGGAGGAAGAACA
	Exon 7 SA 1	ACCCTGTATGGGCAGAGTTC	ACCCUGUAUGGGCAGAGUUC
	Exon 7 SA 2	GACCCTGTATGGGCAGAGTT	GACCCUGUAUGGGCAGAGUU

	Exon 8 SA 1	AAGTCCTGGGGGGAGGTCCG	AAGUCCUGGGGGGAGGUCCG
	Exon 8 SA 2	GAAGTCCTGGGGGGAGGTCC	GAAGUCCUGGGGGGAGGUCC
	Exon 8 SA 3	GGAAGTCCTGGGGGGAGGTC	GGAAGUCCUGGGGGGAGGUC
	Exon 9 SA	AGGTTCCTCTGCAAAAGGAC	AGGUUCCUCUGCAAAAGGAC
	Exon 9 SD 1	GTTACCTTGATGACCGGCGG	GUUACCUUGAUGACCGGCGG
	Exon 9 SD 2	TGTGTTACCTTGATGACCGG	UGUGUUACCUUGAUGACCGG
	Exon 10 SA 1	CACCTGTGGGACGCCCAGGA	CACCUGUGGGACGCCCAGGA
	Exon 10 SA 2	ATTCCACCTGTGGGACGCCC	AUUCCACCUGUGGGACGCCC
	Exon 10 SD	CCTCACCTGCGGGCAGCATC	CCUCACCUGCGGGCAGCAUC
	Exon 11 SA 1	ATGCCACCTGTGGAAAGAAT	AUGCCACCUGUGGAAAGAAU
	Exon 11 SA 2	GATGCCACCTGTGGAAAGAA	GAUGCCACCUGUGGAAAGAA
	Exon 11 SD	TTTACCTTGTGGCCGGGCTC	UUUACCUUGUGGCCGGGCUC
	Exon 12 SA 1	TTTCTGGGAGAGGGGCAGGT	UUUCUGGGAGAGGGGCAGGU
	Exon 12 SA 2	TCCTTTTCTGGGAGAGGGGC	UCCUUUUCUGGGAGAGGGGC
	Exon 14 SA	CTCACCTGGAGGGAAAGCAG	CUCACCUGGAGGGAAAGCAG
	Exon 14 SD 1	TTACCTGTGCTTGCTTCTCT	UUACCUGUGCUUGCUUCUCU
	Exon 14 SD 2	GTTACCTGTGCTTGCTTCTC	GUUACCUGUGCUUGCUUCUC
	Exon 15 SA 1	AAGACCTAGGATTCGGGGAA	AAGACCUAGGAUUCGGGGAA
	Exon 15 SA 2	AAAGACCTAGGATTCGGGGA	AAAGACCUAGGAUUCGGGGA
	Exon 15 SD 1	ACCCACCTGTAGCAGTGGCC	ACCCACCUGUAGCAGUGGCC
	Exon 15 SD 2	GACCCACCTGTAGCAGTGGC	GACCCACCUGUAGCAGUGGC
	Exon 16 SA 1	AGCCTGGATGGAGAGAAACC	AGCCUGGAUGGAGAGAAACC
	Exon 16 SA 2	CAGCCTGGATGGAGAGAAAC	CAGCCUGGAUGGAGAGAAAC
	Exon 17 SD 1	CCTTACCTTTCTGCTTTCTG	CCUUACCUUUCUGCUUUCUG
	Exon 17 SD 2	GCCTTACCTTTCTGCTTTCT	GCCUUACCUUUCUGCUUUCU
	Exon 17 SD 3	GGCCTTACCTTTCTGCTTTC	GGCCUUACCUUUCUGCUUUC
	Exon 18 SA 1	CACCCTGGAGATGGAGGTGG	CACCCUGGAGAUGGAGGUGG
	Exon 18 SA 2	CTTCACCCTGGAGATGGAGG	CUUCACCCUGGAGAUGGAGG
	Exon 19 SD	ACAAACCCAGCAGCACCATG	ACAAACCCAGCAGCACCAUG
	Exon 20 SA 1	GCGCCTGGGGAACGAAGAGG	GCGCCUGGGGAACGAAGAGG
	Exon 20 SA 2	GGCGCCTGGGGAACGAAGAG	GGCGCCUGGGGAACGAAGAG
	Exon 20 SA 3	CGGCGCCTGGGGAACGAAGA	CGGCGCCUGGGGAACGAAGA
	Exon 20 SA 4	ACGGCGCCTGGGGAACGAAG	ACGGCGCCUGGGGAACGAAG
	Exon 20 SD	CCGTACCTGCGGTGGTCAGC	CCGUACCUGCGGUGGUCAGC
	Exon 23 SA	AGACGCCTGGGGGCCGGGGG	AGACGCCUGGGGGCCGGGGG
	Exon 24 SD	ACTCACAGAACACGCTGGCC	ACUCACAGAACACGCUGGCC
	Exon 25 SD	CACCTACCTGCCCTTCCAGC	CACCUACCUGCCCUUCCAGC
	Exon 26 SA	AAGACCTGATGAGAGACCAC	AAGACCUGAUGAGAGACCAC
	Exon 28 SA	GCAGGTCTGCAGGGGAGGGA	GCAGGUCUGCAGGGGAGGGA
ELOB (TCEB2)	Ex2 SA1 (Pos 6)	ACGTCCTGGGGGCGGCGGGC	ACGUCCUGGGGGCGGCGGGC
	Ex3 SA1 (Pos 7)	GTCATCCTGAGGAGAGAAGC	GUCAUCCUGAGGAGAGAAGC
	Ex4 SD1 (Pos 4)	TCACTGCACGGCTTGTTCAT	UCACUGCACGGCUUGUUCAU

	Ex5 SA1 (Pos 8)	ATGCAGGCTATGGGGGTGGG	AUGCAGGCUAUGGGGGUGGG
	Ex5 SA2 (Pos 9)	CATGCAGGCTATGGGGGTGG	CAUGCAGGCUAUGGGGGUGG
ENTPD1 (CD39)	Isoform ATG	GTCTTTTTCATTGACTAGAA	GUCUUUUUCAUUGACUAGAA
	Exon 2 SD	TGACTTACCTTAACGTTTTC	UGACUUACCUUAACGUUUUC
	Exon 3 STOP	TGCATCAAGTAGAAGAATGC	UGCAUCAAGUAGAAGAAUGC
	Isoform 2 ATG	CTCTTTCCATGCAATCAGTC	CUCUUUCCAUGCAAUCAGUC
	Exon 4 STOP	GTTCAGAAAGTAAATGAAAT	GUUCAGAAAGUAAAUGAAAU
	Isoform 3 ATG	TCACTTTCCATCCTGTACAA	UCACUUUCCAUCCUGUACAA
	Exon 5 STOP	GGCCAAGAGGAAGGTGCCTA	GGCCAAGAGGAAGGUGCCUA
	Exon 6 STOP 1	GCTCTGCAATTTCGCCTCTA	GCUCUGCAAUUUCGCCUCUA
	Exon 6 STOP 2	ACTCTGGCAGAAACTGGCCA	ACUCUGGCAGAAACUGGCCA
	Exon 6 SD	TATACTTGCCTGAATGTCCT	UAUACUUGCCUGAAUGUCCU
	Exon 7 SD	AACTTACCCCAAAATCCCCC	AACUUACCCCAAAAUCCCCC
	Exon 8 STOP	CTGTGCTCAGCCTTGGGAGG	CUGUGCUCAGCCUUGGGAGG
	Exon 9 SA 4	TTTTATCTAGAAGTGAAGTG	UUUUAUCUAGAAGUGAAGUG
	Exon 9 SA 3	TTTATCTAGAAGTGAAGTGA	UUUAUCUAGAAGUGAAGUGA
	Exon 9 SA 2	TTATCTAGAAGTGAAGTGAG	UUAUCUAGAAGUGAAGUGAG
	Exon 9 SA 1	TATCTAGAAGTGAAGTGAGG	UAUCUAGAAGUGAAGUGAGG
	Exon 9 SD	AAATTACCTTGCCAATGAAA	AAAUUACCUUGCCAAUGAAA
FADD	Exon 1 SD	CCCACCTTCTTCCCCAGGCG	CCCACCUUCUUCCCCAGGCG
	Exon 1 STOP	GAGCAGAACGACCTGGAGCC	GAGCAGAACGACCUGGAGCC
	Exon 2 STOP 1	GTCCTGCCAGATGAACCTGG	GUCCUGCCAGAUGAACCUGG
	Exon 2 STOP 2	CCTGGTACAAGAGGTTCAGC	CCUGGUACAAGAGGUUCAGC
	Exon 2 STOP 3	GTGACCTCCAGAACAGGAGT	GUGACCUCCAGAACAGGAGU
	Exon 2 STOP 4	TGACCTCCAGAACAGGAGTG	UGACCUCCAGAACAGGAGUG
	Exon 2 STOP 5	GTTCCATGACATCGGGGACA	GUUCCAUGACAUCGGGGACA
IL6	Exon 1 ATG	GGAGTTCATAGCTGGGCTCC	GGAGUUCAUAGCUGGGCUCC
	Exon 2 SD	CCTACCCACCTCCTTTCTCA	CCUACCCACCUCCUUUCUCA
	Exon 3 SD	GTACCTCATTGAATCCAGAT	GUACCUCAUUGAAUCCAGAU
	Exon 3 STOP	CTTCCAATCTGGATTCAATG	CUUCCAAUCUGGAUUCAAUG
	Exon 4 SA 1	CTCCTAAGAGGAAAGATGGT	CUCCUAAGAGGAAAGAUGGU
	Exon 4 SA 2	TCTCCTAAGAGGAAAGATGG	UCUCCUAAGAGGAAAGAUGG
	Exon 4 SA 3	AAGTCTCCTAAGAGGAAAGA	AAGUCUCCUAAGAGGAAAGA
	Exon 4 SD	CCACCTTTTTCTGCAGGAAC	CCACCUUUUUCUGCAGGAAC
	Exon 5 STOP 1	AGCTGCAGGCACAGAACCAG	AGCUGCAGGCACAGAACCAG
	Exon 5 STOP 2	GGCACAGAACCAGTGGCTGC	GGCACAGAACCAGUGGCUGC
	Exon 5 STOP 3	GTTCCTGCAGTCCAGCCTGA	GUUCCUGCAGUCCAGCCUGA
	Iso 1 Exon 2 SD	TGGGGGTACTGGGGCAGGGA	UGGGGGUACUGGGGCAGGGA
	Iso 1 Exon 4 STOP	ATTCCCTCAACTTGGTGTGG	AUUCCCUCAACUUGGUGUGG
IL6R	Exon 1 ATG	CGGCCAGCATGCTTCCTCCT	CGGCCAGCAUGCUUCCUCCU
	Exon 4 SA	TGACTGTTAGACACAAAACA	UGACUGUUAGACACAAAACA
	Exon 4 STOP 1	CTCCTGCCAGTTAGCAGTCC	CUCCUGCCAGUUAGCAGUCC

	Exon 4 STOP 2	ACTCAAACCTTTCAGGGTTG	ACUCAAACCUUUCAGGGUUG
	Exon 5 STOP	CCTGGCAAGACCCCCACTCC	CCUGGCAAGACCCCCACUCC
	Exon 6 SA	TTGACCTGAGGGCGGGGGCA	UUGACCUGAGGGCGGGGGCA
	Exon 6 STOP 1	CGTGGTGCAGCTTCGTGCCC	CGUGGUGCAGCUUCGUGCCC
	Exon 6 STOP 2	ACCTGTCCAAGGCGTGCCCA	ACCUGUCCAAGGCGUGCCCA
	Exon 7 STOP	CTGGGTCCCAAATGCCACCC	CUGGGUCCCAAAUGCCACCC
	Exon 8 SA	GGATTCTGCGGACAGAAGAA	GGAUUCUGCGGACAGAAGAA
	Exon 8 SD	GGAGCTCACCTGCATGGGGG	GGAGCUCACCUGCAUGGGGG
IL10	Exon 2 SA	TTTGCTGCAGGAAGAACAAA	UUUGCUGCAGGAAGAACAAA
	Exon2 STOP	GCAGCAAATGAAGGATCAGC	GCAGCAAAUGAAGGAUCAGC
	Exon 3 SA	ACCCTAAGGGCAGGAGCCAA	ACCCUAAGGGCAGGAGCCAA
	Exon 3 STOP 1	GATGATCCAGTTTTACCTGG	GAUGAUCCAGUUUUACCUGG
	Exon 3 STOP 2	GAACCAAGACCCAGACATCA	GAACCAAGACCCAGACAUCA
IL10RA	Ex 5 STOP (pos 6)	CCAGGCAGTGTGAGTCAGCT	CCAGGCAGUGUGAGUCAGCU
	Ex 6 STOP (pos 9)	TGGCCCTCCAGCTGTATGTG	UGGCCCUCCAGCUGUAUGUG
	Ex2 STOP (pos 8)	CTTCAAACCACACAGACGGA	CUUCAAACCACACAGACGGA
	Ex2 STOP (pos 8)	TGGGTGTCCAGTGGAGGATG	UGGGUGUCCAGUGGAGGAUG
	Ex2 STOP (pos 9)	GCTTCAAACCACACAGACGG	GCUUCAAACCACACAGACGG
	Ex3 SA (pos 8)	TCCATACCTGAGGAGATACC	UCCAUACCUGAGGAGAUACC
	Ex3 STOP (pos 8)	TGACGGTCCAGTTGGAGTGC	UGACGGUCCAGUUGGAGUGC
	Ex4 SD (pos 8)	CCCCATACCGTGAAGTTTCC	CCCCAUACCGUGAAGUUUCC
	Ex5 SD (pos 8)	TGACTCACACTGCCTGGTGA	UGACUCACACUGCCUGGUGA
	Ex5 SD (pos 9)	CTGACTCACACTGCCTGGTG	CUGACUCACACUGCCUGGUG
	Ex5 STOP (pos 7)	ACCAGGCAGTGTGAGTCAGC	ACCAGGCAGUGUGAGUCAGC
	Ex7 SA (pos 9)	TTGAAGAGCTGGGGAAGAGA	UUGAAGAGCUGGGGAAGAGA
	Ex7 STOP (pos 5)	AGCTCCAGTCAGATATTCCC	AGCUCCAGUCAGAUAUUCCC
	Ex7 STOP (pos 5)	AGTTCAAAACTCTGAGGGCC	AGUUCAAAACUCUGAGGGCC
	Ex7 STOP (pos 6)	TAGTTCAAAACTCTGAGGGC	UAGUUCAAAACUCUGAGGGC
	Ex7 STOP (pos 7)	CTGGTTCCACTGTCCCGTTG	CUGGUUCCACUGUCCCGUUG
	Ex7 STOP (pos 8)	GCTGGTTCCACTGTCCCGTT	GCUGGUUCCACUGUCCCGUU
	Ex7 STOP (pos 9)	TGGCATTCCAGGGTTACCTG	UGGCAUUCCAGGGUUACCUG
	Ex7 STOP (pos 9)	GGCTGGTTCCACTGTCCCGT	GGCUGGUUCCACUGUCCCGU
IRF4	Exon 1 SD (pos 9)	GCCGGAGACCTTGAAGAGCG	GCCGGAGACCUUGAAGAGCG
	Exon 1 STOP 1 (pos 7)	GGCGGCCGAGGCGGAGAGTT	GGCGGCCGAGGCGGAGAGUU
	Exon 1 STOP 2 (pos 8/9)	CGTTCTCCCACACCAGCCCG	CGUUCUCCCACACCAGCCCG
	Exon 1 STOP 3 (pos 5)	CGCGGTTGTAGTCCTGCTTG	CGCGGUUGUAGUCCUGCUUG
	Exon 2 SD 1	CTACCTTTTTTGGCTCCCTC	CUACCUUUUUUGGCUCCCUC
	Exon 2 STOP 1	GTCTTCCAGGTGGGAGGGTC	GUCUUCCAGGUGGGAGGGUC
	Exon 3 SA 1	CTCCTACATGTTTGGGGAAA	CUCCUACAUGUUUGGGGAAA
	Exon 3 SD 1	ATACCTGGGCTGGGAGCGAA	AUACCUGGGCUGGGAGCGAA
	Exon 3 SD 2	CATACCTGGGCTGGGAGCGA	CAUACCUGGGCUGGGAGCGA
	Exon 3 STOP 1	AGCCAAGCAGCTCACCCTGG	AGCCAAGCAGCUCACCCUGG

	Exon 3 STOP 2	CCCAGCCCAGGTATGGTGGA	CCCAGCCCAGGUAUGGUGGA
	Exon 4 SA 1	CTGCTAAAGGAGTGCAGGAG	CUGCUAAAGGAGUGCAGGAG
	Exon 4 SD 1	TCCTTACCATTTTCACAAGC	UCCUUACCAUUUUCACAAGC
	Exon 4 SD 2	CCTTACCATTTTCACAAGCT	CCUUACCAUUUUCACAAGCU
	Exon 4 STOP 1	AGTCCCTCCAGCTTCGGTCG	AGUCCCUCCAGCUUCGGUCG
	Exon 4 STOP 2	GTCCCTCCAGCTTCGGTCGA	GUCCCUCCAGCUUCGGUCGA
	Exon 4 STOP 3	TCCCTCCAGCTTCGGTCGAG	UCCCUCCAGCUUCGGUCGAG
	Exon 4 STOP 4	TACCAATGTCCCATGACGTT	UACCAAUGUCCCAUGACGUU
	Exon 4 STOP 5	GCCTTGCCAGTGGTGGCCGC	GCCUUGCCAGUGGUGGCCGC
	Exon 4 STOP 6	CCTTGCCAGTGGTGGCCGCG	CCUUGCCAGUGGUGGCCGCG
	Exon 5 SD 1	GCACTCACCTGAGAACGCCA	GCACUCACCUGAGAACGCCA
	Exon 6 SA 1	AGTCTGCAAACACAGAGCTC	AGUCUGCAAACACAGAGCUC
	Exon 6 STOP 1	TGTGCCAGAGCAGGATCTAC	UGUGCCAGAGCAGGAUCUAC
	Exon 6 STOP 2	GTGCCAGAGCAGGATCTACT	GUGCCAGAGCAGGAUCUACU
	Exon 6 STOP 3	GACACACAGCAGTTCTTGTC	GACACACAGCAGUUCUUGUC
	Exon 7 STOP 1	CTGCAAGCGTTTGCTCACCA	CUGCAAGCGUUUGCUCACCA
	Exon 7 STOP 2	TTCCAGGTGACTCTATGCTT	UUCCAGGUGACUCUAUGCUU
IRF8	Ex1 SA (Pos 7)	CCTTCTCATGGCAGGTGTCC	CCUUCUCAUGGCAGGUGUCC
	Ex1 SD (Pos 5)	CCCACAGATTCAGGGACTCC	CCCACAGAUUCAGGGACUCC
	Ex1 SD (Pos 6)	ACCCACAGATTCAGGGACTC	ACCCACAGAUUCAGGGACUC
	Ex2 SD (Pos 9)	GCTCTTTACCTTAAAAATGG	GCUCUUUACCUUAAAAAUGG
	Ex3 SD (Pos 4)	TTACATTTTTGCTCTTCCTC	UUACAUUUUUGCUCUUCCUC
	Ex4 SA (Pos 6)	GAAGGCTGCACAGTCAGGGG	GAAGGCUGCACAGUCAGGGG
	Ex4 SA (Pos5)	AAGGCTGCACAGTCAGGGGA	AAGGCUGCACAGUCAGGGGA
	Ex4 SA (Pos9)	ACAGAAGGCTGCACAGTCAG	ACAGAAGGCUGCACAGUCAG
	Ex5 SA1 (Pos 7)	CACCATCTGGGAGAATGCTG	CACCAUCUGGGAGAAUGCUG
	Ex5 SA2 (Pos 9)	GGAGAATGCTGTGGACAAGA	GGAGAAUGCUGUGGACAAGA
	Ex5 SD1 (Pos 9)	GGTACAGACCTCGGAAGAAC	GGUACAGACCUCGGAAGAAC
	Ex5 SD2 (Pos 5)	CAGACCTCGGAAGAACTGGC	CAGACCUCGGAAGAACUGGC
	Ex6 SA1 (Pos 9)	CTGCAGCTCTGGAATGACAC	CUGCAGCUCUGGAAUGACAC
JUNB	Ex1 SA1 (Pos 7)	GCACATCCGGGCGGCCCAGG	GCACAUCCGGGCGGCCCAGG
	Ex1 STOP1 (Pos 9)	GGATACGGCCGGGCCCCTGG	GGAUACGGCCGGGCCCCUGG
	Ex1 STOP2 (Pos 7)	TCTGGTCAGGGCTCGGACAC	UCUGGUCAGGGCUCGGACAC
	Ex1 STOP2 (Pos4)	GGACAGTACTTTTACCCCCG	GGACAGUACUUUUACCCCCG
	Ex1 STOP3 (Pos 4)	GAGCAGGAGGGCTTCGCCGA	GAGCAGGAGGGCUUCGCCGA
	Ex1 STOP4 (Pos 4)	GCGCAGCTGGGCTTGGGCCG	GCGCAGCUGGGCUUGGGCCG
	Ex1 STOP6 (Pos 8)	GGAACCGCAGACCGTGCCGG	GGAACCGCAGACCGUGCCGG
	EX1 STOP7 (Pos5)	CAGCCGGGACGCCACGCCGC	CAGCCGGGACGCCACGCCGC
	Ex1 STOP8 (Pos 5)	AGACCAAGAGCGCATCAAAG	AGACCAAGAGCGCAUCAAAG
	Ex1 STOP9 (Pos 5)	CAAGCGGCTGCGGAACCGGC	CAAGCGGCUGCGGAACCGGC
	Ex1 STOP10 (Pos 8)	GCGGCTGCGGAACCGGCTGG	GCGGCUGCGGAACCGGCUGG
LAIR-1	Exon 1 ATG	TCTCTTTCCATCTTCTGTCG	UCUCUUUCCAUCUUCUGUCG

(CD305	Iso 1 Ex 2 SD 2	ATGACTTACCCTCCTGCGTG	AUGACUUACCCUCCUGCGUG
	Iso 1 Ex 2 SD 1	CTTACCCTCCTGCGTGTGGA	CUUACCCUCCUGCGUGUGGA
	Exon 2 STOP	TGGGGTTCAAACATTCCGCC	UGGGGUUCAAACAUUCCGCC
	Exon 3 SA	AGCTTTCTGTAAACAGGGGC	AGCUUUCUGUAAACAGGGGC
	Exon 3 SD 1	TACTGACCAGCTGAGGAGCC	UACUGACCAGCUGAGGAGCC
	Exon 3 SD 2	CTACTGACCAGCTGAGGAGC	CUACUGACCAGCUGAGGAGC
	Exon 5 SA	CATCTAAGAAAGACAGAAAC	CAUCUAAGAAAGACAGAAAC
	Exon 6 SA 1	GGGGGCCCTAAGGACAGTCG	GGGGGCCCUAAGGACAGUCG
	Exon 6 SA 2	GGGGGGCCCTAAGGACAGTC	GGGGGGCCCUAAGGACAGUC
	Exon 8 SA	TGTCTTGGGGAGAAAATACA	UGUCUUGGGGAGAAAAUACA
	Exon 9 SA	GGCCTAAGAGGGAGAGACCC	GGCCUAAGAGGGAGAGACCC
LDHA	Ex0 SD1 (Pos 7)	CCCATACCTTAGCGTGGAAA	CCCAUACCUUAGCGUGGAAA
	Ex4 STOP (Pos 4)	ACCCACCCATGACAGCTTAA	ACCCACCCAUGACAGCUUAA
	Ex6 SD1 (Pos 9)	TAGACCTACCTTAATCATGG	UAGACCUACCUUAAUCAUGG
LIF	Ex2 SA1 (Pos9)	CACAACTCCTGGGGACAGTC	CACAACUCCUGGGGACAGUC
	Ex2 SD1 (Pos7)	AACTTACATAGAGAATAAAG	AACUUACAUAGAGAAUAAAG
	Ex2 SD2 (Pos 6)	ACTTACATAGAGAATAAAGA	ACUUACAUAGAGAAUAAAGA
	Ex2 STOP1 (Pos 5)	GAACCAGATCAGGAGCCAAC	GAACCAGAUCAGGAGCCAAC
	Ex4 SA1 (Pos 8)	CTGTGTACTGAGGGGCAGAA	CUGUGUACUGAGGGGCAGAA
	Ex4 SA2 (Pos9)	GCTGTGTACTGAGGGGCAGA	GCUGUGUACUGAGGGGCAGA
	Ex4 SA3 (Pos5)	TGTACTGAGGGGCAGAAGGG	UGUACUGAGGGGCAGAAGGG
LYN	Ex1 STOP1 (Pos8)	GGTGCTCCCAGGAGCGGAGG	GGUGCUCCCAGGAGCGGAGG
	Ex4 STOP 1 (Pos8)	AGAGGAACAAGGAGACATTG	AGAGGAACAAGGAGACAUUG
	Ex5 SD1 (Pos 5)	GACTCACTCTTCTGTTTCTA	GACUCACUCUUCUGUUUCUA
	Ex6 STOP1 (Pos 7)	GAAAGGCAGCTTTTGGCACC	GAAAGGCAGCUUUUGGCACC
	Ex8 STOP1 (Pos 4)	AGCCACAGAAGCCATGGGAT	AGCCACAGAAGCCAUGGGAU
	Ex8 STOP2 (Pos 5)	CCCCCGGGAGTCCATCAAGT	CCCCCGGGAGUCCAUCAAGU
	Ex9 SA1 (Pos 5)	TAACCTAGGAAGAAAAAAGA	UAACCUAGGAAGAAAAAAGA
	Ex9 SD1 (Pos 5)	CTCACCCTTGGCCATGTACT	CUCACCCUUGGCCAUGUACU
	Ex10 SA1 (Pos 7)	CCTGAACTGGAGGTGAACAA	CCUGAACUGGAGGUGAACAA
	Ex11 SA1 (Pos 7)	TGCAATCTGAAAACAGAAAT	UGCAAUCUGAAAACAGAAAU
	Ex12 SA1 (Pos4)	CACCTAAGGAAGAAGATATG	CACCUAAGGAAGAAGAUAUG
	Ex13 STOP1 (Pos6)	AGCAGCAGCCTTAGAGCACA	AGCAGCAGCCUUAGAGCACA
MAP4K4	Ex1 SD1 (Pos 7)	CACTCACCCGCAGGGAGGAG	CACUCACCCGCAGGGAGGAG
	Ex2 SA1 (Pos 7)	AGGATCCTGGAGAGGAAGGA	AGGAUCCUGGAGAGGAAGGA
	Ex2 SA2 (Pos 6)	GGATCCTGGAGAGGAAGGAG	GGAUCCUGGAGAGGAAGGAG
	Ex4 STOP1 (Pos7)	GATGACCAACTCTGGGTAGG	GAUGACCAACUCUGGGUAGG
	Ex5 SD1 (Pos 6)	CCTTACCCTCAGGATTTCTC	CCUUACCCUCAGGAUUUCUC
	Ex7 STOP1 (Pos 9)	GTGATTCACCGGGATATCAA	GUGAUUCACCGGGAUAUCAA
	Ex8 SA1 (Pos 4)	CAACTGTGGGAGGAAGAAAA	CAACUGUGGGAGGAAGAAAA
	Ex8 SD1 (Pos 9)	GCCTCTTACTCTGTAATCAT	GCCUCUUACUCUGUAAUCAU
	Ex8 SD2 (Pos 6)	TCTTACTCTGTAATCATAGG	UCUUACUCUGUAAUCAUAGG

	Ex10 SA1 (Pos 7)	AGAGAGCTGGGGAGAGGAGA	AGAGAGCUGGGGAGAGGAGA
	Ex12 STOP1 (Pos 6)	CTGAACAGGAAGGAGAGCCA	CUGAACAGGAAGGAGAGCCA
	Ex13 SA1 (Pos 7)	ATGGAACTGTTGGAAAAAGC	AUGGAACUGUUGGAAAAAGC
	Ex13 STOP2 (Pos 6)	TTGAGCAGCAGAAAGAACAG	UUGAGCAGCAGAAAGAACAG
	Ex14 STOP1 (Pos 9)	GGAGAGAGCGGGAAGCTAGA	GGAGAGAGCGGGAAGCUAGA
	Ex15 STOP1 (Pos 5)	CCTTCAGCAGCAGCTGCTCC	CCUUCAGCAGCAGCUGCUCC
	Ex16 SA1 (Pos 8)	GGCACTCCTTGGAGAGGGAG	GGCACUCCUUGGAGAGGGAG
	Ex16 STOP1 (Pos 7)	CTCCCGCCATCGGCACTCCT	CUCCCGCCAUCGGCACUCCU
	Ex17 STOP1 (Pos 7)	GAAGCCCAGTCTAAGCAGAC	GAAGCCCAGUCUAAGCAGAC
	Ex17 SD1 (Pos 8)	GTCGCTACCTGTGGCTCCGC	GUCGCUACCUGUGGCUCCGC
	Ex18 STOP1 (Pos 5)	TAGACCAAGCCTTTTGGGTA	UAGACCAAGCCUUUUGGGUA
	Ex18 STOP1 (Pos 4)	GGGCAGCAGAATAGCCAGGC	GGGCAGCAGAAUAGCCAGGC
	Ex19 STOP1 (Pos 9)	AGGCTTCTGTGGGAGAGAGT	AGGCUUCUGUGGGAGAGAGU
	Ex19 STOP2 (Pos 8)	TGGGTCTCAGAGTGGCTCCG	UGGGUCUCAGAGUGGCUCCG
	Ex20 SA1 (Pos7)	ATGATGCTGTTGGGTTCAAA	AUGAUGCUGUUGGGUUCAAA
	Ex20 SA2 (Pos 6)	TGATGCTGTTGGGTTCAAAA	UGAUGCUGUUGGGUUCAAAA
	Ex20 SD1 (Pos 8)	ATCCTTACAGCAGGCTTGAG	AUCCUUACAGCAGGCUUGAG
	Ex20 SD2 (Pos (9)	AATCCTTACAGCAGGCTTGA	AAUCCUUACAGCAGGCUUGA
	Ex24 STOP1 (Pos 8)	TAGGCCACAGAGTGACACCC	UAGGCCACAGAGUGACACCC
	Ex25 STOP (Pos 8)	CCGAAGACGATTTCAACAAA	CCGAAGACGAUUUCAACAAA
	Ex26 STOP1 (Pos 4)	AAGCAGGGATGGACAACCGT	AAGCAGGGAUGGACAACCGU
	Ex30 STOP1 (Pos 5)	TCCCCAGCCCATTGTCTGAT	UCCCCAGCCCAUUGUCUGAU
	Ex30 STOP2 (Pos 7)	GAGATCCGATCTGTGGAAAC	GAGAUCCGAUCUGUGGAAAC
MAPK14	Ex1 SD1 (Pos 7)	CACTCACCACACAGAGCCAT	CACUCACCACACAGAGCCAU
	Ex1 STOP1 (Pos 5)	TTACCAGAACCTGTCTCCAG	UUACCAGAACCUGUCUCCAG
	Ex2 SA1 (Pos 8)	AGCAGCACTAAGGAGAAAAA	AGCAGCACUAAGGAGAAAAA
	Ex2 SA2 (Pos 7)	GCAGCACTAAGGAGAAAAAA	GCAGCACUAAGGAGAAAAAA
	Ex5 SA1 (Pos 6)	AGGTCCTAGGAAGCAAATAC	AGGUCCUAGGAAGCAAAUAC
	Ex5 SA2 (Pos 5)	GGTCCTAGGAAGCAAATACA	GGUCCUAGGAAGCAAAUACA
	Ex6 SA1 (Pos 7)	CAGAATCTAAAGGGCAGAAG	CAGAAUCUAAAGGGCAGAAG
	Ex6 STOP1 (Pos 4)	ATCCAGTTCAGCATGATCTC	AUCCAGUUCAGCAUGAUCUC
	Ex7 SD1 (Pos 5)	AATACCTCAGTTGCCGGTGC	AAUACCUCAGUUGCCGGUGC
	Ex7 STOP1 (Pos 9)	CCCACTGACCAAATATCAAC	CCCACUGACCAAAUAUCAAC
	Ex8 SA1 (Pos 4)	TATCTAATGGTGGACCATAA	UAUCUAAUGGUGGACCAUAA
	Ex9 SA1 (Pos 5)	ATATCTTAGATGCCTAGTCA	AUAUCUUAGAUGCCUAGUCA
	Ex9 STOP1 (Pos 4)	CAGCAGATTATGCGTCTGAC	CAGCAGAUUAUGCGUCUGAC
	Ex10 SA1 (Pos 9)	TTTCTTGCCTGAAAAAACAA	UUUCUUGCCUGAAAAAACAA
	Ex11 SA1 (Pos 4)	CAGCTATCAGTACCATAGAC	CAGCUAUCAGUACCAUAGAC
	Ex11 STOP1 (Pos 6)	ATGATCAGTCCTTTGAAAGC	AUGAUCAGUCCUUUGAAAGC
	Ex12 SA1 (Pos 8)	GTCAGGCCTAGAAATTGGGA	GUCAGGCCUAGAAAUUGGGA
MEF2D	Ex2 SA1 (Pos 8)	AAGTCACCTGCAGAGAAGGA	AAGUCACCUGCAGAGAAGGA
	Ex2 SA2 (Pos 7)	AGTCACCTGCAGAGAAGGAT	AGUCACCUGCAGAGAAGGAU

	Ex2 SD1 (Pos 9)	TGGGCCCACCTCGATGATGT	UGGGCCCACCUCGAUGAUGU
	Ex3 STOP1 (Pos 5)	GGAACAGAGCCCCCTGCTGG	GGAACAGAGCCCCCUGCUGG
	Ex3 STOP2 (Pos 8)	CAAGTACCGACGCGCCAGCG	CAAGUACCGACGCGCCAGCG
	Ex4 SA1 (Pos 6)	AGCGCCTGGGGGGAAGGGGC	AGCGCCUGGGGGGAAGGGGC
	Ex4 STOP1 (Pos 8)	AATGATGCAGAGTTATAGAC	AAUGAUGCAGAGUUAUAGAC
	Ex5 SA1 (Pos 8)	CAGTTGACTAGACAGAAAGA	CAGUUGACUAGACAGAAAGA
	Ex5 SA2 (Pos 5)	TTGACTAGACAGAAAGATGG	UUGACUAGACAGAAAGAUGG
	Ex5 SA3 (Pos 4)	TGACTAGACAGAAAGATGGA	UGACUAGACAGAAAGAUGGA
	Ex5 STOP1 (Pos 6)	CCCAGCAGCCAGCACTACAG	CCCAGCAGCCAGCACUACAG
	Ex6 SD1 (Pos 9)	TGCACTCACCAACAGGGCTG	UGCACUCACCAACAGGGCUG
	Ex6 SD2 (Pos 5)	CTCACCAACAGGGCTGGGGC	CUCACCAACAGGGCUGGGGC
	Ex7 STOP1 (Pos 6)	ACCTGCGAGTCATCACTTCC	ACCUGCGAGUCAUCACUUCC
	Ex9 SD1 (Pos 5)	CTCACCTGTGTTGTAGGCAG	CUCACCUGUGUUGUAGGCAG
	Ex9 SD2 (Pos 4)	TCACCTGTGTTGTAGGCAGT	UCACCUGUGUUGUAGGCAGU
	Ex10 STOP1 (Pos 5)	ACCTCAGCAACAGTCCCACC	ACCUCAGCAACAGUCCCACC
	Ex11 SA1 (Pos 7)	CCCGGGCTGGAGGCAGGCAA	CCCGGGCUGGAGGCAGGCAA
	Ex12 SA1 (Pos 5)	ATGTCTGTGAAGAGAGGAGA	AUGUCUGUGAAGAGAGGAGA
MGAT5	Ex2 STOP1 (Pos 5)	TTTGCAGCGCATTGGCAAGT	UUUGCAGCGCAUUGGCAAGU
	Ex6 STOP1 (Pos 6)	TCCGGCGAATGGCTGACGCA	UCCGGCGAAUGGCUGACGCA
	Ex7 SA1 (Pos 8)	CGAGGACCTGGAAAACAAAG	CGAGGACCUGGAAAACAAAG
	Ex8 SD1 (Pos 7)	CACTTACTGGTAATGAACCC	CACUUACUGGUAAUGAACCC
	Ex9 STOP1 (Pos 4)	CTCCGAGTCCTTGATTCATT	CUCCGAGUCCUUGAUUCAUU
	Ex9 STOP2 (Pos 8)	CAGATTCCATTTTCCCCAAG	CAGAUUCCAUUUUCCCCAAG
	Ex9 STOP3 (Pos9)	TCAGATTCCATTTTCCCCAA	UCAGAUUCCAUUUUCCCCAA
	Ex10 STOP1 (Pos 7)	GAAGGCCATGCCTGGAACAC	GAAGGCCAUGCCUGGAACAC
	Ex11 SD1 (Pos 4)	TTACCTTGGTTTCTCGAAGA	UUACCUUGGUUUCUCGAAGA
	Ex12 SD1 (Pos 8)	ATGCTTACCTCTCTCAGAGT	AUGCUUACCUCUCUCAGAGU
	Ex13 STOP1 (Pos 8)	CAACAATCAGGAGGAAGTAG	CAACAAUCAGGAGGAAGUAG
	Ex15 STOP1 (Pos 5)	GTGGCCACATCACTTGCCCA	GUGGCCACAUCACUUGCCCA
	Ex15 STOP2 (Pos 8)	GCCCGGGCAGTCCTGCAAGC	GCCCGGGCAGUCCUGCAAGC
	Ex16 SA1 (Pos 5)	GTACCTGAAGAGGAAGAGAA	GUACCUGAAGAGGAAGAGAA
	Ex16 STOP2 (Pos 6)	CCCTGCCGGGACTTCATCAA	CCCUGCCGGGACUUCAUCAA
NT5E (CD73)	Ex1 SD1 (Pos 4)	TTACCATGGCATCGTAGCGC	UUACCAUGGCAUCGUAGCGC
	Ex1 STOP (Pos 4)	AGGCCACAGCACCGCGCCCA	AGGCCACAGCACCGCGCCCA
	Ex3 STOP1 (Pos 5)	CGCTCAGAAAGTGAGGGGTG	CGCUCAGAAAGUGAGGGGUG
	Ex4 STOP1 (Pos 7)	GTAGTCCAGGCCTATGCTTT	GUAGUCCAGGCCUAUGCUUU
	Ex5 SA1 (Pos 4)	GATCTAGAAGAAAGAAAAGA	GAUCUAGAAGAAAGAAAAGA
	Ex5 SD1 (Pos 5)	TTACCATTGCATCACAAATC	UUACCAUUGCAUCACAAAUC
	Ex7 SD1 (Pos 9)	GTGACTTACCGCCCACCTGC	GUGACUUACCGCCCACCUGC
	Ex7 STOP1 (Pos 4)	CTCCCAGGTAATTGTGCCTG	CUCCCAGGUAAUUGUGCCUG
	Ex8 STOP1 (Pos 8)	AGGTGACCAAGATATCAACG	AGGUGACCAAGAUAUCAACG
	Ex8 STOP2 (Pos 4)	GGTCGGATCAAGTTTTCCAC	GGUCGGAUCAAGUUUUCCAC

ODC1	Ex2 SA1 (Pos 9)	TTATCATCCTGAAACAAGAG	UUAUCAUCCUGAAACAAGAG
	Ex3 SA1 (Pos 7)	TTCAGTCTGAAAAAGAAGAG	UUCAGUCUGAAAAAGAAGAG
	Ex7 STOP1 (Pos 7)	AAGGAACAGACGGGCTCTGA	AAGGAACAGACGGGCUCUGA
	Ex8 SD1 (Pos 8)	GAAATTACCTTTTGCAGAAG	GAAAUUACCUUUUGCAGAAG
	Ex8 SD2 (Pos 9)	AGAAATTACCTTTTGCAGAA	AGAAAUUACCUUUUGCAGAA
	Ex10 SA1 (Pos 6)	GTTGCCTGAGAAAGAAAAAG	GUUGCCUGAGAAAGAAAAAG
	Ex10 STOP1 (Pos 6)	CTCTCCCAGGCACAAGACAC	CUCUCCCAGGCACAAGACAC
OTULINL (FAM105A)	Exon 1 ATG	CCGCCATGCCGGCCGCGCTG	CCGCCAUGCCGGCCGCGCUG
	Exon 2 STOP	GAAGTGACCAAGTTCACTCC	GAAGUGACCAAGUUCACUCC
	Exon 3 SA 2	CAATCCACCTGAAAGATAAA	CAAUCCACCUGAAAGAUAAA
	Exon 3 SA 1	AATCCACCTGAAAGATAAAA	AAUCCACCUGAAAGAUAAAA
	Exon 4 STOP	GTTATTTCAGATATTCAGCC	GUUAUUUCAGAUAUUCAGCC
	Exon 5 STOP 1	TGTTTTCACAAGGTTGTAAT	UGUUUUCACAAGGUUGUAAU
	Exon 5 STOP 2	TGGATTCAGCAGTACAGTTT	UGGAUUCAGCAGUACAGUUU
	Exon 5 STOP 3	AAAACACAGGTAAGTGTTTG	AAAACACAGGUAAGUGUUUG
	Exon 5 STOP 4	AAACACAGGTAAGTGTTTGC	AAACACAGGUAAGUGUUUGC
	Exon 5 STOP 5	AACACAGGTAAGTGTTTGCG	AACACAGGUAAGUGUUUGCG
	Exon 6 STOP 1	TGAACAAATGAAGACTAAAA	UGAACAAAUGAAGACUAAAA
	Exon 6 STOP 2	ACTAGAGCAGGTAACCGGGG	ACUAGAGCAGGUAACCGGGG
	Exon 7 STOP	ATCTCCGGCCAGTCCCTGAG	AUCUCCGGCCAGUCCCUGAG
PAG1	E1 STOP1 (Pops 9)	GACAGATGCAGATCACCCTG	GACAGAUGCAGAUCACCCUG
	Ex4 STOP1 (pos 9)	AGGAAGTCCAGACATCGGCC	AGGAAGUCCAGACAUCGGCC
	Ex4 STOp2 (Pos 9)	TGGATTCCCAGGACAGCACA	UGGAUUCCCAGGACAGCACA
	Ex4 SD1 (Pos 6)	GCCCACCTTGTTAGTTTCAC	GCCCACCUUGUUAGUUUCAC
	Ex4 STOP4 (Pos 8)	AAACCTTCAGGAGAAGGAAG	AAACCUUCAGGAGAAGGAAG
	Ex6 SA1 (Pos 9)	GCTGAGATCTAGGAGACAAA	GCUGAGAUCUAGGAGACAAA
	Ex6 STOP1 (Pos 6)	GATACAGACTCTCAACAGAG	GAUACAGACUCUCAACAGAG
	Ex6 STOP2 (Pos 6)	CCATTCAAGGGGACCCACAG	CCAUUCAAGGGGACCCACAG
	Ex6 STOP3 (Pos 5)	GGGGCAGTCGCTTACAGTTC	GGGGCAGUCGCUUACAGUUC
PDIA3	Ex1 SD1 (Pos 6)	ACTTACCCTCTGACTTCATA	ACUUACCCUCUGACUUCAUA
	Ex6 STOP1 (Pos 7)	ACAGAGCAAAAAATGACCAG	ACAGAGCAAAAAAUGACCAG
	Ex7 SD1 (Pos 7)	TTACCTGTTTCTCCAGTAGT	UUACCUGUUUCUCCAGUAGU
	Ex9 SA1 (Pos 5)	ACGCCTACAATTGGAAAACA	ACGCCUACAAUUGGAAAACA
	Ex9 SA2 (Pos 4)	CGCCTACAATTGGAAAACAA	CGCCUACAAUUGGAAAACAA
	Ex9 STOP1 (Pos 7)	TTCCTGCAGGATTACTTTGA	UUCCUGCAGGAUUACUUUGA
	Ex11 SA1 (Pos 8)	TGCTGAGCTGTTAATAAAAC	UGCUGAGCUGUUAAUAAAAC
	Ex12 SD1 (Pos 8)	TCACTTACTTCATATTTCTT	UCACUUACUUCAUAUUUCUU
PHD1 (EGLN2)	Ex1 STOP1 (Pos 8)	CCAGCCGCAGCCCCTAAGTC	CCAGCCGCAGCCCCUAAGUC
	Ex1 STOP2 (Pos 5)	TGGCCGGGCCAGGATGGGAG	UGGCCGGGCCAGGAUGGGAG
	Ex1 STOP3 (Pos 6)	ACGGGCAGCTAGTGAGCCAG	ACGGGCAGCUAGUGAGCCAG
	Ex1 STOP4 (Pos 5)	CGGGCAGCTAGTGAGCCAGA	CGGGCAGCUAGUGAGCCAGA
	Ex2 SA1 (Pos 4)	GGCCTGGCAGGGATGGAGGG	GGCCUGGCAGGGAUGGAGGG

	Ex2 SA2 (Pos 7)	CATGGCCTGGCAGGGATGGA	CAUGGCCUGGCAGGGAUGGA
	Ex2 SA3 (Pos 8)	CCATGGCCTGGCAGGGATGG	CCAUGGCCUGGCAGGGAUGG
	Ex2 STOP1 (Pos 8)	TAACGTCCCAGTTCTGATTC	UAACGUCCCAGUUCUGAUUC
	Ex2 STOP2 (Pos 5)	GAATCAGAACTGGGACGTTA	GAAUCAGAACUGGGACGUUA
	Ex3 SA1 (Pos 6)	TGCACCTGGGGGCAGGCCAA	UGCACCUGGGGGCAGGCCAA
	Ex3 SD1 (Pos 6)	TCATACCTGGTGGCATAGGC	UCAUACCUGGUGGCAUAGGC
	Ex3 STOP1 (Pos 8)	CCTGCTGCAGATCTTCCCTG	CCUGCUGCAGAUCUUCCCUG
	Ex4 SA1 (Pos 4)	TACCTGGAGACCAGGGTGGT	UACCUGGAGACCAGGGUGGU
	Ex4 SA2 (Pos 8)	GGCGTACCTGGAGACCAGGG	GGCGUACCUGGAGACCAGGG
	Ex5 SA1 (Pos 8)	CCTGATGCTGGGGGTGAGAG	CCUGAUGCUGGGGGUGAGAG
	Ex5 SA2 (Pos 9)	TCCTGATGCTGGGGGTGAGA	UCCUGAUGCUGGGGGUGAGA
PHD2 (EGLN1)	Ex1 STOP1 (Pos 4)	CGGCAGTACTGCGAGCTGTG	CGGCAGUACUGCGAGCUGUG
	Ex1 STOP1 (Pos 5)	CGGACAGCAGATCGGCGACG	CGGACAGCAGAUCGGCGACG
	Ex1 STOP2 (Pos 8)	GCTTCTTCCAGTCCTGACGC	GCUUCUUCCAGUCCUGACGC
	Ex3 SD1 (Pos 6)	TACTACCTTGTAGCATATGC	UACUACCUUGUAGCAUAUGC
	Ex3 STOP1 (Pos 7)	ATACTTCGAATTTTTCCAGA	AUACUUCGAAUUUUUCCAGA
PHD3 (EGLN3)	Ex1 STOP1 (Pos 7)	GTCAAGCAGCTGCACTGCAC	GUCAAGCAGCUGCACUGCAC
	Ex1 STOP2 (Pos 6)	TCAAGCAGCTGCACTGCACC	UCAAGCAGCUGCACUGCACC
	Ex1 STOP3 (Pos 5)	CAAGCAGCTGCACTGCACCG	CAAGCAGCUGCACUGCACCG
	Ex3 SA1 (Pos 5)	GTAGCTGAAAGACACAAAGA	GUAGCUGAAAGACACAAAGA
	Ex3 SA2 (Pos 4)	TAGCTGAAAGACACAAAGAA	UAGCUGAAAGACACAAAGAA
	Ex3 SD1 (Pos 7)	CTATTACCTGGTTGCGTAAG	CUAUUACCUGGUUGCGUAAG
	Ex3 SD2 (Pos 6)	TATTACCTGGTTGCGTAAGA	UAUUACCUGGUUGCGUAAGA
	Ex3 STOP1 (Pos 7)	ATCCTGCGGATATTTCCAGA	AUCCUGCGGAUAUUUCCAGA
	Ex3 STOP2 (Pos 6)	TCCTGCGGATATTTCCAGAG	UCCUGCGGAUAUUUCCAGAG
	Ex5 SA1 (Pos 4)	TCCCTGGGTTGGGGACAGAA	UCCCUGGGUUGGGGACAGAA
PIK3CD	Ex1 SD1 (Pos 4)	ATACCTGCTTGATGGTGCTG	AUACCUGCUUGAUGGUGCUG
	Ex1 STOP1 (Pos 8)	CCTTGGTCCAGAATTCCATG	CCUUGGUCCAGAAUUCCAUG
	Ex2 SA1 (Pos 5)	GCAGCTGGAGGGACAGTCAC	GCAGCUGGAGGGACAGUCAC
	Ex2 STOP1 (Pos 7)	GCGGTGCCACAGCAGCTGGA	GCGGUGCCACAGCAGCUGGA
	Ex2 STOP2 (Pos 8)	CGCGGTGCCACAGCAGCTGG	CGCGGUGCCACAGCAGCUGG
	Ex2 STOP3 (Pos 6)	GGCCCCCAGGTTTGAGCCGA	GGCCCCCAGGUUUGAGCCGA
	Ex2 STOP4 (Pos 7)	AGGCCCCCAGGTTTGAGCCG	AGGCCCCCAGGUUUGAGCCG
	Ex3 SA1 (Pos 6)	CTCTCCTGTGGGGAGGAGGG	CUCUCCUGUGGGGAGGAGGG
	Ex3 SA2 (Pos 9)	AAGCTCTCCTGTGGGGAGGA	AAGCUCUCCUGUGGGGAGGA
	Ex3 SD1 (Pos 5)	CTCACCTGGAACTGGCAGAG	CUCACCUGGAACUGGCAGAG
	Ex3 SD2 (Pos 4)	TCACCTGGAACTGGCAGAGC	UCACCUGGAACUGGCAGAGC
	Ex4 SA1 (Pos 7)	GATGTACTGAGACGGGGTGC	GAUGUACUGAGACGGGGUGC
	Ex4 SA2 (Pos 6)	ATGTACTGAGACGGGGTGCA	AUGUACUGAGACGGGGUGCA
	Ex4 SD1 (Pos 6)	TCTCACCTTCTTCGCAGGAA	UCUCACCUUCUUCGCAGGAA
	Ex4 SD2 (Pos 5)	CTCACCTTCTTCGCAGGAAT	CUCACCUUCUUCGCAGGAAU
	Ex5 SA1 (Pos4)	AGGCTGGGGGCCGGGGAAGC	AGGCUGGGGGCCGGGGAAGC

	Ex5 SD1 (Pos 6)	CCCCACCTTCATCCGCTCGT	CCCCACCUUCAUCCGCUCGU
	Ex6 SA1 (Pos 7)	CACCAGCTGTAGAAGGTGCC	CACCAGCUGUAGAAGGUGCC
	Ex6 SA2 (Pos 8)	CCACCAGCTGTAGAAGGTGC	CCACCAGCUGUAGAAGGUGC
	Ex6 STOP1 (Pos 4)	GTGCAGGCCGGGCTTTTCCA	GUGCAGGCCGGGCUUUUCCA
	Ex7 SA1 (Pos 4)	GTCCTGCAGAAGGACAGGGC	GUCCUGCAGAAGGACAGGGC
	Ex7 SA2 (Pos 8)	GGCAGTCCTGCAGAAGGACA	GGCAGUCCUGCAGAAGGACA
	Ex7 SA3 (Pos 9)	GGGCAGTCCTGCAGAAGGAC	GGGCAGUCCUGCAGAAGGAC
	Ex7 STOP1 (Pos 8)	CAAGGACCAGCTTAAGACCG	CAAGGACCAGCUUAAGACCG
	Ex7 STOP2 (Pos 9)	ACAAGGACCAGCTTAAGACC	ACAAGGACCAGCUUAAGACC
	Ex8 SD1 (Pos 7)	CACTGACCTTCTCCAGGGCG	CACUGACCUUCUCCAGGGCG
	Ex8 SD2 (Pos 8)	CCACTGACCTTCTCCAGGGC	CCACUGACCUUCUCCAGGGC
	Ex9 STOP1 (Pos 9)	AGGCGCATCCACAGCAGCTC	AGGCGCAUCCACAGCAGCUC
	Ex9 STOP2 (Pos 8)	TTTGTTGCAGATCTTGGAGC	UUUGUUGCAGAUCUUGGAGC
	Ex9 STOP3 (Pos 6)	TGTTGCAGATCTTGGAGCTG	UGUUGCAGAUCUUGGAGCUG
	Ex9 STOP3 (Pos 7)	TTGTTGCAGATCTTGGAGCT	UUGUUGCAGAUCUUGGAGCU
	Ex10 SA1 (Pos 6)	AGCTGCTGAGGGGTGTGGGC	AGCUGCUGAGGGGUGUGGGC
	Ex10 SA2 (Pos 7)	CTGCAGCTGCTGAGGGGTGT	CUGCAGCUGCUGAGGGGUGU
	Ex10 SA3 (Pos 8)	GCTGCAGCTGCTGAGGGGTG	GCUGCAGCUGCUGAGGGGUG
	Ex10 STOP1 (Pos 8)	TGTGGCCCAGGTGGGTGGGG	UGUGGCCCAGGUGGGUGGGG
	Ex11 SA1 (Pos 8)	AGAGCATCTGGGGGGAGCCG	AGAGCAUCUGGGGGGAGCCG
	Ex12 SA1 (Pos 8)	ATCGTCCCTGCAGGGAAGGA	AUCGUCCCUGCAGGGAAGGA
	Ex12 SD1 (Pos 5)	GCTACCGGAGGTGCCAGAAA	GCUACCGGAGGUGCCAGAAA
	Ex13 SA1 (Pos 5)	GGAGCTGGAAGGTGAAGGGA	GGAGCUGGAAGGUGAAGGGA
	Ex13 SA2 (Pos 6)	CGGAGCTGGAAGGTGAAGGG	CGGAGCUGGAAGGUGAAGGG
	Ex13 SA3 (Pos 9)	TCTCGGAGCTGGAAGGTGAA	UCUCGGAGCUGGAAGGUGAA
	Ex13 STOP1 (Pos 8)	GATGAAGCAGGTGAGGCCCA	GAUGAAGCAGGUGAGGCCCA
	Ex14 SA1 (Pos 4)	CCCCTGGTGGGCAGATGGGA	CCCCUGGUGGGCAGAUGGGA
	Ex14 SA2 (Pos 5)	CCCCCTGGTGGGCAGATGGG	CCCCCUGGUGGGCAGAUGGG
	Ex14 SA3 (Pos 8)	CTTCCCCCTGGTGGGCAGAT	CUUCCCCCUGGUGGGCAGAU
	Ex14 SA4 (Pos 9)	GCTTCCCCCTGGTGGGCAGA	GCUUCCCCCUGGUGGGCAGA
	Ex14 SD1 (Pos 5)	CTCACCAGACTTCAGCCAGC	CUCACCAGACUUCAGCCAGC
	Ex14 SD2 (Pos 4)	TCACCAGACTTCAGCCAGCA	UCACCAGACUUCAGCCAGCA
	Ex15 SA1 (Pos 4)	ACGCTGCCAGGCCAGAGAGC	ACGCUGCCAGGCCAGAGAGC
	Ex15 SA2 (Pos 5)	CACGCTGCCAGGCCAGAGAG	CACGCUGCCAGGCCAGAGAG
	Ex15 STOP1 (Pos 4)	GATCCACAGGGGCTTCATCT	GAUCCACAGGGGCUUCAUCU
	Ex16 SA1 (Pos 4)	GGTCTGTGCCACCGGCCGGT	GGUCUGUGCCACCGGCCGGU
	Ex16 SA2 (Pos 5)	AGGTCTGTGCCACCGGCCGG	AGGUCUGUGCCACCGGCCGG
	Ex16 SA3 (Pos 8)	CGGAGGTCTGTGCCACCGGC	CGGAGGUCUGUGCCACCGGC
	Ex16 STOP1 (Pos 5)	CCTGCAGATGATCCAGCTCA	CCUGCAGAUGAUCCAGCUCA
	Ex17 SA1 (Pos 4)	ATCCTAGGCAAGGGGGAAGA	AUCCUAGGCAAGGGGGAAGA
	Ex17 SA2 (Pos 5)	CATCCTAGGCAAGGGGGAAG	CAUCCUAGGCAAGGGGGAAG
	Ex18 SD1 (Pos 7)	CCTGTACCTGCCCACTCTCT	CCUGUACCUGCCCACUCUCU

	Ex18 STOP1 (Pos 8)	GAGTGGGCAGGTACAGGGGC	GAGUGGGCAGGUACAGGGGC
	Ex19 SA1 (Pos 6)	AACAGCTGAGGGGAGGGGAG	AACAGCUGAGGGGAGGGGAG
	Ex20 SA1 (Pos 4)	AACCTGCAGGTAGGGGACAG	AACCUGCAGGUAGGGGACAG
	Ex21 SA1 (Pos 4)	GTCCTGCAAACAAATCACAG	GUCCUGCAAACAAAUCACAG
	Ex21 STOP1 (Pos7)	GGTTTTCCAGCTCTCACGGA	GGUUUUCCAGCUCUCACGGA
	Ex21 STOP2 (Pos 8)	TGGTTTTCCAGCTCTCACGG	UGGUUUUCCAGCUCUCACGG
PIKFYVE	Ex2 STOP1 (Pos 7)	GGAGAACAGCAGCCTTTGAG	GGAGAACAGCAGCCUUUGAG
	Ex2 STOP2 (Pos 6)	CTGGTCCAACTTCCACTCAA	CUGGUCCAACUUCCACUCAA
	Ex5 STOP1 (Pos 8)	CTGGCATCCAGTATTGTTTC	CUGGCAUCCAGUAUUGUUUC
	Ex7 SA1 (Pos 8)	AATCAAACTATAAAGAAAAT	AAUCAAACUAUAAAGAAAAU
	Ex7 SD1 (Pos 9)	TCAGTTTACCTATTTCGAGC	UCAGUUUACCUAUUUCGAGC
	Ex9 STOP1 (Pos 6)	GTGTGCAGTTAAAAGACCTG	GUGUGCAGUUAAAAGACCUG
	Ex10 STOP1 (Pos 7)	AGGGCACAAGCTATAGCAAT	AGGGCACAAGCUAUAGCAAU
	Ex11 SA1 (Pos 5)	TACTCTGAAAGGATGAAGAC	UACUCUGAAAGGAUGAAGAC
	Ex11 STOP1 (Pos 7)	ACAGAACAGATAGCTGAAGA	ACAGAACAGAUAGCUGAAGA
	Ex12 SA1 (Pos 4)	AATCTTTTAGTGTTGGGAAG	AAUCUUUUAGUGUUGGGAAG
	Ex12 SA2 (Pos 5)	GAATCTTTTAGTGTTGGGAA	GAAUCUUUUAGUGUUGGGAA
	Ex12 SA3 (Pos 6)	AGAATCTTTTAGTGTTGGGA	AGAAUCUUUUAGUGUUGGGA
	Ex12 SD1 (Pos 7)	CTCCTACCTTTTTGGTCAGC	CUCCUACCUUUUUGGUCAGC
	Ex12 SD2 (Pos 6)	TCCTACCTTTTTGGTCAGCA	UCCUACCUUUUUGGUCAGCA
	Ex14 SD1 (Pos 7)	GCCTTACAGCAACCTCTCCA	GCCUUACAGCAACCUCUCCA
	Ex17 SD1 (Pos 8)	ATTCTTACCTGAAGCACAAT	AUUCUUACCUGAAGCACAAU
PPARa	Ex2 SA1 (Pos 9)	AAGCGTGTCTGGGGAAAAAG	AAGCGUGUCUGGGGAAAAAG
	Ex4 SA1 (Pos 7)	GAATCGCTAGGGTTTGGAGG	GAAUCGCUAGGGUUUGGAGG
	Ex4 SA2 (Pos 8)	CGAATCGCTAGGGTTTGGAG	CGAAUCGCUAGGGUUUGGAG
	Ex4 SA3 (Pos 9)	ACGAATCGCTAGGGTTTGGA	ACGAAUCGCUAGGGUUUGGA
	Ex4 SD1 (Pos 9)	AACACCTACTGGATTGTTAC	AACACCUACUGGAUUGUUAC
	Ex5 STOP1 (Pos 8)	TGGCATCCAGAACAAGGAGG	UGGCAUCCAGAACAAGGAGG
	Ex5 STOP2 (Pos 5)	CATCCAGAACAAGGAGGCGG	CAUCCAGAACAAGGAGGCGG
	Ex6 STOP1 (Pos 9)	CGCTACTGCAGGAGATCTAC	CGCUACUGCAGGAGAUCUAC
	Ex6 STOP2 (Pos 5)	GGTGCAGATCATCAAGAAGA	GGUGCAGAUCAUCAAGAAGA
	Ex6 STOP3 (Pos 8)	CCACCTGCAGAGCAACCACC	CCACCUGCAGAGCAACCACC
PPARd	Ex1 SD1 (Pos 4)	TCACCTGTGTAGCTGCTGGA	UCACCUGUGUAGCUGCUGGA
	Ex1 SD2 (Pos 8)	CTCCTCACCTGTGTAGCTGC	CUCCUCACCUGUGUAGCUGC
	Ex1 STOP (Pos 5)	GCCACAGGAGGAAGCCCCTG	GCCACAGGAGGAAGCCCCUG
	Ex2 SA1 (Pos 7)	AGAGGTCTGCGGACACACGA	AGAGGUCUGCGGACACACGA
	Ex2 STOP1 (Pos 7)	CAACTGCAGATGGGCTGTGA	CAACUGCAGAUGGGCUGUGA
	Ex2 STOP2 (Pos 6)	AACTGCAGATGGGCTGTGAC	AACUGCAGAUGGGCUGUGAC
	Ex3 SA1 (Pos 5)	AGCCCTGAAGCACCAAGAAC	AGCCCUGAAGCACCAAGAAC
	Ex3 STOP1 (pos 5)	CTTCCAGAAGTGCCTGGCAC	CUUCCAGAAGUGCCUGGCAC
	Ex4 SA1 (Pos 7)	GGATAGCTGCACAGGGAAGG	GGAUAGCUGCACAGGGAAGG
	Ex4 SA2 (Pos 8)	CGGATAGCTGCACAGGGAAG	CGGAUAGCUGCACAGGGAAG

	Ex4 SA3 (Pos 9)	ACGGATAGCTGCACAGGGAA	ACGGAUAGCUGCACAGGGAA
	Ex4 SD1 (Pos 9)	AACACTCACCGCCGTGTGGC	AACACUCACCGCCGUGUGGC
	Ex5 SD1 (Pos 5)	CTCACCTCCACACAGAATGA	CUCACCUCCACACAGAAUGA
	Ex5 STOP1 (Pos 5)	GTGGCAGGCAGAGAAGGGGC	GUGGCAGGCAGAGAAGGGGC
	Ex6 SA1 (Pos 8)	GGCCGGTCTGTGGGGACACA	GGCCGGUCUGUGGGGACACA
	Ex6 SA2 (Pos 9)	TGGCCGGTCTGTGGGGACAC	UGGCCGGUCUGUGGGGACAC
	Ex6 SA3 (Pos 4)	CAGCTTGGGGAAGAGGTACT	CAGCUUGGGGAAGAGGUACU
	Ex6 SA4 (Pos 5)	GCAGCTTGGGGAAGAGGTAC	GCAGCUUGGGGAAGAGGUAC
	Ex6 STOP1 (Pos 8)	TCTGCTCCAGGAGATCTACA	UCUGCUCCAGGAGAUCUACA
PRDMI1	Iso 1 ATG	GGTCATGGCCGCCAGACCCT	GGUCAUGGCCGCCAGACCCU
	Exon 2 STOP	GTCCAGTGTCCCAGAATGCC	GUCCAGUGUCCCAGAAUGCC
	Exon 2 SD	AATCACCTCTGAACAATCCC	AAUCACCUCUGAACAAUCCC
	Exon 3 SA	CAGCCTGGAAGAGAAAGGAA	CAGCCUGGAAGAGAAAGGAA
	Exon 3 SD 1	GCTTACCTCTTCACTGTTGG	GCUUACCUCUUCACUGUUGG
	Exon 3 SD 2	GAGGCTTACCTCTTCACTGT	GAGGCUUACCUCUUCACUGU
	Exon 6 SA 3	TTGTGCTGAAATAAAGAAAA	UUGUGCUGAAAUAAAGAAAA
	Exon 6 SA 2	TGTGCTGAAATAAAGAAAAA	UGUGCUGAAAUAAAGAAAAA
	Exon 6 SA 1	GTGCTGAAATAAAGAAAAAG	GUGCUGAAAUAAAGAAAAAG
	Exon 6 SD	CTACCTTCAGATTGGAGAGC	CUACCUUCAGAUUGGAGAGC
	Exon 7 SD	CTGCGCACCTGGCATTCATG	CUGCGCACCUGGCAUUCAUG
	Exon 8 STOP	TTGCAAAGAAACATGGGGAA	UUGCAAAGAAACAUGGGGAA
PRKACA	Ex1 Isoform SD1 (Pos 4)	TGACCGACATTCCATGGCCA	UGACCGACAUUCCAUGGCCA
	Ex2 SA1 (Pos 6)	TTTCACTGAAAGGGAGAGAG	UUUCACUGAAAGGGAGAGAG
	Ex2 SA2 (Pos 7)	CTTTCACTGAAAGGGAGAGA	CUUUCACUGAAAGGGAGAGA
	Ex2 SA3 (Pos 8)	TCTTTCACTGAAAGGGAGAG	UCUUUCACUGAAAGGGAGAG
	Ex3 SA1 (Pos5)	TGTTCTGTGGGCAGAGGGGT	UGUUCUGUGGGCAGAGGGGU
	Ex3 SA2 (Pos 9)	GCTGTGTTCTGTGGGCAGAG	GCUGUGUUCUGUGGGCAGAG
	Ex3 SD1 (Pos 7)	ACCTCACCTTCTGTTTGTCG	ACCUCACCUUCUGUUUGUCG
	Ex4 SA1 (Pos 5)	CCACCTGGGAAGGGAAGGAG	CCACCUGGGAAGGGAAGGAG
	Ex4 SA2 (Pos 6)	ACCACCTGGGAAGGGAAGGA	ACCACCUGGGAAGGGAAGGA
	Ex4 SA3 (Pos 7)	CACCACCTGGGAAGGGAAGG	CACCACCUGGGAAGGGAAGG
	Ex5 SA1 (Pos 4)	GTCCTGTGGGAAGCAGTGGC	GUCCUGUGGGAAGCAGUGGC
	Ex5 SA2 (Pos 8)	AGTTGTCCTGTGGGAAGCAG	AGUUGUCCUGUGGGAAGCAG
	Ex6 SA1 (Pos 5)	CTCACTGATGGGGACAAATG	CUCACUGAUGGGGACAAAUG
	Ex6 SA2 (Pos 6)	GCTCACTGATGGGGACAAAT	GCUCACUGAUGGGGACAAAU
	Ex6 SA3 (Pos 7)	GGCTCACTGATGGGGACAAA	GGCUCACUGAUGGGGACAAA
	Ex6 SD1 (Pos 4)	GCACCTGAATGTAGCCCTGC	GCACCUGAAUGUAGCCCUGC
	Ex7 SA1 (Pos 7)	TCACCTGTGGGCACAAGAAC	UCACCUGUGGGCACAAGAAC
	Ex8 SA1 (Pos 5)	AGCCCTGGAGCAAGATGGGG	AGCCCUGGAGCAAGAUGGGG
	Ex8 SA2 (Pos 6)	TAGCCCTGGAGCAAGATGGG	UAGCCCUGGAGCAAGAUGGG
	Ex8 SA3 (Pos 7)	GTAGCCCTGGAGCAAGATGG	GUAGCCCUGGAGCAAGAUGG
	Ex8 SA4 (Pos 8)	TGTAGCCCTGGAGCAAGATG	UGUAGCCCUGGAGCAAGAUG

	Ex8 SA5 (Pos 9)	TTGTAGCCCTGGAGCAAGAT	UUGUAGCCCUGGAGCAAGAU
	Ex9 SA1 (Pos 4)	CACCTGGAGGAAGGGGTACA	CACCUGGAGGAAGGGGUACA
	Ex9 SD1 (Pos 6)	GCCCACCTTCCTCTGGTAGA	GCCCACCUUCCUCUGGUAGA
PTEN	Ex1 STOP1 (Pos 5)	GGAGCGGGGGGGAGAAGCGG	GGAGCGGGGGGGAGAAGCGG
	Ex1 STOP2 (Pos 4)	AAACAAAAGGAGATATCAAG	AAACAAAAGGAGAUAUCAAG
	Ex2 (SA1 (Pos 5)	ATATCTGAGTACTTTAGTTA	AUAUCUGAGUACUUUAGUUA
	Ex4 SD1 (Pos 5)	CCTACCTCTGCAATTAAATT	CCUACCUCUGCAAUUAAAUU
	Ex5 SD1 (Pos 9)	ATAACTTACCTTTTTGTCTC	AUAACUUACCUUUUUGUCUC
PTPN2	Ex1 STOP1 pos 8)	TCGCTGGCAGCCGCTGTACT	UCGCUGGCAGCCGCUGUACU
	Ex6 SA1 (Pos 4)	ACTCTAAAAAGTGAAAATCA	ACUCUAAAAAGUGAAAAUCA
	Ex9 STOP1 (pos 6)	AGGTGCAGCAGATGAAACAG	AGGUGCAGCAGAUGAAACAG
PTPN6	Ex2 SA1 (Pos 6)	ACCACCTGGGCGGCCCAGGC	ACCACCUGGGCGGCCCAGGC
	Ex2 SA2 (Pos 5)	CCACCTGGGCGGCCCAGGCA	CCACCUGGGCGGCCCAGGCA
	Ex3 SA1 (Pos 9)	CCCCCACCCTGCAGGGCACC	CCCCCACCCUGCAGGGCACC
	Ex3 SA2 (Pos 6)	CCACCCTGCAGGGCACCAGG	CCACCCUGCAGGGCACCAGG
	Ex3 SD1 (pos 9)	AGCCCTCACCTCTCACTAGT	AGCCCUCACCUCUCACUAGU
	Ex4 SA1 (Pos 4)	CACCTAGAGAAGGCAGCGTC	CACCUAGAGAAGGCAGCGUC
	Ex5 SA1 (Pos 4)	ACCCTGGGGGGAGCCAAATT	ACCCUGGGGGGAGCCAAAUU
	Ex7 SA1 (Pos 7)	CAAACTCTGAGATGTGGGTG	CAAACUCUGAGAUGUGGGUG
	Ex7 SA1 (Pos 8)	GCAAACTCTGAGATGTGGGT	GCAAACUCUGAGAUGUGGGU
	Ex8 SA1 (Pos 5)	TCAACTGGGAGTGGGCGGAG	UCAACUGGGAGUGGGCGGAG
	Ex8 SD1 (Pos 9)	ACTGCTGACCTTGATGTAGT	ACUGCUGACCUUGAUGUAGU
	Ex9 SA1 (Pos 8)	GCTGGTTCTGGACGCAAGCG	GCUGGUUCUGGACGCAAGCG
	Ex10 SA1 (Pos 6)	TTGTTCTGGAAAGGGAGGGT	UUGUUCUGGAAAGGGAGGGU
	Ex10 SA2 (Pos 5)	TGTTCTGGAAAGGGAGGGTC	UGUUCUGGAAAGGGAGGGUC
	Ex10 SD1 (Pos 9)	GCCACTCACATTGTCCAGCG	GCCACUCACAUUGUCCAGCG
	Ex11 SA1 (Pos 5)	CTCCCTAAGCCGAGGACATA	CUCCCUAAGCCGAGGACAUA
	Ex11 SA2 (Pos 6)	TCTCCCTAAGCCGAGGACAT	UCUCCCUAAGCCGAGGACAU
	Ex11 SD1 (Pos 4)	TCACCTGCAGTGCACGATGA	UCACCUGCAGUGCACGAUGA
	Ex12 SA1 (Pos 7)	CCGGCGCTGGGGAAAGACGG	CCGGCGCUGGGGAAAGACGG
	Ex12 SA2 (Pos 8)	GCCGGCGCTGGGGAAAGACG	GCCGGCGCUGGGGAAAGACG
	Ex12 SA3 (Pos 9)	TGCCGGCGCTGGGGAAAGAC	UGCCGGCGCUGGGGAAAGAC
	Ex14 SD1 (Pos 7)	CACTCACTTGGACGAGGTGC	CACUCACUUGGACGAGGUGC
	Ex14 SD2 (Pos 8)	CCACTCACTTGGACGAGGTG	CCACUCACUUGGACGAGGUG
	Ex15 SA1 (Pos 4)	TGTCTGCAGCCGGGTGCAGG	UGUCUGCAGCCGGGUGCAGG
	Ex15 SA2 (Pos 5)	GTGTCTGCAGCCGGGTGCAG	GUGUCUGCAGCCGGGUGCAG
	Ex15 SA3 (Pos 6)	TGTGTCTGCAGCCGGGTGCA	UGUGUCUGCAGCCGGGUGCA
	Ex15 SD1 (Pos 6)	CACCGCTCACTTCCTCTTGA	CACCGCUCACUUCCUCUUGA
	Ex15 SD2 (Pos 7)	GCACCGCTCACTTCCTCTTG	GCACCGCUCACUUCCUCUUG
PTPN11	Ex3 SA1 (Pos 7)	TTTCTTCTAAAATAGTCCAT	UUUCUUCUAAAAUAGUCCAU
	Ex3 SD1 (Pos 9)	GTTACTGACCTTTCAGAGGT	GUUACUGACCUUUCAGAGGU
	Ex13 SD1 (Pos 9)	ACTACTTACTCTGCACAGGG	ACUACUUACUCUGCACAGGG

RASA2	Ex2 SA1 (Pos 4)	GACCTAAAATATAAAAAATT	GACCUAAAAUAUAAAAAAUU
	Ex5 SD1 (pos 6)	ATTTACCTGAACCTCTGAAT	AUUUACCUGAACCUCUGAAU
	Ex6 SD1 (Pos 5)	CTTACTGTACAACAAGCTGC	CUUACUGUACAACAAGCUGC
	Ex10 SA1 (Pos 6)	CGATCCTGAAAATTGAAAAC	CGAUCCUGAAAAUUGAAAAC
	Ex10 SD1 (Pos 7)	CCCTTACCAGGCTTGATGAG	CCCUUACCAGGCUUGAUGAG
	Ex12 SA1 (Pos 9)	GATATTGGCTAAATACAGAA	GAUAUUGGCUAAAUACAGAA
	Ex13 SD1 (Pos 8)	TTCTGTACCTCATCAAGAAT	UUCUGUACCUCAUCAAGAAU
	Ex15 SA1 (Pos 6)	TTCTCCTGCAGGATTTAAAA	UUCUCCUGCAGGAUUUAAAA
	Ex16 SA1 (Pos 7)	GGTCATCTGCAGGAAAAAAA	GGUCAUCUGCAGGAAAAAAA
	Ex19 SA1 (Pos 7)	CAAGAACTAAATGGGGAAAT	CAAGAACUAAAUGGGGAAAU
SIGLEC15	Ex3 SA1 (Pos 8)	GGCGAGCCTGAGGGCGGGGC	GGCGAGCCUGAGGGCGGGGC
	Ex3 SA2 (Pos 9)	TGGCGAGCCTGAGGGCGGGG	UGGCGAGCCUGAGGGCGGGG
	Ex3 SD1 (Pos 6)	CCTCGCCTGTCACGTGCAGC	CCUCGCCUGUCACGUGCAGC
	Ex3 STOP1 (pos 6)	GCGCGCCAGATGGCCGTCAG	GCGCGCCAGAUGGCCGUCAG
	Ex3 STOP2 (Pos 9)	TGCATGGACCAGCGCTGCGC	UGCAUGGACCAGCGCUGCGC
	Ex3 STOP3 (Pos 8)	GTCCATGCAGGTGCCACCCG	GUCCAUGCAGGUGCCACCCG
	Ex3 STOP4 (Pos 5)	AGCGCAGCGCTGGTCCATGC	AGCGCAGCGCUGGUCCAUGC
	Ex4 SD (Pos 9)	CGCACCCACCTGGGCGGCGG	CGCACCCACCUGGGCGGCGG
	Ex4 SD1 (Pos 6)	CGCGGCTGCAGGGGAGAAGG	CGCGGCUGCAGGGGAGAAGG
	Ex4 SD1 (Pos 8)	GCACCCACCTGGGCGGCGGC	GCACCCACCUGGGCGGCGGC
	Ex4 SD1 (Pos 9)	CGGCGCGGCTGCAGGGGAGA	CGGCGCGGCUGCAGGGGAGA
	Ex4 SD3 (Pos 5)	GCGGCTGCAGGGGAGAAGGC	GCGGCUGCAGGGGAGAAGGC
	Ex4 STOP1 (Pos 9)	GGGCCGGACCAGGCGAGGGC	GGGCCGGACCAGGCGAGGGC
	Ex4 STOP2 (Pos 6)	CCGGACCAGGCGAGGGCGGG	CCGGACCAGGCGAGGGCGGG
	Ex6 STOP1 (Pos 9)	ATTTGAGCCAGATGAACCCC	AUUUGAGCCAGAUGAACCCC
SLA	Ex2 SD1 (pos 4)	TTACCCTCCGGGTTGGGCAG	UUACCCUCCGGGUUGGGCAG
	Ex2 SD2 (Pos 5)	CTTACCCTCCGGGTTGGGCA	CUUACCCUCCGGGUUGGGCA
	Ex3 SA1 (Pos 4))	GTCCTGGGGAAACAAAGGCA	GUCCUGGGGAAACAAAGGCA
	Ex3 SA2 (pos 9)	ATCCAGTCCTGGGGAAACAA	AUCCAGUCCUGGGGAAACAA
	Ex4 SD1 (Pos 7)	TACTCACCCATGGTAAACTC	UACUCACCCAUGGUAAACUC
	Ex6 SA1 (Pos 8)	TAAAACCCTGCAGGAGGTGG	UAAAACCCUGCAGGAGGUGG
	Ex8 SA1 (Pos 4)	AGTCTGTGGGCCAGAAGAAA	AGUCUGUGGGCCAGAAGAAA
SLAMF7	Ex1 SD1 (Pos 6)	ACTCACCTGTGAGCTGCCAA	ACUCACCUGUGAGCUGCCAA
	Ex1 STOP1 (Pos 5)	GCTGCCAAAGGATATAGATG	GCUGCCAAAGGAUAUAGAUG
	Ex1 STOP2 (Pos 4)	CTGCCAAAGGATATAGATGA	CUGCCAAAGGAUAUAGAUGA
	Ex3 SD1 (Pos 5)	GTCACCTTCACAGAGCTTCC	GUCACCUUCACAGAGCUUCC
	Ex3 SD2 (Pos 7)	CAGCACCTTCAGAGAATGGG	CAGCACCUUCAGAGAAUGGG
	Ex3 SD3 (Pos 4)	CACCTTCAGAGAATGGGTGG	CACCUUCAGAGAAUGGGUGG
	Ex4 SD1 (Pos 8)	ATGTACTCTAAAAGCAAGTT	AUGUACUCUAAAAGCAAGUU
	Ex4 SD2 (Pos 9)	AATGTACTCTAAAAGCAAGT	AAUGUACUCUAAAAGCAAGU
SOCS1	Ex1 STOP1 (Pos 9)	CGCTGCGCCAGCGCCGCGTG	CGCUGCGCCAGCGCCGCGUG
	Ex1 STOP2 (Pos 7)	GGGCCCCCAGTAGAATCCGC	GGGCCCCCAGUAGAAUCCGC

STK4	Ex1 SD1 (Pos 9)	CTTACCTACCTCCCAATGTC	CUUACCUACCUCCCAAUGUC
	Ex1 SA2 (Pos 7)	CTGCATCTACAGTAATCTGA	CUGCAUCUACAGUAAUCUGA
	Ex5 SA1 (Pos 9)	ATGGTATCCTAAAATAGAAA	AUGGUAUCCUAAAAUAGAAA
	Ex6 SD1 (Pos 6)	TCATACCTGCAGGAGCTGAG	UCAUACCUGCAGGAGCUGAG
	Ex9 SA1 (Pos 5)	TAAGCTAGAAGAGAAGTGGA	UAAGCUAGAAGAGAAGUGGA
	Ex9 SA2 (Pos 9)	CTCTTAAGCTAGAAGAGAAG	CUCUUAAGCUAGAAGAGAAG
SUV39H1	Ex2 SA1 (Pos 9)	GCTGCAGCCTGGATCAAGCG	GCUGCAGCCUGGAUCAAGCG
	Ex2 STOP1 (Pos 5)	GCTGCAGGACCTGTGCCGCC	GCUGCAGGACCUGUGCCGCC
	Ex3 SA1 (Pos 4)	TTCCTGTTGGGGGTGGGTAG	UUCCUGUUGGGGGUGGGUAG
	Ex3 SA2 (Pos 5)	GTTCCTGTTGGGGGTGGGTA	GUUCCUGUUGGGGGUGGGUA
	Ex3 SA3 (Pos 5)	TGTTCCTGTTGGGGGTGGGT	UGUUCCUGUUGGGGGUGGGU
	Ex3 STOP1 (Pos8)	ACAGGAACAGGAATATTACC	ACAGGAACAGGAAUAUUACC
	Ex3 STOP2 (Pos 6)	GATATCCACGCCATTTCACC	GAUAUCCACGCCAUUUCACC
TMEM222	Ex1 SD1 (Pos 6)	ACTCACGTGAGCACCGGGAT	ACUCACGUGAGCACCGGGAU
	Ex1 SD2 (Pos 7)	GACTCACGTGAGCACCGGGA	GACUCACGUGAGCACCGGGA
	Ex2 SA1 (Pos 4)	CACCTGTGAGGAAAAGGACG	CACCUGUGAGGAAAAGGACG
	Ex2 SA2 (Pos 5)	CCACCTGTGAGGAAAAGGAC	CCACCUGUGAGGAAAAGGAC
	Ex2 SD1 (Pos 7)	GACTCACTGAGACAAAGTAG	GACUCACUGAGACAAAGUAG
	Ex2 SA3 (Pos 6)	ACCACCTGTGAGGAAAAGGA	ACCACCUGUGAGGAAAAGGA
	Ex2 SD2 (Pos 8)	GGACTCACTGAGACAAAGTA	GGACUCACUGAGACAAAGUA
	Ex2 SD3 (Pos 9)	GGGACTCACTGAGACAAAGT	GGGACUCACUGAGACAAAGU
	Ex3 SD1 (Pos 4)	TTACTTGGCAGGCTTTCCAA	UUACUUGGCAGGCUUUCCAA
	Ex4 SA1 (Pos 6)	CAGTACCTGGGGGGAGAGAA	CAGUACCUGGGGGGAGAGAA
	Ex4 SA1 (Pos 7)	CCAGTACCTGGGGGGAGAGA	CCAGUACCUGGGGGGAGAGA
	Ex5 SA1 (Pos 6)	TTGTGCTGGAGGCACCAGAA	UUGUGCUGGAGGCACCAGAA
	Ex6 SA2 (Pos 7)	ATTGTGCTGGAGGCACCAGA	AUUGUGCUGGAGGCACCAGA
	Ex7 SA1 (Pos 8)	CCCAACGCTGACAGAGAGAA	CCCAACGCUGACAGAGAGAA
TNFAIP3	Ex2 SA1 (Pos 6)	GTCACCTGAGGACAGAAAGG	GUCACCUGAGGACAGAAAGG
	Ex2 SA2 (Pos 9)	GCCGTCACCTGAGGACAGAA	GCCGUCACCUGAGGACAGAA
	Ex3 SA1 (Pos 9)	TTCCAGTTCTAAGGGGAGCG	UUCCAGUUCUAAGGGGAGCG
	Ex3 SD1 (pos 4)	TCACCTGAAATGACAATGAT	UCACCUGAAAUGACAAUGAU
	Ex6 SA1 (pos 9)	CATCCAACCTGAAGACCAAA	CAUCCAACCUGAAGACCAAA
	Ex8 SA1 (Pos 6)	GATCTCTGAGTGGAAAGAAC	GAUCUCUGAGUGGAAAGAAC
TNFRSF8 (CD30)	Exon 1 ATG	AGGACGCGCATCCCCGGGGC	AGGACGCGCAUCCCCGGGGC
	Exon 1 STOP 2	TCCCACAGGTAAGCGGGTGA	UCCCACAGGUAAGCGGGUGA
	Exon 1 STOP 1	GGCGCTACGAGCCTTCCCAC	GGCGCUACGAGCCUUCCCAC
	Exon 2 SD	CCCACTCACCCATGGGGCAG	CCCACUCACCCAUGGGGCAG
	Exon 3 STOP	CGACACAGCAGTGCCCACAG	CGACACAGCAGUGCCCACAG
	Exon 4 SA 3	GTCGTCTAAGGGACACAGAC	GUCGUCUAAGGGACACAGAC
	Exon 4 SA 2	TCGTCTAAGGGACACAGACA	UCGUCUAAGGGACACAGACA
	Exon 4 SA 1	CGTCTAAGGGACACAGACAG	CGUCUAAGGGACACAGACAG
	Exon 6 STOP	CCCCCAGGCCAAGCCCACCC	CCCCCAGGCCAAGCCCACCC

	Exon 6 SD	AAATTACCTGGATCTGAACT	AAAUUACCUGGAUCUGAACU
	Exon 8 SA	TCATCTAAGGGACACAGATG	UCAUCUAAGGGACACAGAUG
	Exon 10 SA	GTGCTGCGGGGAGAAGCCCA	GUGCUGCGGGGAGAAGCCCA
	Exon 10 SD 2	ACCATTACCTGCATCCAGAA	ACCAUUACCUGCAUCCAGAA
	Exon 10 SD 1	CCATTACCTGCATCCAGAAC	CCAUUACCUGCAUCCAGAAC
	Exon 10 STOP	CCCCACTCAGAGCTTGCTGG	CCCCACUCAGAGCUUGCUGG
	Exon 11 STOP 1	ACCCAGAAGAGCACTGGCCC	ACCCAGAAGAGCACUGGCCC
	Exon 11 STOP 2	AGGATCACCCAGAAGAGCAC	AGGAUCACCCAGAAGAGCAC
	Exon 12 SA 2	TGGAGCTCTGAAACGACACC	UGGAGCUCUGAAACGACACC
	Exon 12 SA 1	AGCTCTGAAACGACACCAGG	AGCUCUGAAACGACACCAGG
	Exon 12 SD 2	CTCACCCACAAGCTCTAGCT	CUCACCCACAAGCUCUAGCU
	Exon 12 SD 1	TCACCCACAAGCTCTAGCTT	UCACCCACAAGCUCUAGCUU
	Exon 13 SD 2	ACTTACCGTTGAGCTCCTCC	ACUUACCGUUGAGCUCCUCC
	Exon 13 SD 1	CTTACCGTTGAGCTCCTCCT	CUUACCGUUGAGCUCCUCCU
	Exon 14 SA 1	AGCTGCTGTGGGACGGGAAT	AGCUGCUGUGGGACGGGAAU
	Exon 14 SA 2	CAGCTGCTGTGGGACGGGAA	CAGCUGCUGUGGGACGGGAA
	iso exon 11 SA	CTCCTCAGCTGCTGTGGGAC	CUCCUCAGCUGCUGUGGGAC
	Exon 14 STOP	GCCGCTGCAGGATGCCAGCC	GCCGCUGCAGGAUGCCAGCC
	Exon 14 SD	TGACTCACCAATCTTGTTAT	UGACUCACCAAUCUUGUUAU
	Exon 15 SA 1	TCTCTGCAAGGCAAAAGGAT	UCUCUGCAAGGCAAAAGGAU
	Exon 15 SA 2	TTCTCTGCAAGGCAAAAGGA	UUCUCUGCAAGGCAAAAGGA
TNFRSF10B	Ex1 SD1 (Pos 6)	ACTCACCAACAGCAGGACCG	ACUCACCAACAGCAGGACCG
	Ex2 SA1 (Pos 6)	GAGACCTGTGGGGACAAAGC	GAGACCUGUGGGGACAAAGC
	Ex 3 SD1 (Pos 5)	CTCACCCTGTGCGGCACTTC	CUCACCCUGUGCGGCACUUC
	Ex4 SA1 (Pos 5)	ACACCTGGGTACACACACAG	ACACCUGGGUACACACACAG
	Ex6 SA1 (Pos 7)	TGAGCTCTGGAAAAAGACAT	UGAGCUCUGGAAAAAGACAU
	Ex8 SA1 (pos 8)	CCGGTTCCTGTAACACATAG	CCGGUUCCUGUAACACAUAG
	Ex8 SA2 (Pos 7)	CGGTTCCTGTAACACATAGT	CGGUUCCUGUAACACAUAGU
TOX	Exon 1 SD 3	GTTCACCTTGTTGCAATAGT	GUUCACCUUGUUGCAAUAGU
	Exon 1 SD 2	TTCACCTTGTTGCAATAGTA	UUCACCUUGUUGCAAUAGUA
	Exon 1 SD 1	TCACCTTGTTGCAATAGTAG	UCACCUUGUUGCAAUAGUAG
	Exon 4 STOP	TCACAGCTAAGTGCTCAACT	UCACAGCUAAGUGCUCAACU
	Exon 5 STOP 1	TGATACTCAGGCCGCCATCA	UGAUACUCAGGCCGCCAUCA
	Exon 5 STOP 2	GATACTCAGGCCGCCATCAA	GAUACUCAGGCCGCCAUCAA
	Exon 5 STOP 3	GAGAAGAGCAAAAACAGGTA	GAGAAGAGCAAAAACAGGUA
	Exon 5 STOP 4	AGAAGAGCAAAAACAGGTAA	AGAAGAGCAAAAACAGGUAA
	Exon 6 SD	CGTTACCTTGGATACAAGGC	CGUUACCUUGGAUACAAGGC
	Exon 7 STOP 1	TCACCATGCAGCAGCCCCTT	UCACCAUGCAGCAGCCCCUU
	Exon 7 STOP 2	TGGGAACCAGCTCCCCATGC	UGGGAACCAGCUCCCCAUGC
	Exon 7 STOP 3	CATGCAGCAAGTAAGTGCAA	CAUGCAGCAAGUAAGUGCAA
	Exon 7 SD 2	TGCACTTACTTGCTGCATGG	UGCACUUACUUGCUGCAUGG
	Exon 7 SD 1	GCACTTACTTGCTGCATGGT	GCACUUACUUGCUGCAUGGU

	Exon 8 STOP 1	AGCTGCACAAGTTGTCACCC	AGCUGCACAAGUUGUCACCC
	Exon 8 STOP 2	CTCCCCCACAACCGGTGGAC	CUCCCCCACAACCGGUGGAC
	Exon 8 STOP 6	TTATTCCAGTCCACCGGTTG	UUAUUCCAGUCCACCGGUUG
	Exon 8 STOP 5	TATTCCAGTCCACCGGTTGT	UAUUCCAGUCCACCGGUUGU
	Exon 8 STOP 4	ATTCCAGTCCACCGGTTGTG	AUUCCAGUCCACCGGUUGUG
	Exon 8 STOP 3	TTCCAGTCCACCGGTTGTGG	UUCCAGUCCACCGGUUGUGG
TOX2	Exon 1 ATG 1	GTCCATGGCGGGCGCGGCGG	GUCCAUGGCGGGCGCGGCGG
	Exon 1 ATG 2	GACGTCCATGGCGGGCGCGG	GACGUCCAUGGCGGGCGCGG
	Exon 2 STOP 1	TATGCAGCAGACTCGCACAG	UAUGCAGCAGACUCGCACAG
	Exon 2 STOP 2	TTTCCGCAGAAGGTAAGCAG	UUUCCGCAGAAGGUAAGCAG
	Exon 3 SA 2	TCAAACTAGAATAGAGAGAG	UCAAACUAGAAUAGAGAGAG
	Exon 3 SA 1	CAAACTAGAATAGAGAGAGA	CAAACUAGAAUAGAGAGAGA
	Exon 3 SD	CACCCACCTGGCTGGTTGAC	CACCCACCUGGCUGGUUGAC
	Exon 5 SD	TGGATCTAAGAGAGGAGGAC	UGGAUCUAAGAGAGGAGGAC
	Exon 5 STOP 1	GATCCAGGAGATGGTCCACT	GAUCCAGGAGAUGGUCCACU
	Exon 5 STOP 2	GTCCCAGCTCATCTCGCAGA	GUCCCAGCUCAUCUCGCAGA
	Exon 5 STOP 3	TCATCTCGCAGATGGGCATC	UCAUCUCGCAGAUGGGCAUC
	Exon 5 STOP 4	CTCCACTCAGGAAGAGGAGT	CUCCACUCAGGAAGAGGAGU
	Exon 6 STOP 1	TGAGCCGCAGAAGCCTGTGT	UGAGCCGCAGAAGCCUGUGU
	Exon 6 STOP 2	AGACACTCAGGCCGCCATCA	AGACACUCAGGCCGCCAUCA
	Exon 6 STOP 3	GACACTCAGGCCGCCATCAA	GACACUCAGGCCGCCAUCAA
	Exon 8 STOP 1	GACCTGCAGGCCTTCCGCAG	GACCUGCAGGCCUUCCGCAG
	Exon 8 STOP 2	ACCTGCAGGCCTTCCGCAGT	ACCUGCAGGCCUUCCGCAGU
	Exon 8 STOP 3	CCTGCAGGCCTTCCGCAGTG	CCUGCAGGCCUUCCGCAGUG
	Exon 8 SD	CTGCTTACCTGTGGCCCTGG	CUGCUUACCUGUGGCCCUGG
	Exon 9 SA 2	GGAAGTCCTACAGAGTGGGA	GGAAGUCCUACAGAGUGGGA
	Exon 9 SA 1	AGTCCTACAGAGTGGGAAGG	AGUCCUACAGAGUGGGAAGG
	Exon 9 STOP 2	GTCCCAGTCCCCGCTGCTGG	GUCCCAGUCCCCGCUGCUGG
	Exon 9 STOP 1	TCCCAGTCCCCGCTGCTGGT	UCCCAGUCCCCGCUGCUGGU
	Exon 9 STOP 3	GCTGTCCCAGTCCCCGCTGC	GCUGUCCCAGUCCCCGCUGC
	Exon 10 SA 1	CAGGCTGTGAGAGAGAGGAG	CAGGCUGUGAGAGAGAGGAG
	Exon 10 SA 2	GCAGGCTGTGAGAGAGAGGA	GCAGGCUGUGAGAGAGAGGA
	Exon 10 SA 3	AGCAGGCTGTGAGAGAGAGG	AGCAGGCUGUGAGAGAGAGG
UBASH3A	Ex1 SD1 (Pos 7)	TGCCATCTCTTCCTGCCCTT	UGCCAUCUCUUCCUGCCCUU
	Ex1 SD1 (Pos8)	GTACTCACGCGGTGTGCACC	GUACUCACGCGGUGUGCACC
	Ex5 SA1 (Pos 9)	GGAAGTGCCTGGGTGAGGAC	GGAAGUGCCUGGGUGAGGAC
	Ex7 SA1 (Pos 6)	AGGGTCTAGAAAAGAGGCAA	AGGGUCUAGAAAAGAGGCAA
	Ex7 SA2 (Pos 5)	GGGTCTAGAAAAGAGGCAAA	GGGUCUAGAAAAGAGGCAAA
	Ex7 SA3 (Pos 4)	GGTCTAGAAAAGAGGCAAAG	GGUCUAGAAAAGAGGCAAAG
	Ex9 SA1 (Pos 7)	GGTAGCCTGGGGGGTGGGGC	GGUAGCCUGGGGGGUGGGGC
	Ex11 SA1 (Pos 4)	CCCCTGGAAAATAGTGAAAA	CCCCUGGAAAAUAGUGAAAA
	Ex11 SD1 (Pos 5)	CTGACCTTCCAGGATGAGTT	CUGACCUUCCAGGAUGAGUU

	Ex11 SD2 Pos (7)	GAGGTTCTCACTGACCTTCC	GAGGUUCUCACUGACCUUCC
	Ex13 SA1 (Pos 7)	GCGGGCCTGGAAGGATGAGA	GCGGGCCUGGAAGGAUGAGA
	Ex14 SD1 (Pos 6)	GCGTACCTTTCTCACGAGTT	GCGUACCUUUCUCACGAGUU
	Ex14 SD2 (Pos 7)	CGCGTACCTTTCTCACGAGT	CGCGUACCUUUCUCACGAGU
VHL	Ex1 SD1 (Pos 9)	GGCCCGTACCTCGGTAGCTG	GGCCCGUACCUCGGUAGCUG
	Ex1 STOP1 (Pos 4)	GTCCCAGTTCTCCGCCCTCC	GUCCCAGUUCUCCGCCCUCC
	Ex2 STOP1 (Pos 9)	GGAACAAGCCAGGGTCATGT	GGAACAAGCCAGGGUCAUGU
	Ex2 STOP2 (Pos 9)	CAACCCCTCCATCTCCCAGC	CAACCCCUCCAUCUCCCAGC
	Ex3 SA1 (Pos 5)	TGACCTATCGGGACAAGCAA	UGACCUAUCGGGACAAGCAA
	Ex3 SD1 (Pos 5)	AGTACCTGGCAGTGTGATAT	AGUACCUGGCAGUGUGAUAU
	Ex4 STOP1 (Pos 6)	ATGTGCAGAAAGACCTGGAG	AUGUGCAGAAAGACCUGGAG
XBP1	Ex1 STOP1 (Pos 4)	GGGCAGCCCGCCTCCGCCGC	GGGCAGCCCGCCUCCGCCGC
	Ex1 STOP2 (Pos 5)	CGGCCAGGCCCTGCCGCTCA	CGGCCAGGCCCUGCCGCUCA

Methods of using fusion proteins comprising a cytidine or adenosine deaminase and a Cas9 domain

[00771] Some aspects of this disclosure provide methods of using the fusion proteins, or complexes provided herein. For example, some aspects of this disclosure provide methods comprising contacting a DNA molecule with any of the fusion proteins provided herein, and with at least one guide RNA, wherein the guide RNA is about 15-100 nucleotides long and comprises a sequence of at least 10 contiguous nucleotides that is complementary to a target sequence. In some embodiments, the 3' end of the target sequence is immediately adjacent to a canonical PAM sequence (NGG). In some embodiments, the 3' end of the target sequence is not immediately adjacent to a canonical PAM sequence (NGG). In some embodiments, the 3' end of the target sequence is immediately adjacent to an AGC, GAG, TTT, GTG, or CAA sequence. In some embodiments, the 3' end of the target sequence is immediately adjacent to an NGA, NGCG, NGN, NNGRRT, NNNRRT, NGCG, NGCN, NGTN, NGTN, NGTN, or 5' (TTTV) sequence.

[00772] In some embodiments, a fusion protein of the invention is used for mutagenizing a target of interest. In particular, a cytidine deaminase or adenosine deaminase nucleobase editor described herein is capable of making multiple mutations within a target sequence. These mutations may affect the function of the target. For example, when a cytidine deaminase or adenosine deaminase nucleobase editor is used to target a regulatory region the function of the regulatory region is altered and the expression of the downstream protein is reduced.

[00773] It will be understood that the numbering of the specific positions or residues in the respective sequences depends on the particular protein and numbering scheme used. Numbering might be different, e.g., in precursors of a mature protein and the mature protein itself, and differences in sequences from species to species may affect numbering. One of skill in the art will be able to identify the respective residue in any homologous protein and in the respective encoding nucleic acid by methods well known in the art, e.g., by sequence alignment and determination of homologous residues.

[00774] It will be apparent to those of skill in the art that in order to target any of the fusion proteins comprising a Cas9 domain and a cytidine or adenosine deaminase, as disclosed herein, to a target site, *e.g.*, a site comprising a mutation to be edited, it is typically necessary to co-express the fusion protein together with a guide RNA, *e.g.*, an sgRNA. As explained in more detail elsewhere herein, a guide RNA typically comprises a tracrRNA framework allowing for Cas9 binding, and a guide sequence, which confers sequence specificity to the Cas9:nucleic acid editing enzyme/domain fusion protein. Alternatively, the guide RNA and tracrRNA may be provided separately, as two nucleic acid molecules. In some embodiments, the guide RNA comprises a structure, wherein the guide sequence comprises a sequence that is complementary to the target sequence. The guide sequence is typically 20 nucleotides long. The sequences of suitable guide RNAs for targeting Cas9:nucleic acid editing enzyme/domain fusion proteins to specific genomic target sites will be apparent to those of skill in the art based on the instant disclosure. Such suitable guide RNA sequences typically comprise guide sequences that are complementary to a nucleic sequence within 50 nucleotides upstream or downstream of the target nucleotide to be edited. Some exemplary guide RNA sequences suitable for targeting any of the provided fusion proteins to specific target sequences are provided herein.

Base Editor Efficiency

[00775] Some aspects of the disclosure are based on the recognition that any of the base editors provided herein can modify a specific nucleotide base without generating a sizable proportion of indels. An "indel", as used herein, refers to the insertion or deletion of a nucleotide base within a nucleic acid. Such insertions or deletions can lead to frame shift mutations within a coding region of a gene. In some embodiments, it is desirable to generate base editors that efficiently modify (*e.g.* mutate) a specific nucleotide within a nucleic acid, without generating a large number of insertions or deletions (*i.e.*, indels) in the nucleic acid. In some embodiments, it is desirable to generate base editors that efficiently modify (*e.g.*

mutate or methylate) a specific nucleotide within a nucleic acid, without generating a large number of insertions or deletions (*i.e.*, indels) in the nucleic acid. In certain embodiments, any of the base editors provided herein can generate a greater proportion of intended modifications (*e.g.*, methylations) versus indels. In certain embodiments, any of the base editors provided herein can generate a greater proportion of intended modifications (*e.g.*, mutations) versus indels. In some embodiments, the base editors provided herein are capable of generating a ratio of intended mutations to indels that is greater than 1:1. In some embodiments, the base editors provided herein are capable of generating a ratio of intended mutations to indels that is at least 1.5:1, at least 2:1, at least 2.5:1, at least 3:1, at least 3.5:1, at least 4:1, at least 4.5:1, at least 5:1, at least 5.5:1, at least 6:1, at least 6.5:1, at least 7:1, at least 7.5:1, at least 8:1, at least 10:1, at least 12:1, at least 15:1, at least 20:1, at least 25:1, at least 30:1, at least 40:1, at least 50:1, at least 100:1, at least 200:1, at least 300:1, at least 400:1, at least 500:1, at least 600:1, at least 700:1, at least 800:1, at least 900:1, or at least 1000:1, or more. The number of intended mutations and indels may be determined using any suitable method.

[00776] In some embodiments, the base editors provided herein can limit formation of indels in a region of a nucleic acid. In some embodiments, the region is at a nucleotide targeted by a base editor or a region within 2, 3, 4, 5, 6, 7, 8, 9, or 10 nucleotides of a nucleotide targeted by a base editor. In some embodiments, any of the base editors provided herein can limit the formation of indels at a region of a nucleic acid to less than 1%, less than 1.5%, less than 2%, less than 2.5%, less than 3%, less than 3.5%, less than 4%, less than 4.5%, less than 5%, less than 6%, less than 7%, less than 8%, less than 9%, less than 10%, less than 12%, less than 15%, or less than 20%. The number of indels formed at a nucleic acid region may depend on the amount of time a nucleic acid (*e.g.*, a nucleic acid within the genome of a cell) is exposed to a base editor. In some embodiments, a number or proportion of indels is determined after at least 1 hour, at least 2 hours, at least 6 hours, at least 12 hours, at least 24 hours, at least 36 hours, at least 48 hours, at least 3 days, at least 4 days, at least 5 days, at least 7 days, at least 10 days, or at least 14 days of exposing a nucleic acid (*e.g.*, a nucleic acid within the genome of a cell) to a base editor.

[00777] Some aspects of the disclosure are based on the recognition that any of the base editors provided herein are capable of efficiently generating an intended mutation in a nucleic acid (*e.g.* a nucleic acid within a genome of a subject) without generating a considerable number of unintended mutations. In some embodiments, an intended mutation is a mutation that is generated by a specific base editor bound to a gRNA, specifically designed to generate

the intended mutation. In some embodiments, the intended mutation is a mutation that generates a stop codon, for example, a premature stop codon within the coding region of a gene. In some embodiments, the intended mutation is a mutation that eliminates a stop codon. In some embodiments, the intended mutation is a mutation that alters the splicing of a gene. In some embodiments, the intended mutation is a mutation that alters the regulatory sequence of a gene (*e.g.*, a gene promotor or gene repressor). In some embodiments, any of the base editors provided herein are capable of generating a ratio of intended mutations to unintended mutations (*e.g.,* intended mutations:unintended mutations) that is greater than 1:1. In some embodiments, any of the base editors provided herein are capable of generating a ratio of intended mutations to unintended mutations that is at least 1.5:1, at least 2:1, at least 2.5:1, at least 3:1, at least 3.5:1, at least 4:1, at least 4.5:1, at least 5:1, at least 5.5:1, at least 6:1, at least 6.5:1, at least 7:1, at least 7.5:1, at least 8:1, at least 10:1, at least 12:1, at least 15:1, at least 20:1, at least 25:1, at least 30:1, at least 40:1, at least 50:1, at least 100:1, at least 150:1, at least 200:1, at least 250:1, at least 500:1, or at least 1000:1, or more. It should be appreciated that the characteristics of the base editors described in the "*Base Editor Efficiency*" section, herein, may be applied to any of the fusion proteins, or methods of using the fusion proteins provided herein.

[00778] A base editing is often referred to as a "modification", such as, a genetic modification, a gene modification and modification of the nucleic acid sequence and is clearly understandable based on the context that the modification is a base editing modification. A base editing modification is therefore a modification at the nucleotide base level, for example as a result of the deaminase activity discussed throughout the disclosure, which then results in a change in the gene sequence, and may affect the gene product. In essence therefore, the gene editing modification described herein may result in a modification of the gene, structurally and/or functionally, wherein the expression of the gene product may be modified, for example, the expression of the gene is knocked out; or conversely, enhanced, or, in some circumstances, the gene function or activity may be modified. Using the methods disclosed herein, a base editing efficiency may be determined as the knockdown efficiency of the gene in which the base editing is performed, wherein the base editing is intended to knockdown the expression of the gene. A knockdown level may be validated quantitatively by determining the expression level by any detection assay, such as assay for protein expression level, for example, by flow cytometry; assay for detecting RNA expression such as quantitative RT-PCR, northern blot analysis, or any other suitable assay such as pyrosequencing; and may be validated qualitatively by nucleotide sequencing reactions.

[00779] In some embodiments, the modification, e.g., single base edit results in at least 10% reduction of the gene targeted expression. In some embodiments, the base editing efficiency may result in at least 10% reduction of the gene targeted expression. In some embodiments, the base editing efficiency may result in at least 20% reduction of the gene targeted expression. In some embodiments, the base editing efficiency may result in at least 30% reduction of the gene targeted expression. In some embodiments, the base editing efficiency may result in at least 40% reduction of the gene targeted expression. In some embodiments, the base editing efficiency may result in at least 50% reduction of the gene targeted expression. In some embodiments, the base editing efficiency may result in at least 60% reduction of the targeted gene expression. In some embodiments, the base editing efficiency may result in at least 70% reduction of the targeted gene expression. In some embodiments, the base editing efficiency may result in at least 80% reduction of the targeted gene expression. In some embodiments, the base editing efficiency may result in at least 90% reduction of the targeted gene expression. In some embodiments, the base editing efficiency may result in at least 91% reduction of the targeted gene expression. In some embodiments, the base editing efficiency may result in at least 92% reduction of the targeted gene expression. In some embodiments, the base editing efficiency may result in at least 93% reduction of the targeted gene expression. In some embodiments, the base editing efficiency may result in at least 94% reduction of the targeted gene expression. In some embodiments, the base editing efficiency may result in at least 95% reduction of the targeted gene expression. In some embodiments, the base editing efficiency may result in at least 96% reduction of the targeted gene expression . In some embodiments, the base editing efficiency may result in at least 97% reduction of the targeted gene expression. In some embodiments, the base editing efficiency may result in at least 98% reduction of the targeted gene expression. In some embodiments, the base editing efficiency may result in at least 99% reduction of the targeted gene expression. In some embodiments, the base editing efficiency may result in knockout (100% knockdown of the gene expression) of the gene that is targeted.

[00780] In some embodiments, targeted modifications, e.g., single base editing, are used simultaneously to target at least 4, 5, 6, 7, 8, 9, 10, 11, 12 13, 14, 15, 16, 17 ,18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49 or 50 different endogenous sequences for base editing with different guide RNAs. In some embodiments, targeted modifications, e.g. single base editing, are used to sequentially target at least 4, 5, 6, 7, 8, 9, 10, 11, 12 13, 14, 15, 16, 17 ,18, 19, 20, 21, 22, 23,

24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49 50, or more different endogenous gene sequences for base editing with different guide RNAs.

[00781] In some embodiments, a single gene delivery event (e.g., by transduction, transfection, electroporation or any other method) can be used to target base editing of 5 sequences within a cell's genome. In some embodiments, a single gene delivery event can be used to target base editing of 6 sequences within a cell's genome. In some embodiments, a single gene delivery event can be used to target base editing of 7 sequences within a cell's genome. In some embodiments, a single electroporation event can be used to target base editing of 8 sequences within a cell's genome. In some embodiments, a single gene delivery event can be used to target base editing of 9 sequences within a cell's genome. In some embodiments, a single gene delivery event can be used to target base editing of 10 sequences within a cell's genome. In some embodiments, a single gene delivery event can be used to target base editing of 20 sequences within a cell's genome. In some embodiments, a single gene delivery event can be used to target base editing of 30 sequences within a cell's genome. In some embodiments, a single gene delivery event can be used to target base editing of 40 sequences within a cell's genome. In some embodiments, a single gene delivery event can be used to target base editing of 50 sequences within a cell's genome.

[00782] In some embodiments, the method described herein, for example, the base editing methods has minimum to no off-target effects.

[00783] In some embodiments, the base editing method described herein results in at least 50% of a cell population that have been successfully edited (i.e., cells that have been successfully engineered). In some embodiments, the base editing method described herein results in at least 55% of a cell population that have been successfully edited. In some embodiments, the base editing method described herein results in at least 60% of a cell population that have been successfully edited. In some embodiments, the base editing method described herein results in at least 65% of a cell population that have been successfully edited. In some embodiments, the base editing method described herein results in at least 70% of a cell population that have been successfully edited. In some embodiments, the base editing method described herein results in at least 75% of a cell population that have been successfully edited. In some embodiments, the base editing method described herein results in at least 80% of a cell population that have been successfully edited. In some embodiments, the base editing method described herein results in at least 85% of a cell population that have been successfully edited. In some embodiments, the base editing method described herein

results in at least 90% of a cell population that have been successfully edited. In some embodiments, the base editing method described herein results in at least 95% of a cell population that have been successfully edited. In some embodiments, the base editing method described herein results in about 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% of a cell population that have been successfully edited.

[00784] In some embodiments, the live cell recovery following a base editing intervention is greater than at least 60%, 70%, 80%, 90% of the starting cell population at the time of the base editing event. In some embodiments, the live cell recovery as described above is about 70%. In some embodiments, the live cell recovery as described above is about 75%. In some embodiments, the live cell recovery as described above is about 80%. In some embodiments, the live cell recovery as described above is about 85%. In some embodiments, the live cell recovery as described above is about 90%, or about 91%, 92%, 93%, 94% 95%, 96%, 97%, 98%, or 99%, or 100% of the cells in the population at the time of the base editing event.

[00785] In some embodiments the engineered cell population can be further expanded in vitro by about 2 fold, about 3-fold, about 4-fold, about 5-fold, about 6-fold, about 7-fold, about 8-fold, about 9-fold, about 10-fold, about 15-fold, about 20-fold, about 25-fold, about 30-fold, about 35-fold, about 40-fold, about 45-fold, about 50-fold, or about 100-fold.

Methods for Editing Nucleic Acids

[00786] Some aspects of the disclosure provide methods for editing a nucleic acid. In some embodiments, the method is a method for editing a nucleobase of a nucleic acid (*e.g.*, a base pair of a double-stranded DNA sequence). In some embodiments, the method comprises the steps of: a) contacting a target region of a nucleic acid (*e.g.*, a double-stranded DNA sequence) with a complex comprising a base editor (*e.g.*, a Cas9 domain fused to a cytidine or adenosine deaminase) and a guide nucleic acid (*e.g.*, gRNA), wherein the target region comprises a targeted nucleobase pair, b) inducing strand separation of said target region, c) converting a first nucleobase of said target nucleobase pair in a single strand of the target region to a second nucleobase, and d) cutting no more than one strand of said target region, where a third nucleobase complementary to the first nucleobase base is replaced by a fourth nucleobase complementary to the second nucleobase. In some embodiments, the method results in less than 20% indel formation in the nucleic acid. It should be appreciated that in some embodiments, step b is omitted. In some embodiments, the method results in less than 19%, 18%, 16%, 14%, 12%, 10%, 8%, 6%, 4%, 2%, 1%, 0.5%, 0.2%, or less than 0.1% indel formation. In some embodiments, the method further comprises replacing the second

nucleobase with a fifth nucleobase that is complementary to the fourth nucleobase, thereby generating an intended edited base pair (*e.g.*, C•G to T•A). In some embodiments, at least 5% of the intended base pairs are edited. In some embodiments, at least 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, or 50% of the intended base pairs are edited.

[00787] In some embodiments, the ratio of intended products to unintended products in the target nucleotide is at least 2:1, 5:1, 10:1, 20:1, 30:1, 40:1, 50:1, 60:1, 70:1, 80:1, 90:1, 100:1, or 200:1, or more. In some embodiments, the ratio of intended mutation to indel formation is greater than 1:1, 10:1, 50:1, 100:1, 500:1, or 1000:1, or more. In some embodiments, the cut single strand (nicked strand) is hybridized to the guide nucleic acid. In some embodiments, the cut single strand is opposite to the strand comprising the first nucleobase. In some embodiments, the base editor comprises a Cas9 domain. In some embodiments, the base editor protects or binds the non-edited strand. In some embodiments, the base editor comprises nickase activity. In some embodiments, the intended edited base pair is upstream of a PAM site. In some embodiments, the intended edited base pair is 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 nucleotides upstream of the PAM site. In some embodiments, the intended edited base pair is downstream of a PAM site. In some embodiments, the intended edited base pair is 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 nucleotides downstream stream of the PAM site. In some embodiments, the method does not require a canonical (*e.g.,* NGG) PAM site. In some embodiments, the nucleobase editor comprises a linker. In some embodiments, the linker is 1-25 amino acids in length. In some embodiments, the linker is 5-20 amino acids in length. In some embodiments, linker is 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 amino acids in length. In one embodiment, the linker is 32 amino acids in length. In another embodiment, a "long linker" is at least about 60 amino acids in length. In other embodiments, the linker is between about 3-100 amino acids in length. In some embodiments, the target region comprises a target window, wherein the target window comprises the target nucleobase pair. In some embodiments, the target window comprises 1-10 nucleotides. In some embodiments, the target window is 1-9, 1-8, 1-7, 1-6, 1-5, 1-4, 1-3, 1-2, or 1 nucleotides in length. In some embodiments, the target window is 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 nucleotides in length. In some embodiments, the intended edited base pair is within the target window. In some embodiments, the target window comprises the intended edited base pair. In some embodiments, the method is performed using any of the base editors provided herein. In some embodiments, a target window is a methylation window.

[00788] In some embodiments, the disclosure provides methods for editing a nucleotide. In some embodiments, the disclosure provides a method for editing a nucleobase pair of a double-stranded DNA sequence. In some embodiments, the method comprises a) contacting a target region of the double-stranded DNA sequence with a complex comprising a base editor and a guide nucleic acid (*e.g.*, gRNA), where the target region comprises a target nucleobase pair, b) inducing strand separation of said target region, c) converting a first nucleobase of said target nucleobase pair in a single strand of the target region to a second nucleobase, d) cutting no more than one strand of said target region, wherein a third nucleobase complementary to the first nucleobase base is replaced by a fourth nucleobase complementary to the second nucleobase, and the second nucleobase is replaced with a fifth nucleobase that is complementary to the fourth nucleobase, thereby generating an intended edited base pair, wherein the efficiency of generating the intended edited base pair is at least 5%. It should be appreciated that in some embodiments, step b is omitted. In some embodiments, at least 5% of the intended base pairs are edited. In some embodiments, at least 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, or 50% of the intended base pairs are edited. In some embodiments base editing by a method described herein may have a base conversion efficiency of at least 10% at any particular gene site. In some embodiments, base editing by a method described herein may have a base conversion efficiency of at least 20%, at least 25%, at least 30%, at least 35%, at least 40%, at least 45%, at least 50% at least 55% or at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, or at least 95%, 96%, 97%, 98% or at least 99% at any particular gene site. In some embodiments base editing by a method described herein may have a base conversion efficiency of at least 70% at any particular gene site. In some embodiments base editing by a method described herein may have a base conversion efficiency of at least 80% at any particular gene site. In some embodiments base editing by a method described herein may have a base conversion efficiency of at least 90% at any particular gene site.

[00789] In some embodiments, the method causes less than 19%, 18%, 16%, 14%, 12%, 10%, 8%, 6%, 4%, 2%, 1%, 0.5%, 0.2%, or less than 0.1% indel formation. In some embodiments, the ratio of intended product to unintended products at the target nucleotide is at least 2:1, 5:1, 10:1, 20:1, 30:1, 40:1, 50:1, 60:1, 70:1, 80:1, 90:1, 100:1, or 200:1, or more. In some embodiments, the ratio of intended mutation to indel formation is greater than 1:1, 10:1, 50:1, 100:1, 500:1, or 1000:1, or more. In some embodiments, the cut single strand is hybridized to the guide nucleic acid. In some embodiments, the cut single strand is opposite to the strand comprising the first nucleobase. In some embodiments, the nucleobase editor

comprises nickase activity. In some embodiments, the intended edited base pair is upstream of a PAM site. In some embodiments, the intended edited base pair is 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 nucleotides upstream of the PAM site. In some embodiments, the intended edited base pair is downstream of a PAM site. In some embodiments, the intended edited base pair is 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 nucleotides downstream stream of the PAM site. In some embodiments, the method does not require a canonical (e.g., NGG) PAM site. In some embodiments, the nucleobase editor comprises a linker. In some embodiments, the linker is 1-25 amino acids in length. In some embodiments, the linker is 5-20 amino acids in length. In some embodiments, the linker is 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 amino acids in length. e.g., In some embodiments, the target region comprises a target window, wherein the target window comprises the target nucleobase pair. In some embodiments, the target window comprises 1-10 nucleotides. In some embodiments, the target window is 1-9, 1-8, 1-7, 1-6, 1-5, 1-4, 1-3, 1-2, or 1 nucleotides in length. In some embodiments, the target window is 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 nucleotides in length. In some embodiments, the intended edited base pair occurs within the target window. In some embodiments, the target window comprises the intended edited base pair. In some embodiments, the nucleobase editor is any one of the base editors provided herein.

Nucleic Acid-Based Delivery of Cytidine or Adenosine Deaminase Nucleobase Editor

[00790] Nucleic acids encoding a cytidine or adenosine deaminase nucleobase editor according to the present disclosure can be administered to subjects or delivered into cells by art-known methods or as described herein. For example, cytidine or adenosine deaminase nucleobase editors can be delivered by, e.g., vectors (e.g., viral or non-viral vectors), non-vector based methods (e.g., using naked DNA or DNA complexes), or a combination thereof.

[00791] Nucleic acids encoding cytidine or adenosine deaminase nucleobase editors can be delivered directly to cells as naked DNA or RNA, for instance by means of transfection or electroporation, or can be conjugated to molecules (e.g., N-acetylgalactosamine) promoting uptake by the target cells. Nucleic acid vectors, such as the vectors can also be used. In particular embodiments, a polynucleotide, e.g. a mRNA encoding a base editor or a functional component thereof may be co-electroporated with a combination of multiple guide RNAs as described herein.

[00792] Nucleic acid vectors can comprise one or more sequences encoding a domain of a fusion protein described herein. A vector can also comprise a sequence encoding a signal

peptide (e.g., for nuclear localization, nucleolar localization, or mitochondrial localization), associated with (e.g., inserted into or fused to) a sequence coding for a protein. As one example, a nucleic acid vectors can include a Cas9 coding sequence that includes one or more nuclear localization sequences (e.g., a nuclear localization sequence from SV40), and one or more deaminases.

[00793] The nucleic acid vector can also include any suitable number of regulatory/control elements, e.g., promoters, enhancers, introns, polyadenylation signals, Kozak consensus sequences, or internal ribosome entry sites (IRES). These elements are well known in the art.

[00794] Nucleic acid vectors according to this disclosure include recombinant viral vectors. Exemplary viral vectors are set forth herein above. Other viral vectors known in the art can also be used. In addition, viral particles can be used to deliver genome editing system components in nucleic acid and/or peptide form. For example, "empty" viral particles can be assembled to contain any suitable cargo. Viral vectors and viral particles can also be engineered to incorporate targeting ligands to alter target tissue specificity.

[00795] In addition to viral vectors, non-viral vectors can be used to deliver nucleic acids encoding genome editing systems according to the present disclosure. One important category of non-viral nucleic acid vectors are nanoparticles, which can be organic or inorganic. Nanoparticles are well known in the art. Any suitable nanoparticle design can be used to deliver genome editing system components or nucleic acids encoding such components. For instance, organic (e.g. lipid and/or polymer) nanoparticles can be suitable for use as delivery vehicles in certain embodiments of this disclosure. Exemplary lipids for use in nanoparticle formulations, and/or gene transfer are shown in Table 9 (below).

[00796] Table 9.

Lipids Used for Gene Transfer

Lipid	Abbreviation	Feature
1,2-Dioleoyl-sn-glycero-3-phosphatidylcholine	DOPC	Helper
1,2-Dioleoyl-sn-glycero-3-phosphatidylethanolamine	DOPE	Helper
Cholesterol		Helper
N-[1-(2,3-Dioleyloxy)prophyl]N,N,N-trimethylammonium chloride	DOTMA	Cationic
1,2-Dioleoyloxy-3-trimethylammonium-propane	DOTAP	Cationic
Dioctadecylamidoglycylspermine	DOGS	Cationic
N-(3-Aminopropyl)-N,N-dimethyl-2,3-bis(dodecyloxy)-1-propanaminium bromide	GAP-DLRIE	Cationic
Cetyltrimethylammonium bromide	CTAB	Cationic
6-Lauroxyhexyl ornithinate	LHON	Cationic
1-(2,3-Dioleoyloxypropyl)-2,4,6-trimethylpyridinium	2Oc	Cationic

Lipids Used for Gene Transfer

Lipid	Abbreviation	Feature
2,3-Dioleyloxy-N-[2(sperminecarboxamido-ethyl]-N,N-dimethyl-1-propanaminium trifluoroacetate	DOSPA	Cationic
1,2-Dioleyl-3-trimethylammonium-propane	DOPA	Cationic
N-(2-Hydroxyethyl)-N,N-dimethyl-2,3-bis(tetradecyloxy)-1-propanaminium bromide	MDRIE	Cationic
Dimyristooxypropyl dimethyl hydroxyethyl ammonium bromide	DMRI	Cationic
3β-[N-(N',N'-Dimethylaminoethane)-carbamoyl]cholesterol	DC-Chol	Cationic
Bis-guanidium-tren-cholesterol	BGTC	Cationic
1,3-Diodeoxy-2-(6-carboxy-spermyl)-propylamide	DOSPER	Cationic
Dimethyloctadecylammonium bromide	DDAB	Cationic
Dioctadecylamidoglicylspermidin	DSL	Cationic
rac-[(2,3-Dioctadecyloxypropyl)(2-hydroxyethyl)]-dimethylammonium chloride	CLIP-1	Cationic
rac-[2(2,3-Dihexadecyloxypropyl-oxymethyloxy)ethyl]trimethylammoniun bromide	CLIP-6	Cationic
Ethyldimyristoylphosphatidylcholine	EDMPC	Cationic
1,2-Distearyloxy-N,N-dimethyl-3-aminopropane	DSDMA	Cationic
1,2-Dimyristoyl-trimethylammonium propane	DMTAP	Cationic
O,O'-Dimyristyl-N-lysyl aspartate	DMKE	Cationic
1,2-Distearoyl-sn-glycero-3-ethylpho sphocholine	DSEPC	Cationic
N-Palmitoyl D-erythro-sphingosyl carbamoyl-spermine	CCS	Cationic
N-t-Butyl-N0-tetradecyl-3-tetradecylaminopropionamidine	diC14-amidine	Cationic
Octadecenolyoxy[ethyl-2-heptadecenyl-3 hydroxyethyl] imidazolinium chloride	DOTIM	Cationic
N1 -Cholesteryloxycarbonyl-3,7-diazanonane-1,9-diamine	CDAN	Cationic
2-(3-[Bis(3-amino-propyl)-amino]propylamino)-N-ditetradecylcarbamoylme-ethyl-acetamide	RPR209120	Cationic
1,2-dilinoleyloxy-3-dimethylaminopropane	DLinDMA	Cationic
2,2-dilinoleyl-4-dimethylaminoethyl-[1,3]-dioxolane	DLin-KC2-DMA	Cationic
dilinoleyl-methyl-4-dimethylaminobutyrate	DLin-MC3-DMA	Cationic

[00797] Table 10 lists exemplary polymers for use in gene transfer and/or nanoparticle formulations.

Table 10

Polymers Used for Gene Transfer

Polymer	Abbreviation
Poly(ethylene)glycol	PEG
Polyethylenimine	PEI
Dithiobis (succinimidylpropionate)	DSP
Dimethyl-3,3'-dithiobispropionimidate	DTBP
Poly(ethylene imine)biscarbamate	PEIC
Poly(L-lysine)	PLL
Histidine modified PLL	
Poly(N-vinylpyrrolidone)	PVP
Poly(propylenimine)	PPI
Poly(amidoamine)	PAMAM
Poly(amidoethylenimine)	SS-PAEI
Triethylenetetramine	TETA
Poly(β-aminoester)	
Poly(4-hydroxy-L-proline ester)	PHP
Poly(allylamine)	
Poly(α-[4-aminobutyl]-L-glycolic acid)	PAGA
Poly(D,L-lactic-co-glycolic acid)	PLGA
Poly(N-ethyl-4-vinylpyridinium bromide)	
Poly(phosphazene)s	PPZ
Poly(phosphoester)s	PPE
Poly(phosphoramidate)s	PPA
Poly(N-2-hydroxypropylmethacrylamide)	pHPMA
Poly (2-(dimethylamino)ethyl methacrylate)	pDMAEMA
Poly(2-aminoethyl propylene phosphate)	PPE-EA
Chitosan	
Galactosylated chitosan	
N-Dodacylated chitosan	
Histone	
Collagen	
Dextran-spermine	D-SPM

[00798] The following Table 11 summarizes delivery methods for a polynucleotide encoding a fusion protein described herein.

Table 11

Delivery	Vector/Mode	Delivery into Non-Dividing Cells	Duration of Expression	Genome Integration	Type of Molecule Delivered
Physical	(*e.g.*, electroporation, particle gun, Calcium Phosphate transfection	YES	Transient	NO	Nucleic Acids and Proteins
Viral	Retrovirus	NO	Stable	YES	RNA
	Lentivirus	YES	Stable	YES/NO with modification	RNA
	Adenovirus	YES	Transient	NO	DNA
	Adeno-Associated Virus (AAV)	YES	Stable	NO	DNA
	Vaccinia Virus	YES	Very Transient	NO	DNA
	Herpes Simplex Virus	YES	Stable	NO	DNA
Non-Viral	Cationic Liposomes	YES	Transient	Depends on what is delivered	Nucleic Acids and Proteins
	Polymeric Nanoparticles	YES	Transient	Depends on what is delivered	Nucleic Acids and Proteins
Biological Non-Viral Delivery Vehicles	Attenuated Bacteria	YES	Transient	NO	Nucleic Acids
	Engineered Bacteriophages	YES	Transient	NO	Nucleic Acids
	Mammalian Virus-like Particles	YES	Transient	NO	Nucleic Acids
	Biological liposomes: Erythrocyte Ghosts and Exosomes	YES	Transient	NO	Nucleic Acids

[00799] In particular embodiments, a fusion protein of the invention is encoded by a polynucleotide present in a viral vector (e.g., adeno-associated virus (AAV), AAV3, AAV3b, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAVrh8, AAV10, and variants thereof), or a suitable capsid protein of any viral vector. Thus, in some aspects, the disclosure relates to the viral delivery of a fusion protein. Examples of viral vectors include retroviral vectors (e.g. Maloney murine leukemia virus, MML-V), adenoviral vectors (e.g. AD100), lentiviral vectors (HIV and FIV-based vectors), herpesvirus vectors (e.g. HSV-2).

[00800] In one embodiment, inteins are utilized to join fragments or portions of a cytidine or adenosine deaminase base editor protein that is grafted onto an AAV capsid protein. As used herein, "intein" refers to a self-splicing protein intron (e.g., peptide) that ligates flanking N-terminal and C-terminal exteins (e.g., fragments to be joined). The use of certain inteins for joining heterologous protein fragments is described, for example, in Wood et al., J. Biol. Chem. 289(21); 14512-9 (2014). For example, when fused to separate protein fragments, the inteins IntN and IntC recognize each other, splice themselves out and simultaneously ligate the flanking N- and C-terminal exteins of the protein fragments to which they were fused, thereby reconstituting a full-length protein from the two protein fragments. Other suitable inteins will be apparent to a person of skill in the art.

[00801] A fragment of a fusion protein of the invention can vary in length. In some embodiments, a protein fragment ranges from 2 amino acids to about 1000 amino acids in length. In some embodiments, a protein fragment ranges from about 5 amino acids to about 500 amino acids in length. In some embodiments, a protein fragment ranges from about 20 amino acids to about 200 amino acids in length. In some embodiments, a protein fragment ranges from about 10 amino acids to about 100 amino acids in length. Suitable protein fragments of other lengths will be apparent to a person of skill in the art.

[00802] In some embodiments, a portion or fragment of a nuclease (e.g., a fragment of a deaminase, such as cytidine or adenosine deaminase, or a fragment of *Cas9*) is fused to an intein. The nuclease can be fused to the N-terminus or the C-terminus of the intein. In some embodiments, a portion or fragment of a fusion protein is fused to an intein and fused to an AAV capsid protein. The intein, nuclease and capsid protein can be fused together in any arrangement (e.g., nuclease-intein-capsid, intein-nuclease-capsid, capsid-intein-nuclease, etc.). In some embodiments, the N-terminus of an intein is fused to the C-terminus of a fusion protein and the C-terminus of the intein is fused to the N-terminus of an AAV capsid protein.

[00803] In some aspects, the methods described herein for editing specific genes in an immune cell can be used to genetically modify a CAR-T cell. Such CAR-T cells, and

methods to produce such CAR-T cells are described in International Application Nos. PCT/US2016/060736, PCT/US2016/060734, PCT/US2016/034873, PCT/US2015/040660, PCT/EP2016/055332, PCT/IB2015/058650, PCT/EP2015/067441, PCT/EP2014/078876, PCT/EP2014/059662, PCT/IB2014/061409, PCT/US2016/019192, PCT/US2015/059106, PCT/US2016/052260, PCT/US2015/020606, PCT/US2015/055764, PCT/CN2014/094393, PCT/US2017/059989, PCT/US2017/027606, and PCT/US2015/064269, the contents of each is hereby incorporated in its entirety.

Pharmaceutical Compositions

[00804] In some aspects, the present invention provides a pharmaceutical composition comprising a genetically modified immune cell of the present invention. More specifically, provided herein are pharmaceutical compositions comprising a genetically modified immune cell, or a population of such immune cells, expressing a chimeric antigen receptor, wherein said modified immune cell, or a population thereof, has at least one edited gene edited to enhance the function of the modified immune cell or to reduce immunosuppression or inhibition of the modified immune cell, wherein expression of the edited gene is either knocked out or knocked down. In some embodiments the at least one edited gene is TRAC, B2M, PDCD1, CBLB, TGFBR2, ZAP70, NFATc1, TET2, or combination thereof.

[00805] The pharmaceutical compositions of the present invention can be prepared in accordance with known techniques. See, e.g., Remington, The Science And Practice of Pharmacy (21st ed. 2005). In general, the immune cell, or population thereof is admixed with a suitable carrier prior to administration or storage, and in some embodiments, the pharmaceutical composition further comprises a pharmaceutically acceptable carrier. Suitable pharmaceutically acceptable carriers generally comprise inert substances that aid in administering the pharmaceutical composition to a subject, aid in processing the pharmaceutical compositions into deliverable preparations, or aid in storing the pharmaceutical composition prior to administration. Pharmaceutically acceptable carriers can include agents that can stabilize, optimize or otherwise alter the form, consistency, viscosity, pH, pharmacokinetics, solubility of the formulation. Such agents include buffering agents, wetting agents, emulsifying agents, diluents, encapsulating agents, and skin penetration enhancers. For example, carriers can include, but are not limited to, saline, buffered saline, dextrose, arginine, sucrose, water, glycerol, ethanol, sorbitol, dextran, sodium carboxymethyl cellulose, and combinations thereof.

[00806] In addition to the modified immune cell, or population thereof, and the carrier, the pharmaceutical compositions of the present invention can include at least one additional therapeutic agent useful in the treatment of disease. For example, some embodiments of the pharmaceutical composition described herein further comprise a chemotherapeutic agent. In some embodiments, the pharmaceutical composition further comprises a cytokine peptide or a nucleic acid sequence encoding a cytokine peptide. In some embodiments, the pharmaceutical compositions comprising the modified immune cell or population thereof can be administered separately from an additional therapeutic agent.

[00807] The pharmaceutical compositions of the present invention can be used to treat any disease or condition that is responsive to autologous or allogeneic immune cell immunotherapy. For example, the pharmaceutical compositions, in some embodiments are useful in the treatment of neoplasia. In some embodiments, the neoplasia is a hematological cancer. In some embodiments, the hematological cancer is a B cell cancer, and in some embodiments, the B cell cancer is multiple myeloma. In some embodiments, the B cell cancer is relapsed of relapsed/refractory multiple myeloma.

[00808] One consideration concerning the therapeutic use of genetically modified immune cells of the invention is the quantity of cells necessary to achieve an optimal or satisfactory effect. The quantity of cells to be administered may vary for the subject being treated. In one embodiment, between 10^4 to 10^{10}, between 10^5 to 10^9, or between 10^6 and 10^8 genetically modified immunoresponsive cells of the invention are administered to a human subject. In some embodiments, at least about 1 x 10^8, 2 x 10^8, 3 x 10^8, 4 x 10^8, and 5 x 10^8 genetically modified immune cells of the invention are administered to a human subject. Determining the precise effective dose may be based on factors for each individual subject, including their size, age, sex, weight, and condition. Dosages can be readily ascertained by those skilled in the art from this disclosure and the knowledge in the art.

[00809] The skilled artisan can readily determine the number of cells and amount of optional additives, vehicles, and/or carriers in compositions and to be administered in methods of the invention. Typically, additives (in addition to the active immune cell(s)) are present in an amount of 0.001 to 50 % (weight) solution in phosphate buffered saline, and the active ingredient is present in the order of micrograms to milligrams, such as about 0.0001 to about 5 wt%, preferably about 0.0001 to about 1 wt%, still more preferably about 0.0001 to about 0.05 wt% or about 0.001 to about 20 wt%, preferably about 0.01 to about 10 wt%, and still more preferably about 0.05 to about 5 wt %. Of course, for any composition to be administered to an animal or human, and for any particular method of administration, it is

preferred to determine therefore: toxicity, such as by determining the lethal dose (LD) and LD50 in a suitable animal model (e.g., a rodent such as a mouse); and, the dosage of the composition(s), concentration of components therein, and the timing of administering the composition(s), which elicit a suitable response. Such determinations do not require undue experimentation from the knowledge of the skilled artisan, this disclosure and the documents cited herein. And, the time for sequential administrations can be ascertained without undue experimentation.

[00810] In one embodiment, the method and compositions described herein may be used in generating engineered T cells that express a CAR and may have one or more base edited modifications, such that the engineered T cell can mount a specific immune response against the target. The CAR may be specifically directed towards an antigen target, the antigen may be presented by a cell in a host. In some embodiments, the immune response encompasses cytotoxicity. In some embodiments, the engineered T cell has enhanced cytotoxic response against its target. In some embodiments, the engineered T cell induces an enhanced cytotoxic response against its target as compared to a non-engineered T cell. In some embodiments, the engineered T cell exhibits an enhanced cytotoxic response by at least 1.1-fold, 1.5-fold, 2-fold, 3-fold, 4-fold, 5-fold, 6-fold, 7-fold, 8-fold, 9-fold, 10-fold, 11-fold, 12-fold, 13-fold, 14-fold, 15-fold, 16-fold, 17-fold, 18-fold, 19-fold, 20-fold or more compared to a non-engineered cell. In some embodiments, the engineered T cell can kill at least 10%, at least 20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 100%, at least 200%, at least 500% or at least 1000% more target cells than a non-engineered cell. In some embodiments, the T cell can induce higher memory response. In some embodiments, the T cell can induce lower levels of inflammatory cytokines than a non-engineered cell, that is, the engineered cell does not cause a cytokine storm response. In some embodiments, the engineered T cell is administered to an allogenic host, wherein the engineered T cell has no rejection by the host. In some embodiments, the allogenic T cell induces negligible or minimum rejection by the host.

Methods of Treatment

[00811] Some aspects of the present invention provide methods of treating a subject in need, the method comprising administering to a subject in need an effective therapeutic amount of a pharmaceutical composition as described herein. More specifically, the methods of treatment comprise administering to a subject in need thereof a pharmaceutical composition comprising a population of modified immune cells expressing a chimeric receptor and having at least one edited gene, wherein the at least one edited gene enhances

the function or reduces the immunosuppression or inhibition of the modified immune cell, and wherein expression of the at least one edited gene is either knocked out or knocked down. In some embodiments, the method of treatment is an autologous immune cell therapy. In other embodiments, the method of treatment is an allogeneic immune cell therapy.

[00812] In certain embodiments, the specificity of an immune cell is redirected to a marker expressed on the surface of a diseased or altered cell in a subject by genetically modifying the immune cell to express a chimeric antigen receptor contemplated herein. In some embodiments, the method of treatment comprises administering to a subject an immune cell as described herein, wherein the immune cell has been genetically modified to redirect its specificity to a marker expressed on a neoplastic cell. In some embodiments, the neoplasia is a B cell cancer; for example, a B cell cancer such as a lymphoma, leukemia, or a myeloma, for example, multiple myeloma. Thus, some embodiments of the present disclosure provide a method of treating a neoplasia in a subject. In some embodiments, the neoplasia being treated is a B cell cancer. In some embodiments, the B cell cancer is a lymphoma, leukemia, or multiple myeloma.

[00813] Some embodiments of the methods of treating a neoplasia in a subject comprise administering to the subject an immune cell as described herein and one or more additional therapeutic agents. For example, the immune cell of the present invention can be co-administered with a cytokine. In some embodiments, the cytokine is IL-2, IFN-á, IFN-ã, or a combination thereof. In some embodiments, the immune cell is co-administered with a chemotherapeutic agent. The chemotherapeutic can be cyclophosphamide, doxorubicin, vincristine, prednisone, or rituximab, or a combination thereof. Other chemotherapeutics include obinutuzumab, bendamustine, chlorambucil, cyclophosphamide, ibrutinib, methotrexate, cytarabine, dexamethasone, cisplatin, bortezomib, fludarabine, idelalisib, acalabrutinib, lenalidomide, venetoclax, cyclophosphamide, ifosfamide, etoposide, pentostatin, melphalan, carfilzomib, ixazomib, panobinostat, daratumumab, elotuzumab, thalidomide, lenalidomide, or pomalidomide, or a combination thereof. "Co-administered" refers to administering two or more therapeutic agents or pharmaceutical compositions during a course of treatment. Such co-administration can be simultaneous administration or sequential administration. Sequential administration of a later-administered therapeutic agent or pharmaceutical composition can occur at any time during the course of treatment after administration of the first pharmaceutical composition or therapeutic agent.

[00814] In some embodiments of the present invention, an administered immune cell proliferates *in vivo* and can persist in the subject for an extended period of time. Immune

cells of the present invention, in some embodiments can mature into memory immune cells and remain in circulation within the subject, thereby generating a population of cells able to actively respond to recurrence of a diseased or altered cell expressing the marker recognized by the chimeric antigen receptor.

[00815] Administration of the pharmaceutical compositions contemplated herein may be carried out using conventional techniques including, but not limited to, infusion, transfusion, or parenterally. In some embodiments, parenteral administration includes infusing or injecting intravascularly, intravenously, intramuscularly, intraarterially, intrathecally, intratumorally, intradermally, intraperitoneally, transtracheally, subcutaneously, subcuticularly, intraarticularly, subcapsularly, subarachnoidly and intrasternally.

[00816] *Kits, Vectors, Cells*

[00817] The invention also provides kits comprising a nucleic acid construct comprising a nucleotide sequence encoding a cytidine or adenosine deaminase nucleobase editor at least two guide RNAs, each guide RNA having a nucleic acid sequence at least 85% complementary to a nucleic acid sequence of gene encoding TRAC, B2M, PD1, CBLB, and/or CTLA4. In some embodiments, the nucleotide sequence encoding the cytidine or adenosine deaminase comprises a heterologous promoter that drives expression of the cytidine or adenosine deaminase nucleobase editor.

[00818] Some aspects of this disclosure provide kits comprising a nucleic acid construct, comprising (a) a nucleotide sequence encoding (a) a Cas9 domain fused to a cytidine or adenosine deaminase as provided herein; and (b) a heterologous promoter that drives expression of the sequence of (a).

[00819] Some aspects of this disclosure provide kits for the treatment of a neoplasia comprising a modified immune cell or immune cell having reduced immunogenicity and enhanced anti-neoplasia activity, the immune or immune cell comprising a mutation in a TRAC, B2M, PD1, CBLB, and/or CTLA4 polypeptide, or a combination thereof. In some embodiments, the modified immune cell further comprises a chimeric antigen receptor having an affinity for a marker associated with the neoplasia. The neoplasia treatment kits comprise written instructions for using the modified immune cells in the treatment of the neoplasia.

[00820] The practice of the present invention employs, unless otherwise indicated, conventional techniques of molecular biology (including recombinant techniques), microbiology, cell biology, biochemistry and immunology, which are well within the purview of the skilled artisan. Such techniques are explained fully in the literature, such as,

"Molecular Cloning: A Laboratory Manual", second edition (Sambrook, 1989); "Oligonucleotide Synthesis" (Gait, 1984); "Animal Cell Culture" (Freshney, 1987); "Methods in Enzymology" "Handbook of Experimental Immunology" (Weir, 1996); "Gene Transfer Vectors for Mammalian Cells" (Miller and Calos, 1987); "Current Protocols in Molecular Biology" (Ausubel, 1987); "PCR: The Polymerase Chain Reaction", (Mullis, 1994); "Current Protocols in Immunology" (Coligan, 1991). These techniques are applicable to the production of the polynucleotides and polypeptides of the invention, and, as such, may be considered in making and practicing the invention. Particularly useful techniques for particular embodiments will be discussed in the sections that follow.

[00821] The following examples are put forth so as to provide those of ordinary skill in the art with a complete disclosure and description of how to make and use the assay, screening, and therapeutic methods of the invention, and are not intended to limit the scope of what the inventors regard as their invention.

EXAMPLES

[00822] Example 1: Disruption of Splice Sites and Introduction of Stop Codons in Genes Expressed in Immune Cells

[00823] A nucleobase editor, BE4, was used to disrupt splice sites and insert stop codons into a subset of genes expressed in immune cells. A plasmid construct, pCMV_BE4max, encodes BE4, which comprises an APOBEC-1 cytidine deaminase domain having cytidine deaminase activity, a Cas9 domain comprising a D10A mutation and having nicknase activity, and two uracil DNA glycosylase inhibitor (UGI) domains. UGI is an 83-amino acid residue protein derived from *Bacillus subtilis* bacteriophage PBS1 that potently blocks to edit the splice sites of certain genes expressed in immune cells. BE4 further comprises N-terminal and C-terminal nuclear localization signals (NLSs).

[00824] >pCMV_BE4max

ATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCA
TTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTAT
TAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGG
ATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGG
AGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCG
CCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCA
GAGCTGGTTTAGTGAACCGTCAGATCCGCTAGAGATCCGCGGCCGCTAATACGA

CTCACTATAGGGAGAGCCGCCACCATGAAACGGACAGCCGACGGAAGCGAGTTC
GAGTCACCAAAGAAGAAGCGGAAAGTCTCCTCAGAGACTGGGCCTGTCGCCGTC
GATCCAACCCTGCGCCGCCGGATTGAACCTCACGAGTTTGAAGTGTTCTTTGACC
CCCGGGAGCTGAGAAAGGAGACATGCCTGCTGTACGAGATCAACTGGGGAGGCA
GGCACTCCATCTGGAGGCACACCTCTCAGAACACAAATAAGCACGTGGAGGTGA
ACTTCATCGAGAAGTTTACCACAGAGCGGTACTTCTGCCCCAATACCAGATGTAG
CATCACATGGTTTCTGAGCTGGTCCCCTTGCGGAGAGTGTAGCAGGGCCATCACC
GAGTTCCTGTCCAGATATCCACACGTGACACTGTTTATCTACATCGCCAGGCTGT
ATCACCACGCAGACCCAAGGAATAGGCAGGGCCTGCGCGATCTGATCAGCTCCG
GCGTGACCATCCAGATCATGACAGAGCAGGAGTCCGGCTACTGCTGGCGGAACT
TCGTGAATTATTCTCCTAGCAACGAGGCCCACTGGCCTAGGTACCCACACCTGTG
GGTGCGCCTGTACGTGCTGGAGCTGTATTGCATCATCCTGGGCCTGCCCCCTTGT
CTGAATATCCTGCGGAGAAAGCAGCCCCAGCTGACCTTCTTTACAATCGCCCTGC
AGTCTTGTCACTATCAGAGGCTGCCACCCCACATCCTGTGGGCCACAGGCCTGAA
GTCTGGAGGATCTAGCGGAGGATCCTCTGGCAGCGAGACACCAGGAACAAGCGA
GTCAGCAACACCAGAGAGCAGTGGCGGCAGCAGCGGCGGCAGCGACAAGAAGT
ACAGCATCGGCCTGGCCATCGGCACCAACTCTGTGGGCTGGGCCGTGATCACCG
ACGAGTACAAGGTGCCCAGCAAGAAATTCAAGGTGCTGGGCAACACCGACCGGC
ACAGCATCAAGAAGAACCTGATCGGAGCCCTGCTGTTCGACAGCGGCGAAACAG
CCGAGGCCACCCGGCTGAAGAGAACCGCCAGAAGAAGATACACCAGACGGAAG
AACCGGATCTGCTATCTGCAAGAGATCTTCAGCAACGAGATGGCCAAGGTGGAC
GACAGCTTCTTCCACAGACTGGAAGAGTCCTTCCTGGTGGAAGAGGATAAGAAG
CACGAGCGGCACCCCATCTTCGGCAACATCGTGGACGAGGTGGCCTACCACGAG
AAGTACCCCACCATCTACCACCTGAGAAAGAAACTGGTGGACAGCACCGACAAG
GCCGACCTGCGGCTGATCTATCTGGCCCTGGCCCACATGATCAAGTTCCGGGGCC
ACTTCCTGATCGAGGGCGACCTGAACCCCGACAACAGCGACGTGGACAAGCTGT
TCATCCAGCTGGTGCAGACCTACAACCAGCTGTTCGAGGAAAACCCCATCAACG
CCAGCGGCGTGGACGCCAAGGCCATCCTGTCTGCCAGACTGAGCAAGAGCAGAC
GGCTGGAAAATCTGATCGCCCAGCTGCCCGGCGAGAAGAAGAATGGCCTGTTCG
GAAACCTGATTGCCCTGAGCCTGGGCCTGACCCCCAACTTCAAGAGCAACTTCGA
CCTGGCCGAGGATGCCAAACTGCAGCTGAGCAAGGACACCTACGACGACGACCT
GGACAACCTGCTGGCCCAGATCGGCGACCAGTACGCCGACCTGTTTCTGGCCGCC
AAGAACCTGTCCGACGCCATCCTGCTGAGCGACATCCTGAGAGTGAACACCGAG
ATCACCAAGGCCCCCCTGAGCGCCTCTATGATCAAGAGATACGACGAGCACCAC

CAGGACCTGACCCTGCTGAAAGCTCTCGTGCGGCAGCAGCTGCCTGAGAAGTAC
AAAGAGATTTTCTTCGACCAGAGCAAGAACGGCTACGCCGGCTACATTGACGGC
GGAGCCAGCCAGGAAGAGTTCTACAAGTTCATCAAGCCCATCCTGGAAAAGATG
GACGGCACCGAGGAACTGCTCGTGAAGCTGAACAGAGAGGACCTGCTGCGGAA
GCAGCGGACCTTCGACAACGGCAGCATCCCCCACCAGATCCACCTGGGAGAGCT
GCACGCCATTCTGCGGCGGCAGGAAGATTTTTACCCATTCCTGAAGGACAACCG
GGAAAAGATCGAGAAGATCCTGACCTTCCGCATCCCCTACTACGTGGGCCCTCTG
GCCAGGGGAAACAGCAGATTCGCCTGGATGACCAGAAAGAGCGAGGAAACCAT
CACCCCCTGGAACTTCGAGGAAGTGGTGGACAAGGGCGCTTCCGCCCAGAGCTT
CATCGAGCGGATGACCAACTTCGATAAGAACCTGCCCAACGAGAAGGTGCTGCC
CAAGCACAGCCTGCTGTACGAGTACTTCACCGTGTATAACGAGCTGACCAAAGT
GAAATACGTGACCGAGGGAATGAGAAAGCCCGCCTTCCTGAGCGGCGAGCAGA
AAAAGGCCATCGTGGACCTGCTGTTCAAGACCAACCGGAAAGTGACCGTGAAGC
AGCTGAAAGAGGACTACTTCAAGAAAATCGAGTGCTTCGACTCCGTGGAAATCT
CCGGCGTGGAAGATCGGTTCAACGCCTCCCTGGGCACATACCACGATCTGCTGA
AAATTATCAAGGACAAGGACTTCCTGGACAATGAGGAAAACGAGGACATTCTGG
AAGATATCGTGCTGACCCTGACACTGTTTGAGGACAGAGAGATGATCGAGGAAC
GGCTGAAAACCTATGCCCACCTGTTCGACGACAAAGTGATGAAGCAGCTGAAGC
GGCGGAGATACACCGGCTGGGGCAGGCTGAGCCGGAAGCTGATCAACGGCATCC
GGGACAAGCAGTCCGGCAAGACAATCCTGGATTTCCTGAAGTCCGACGGCTTCG
CCAACAGAAACTTCATGCAGCTGATCCACGACGACAGCCTGACCTTTAAAGAGG
ACATCCAGAAAGCCCAGGTGTCCGGCCAGGGCGATAGCCTGCACGAGCACATTG
CCAATCTGGCCGGCAGCCCCGCCATTAAGAAGGGCATCCTGCAGACAGTGAAGG
TGGTGGACGAGCTCGTGAAAGTGATGGGCCGGCACAAGCCCGAGAACATCGTGA
TCGAAATGGCCAGAGAGAACCAGACCACCCAGAAGGGACAGAAGAACAGCCGC
GAGAGAATGAAGCGGATCGAAGAGGGCATCAAAGAGCTGGGCAGCCAGATCCT
GAAAGAACACCCCGTGGAAAACACCCAGCTGCAGAACGAGAAGCTGTACCTGTA
CTACCTGCAGAATGGGCGGGATATGTACGTGGACCAGGAACTGGACATCAACCG
GCTGTCCGACTACGATGTGGACCATATCGTGCCTCAGAGCTTTCTGAAGGACGAC
TCCATCGACAACAAGGTGCTGACCAGAAGCGACAAGAACCGGGGCAAGAGCGA
CAACGTGCCCTCCGAAGAGGTCGTGAAGAAGATGAAGAACTACTGGCGGCAGCT
GCTGAACGCCAAGCTGATTACCCAGAGAAAGTTCGACAATCTGACCAAGGCCGA
GAGAGGCGGCCTGAGCGAACTGGATAAGGCCGGCTTCATCAAGAGACAGCTGGT
GGAAACCCGGCAGATCACAAAGCACGTGGCACAGATCCTGGACTCCCGGATGAA

CACTAAGTACGACGAGAATGACAAGCTGATCCGGGAAGTGAAAGTGATCACCCT
GAAGTCCAAGCTGGTGTCCGATTTCCGGAAGGATTTCCAGTTTTACAAAGTGCGC
GAGATCAACAACTACCACCACGCCCACGACGCCTACCTGAACGCCGTCGTGGGA
ACCGCCCTGATCAAAAAGTACCCTAAGCTGGAAAGCGAGTTCGTGTACGGCGAC
TACAAGGTGTACGACGTGCGGAAGATGATCGCCAAGAGCGAGCAGGAAATCGG
CAAGGCTACCGCCAAGTACTTCTTCTACAGCAACATCATGAACTTTTTCAAGACC
GAGATTACCCTGGCCAACGGCGAGATCCGGAAGCGGCCTCTGATCGAGACAAAC
GGCGAAACCGGGGAGATCGTGTGGGATAAGGGCCGGGATTTTGCCACCGTGCGG
AAAGTGCTGAGCATGCCCCAAGTGAATATCGTGAAAAAGACCGAGGTGCAGACA
GGCGGCTTCAGCAAAGAGTCTATCCTGCCCAAGAGGAACAGCGATAAGCTGATC
GCCAGAAAGAAGGACTGGGACCCTAAGAAGTACGGCGGCTTCGACAGCCCCACC
GTGGCCTATTCTGTGCTGGTGGTGGCCAAAGTGGAAAAGGGCAAGTCCAAGAAA
CTGAAGAGTGTGAAAGAGCTGCTGGGGATCACCATCATGGAAAGAAGCAGCTTC
GAGAAGAATCCCATCGACTTTCTGGAAGCCAAGGGCTACAAAGAAGTGAAAAAG
GACCTGATCATCAAGCTGCCTAAGTACTCCCTGTTCGAGCTGGAAAACGGCCGG
AAGAGAATGCTGGCCTCTGCCGGCGAACTGCAGAAGGGAAACGAACTGGCCCTG
CCCTCCAAATATGTGAACTTCCTGTACCTGGCCAGCCACTATGAGAAGCTGAAGG
GCTCCCCCGAGGATAATGAGCAGAAACAGCTGTTTGTGGAACAGCACAAGCACT
ACCTGGACGAGATCATCGAGCAGATCAGCGAGTTCTCCAAGAGAGTGATCCTGG
CCGACGCTAATCTGGACAAAGTGCTGTCCGCCTACAACAAGCACCGGGATAAGC
CCATCAGAGAGCAGGCCGAGAATATCATCCACCTGTTTACCCTGACCAATCTGGG
AGCCCCTGCCGCCTTCAAGTACTTTGACACCACCATCGACCGGAAGAGGTACACC
AGCACCAAAGAGGTGCTGGACGCCACCCTGATCCACCAGAGCATCACCGGCCTG
TACGAGACACGGATCGACCTGTCTCAGCTGGGAGGTGACAGCGGCGGGAGCGGC
GGGAGCGGGGGGAGCACTAATCTGAGCGACATCATTGAGAAGGAGACTGGGAA
ACAGCTGGTCATTCAGGAGTCCATCCTGATGCTGCCTGAGGAGGTGGAGGAAGT
GATCGGCAACAAGCCAGAGTCTGACATCCTGGTGCACACCGCCTACGACGAGTC
CACAGATGAGAATGTGATGCTGCTGACCTCTGACGCCCCCGAGTATAAGCCTTGG
GCCCTGGTCATCCAGGATTCTAACGGCGAGAATAAGATCAAGATGCTGAGCGGA
GGATCCGGAGGATCTGGAGGCAGCACCAACCTGTCTGACATCATCGAGAAGGAG
ACAGGCAAGCAGCTGGTCATCCAGGAGAGCATCCTGATGCTGCCCGAAGAAGTC
GAAGAAGTGATCGGAAACAAGCCTGAGAGCGATATCCTGGTCCATACCGCCTAC
GACGAGAGTACCGACGAAAATGTGATGCTGCTGACATCCGACGCCCCAGAGTAT
AAGCCCTGGGCTCTGGTCATCCAGGATTCCAACGGAGAGAACAAAATCAAAATG

CTGTCTGGCGGCTCAAAAAGAACCGCCGACGGCAGCGAATTCGAGCCCAAGAAG
AAGAGGAAAGTCTAACCGGTCATCATCACCATCACCATTGAGTTTAAACCCGCTG
ATCAGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCG
TGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGA
GGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTG
GGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGA
TGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCTCGATACC
GTCGACCTCTAGCTAGAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGA
AATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGT
AAAGCCTAGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCAC
TGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCA
ACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACT
GACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGG
CGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAG
CAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTT
TCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGA
GGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCT
CCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTT
CTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTT
CGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCC
CGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACAC
GACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTAT
GTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGA
AGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAG
TTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGT
TTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGAT
CTTTTCTACGGGGTCTGACACTCAGTGGAACGAAAACTCACGTTAAGGGATTTTG
GTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAA
GTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATG
CTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTG
CCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCC
CAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCA
ATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCC
GCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAG

TTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTC
GTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACA
TGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTG
TCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAA
TTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAA
CCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTC
AATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGG
AAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGT
TCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAG
CGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAA
GGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAG
CATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAA
AATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTC
GACGGATCGGGAGATCGATCTCCCGATCCCCTAGGGTCGACTCTCAGTACAATCT
GCTCTGATGCCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTTGGAGGT
CGCTGAGTAGTGCGCGAGCAAAATTTAAGCTACAACAAGGCAAGGCTTGACCGA
CAATTGCATGAAGAATCTGCTTAGGGTTAGGCGTTTTGCGCTGCTTCGCGATGTA
CGGGCCAGATATACGCGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAA
TTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTAC
GGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAAT
AATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGG
GTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATC

[00825] To ascertain the effectiveness of BE4 in knocking down or out protein expression in immune cells, a first population of immune cells was co-transfected with mRNA encoding BE4 and an sgRNA that targeted the C base complementary to the G base of the donor or acceptor splice site of TRAC exon 1, TRAC exon 3, or B2M exon 1, depending on the specific target site. mRNA was produced by *in vitro* transcription, (TriLin Biotechnologies). Briefly, 4 microgm of BE4 mRNA and 2 microgm of synthetic gRNA were electroporated into 1M CD3+ T cells (Nucleofector™ Platform, Lonza Bioscience). The cells were then cultured for 3 days to allow sufficient time for base-editing. For comparison, a second population of immune cells was co-transfected with mRNA encoding a Cas9 nuclease and sgRNA that target the G base of the donor splice site of B2M exon 1. No discernible difference between BE4 editing and the Cas9 editing was observed, and the knockdown for

each edited gene was greater than 90%, whereas unelectroporated control cells had no significant knockdown (FIG. 2).

[00826] It was hypothesized that the genetic modifications responsible for the observed knockdown of the targeted genes would differ if the cells were transfected with mRNA encoding BE4, which catalyzes single strand nicks, or with the the Cas9 nuclease that catalyzes double-strand breaks. To test this hypothesis, immune cells were co-transfected with either 2 microgm BE4/1 microgm sgRNA (medium) or 4 microgm BE4/2 microgm sgRNA (high) RNA encoding the BE4 base editor and sgRNA that target the G base of the donor splice site of the B2M exon 1. After incubation for 3, 5, and 7 days, DNA was collected and sequenced. Referring to FIG. 3, the majority of base edits revealed disruption of only the splice site and in the manner expected (i.e., C to T transition in the antisense strand was incorporated, resulting in a G to A transition in the sense strand). These results contrasted with those obtained from cells transfected with a Cas9 nuclease, which show that most edits in the Cas9 transfected cells were indels (FIG. 3).

[00827] Disruption of splice site and the introduction of stop codons can be effective in knocking down expression of a target gene. BE4-mediated editing of the splice acceptor in TRAC exon 3 and the splice donor in B2M exon 1 and PDCD1 exon 1 resulted in reduced expression of the full-length proteins (FIGs. 4 and 5). The BE4-mediated changes observed in the splice site were C to T transitions, although indels and C to G transversions were also observed. Insertion of an ochre stop codon into exon 2 of the PDCD1 gene, in which consecutive cytidine residues in the exon were targeted and edited to thymidine residues, also resulted in significant knock down of gene expression, albeit a lesser reduction than that seen for the TRAC and B2M genes (FIG. 4). These results further suggest that BE4-mediated single or consecutive cytidine base editing of genes expressed in immune cells results in efficient reduction of gene expression.

[00828] Example 2: In silico analysis of spice site disruption and stop codon insertion

[00829] To determine if designed gRNA would bind to off-site targets, the nucleic acid sequences of the gRNAs were analyzed using CAS-OFFinder. Referring to FIG. 6, an "X" bulge type indicates that the gRNA aligns with the genomic DNA and any discrepancy is a mismatch. As the number of mismatches increases from one to four, the potential off-site binding increases. For example, results for the TRAC exon 3 splice acceptor show that when there are three mismatches, there are 26 offsite binding possibilities, while there are 164 with four mismatches.

[00830] If the gRNA has a bulge, wherein the gRNA has twenty base pairs, but aligns with nineteen base pairs of genomic DNA, a bulge results. Again referring to FIG. 6, when the TRAC exon 3 splice acceptor gRNA has a bulge of one base pair, the number of offsite binding possibilities increases with increasing mismatches; however, the number of possibilities is significantly lower than when there is no bulge (i.e., when the bulge size is zero).

[00831] Example 3: Multiplex Base Editing in Immune Cells

[00832] To determine if BE4 could mediate base editing of multiple genes to generate a multi-knockdown cell, immune cells were co-transfected with mRNA encoding a BE4 base editor along with sgRNA that target specific sites in B2M, TRAC, PD1, or in combinations thereof. Referring to FIG. 7, the BE4 system elicited effective knockdown, as measured by flow cytometry, to identify the percentage of cells with decreased protein production in single, double, and triple gene edits. The cells were gated on B2M and CD3 expression, with CD3 expression serving as a proxy for TRAC expression. Because PD1 staining is inefficient, direct measurement of cells expressing this protein was not performed. No differences were observed between cell populations with single, double, and triple gene edits, and immune cells modified to knock-down expression of B2M, TRAC, and PD1 (a triple gene edit) are detectably distinct from non-modified control immune cell (FIG. 8).

[00833] The modifications to the genes responsible for the decreased protein expression are summarized in FIG. 9. Specifically, and similarly to the mechanism resulting in decreased expression in single gene modification described in Example 1, C to T transitions constitute the vast number of edits observed in the modified B2M single modified gene cell population and in the B2M+PD1, B2M+TRAC, and B2M+TRAC+PD1 multiple modified genes cell populations. Indels and transversions constitute an insignificant minority of observed genetic changes in the edited genes.

[00834] Thus, concurrent modification of three genetic loci by base editing produced highly efficient gene knockouts with no detectable translocation events as assessed by Uni-Directional Targeted Sequencing (UDiTaS; Giannoukos et al., BMC Genomics. 2018 Mar 21;19(1):212. doi: 10.1186/s12864-018-4561-9). Additionally, translocations were not detected in BE4-edited genes. A droplet digital polymerase chain reaction (ddPCR) strategy (FIG. 10) was employed to detect translocations between the B2M, TRAC, and PD1 BE4-edited genes. DNA extracted from cells modified with BE4 or Cas9 to generate B2M+TRAC+PD1 edits was analyzed with next generation sequencing (NGS) using a

QX200 droplet digital instrument (Bio-Rad) to determine the exact sequence of the BE4 and Cas9 edits. As shown on the left panel of FIG. 11, the B2M, TRAC, and PD1 genes were modified in most cells. ddPCR analysis showed that translocations were not present in the BE4-edited cells, but were observed in approximately 1.7% of the Cas9-edited cells (FIG. 11, right panel). Table 12 further illustrates that translocations were not observed in the BE4-edited cells.

[00835] Table 12

Base Editor	Translocation	Control amplicon droplets	Experimental amplicon droplets
Cas9 nuclease	*B2M-TRAC*	61,206	585
	B2M-PDCD1	55,970	291
	PDCD1-TRAC	59,600	112
BE4	*B2M-TRAC*	90,717	0
	B2M-PDCD1	89,028	0
	PDCD1-TRAC	83,501	0

[00836] Example 4: BE4-mediated Editing of Cbl Proto-Oncogene B (CBLB)

[00837] Cbl-b is a T cell receptor (TCR) signaling protein that negatively regulates TCR complex signaling (FIG. 12). Because T cells have a lower activation threshold when Cbl-b signaling is inhibited, knocking out or down this gene could significantly improve the effectiveness of a T cell or a T cell expressing a CAR. To determine if the Cbl-b gene was susceptible cytidine deamination mediated modification, cells were co-transfected with mRNA encoding a BE4 and sgRNA that target the splice site acceptor of exon 8 and 16, the splice site donor of exons 8, 10, 11, and 12, or that would promote the insertion of a STOP codon in exons 1, 4, and 8. Resulting cells were analyzed with flow cytometry.

[00838] Referring to FIG. 13, disruption of the splice site donor of exon 12 and the splice site acceptor of exon 8 resulted in the greatest reduction of Cbl-b expression (67.2% and 57.4%, respectively). And of the cells transfected with the exon 8 splice site acceptor and the exon 12 splice site donor sgRNAs, slightly more than 60% of the cells were edited successfully (FIG. 13, bar graph).

[00839] Example 5: Cas12b Nuclease Characterization in Immune Cells

[00840] Cas12b/c2c1 site specifically targets and cleaves both strands of a double stranded nucleic acid molecule. Two different Cas12b/c2c1 proteins, BhCas12b and BvCas12b, were characterized by determining the propensity the enzymes for mediating indels in the target nucleic acid molecule. mRNA encoding the Cas12b/c2c1 proteins was electroporated into T cells along with guide RNAs specific for a locus in the GRIN2B gene and for a locus in the DNMT1 gene. The cells were cultured for 3-5 days, followed by isolation of cellular DNA. Indel rates were determined by Next Generation Sequencing. Referring to FIG. 14, DNA isolated from cells treated with the BhCas12b protein had a much higher percentage (approximately 75%) of indels in the GRIN2B gene than did the DNA isolated from cells treated with the BvCas12b protein (approximately 20%). Indels in the DNMT1 gene were also observed at a higher rate in the DNA isolated from cells treated with BhCas12b (approximately 20%) than observed in the DNA isolated from cells treated with BvCas12b (approximately 0%).

[00841] The BhCas12b (V4) protein was used to disrupt the TRAC gene. T cells were transduced via electroporation with the mRNA encoding the BhCas12b (V4) protein along with guide RNAs specific for loci in the GRIN2B, DNMT1, and TRAC genes. 96 hours post-electroporation, cells were assessed using fluorescence assisted cell sorting (FACS) analysis, with cells being gated for CD3 (a proxy for TRAC). Referring to FIG. 15, approximately 95% of T cells transduced with a plasmid encoding GFP or with BhCas12b (V4) and guide RNAs specific for GRIN2B or DNMT1 were CD3+. Those cells transduced to express BhCas12b (V4) and guide RNAs specific for loci in the TRAC gene were less likely to be CD3+ (approximately 2% to approximately 50%, depending on the guide RNA used). Three of the eleven TRAC guide RNAs tested led to approximately 100% BhCas12b (V4)-mediated indels.

[00842] Example 6: CAR-P2A-mCherry lentivirus expression characterization

[00843] Cells were transduced to express a chimeric antigen receptor (CAR) using the CAR-P2A-mCherry lentivirus and analyzed for CAR expression using fluorescence assisted cell sorting (FACS). Cells were unstained, incubated with a BCMA protein conjugated to R-phycoerythrin (PE) or fluorescein isothiocyanate (FITC). Because BCMA is the CAR's target antigen, cells expressing the CAR will bind dye-labeled BCMA. Referring to FIG. 16, for cells that were not stained, FACS analysis only detected the presence of mCherry in the transduced sample, with some spillover into the PE channel. The BCMA-PE channel shows a highly positive signal beyond what was seen in the spillover, and these results were

confirmed in cells incubated with BCMA-FITC. The dye-labeled BCMA protein detection results suggest almost identical expression of the CAR as that seen for mCherry. Referring to FIG. 17, 85% CAR expression was detected via FACS analysis in cells transduced with a poly(1,8-octanediol citrate) (POC) lentiviral vector.

[00844] Example 7: BE4 produces efficient, durable gene knockout with high product purity

[00845] BE4 mediates base editing of multiple genes to generate a multi-knockdown cell. Immune cells were co-transfected with mRNA encoding a BE4 base editor along with sgRNA that target specific sites in B2M, TRAC, PD1, or in combinations thereof. As shown by sequencing data, base editing was efficient at modifying cells and durable up to at least 7 days (FIG. 18). High product purity was observed, as C to T transitions constituted the vast number of edits observed. Indels and C-to-G and C-to-A transversions constituted an insignificant minority of observed genetic changes in the edited genes. Base editing was also as efficient as spCas9 nuclease at generating desired modifications.

[00846] The BE4 system elicited effective knockdown as measured by flow cytometry, which identifies the percentage of cells with decreased surface expression (FIG. 19A). Cells gated on B2M expression displayed loss of B2M protein on the cell surface. As measured by flow cytometry, base editing was also as efficient as spCas9 nuclease at generating B2M protein knockout.

[00847] Example 8: Orthogonal translocation detection assay cannot detect BE4-induced rearrangements in triple-edited T cells

[00848] Immune cells were co-transfected with mRNA encoding a BE4 base editor along with sgRNAs that targeted specific sites in B2M, TRAC, and PD1. The triple-edited T cells were evaluated using a translocation detection assay that was capable of detecting specific translocations that were undesirable between B2M, TRAC, and PD1 target genes (FIG. 20). Notably, none of these specific translocations were detected in any of the BE4-edited genes (Table 13). In contrast, Cas9-treated cells displayed low, but detectable levels of the translocations. Thus, multiplex editing of T cells using the BE4 base editor did not generate translocations in contrast to multiplex editing using Cas9 nuclease.

[00849] Table 13

Type	Mock (%)	BE4-treated (%)	Cas9-treated (%)
On-target modification (B2M/TRAC/PDCD1)	0	89.9 / 97.9 / 89.1	53.0 / 77.2 / 55.2
B2M-A / TRAC-A	0	0	0.925
B2M-A / TRAC-B	0	0	0.353
B2M-A / PDCD1-A	0	0	1.647
B2M-A / PDCD1-B	0	0	0.508
B2M-B / TRAC-A*	0	0	0.505

$LLoD_{BE4} = 0.1\%$

**B2M-B* only measurable in this experiment if translocation includes a local rearrangement at the *B2M* locus

[00850] Example 9: Multiplexed base editing does not significantly impair cell expansion

[00851] An extensive guide screen was performed across B2M, TRAC, and PD1 targets with both BE4 and spCas9 sgRNAs. Guides were selected for high editing efficiency and expansion based on single-plex test. Final cell yields compared between 1, 2 and 3 edits using BE4 and spCas9 and were normalized to electroporation only control. BE4 edited cells with the desired edits displayed high yields when up to 3 edits were made (FIG. 21). In contrast, spCas9 edited cells showed reduced yields when increasing numbers of multiplex edits were made. Thus, multiplexed base edited cells maintained high cell expansion even when up to 3 edits were being made. Thus, BE4 generated multiplex-edited T cells with no detectable genomic rearrangements while also maintaining high cell expansion compared to spCas9 treated samples.

[00852] Example 10: BE4 generated triple-edited T cells with similar on-target editing efficiency and cellular phenotype as spCas9

[00853] T cells were co-transfected with mRNA encoding a BE4 base editor along with sgRNAs that target specific sites in B2M, TRAC, and PD1. As shown by sequencing data, base editing was efficient at modifying cells at all three sites (FIG. 22). Modification of the

genes by base editing was similar to that using spCas9 nuclease. Flow cytometry also showed decreased surface expression of B2M and CD3 (FIG. 23, upper panel). Compared to electroporation only control cells, BE4 and Cas9 multiplex edited cells displayed significant reductions of B2M and CD3 protein on the cell surface (>95% $CD3^-/B2M^-$). Although PD1 staining is less efficient, significant reductions (~90%) in PD1 were observed in BE4 and Cas9 multiplex edited cells compared to electroporation only control cells (FIG. 23, lower panel).

[00854] Example 11: BE4 editing does not alter CAR expression or antigen-dependent cell killing

[00855] T cells were co-transfected with mRNA encoding a BE4 base editor along with sgRNAs that target specific sites in B2M, TRAC, and PD1. A chimeric antigen receptor (CAR) targeting BCMA was introduced by integration of a lentiviral vector encoding the anti-BCMA CAR. CAR expression was observed by flow cytometry in BE4 and Cas9 edited cells (FIG. 24), compared to untreated cells that did not receive the lentiviral vector. The CAR-T cells were evaluated for cell killing by nuclear staining of the cells expressing BCMA and detecting loss of nuclear staining, indicating cell death. Antigen dependendent cell killing was observed in cells transduced with the vector and expressing the CAR, including BE4 and Cas9 edited T cells (FIG. 25). In contrast, untreated cells that were not transduced with the vector did not display cell killing activity. Thus, BE4-generated CAR-T cells demonstrated comparable gene disruption, cell phenotype, and antigen-dependent cell killing compared to their nuclease-only counterparts.

[00856] Example 12: Cas12b and BE4 can be paired for highly efficient multiplex editing in T cells

[00857] $CD3^-$, $B2M^-$ T cells were generated using BE4 only or using BE4 and Cas12b. For T cells generated using BE4 only, T cells were co-transfected with mRNA encoding a BE4 base editor along with sgRNAs that target specific sites in B2M and TRAC. For T cells generated using BE4 and Cas12b, T cells were co-transfected with mRNA encoding a BE4 base editor, and an sgRNA that targets a specific site in B2M, mRNA encoding BhCas12b (V4), and a Cas12b sgRNA that targets exon 3 of the TRAC gene, which was used to disrupt the TRAC gene. The resulting T cells were assessed using fluorescence assisted cell sorting (FACS) analysis to detect B2M and CD3 cell surface expression. Knockouts using BE4 only displayed a similar profile to those using BE4 and Cas12b. In particular, a high percentage of

the T cells were $CD3^-$, $B2M^-$: 86% (BE4 only) and 88% (BE4+Cas12b), while the other possible phenotypes $CD3^-$, $B2M^+$; $CD3^+$, $B2M^+$ T cells; and $CD3^+$, $B2M^-$ were represented less in the cell population (FIG. 26). In contrast, electroporation only control showed a population having a high percentage (97.8%) of $CD3^+$ $B2M^+$ cells and a very low percentage of $CD3^-$, $B2M^-$ cells.

[00858] Cas12b was used to generate $CD3^-$, CAR^+ T cells. T cells were co-transfected with mRNA encoding BhCas12b (V4), a Cas12b sgRNA that targets exon 3 of the TRAC gene, and a double-stranded DNA (dsDNA) donor template encoding BCMA02, an anti-BCMA CAR. T cells were assessed using fluorescence assisted cell sorting (FACS) analysis to detect CD3 and BCMA02 cell surface expression. When increasing amounts of Cas12b were introduced into the cell in the presence of the sgRNA, CD3 expression decreased, as seen by a shift in the cell population to the $CD3^-$ quadrant (FIG. 27). When increasing amounts of donor template and were introduced in the cells under the same conditions, a shift to $CD3^-$, CAR^+ quadrant was observed in the cell population.

[00859] Thus, Cas12b can be paired with BE4 to generate multiplex-edited T cells, minimizing genomic rearrangements caused by multiple double-strand breaks.

[00860] Example 13: High efficiency multiplex knockout of eight targets

[00861] In this example, PBMCs were isolated from three donors and activated with soluble CD3 and CD28 antibodies. On day 3 after activation, T cells were electroporated with a reaction mixture including 2 microgm of recombinant BE4 and 1 microgm each of sgRNAs using a LONZA 4D electroporation device. (see Table 10 for sgRNA electroporated). Where indicated, half (1/2) gRNA dose is 0.5 microgm each of sgRNA; and 2x mRNA dose = 4 microgm mRNA with 0.5 microgm of each sgRNA. sgRNA were obtained from Synthego or Agilent.

[00862] Percent knockdown of gene expression was measured by flow cytometry. To determine the base editing efficiency of CIITA gene, HLADR was used as the surrogate protein for staining. These results indicate that efficient and effective multiplex base editing can be successfully performed on a large number of genes simultaneously in single electroporation events.

[00863] Table 14.

Target	Target Sequence
CD3	TTCGTATCTGTAAAACCAAG
CD7	CCTACCTGTCACCAGGACCA
CD52	CTCTTACCTGTACCATAACC
PD1	CACCTACCTAAGAACCATCC
B2M	ACTCACGCTGGATAGCCTCC
CD5	ACTCACCCAGCATCCCCAGC
CIITA	CACTCACCTTAGCCTGAGCA
CD2	CACGCACCTGGACAGCTGAC

[00864] As indicated in FIG.28A and FIG. 28B, knockdown of each of the targeted genes was achieved.

Other Embodiments

[00865] From the foregoing description, it will be apparent that variations and modifications may be made to the invention described herein to adopt it to various usages and conditions. Such embodiments are also within the scope of the following claims.

[00866] The recitation of a listing of elements in any definition of a variable herein includes definitions of that variable as any single element or combination (or subcombination) of listed elements. The recitation of an embodiment herein includes that embodiment as any single embodiment or in combination with any other embodiments or portions thereof.

[00867] All patents and publications mentioned in this specification are herein incorporated by reference to the same extent as if each independent patent and publication was specifically and individually indicated to be incorporated by reference.

What is claimed is:

1. A method for producing a modified immune cell with reduced immunogenicity and/or increased anti-neoplasia activity by multiplexed editing, the method comprising: modifying at least four gene sequences or regulatory elements thereof, at a single target nucleobase in each thereof in an immune cell, thereby generating the modified immune cell with reduced immunogenicity and/or increased anti-neoplasia activity.
2. A method for producing a population of modified immune cells with reduced immunogenicity and/or increased anti-neoplasia activity by multiplexed editing, the method comprising: modifying at least four gene sequences or regulatory elements thereof at a single target nucleobase in each thereof in a population of immune cells, thereby generating the population of modified immune cells with reduced immunogenicity and/or increased anti-neoplasia activity.
3. The method of claim 1 or 2, wherein at least one of the at least four gene sequences is a checkpoint inhibitor gene sequence, an immune response regulation gene sequence, or an immunogenic gene sequence.
4. The method of any one of claims 1-3, wherein the modifying reduces expression of at least one of the at least four gene sequences.
5. The method of any one of claim 1, wherein expression of at least one of the at least four genes is reduced by at least 80% as compared to a control cell without the modification.
6. The method of claim 5, wherein expression of each one of the at least four genes is reduced by at least 80% as compared to a control cell without the modification.
7. The method of claim 2, wherein expression of at least one of the at least four genes is reduced in at least 50% of the population of immune cells.
8. The method of 7, wherein expression of each one of the at least four genes is reduced in at least 50% of the population of immune cells.
9. The method of any one of the preceding claims, wherein the at least four gene sequences comprise a TCR complex gene sequence.
10. The method of claim 9, wherein the at least four gene sequences comprise a TRAC gene sequence.
11. The method of any one of claims 1-8, wherein the at least four gene sequences comprise a check point inhibitor gene sequence.

12. The method of claim 10, wherein the at least four gene sequences comprise a PDCD1 gene sequence.
13. The method of any one of claims 1-8, wherein the at least four gene sequences comprise a T cell marker gene sequence.
14. The method of claim 13, wherein the at least four gene sequences comprise a CD52 gene sequence.
15. The method of claim 13, wherein the at least four gene sequences comprises a CD7 gene sequence.
16. The method of any one of claims 1-15, wherein the at least four gene sequences comprise a TRAC gene sequence, a PDCD1 gene sequence, a CD52 gene sequence, or a CD7 gene sequence.
17. The method of any one of claims 1-16, wherein the at least four sequences comprises a TCR complex gene sequence, a CD7 gene sequence, a CD52 gene sequence ,and a gene sequence selected from the group consisting of a CD2 gene sequence, a CD4 gene sequence, a CD5 gene sequence, a CD7 gene sequence, a CD30 gene sequence, a CD33 gene sequence, a CD52 gene sequence, a CD70 gene sequence, a B2M gene sequence, and a CIITA gene sequence
18. The method of claim any one of claims 1-17, wherein the at least four gene sequences comprise a gene sequence selected from the group consisting of a CD2 gene sequence, a TRAC gene sequence, a CD3 epsilon gene sequence, a CD3 gamma gene sequence, a CD3 delta gene sequence, a TRBC1 gene sequence, a TRBC2 gene sequence, a CD4 gene sequence, a CD5 gene sequence, a CD7 gene sequence, a CD30 gene sequence, a CD33 gene sequence, a CD52 gene sequence, a CD70 gene sequence, a B2M gene sequence, and a CIITA gene sequence.
19. The method of any one of claims 1, 5, and 6, comprising modifying five gene sequences or regulatory elements thereof at a single target nucleobase in each thereof in the immune cell.
20. The method of any one of claims 1, 5, and 6, comprising modifying six gene sequences or regulatory elements thereof at a single target nucleobase in each thereof in the immune cell.
21. The method of any one of claims 1, 5, and 6, comprising modifying seven gene sequences or regulatory elements thereof at a single target nucleobase in each thereof in the immune cell.

22. The method of any one of claims 1, 5, and 6, comprising modifying eight gene sequences or regulatory elements thereof at a single target nucleobase in each thereof in the immune cell.
23. The method of any one of claims 2, 7, and 8, comprising modifying five gene sequences or regulatory elements thereof at a single target nucleobase in each thereof in the population of immune cells.
24. The method of any one of claims 2, 7, and 8, comprising modifying six gene sequences or regulatory elements thereof at a single target nucleobase in each thereof in the population of immune cells.
25. The method of any one of claims 2, 7, and 8, comprising modifying seven gene sequences or regulatory elements thereof at a single target nucleobase in each thereof in the population of immune cells.
26. The method of any one of claims 2, 7, and 8, comprising modifying eight gene sequences or regulatory elements thereof at a single target nucleobase in each thereof in the population of immune cells.
27. The method of any one of claims 19-26, wherein the five, six, seven, or eight gene sequences or regulatory elements thereof are selected from the group consisting of a CD2 gene sequence, a TRAC gene sequence, a CD3 epsilon gene sequence, a CD3 gamma gene sequence, a CD3 delta gene sequence, a TRBC1 gene sequence, a TRBC2 gene sequence, a CD4 gene sequence, a CD5 gene sequence, a CD7 gene sequence, a CD30 gene sequence, a CD33 gene sequence, a CD52 gene sequence, a CD70 gene sequence, a B2M gene sequence, and a CIITA gene sequence.
28. The method of any one of claims 19-27, wherein the five, six, seven, or eight gene sequences or regulatory elements thereof at comprises a CD3 gene sequence, a CD7 gene sequence, a a CD2 gene sequence, a CD5 gene sequence, and a CD52 gene sequence.
29. The method of any one of the preceding claims, wherein the modifying comprises deaminating the single target nucleobase.
30. The method of claim 29, wherein the deaminating is performed by a polypeptide comprising a deaminase.
31. The method of claim 30, wherein the deaminase is associated with a nucleic acid programmable DNA binding protein (napDNAbp) to form a base editor.
32. The method of claim 31, wherein the deaminase is fused to the nucleic acid programmable DNA binding protein (napDNAbp).

33. The method of claim 32, wherein the napDNAbp comprises a Cas9 polypeptide or a portion thereof.
34. The method of claim 33, wherein the napDNAbp comprises a Cas9 nickase or nuclease dead Cas9.
35. The method of any one of claims 30-34, wherein the deaminase is a cytidine deaminase.
36. The method of claim 35, wherein the single target nucleobase is a cytosine (C) and wherein the modification comprises conversion of the C to a thymine (T).
37. The method of claim 36, wherein the base editor further comprises a uracil glycosylase inhibitor.
38. The method of any one of claims 30-34, wherein the deaminase is an adenosine deaminase.
39. The method of claim 38, wherein the single target nucleobase is a adenosine (A) and wherein the modification comprises conversion of the A to a guanine (G).
40. The method of any one of claims 30-39, wherein the modifying comprises contacting the immune cell with a guide nucleic acid sequences.
41. The method of claim 40, wherein the modifying comprises contacting the immune cell with at least four guide nucleic acid sequences, wherein each guide nucleic acid sequence targets the napDNAbp to one of the at least four gene sequences or regulatory elements thereof.
42. The method of claim 40, wherein the guide nucleic acid sequence comprises a sequence selected from guide RNA sequences of table 8A, table 8B, or table 8C.
43. The method of claim 40, wherein the guide nucleic acid sequence comprises a sequence selected from the group consisting of UUCGUAUCUGUAAAACCAAG, CCUACCUGUCACCAGGACCA, CUCUUACCUGUACCAUAACC, CACCUACCUAAGAACCAUCC, ACUCACGCUGGAUAGCCUCC, ACUCACCCAGCAUCCCCAGC, CACUCACCUUAGCCUGAGCA, and CACGCACCUGGACAGCUGAC.
44. The method of any one of claims 1-28, wherein the modifying comprises replacing the single target nucleobase with a different nucleobase by target-primed reverse transcription with a reverse transcriptase and an extended guide nucleic acid sequence.
45. The method of claim 44, wherein the extended guide nucleic acid sequence comprises a reverse transcription template sequence, a reverse transcription primer binding site, or a combination thereof.

46. The method of any one of claims 1-45, wherein the single target nucleobase is in an exon.

47. The method of claim 46, wherein the modifying generates a premature stop codon in the exon.

48. The method of claim 46 or 47, wherein the single target nucleobase is within an exon 1, an exon 2, or an exon 3 of the TRAC gene sequence.

49. The method of claim 46 or 47, wherein the single target nucleobase is within an exon 1, an exon 2, or an exon 5 of the PCDC1 gene sequence.

50. The method of claim 46 or 47, wherein the single target nucleobase is within an exon 1 or an exon 2 of the CD52 gene sequence.

51. The method of claim 46 or 47, wherein the single target nucleobase is within an exon 1, an exon 2, or an exon 3 of the CD7 gene sequence.

52. The method of claim 46 or 47, wherein the single target nucleobase is within an exon 1 or an exon 2 of the B2M gene sequence.

53. The method of claim 46 or 47, wherein the single target nucleobase is within an exon 2, an exon 3, an exon 4, an exon 5, an exon 6, an exon 7, or an exon 8 of the CD5 gene sequence.

54. The method of claim 46 or 47, wherein the single target nucleobase is within an exon 2, an exon 3, an exon 4, or an exon 5 of the CD2 gene sequence.

55. The method of claim 46 or 47, wherein the single target nucleobase is within an exon 1, an exon 2, an exon 4, an exon 7, an exon 8, an exon 9, an exon 10, an exon 11, an exon 12, an exon 14, an exon 15, an exon 18, or an exon 19 of the CIITA gene sequence.

56. The method of any one of claims 1-45, wherein the single target nucleobase is in a splice donor site or a splice acceptor site.

57. The method of claim 50, wherein the single target nucleobase is in a exon 1 splice acceptor site, a exon 1 splice donor site, or a exon 3 splice acceptor site of the TRAC gene sequence.

58. The method of claim 50, wherein the single target nucleobase is in a exon 1 splice acceptor site, a exon 1 splice donor site, an exon 2 splice acceptor site, an exon 3 splice donor site, an exon 4 splice acceptor site, an exon 4 splice donor site, or a exon 5 splice acceptor site of the PDCD1 gene sequence.

59. The method of claim 50, wherein the single target nucleobase is in a exon 1 splice donor site, or an exon 2 splice acceptor site of the CD52 gene sequence.

60. The method of claim 50, wherein the single target nucleobase is in a exon 1 splice donor site, a exon 2 splice donor site, an exon 2 splice acceptor site, or an exon 3 splice acceptor site of the CD7 gene sequence.
61. The method of claim 50, wherein the single target nucleobase is in a exon 1 splice donor site, a exon 2 splice donor site, an exon 2 splice acceptor site, or an exon 3 splice acceptor site of the B2M gene sequence.
62. The method of claim 50, wherein the single target nucleobase is in an exon 3 splice donor site of the CD2 gene sequence.
63. The method of claim 50, wherein the single target nucleobase is in an exon 1 splice donor site, an exon 1 splice acceptor site, an exon 3 splice acceptor site, an exon 3 splice donor site, an exon 4 splice acceptor site, an exon 5 splice donor site, an exon 6 splice acceptor site, an exon 9 splice donor site, an exon 10 splice acceptor site of the CD5 gene sequence.
64. The method of claim 50, wherein the single target nucleobase is in an exon 1 splice donor site, an exon 7 splice donor site, an exon 8 splice acceptor site, an exon 9 slice donor site, an exon 10 splice acceptor site, an exon 11 splice acceptor site, an exon 14 splice acceptor site, an exon 14 splice donor site, an exon 15 splice donor site, an exon 16 splice acceptor site, an exon 16 splice donor site, an exon 17 splice acceptor site, an exon 17 splice donor site, or an exon 19 splice acceptor site of the CIITA gene sequence.
65. The method of any one of claims 4-64, wherein the immune cell is a human cell.
66. The method of claim 65, wherein the immune cell is a cytotoxic T cell, a regulatory T cell, a T helper cell, a dendritic cell, a B cell, or a NK cell.
67. The method of any one of claims 7-66, wherein the population of immune cells are human cells.
68. The method of claim 67, wherein the population of immune cells are cytotoxic T cells, regulatory T cells, T helper cells, dendritic cells, B cells, or NK cells.
69. The method of any one of claims 1-68, wherein the modifying is ex vivo.
70. The method of any one of claims 1-69, wherein the immune cell or the population of immune cells are derived from a single human donor.
71. The method of any one of claims 1-70, further comprising contacting the immune cell or the population of immune cells with a polynucleotide that encodes an exogenous functional chimeric antigen receptor (CAR) or a functional fragment thereof.

72. The method of claim 71, comprising contacting the immune cell or the population of immune cells with a lentivirus comprising the polynucleotide that encodes the CAR.
73. The method of claim 71, comprising contacting the immune cell or the population of immune cells with a napDNAbp and a donor DNA sequence comprising the polynucleotide that encodes the CAR.
74. The method of claim 73, wherein the napDNAbp is a Cas12b.
75. The method of any one of claims 71-74, wherein the CAR specifically binds a marker associated with neoplasia.
76. The method of claim 75, wherein the neoplasia is a T cell cancer, a B cell cancer, a lymphoma, a leukemia, or a multiple myeloma.
77. The method of claim 76, wherein the CAR specifically binds CD7.
78. The method of claim 76, wherein the CAR specifically binds BCMA.
79. The method of any one of claims 2-78, wherein the immune cell or the population of immune cells comprises no detectable translocation.
80. The method of claim 79, wherein at least 50% of the population of immune cells express the CAR.
81. The method of claim 79, wherein at least 50% of the population of immune cells are viable.
82. The method of claim 79,wherein at least 50% of the population of immune cells expand at least 80% of expansion rate of a population of control cells of a same type without the modification.
83. The method of any one of claims 4-79, wherein the modifying generates less than 1% of indels in the immune cell.
84. The method of any one of claims 4-79, wherein the modifying generates less than 5% of non-target edits in the immune cell.
85. The method of any one of claims 4-79, wherein the modifying generates less than 5% of off-target edits in the immune cell.
86. A modified immune cell produced according to the method of any one of claims 4-85.
87. A population of modified immune cells produced according to the method of any one of claims 7-85.
88. A modified immune cell with reduced immunogenicity or increased anti-neoplasia activity, wherein the modified immune cell comprises a single target nucleobase modification in each one of at least four gene sequences or regulatory elements thereof.

89. The modified immune cell of claim 88, wherein each one of the at least four gene sequences is a checkpoint inhibitor gene sequence, an immune response regulation gene sequence,or an immunogenic gene sequence.

90. The modified immune cell of claim 88 or 89, wherein the at least four gene sequences comprise a TCR complex gene sequence.

91. The modified immune cell of claim 90, wherein the at least four gene sequences comprise a TRAC gene sequence.

92. The modified immune cell of claim 88 or 89, wherein the at least four gene sequences comprise a check point inhibitor gene sequence.

93. The modified immune cell of claim 92, wherein the at least four gene sequences comprise a PDCD1 gene sequence.

94. The modified immune cell of claim 88 or 89, wherein the at least four gene sequences comprise a T cell marker gene sequence.

95. The modified immune cell of claim 94, wherein the at least four gene sequences comprise CD52 gene sequence.

96. The modified immune cell of claim 94, wherein the at least four gene sequences comprises a CD7 gene sequence.

97. The modified immune cell of any one of claims 88-96, wherein expression of one of the at least four genes is reduced by at least 80% as compared to a control cell without the modification.

98. The modified immune cell of claim 97, wherein expression of each one of the at least four genes is reduced by at least 90% as compared to a control cell without the modification.

99. The modified immune cell of any one of claims 88-98, wherein the immune cell comprises a modification at a single target nucleobase in each one of five gene sequences or regulatory elements thereof, wherein each one of the five gene sequences is a checkpoint inhibitor gene sequence, an immune response regulation gene sequence, or an immunogenic gene sequence.

100. The modified immune cell of any one of claims 88-98, wherein the immune cell comprises a modification at a single target nucleobase in each one of six gene sequences or regulatory elements thereof, wherein each one of the six gene sequences is a checkpoint inhibitor gene sequence, an immune response regulation gene sequence, or an immunogenic gene sequence.

101. The modified immune cell of any one of claims 88-98, wherein the immune cell comprises a modification at a single target nucleobase in each one of seven gene sequences or regulatory elements thereof, wherein each one of the seven gene sequences is a checkpoint inhibitor gene sequence, an immune response regulation gene sequence or an immunogenic gene sequence.

102. The modified immune cell of any one of claims 88-98, wherein the immune cell comprises a modification at a single target nucleobase in each one of eight gene sequences or regulatory elements thereof, wherein each one of the eight gene sequences is a checkpoint inhibitor gene sequence, an immune response regulation gene sequence, or an immunogenic gene sequence

103. The modified immune cell of any one of claims 99-102, wherein expression of at least one of the five, six, seven or eight genes is reduced by at least 90% as compared to a control cell without the modification.

104. The modified immune cell of any one of claims 99-102, wherein expression of each one of the five, six, seven, or eight genes is reduced by at least 90% as compared to a control cell without the modification.

105. The modified immune cell of any one of claims 99-104, wherein the five, six, seven, or eight gene sequences or regulatory elements thereof comprise a sequence selected from the group consisting of a CD2 gene sequence, a TRAC gene sequence, a CD3 epsilon gene sequence, a CD3 gamma gene sequence, a CD3 delta gene sequence, a TRBC1 gene sequence, a TRBC2 gene sequence, a CD4 gene sequence, a CD5 gene sequence, a CD7 gene sequence, a CD30 gene sequence, a CD33 gene sequence, a CD52 gene sequence, a CD70 gene sequence, a B2M gene sequence, and a CIITA gene sequence.

106. A modified immune cell comprising a single target nucleobase modification in each one of a CD3 gene sequence, a CD5 gene sequence, a CD52 gene sequence, and a CD7 gene sequence, wherein the modified immune cell exhibits reduced immunogenicity or increased anti-neoplasia activity as compared to a control cell of a same type without the modification.

107. The modified immune cell of claim 106, where in the immune cell further comprises a single target nucleobase modification in a CD2 gene sequence,CIITA or a regulatory element of each thereof.

108. The modified immune cell of claim 106, wherein the immune cell comprises a single target nucleobase modification in a TRAC gene sequence, a CD3 epsilon gene

sequence, a CD3 gamma gene sequence, a CD3 delta gene sequence, a TRBC1 gene sequence, or a TRBC2 gene sequence further comprises a single target nucleobase modification in a gene sequence a CD4 gene sequence, a CD30 gene sequence, a CD33 gene sequence, a CD70 gene sequence, a B2M gene sequence, and a CIITA gene sequence or a regulatory element of each thereof.

109. The modified immune cell of claim 107, wherein the immune cell comprises a single nucleobase modification in each one of a TRAC gene sequence, a PDCD1 gene sequence, a CD52 gene sequence, a CD7 gene sequence, a CD2 gene sequence, a CD5 gene sequence, a CIITA gene sequence, and a B2M gene sequence.

110. The modified immune cell of any one of claims 88-109, wherein the immune cell comprises no detectable translocation.

111. The modified immune cell of any one of claims 88-110, wherein the immune cell comprises less than 1% of indels.

112. The modified immune cell of any one of claims 88-110, wherein the immune cell comprises less than 5% of non-target edits.

113. The modified immune cell of any one of claims 88-110, wherein the immune cell comprises less than 5% of off-target edits.

114. The modified immune cell of any one of claims 88-110, wherein the immune cell is a mammalian cell.

115. The modified immune cell of any one of claims 88-110, wherein the immune cell is a human cell.

116. The modified immune cell of any one of claims 88-115, wherein the immune cell is a cytotoxic T cell, a regulatory T cell, a T helper cell, a dendritic cell, a B cell, or a NK cell.

117. The modified immune cell of any one of claims 88-116, wherein the immune cell is in an ex vivo culture.

118. The modified immune cell of any one of claims 88-117, wherein the immune cell is derived from a single human donor.

119. The modified immune cell of any one of claims 88-118, wherein the immune cell further comprises a polynucleotide that encodes an exogenous functional chimeric antigen receptor (CAR) or a functional fragment thereof.

120. The modified immune cell of claim 119, wherein the polynucleotide that encodes the CAR is integrated in the genome of the immune cell.

121. The modified immune cell of claim 119 or 120, wherein the CAR specifically binds a marker associated with neoplasia.

122. The modified immune cell of claim 119 or 120, wherein the neoplasia is a T cell cancer, a B cell cancer, a lymphoma, a leukemia, or a multiple myeloma.

123. The modified immune cell of any one of claims 119-122, wherein the CAR specifically binds CD7.

124. The modified immune cell of any one of claims 119-122, wherein the CAR specifically binds BCMA.

125. The modified immune cell of any one of claims 88-124, wherein the single target nucleobase is in an exon.

126. The modified immune cell of claim 125, wherein the single target nucleobase is within an exon 1, an exon 2, or an exon 3 of the TRAC gene sequence.

127. The modified immune cell of claim 125, wherein the single target nucleobase is within an exon 1, an exon 2, or an exon 5 of the PCDC1 gene sequence.

128. The modified immune cell of claim 125, wherein the single target nucleobase is within an exon 1 or an exon 2 of the CD52 gene sequence.

129. The modified immune cell of claim 125, wherein the single target nucleobase is within an exon 1, an exon 2, or an exon 3 of a CD7 gene sequence.

130. The modified immune cell of any one of claims 88-124, wherein the single target nucleobase is in a splice donor site or a splice acceptor site.

131. The modified immune cell of claim 130, wherein the single target nucleobase is in a exon 1 splice acceptor site, a exon 1 splice donor site, or a exon 3 splice acceptor site of the TRAC gene sequence.

132. The modified immune cell of claim 130, wherein the single target nucleobase is in a exon 1 splice acceptor site, a exon 1 splice donor site, an exon 2 splice acceptor site, an exon 3 splice donor site, an exon 4 splice acceptor site, an exon 4 splice donor site, or a exon 5 splice acceptor site of the PDCD1 gene sequence.

133. The modified immune cell of claim 130, wherein the single target nucleobase is in a exon 1 splice donor site, or an exon 2 splice acceptor site of the CD52 gene sequence.

134. The modified immune cell of claim 130, wherein the single target nucleobase is in a exon 1 splice donor site, a exon 2 splice donor site, an exon 2 splice acceptor site, or an exon 3 splice acceptor site of the CD7 gene sequence.

135. A population of modified immune cells, wherein a plurality of the population of cells comprise a single target nucleobase modification in each one of at least four gene sequences or regulatory elements thereof, and wherein the plurality of the population of cells having the modification exhibit reduced immunogenicity or increased anti-neoplasia activity as compared to a plurality of control cells of a same type without the modification.

136. The population of modified immune cells of claim 135, wherein the plurality of cells comprises at least 50% of the population.

137. The population of modified immune cells of claim 135 or 136, wherein each one of the at least four gene sequences is a checkpoint inhibitor gene sequence, an immune response regulation gene sequence, or an immunogenic gene sequence

138. The population of modified immune cells of claim 137, wherein the at least four gene sequences comprise a TCR component gene sequence, a check point inhibitor gene sequence, or a T cell marker gene sequence.

139. The population of modified immune cells of claim 137, wherein the at least four gene sequences comprise a TRAC gene sequence.

140. The population of modified immune cells of claim 137, wherein the at least four gene sequences comprise a PDCD1 gene sequence.

141. The population of modified immune cells of claim 137, wherein the at least four gene sequences comprise CD52 gene sequence.

142. The population of modified immune cells of claim 137, wherein the at least four gene sequences comprises a CD7 gene sequence.

143. The population of modified immune cells of any one of claims 135-142, wherein expression of at least one of the at least four genes is reduced by at least 80% in the plurality of cells having the modification as compared to a control cell without the modification

144. The population of modified immune cells of claim 143, wherein expression of each one of the at least four genes is reduced by at least 80% in the plurality of cells having the modification as compared to a control cell without the modification.

145. The population of modified immune cells of any one of claims 135-144, wherein the plurality of the population comprises a modification at a single target nucleobase in each one of five gene sequences or regulatory elements thereof, wherein each one of the five gene sequences is a checkpoint inhibitor gene sequence, an immune response regulation gene sequence, or an immunogenic gene sequence.

146. The population of modified immune cells of any one of claims 135-144, wherein the plurality of the population comprises a modification at a single target nucleobase in each one of six gene sequences or regulatory elements thereof, wherein each one of the six sequences is a checkpoint inhibitor gene sequence, an immune response regulation gene sequence, or an immunogenic gene sequence.

147. The population of modified immune cells of any one of claims 135-144, wherein the plurality of the population comprises a modification at a single target nucleobase in each one of seven gene sequences or regulatory elements thereof, wherein each one of the seven gene sequences is a checkpoint inhibitor gene sequence, an immune response regulation gene sequence, or an immunogenic gene sequence.

148. The population of modified immune cells of any one of claims 135-144, wherein the plurality of the population comprises a modification at a single target nucleobase in each one of eight gene sequences or regulatory elements thereof, wherein each one of the eight gene sequences is a checkpoint inhibitor gene sequence, an immune response regulation gene sequence, or an immunogenic gene sequence.

149. The population of modified immune cells of any one of claims 145-148, wherein expression of at least one of the five, six, seven, or eight genes is reduced by at least 90% in the plurality of cells having the modification as compared to a control cell without the modification

150. The population of modified immune cells of any one of claims 145-148, wherein expression of each one of the five, six, seven, or eight genes is reduced by at least 90% in the plurality of cells having the modification as compared to a control cell without the modification

151. The population of modified immune cells of any one of claims 145-148, wherein the five, six, seven, or eight gene sequences or regulatory elements thereof are selected from the group consisting of a CD2 gene sequence, a TRAC gene sequence, a CD3 epsilon gene sequence, a CD3 gamma gene sequence, a CD3 delta gene sequence, a TRBC1 gene sequence, a TRBC2 gene sequence, a CD4 gene sequence, a CD5 gene sequence, a CD7 gene sequence, a CD30 gene sequence, a CD33 gene sequence, a CD52 gene sequence, a CD70 gene sequence, a B2M gene sequence, and a CIITA gene sequence.

152. A population of modified immune cells, wherein a plurality of the population comprise a single target nucleobase modification in each one of a TRAC gene

sequence, a PDCD1 gene sequence, a CD52 gene sequence, and a CD7 gene sequence, and wherein the plurality of the population having the modification exhibit reduced immunogenicity or increased anti-neoplasia activity as compared to a plurality of control cells of a same type without the modification.

153. The population of modified immune cells of claim 152, wherein the plurality of the population further comprises a single target nucleobase modification in a CD2 gene sequence, a CD5 gene sequence, a CIITA gene sequence, a B2M gene sequence, or a regulatory element of each thereof.

154. The population of modified immune cells of claim 152, wherein the plurality of the population further comprises a single target nucleobase modification in a gene sequence of a gene selected from the group consisting of a CD2 gene sequence, a TRAC gene sequence, a CD3 epsilon gene sequence, a CD3 gamma gene sequence, a CD3 delta gene sequence, a TRBC1 gene sequence, a TRBC2 gene sequence, a CD4 gene sequence, a CD5 gene sequence, a CD7 gene sequence, a CD30 gene sequence, a CD33 gene sequence, a CD52 gene sequence, a CD70 gene sequence, a B2M gene sequence, and a CIITA gene sequence or a regulatory element of each thereof.

155. The population of modified immune cells of claim 153, wherein the plurality of the population comprises a single nucleobase modification in each one of a TRAC gene sequence, a PDCD1 gene sequence, a CD52 gene sequence, a CD7 gene sequence, a CD2 gene sequence, a CD5 gene sequence, a CIITA gene sequence, and a B2M gene sequence.

156. The population of modified immune cells of any one of claims 135-155, wherein the plurality of the population comprises no detectable translocation.

157. The population of modified immune cells of any one of claims 135-156, wherein at least 60% of the population of immune cells are viable.

158. The population of modified immune cells of any one of claims 135-156, wherein at least 60% of the population of immune cells expand at least 80% of expansion rate of a population of control cells of a same type without the modification.

159. The population of modified immune cells of any one of claims 135-158, wherein population of immune cells are human cells.

160. The population of modified immune cells of any one of claims 135-159, wherein the population of immune cells are cytotoxic T cells, regulatory T cells, T helper cells, dendritic cells, B cells, or NK cells.

161. The population of modified immune cells of any one of claims 135-160, wherein the population of immune cells are derived from a single human donor.

162. The population of modified immune cells of any one of claims 135-161, wherein the plurality of cells having the modification further comprises a polynucleotide that encodes an exogenous functional chimeric antigen receptor (CAR) or a functional fragment thereof.

163. The population of modified immune cells of claim 162, wherein at least 50% of the population of immune cells express the CAR.

164. The population of modified immune cells of claim 162 or 163, wherein the CAR specifically binds a marker associated with neoplasia.

165. The population of modified immune cells of claim 164, wherein the neoplasia is a T cell cancer, a B cell cancer, a lymphoma, a leukemia, or a multiple myeloma.

166. The population of modified immune cells of claim 165, wherein the CAR specifically binds CD7.

167. The population of modified immune cells of claim 165, wherein the CAR specifically binds BCMA.

168. The population of modified immune cells of claim of any one of claims 135-167, wherein the single target nucleobase is in an exon.

169. The population of modified immune cells of claim 168, wherein the single target nucleobase is within an exon 1, an exon 2, or an exon 3 of the TRAC gene sequence.

170. The population of modified immune cells of claim 168, wherein the single target nucleobase is within an exon 1, an exon 2, or an exon 5 of the PCDC1 gene sequence.

171. The population of modified immune cells of claim 168, wherein the single target nucleobase is within an exon 1 or an exon 2 of the CD52 gene sequence.

172. The population of modified immune cells of claim 168, wherein the single target nucleobase is within an exon 1, an exon 2, or an exon 3 of a CD7 gene sequence.

173. The population of modified immune cells of any one of claims 135-167, wherein the single target nucleobase is in a splice donor site or a splice acceptor site.

174. The population of modified immune cells of claim 173, wherein the single target nucleobase is in a exon 1 splice acceptor site, a exon 1 splice donor site, or a exon 3 splice acceptor site of the TRAC gene sequence.

175. The population of modified immune cells of claim 173, wherein the single target nucleobase is in a exon 1 splice acceptor site, a exon 1 splice donor site, an exon 2 splice acceptor site, an exon 3 splice donor site, an exon 4 splice acceptor site, an exon 4 splice donor site, or a exon 5 splice acceptor site of the PDCD1 gene sequence.

176. The population of modified immune cells of claim 173, wherein the single target nucleobase is in a exon 1 splice donor site, or an exon 2 splice acceptor site of the CD52 gene sequence.

177. The population of modified immune cells of claim 173, wherein the single target nucleobase is in a exon 1 splice donor site, a exon 2 splice donor site, an exon 2 splice acceptor site, or an exon 3 splice acceptor site of the CD7 gene sequence.

178. A composition comprising deaminase and a nucleic acid sequence, wherein the guide nucleic acid sequence comprises a sequence selected from the group consisting of UUCGUAUCUGUAAAACCAAG, CCUACCUGUCACCAGGACCA, CUCUUACCUGUACCAUAACC, CACCUACCUAAGAACCAUCC, ACUCACGCUGGAUAGCCUCC, ACUCACCCAGCAUCCCCAGC, CACUCACCUUAGCCUGAGCA, and CACGCACCUGGACAGCUGAC.

179. The composition of claim 178, wherein the deaminase is associated with a nucleic acid programmable DNA binding protein (napDNAbp) to form a base editor.

180. The composition of claim 179, wherein the napDNAbp comprises a Cas9 nickase or nuclease dead Cas9 and wherein the deaminase is a cytidine deaminase.

181. The composition of claim 180, wherein the base editor further comprises a uracil glycosylase inhibitor.

182. The composition of claim 166, wherein the napDNAbp comprises a Cas9 nickase or nuclease dead Cas9 and wherein the deaminase is a adenosine deaminase.

183. A composition comprising a polymerase and a guide nucleic acid sequence, wherein the guide nucleic acid sequence comprises a sequence selected from the group consisting of the group consisting of UUCGUAUCUGUAAAACCAAG, CCUACCUGUCACCAGGACCA, CUCUUACCUGUACCAUAACC, CACCUACCUAAGAACCAUCC, ACUCACGCUGGAUAGCCUCC, ACUCACCCAGCAUCCCCAGC, CACUCACCUUAGCCUGAGCA, and CACGCACCUGGACAGCUGAC.

184. The composition of claim 170, wherein the polymerase is a reverse transcriptase and wherein the guide nucleic acid sequence is an extended guide

nucleic acid sequence comprising a reverse transcription template sequence, a reverse transcription primer binding site, or a combination thereof.

185. A method for producing a modified immune cell with reduced immunogenicity and/or increased anti-neoplasia activity, the method comprising:

a) modifying a single target nucleobase in a first gene sequence or a regulatory element thereof in an immune cell; and

b) modifying a second gene sequence or a regulatory element thereof in the immune cell with a Cas12 polypeptide, wherein the Cas12 polypeptide generates a site-specific cleavage in the second gene sequence;

wherein each of the first gene and the second gene is a immunogenic gene, a checkpoint inhibitor gene, or an immune response regulation gene,

thereby generating a modified immune cell with reduced immunogenicity and/or increased anti-neoplasia activity.

186. The method of claim 185, further comprising expressing an exogenous functional chimeric antigen receptor (CAR) or a functional fragment thereof in the immune cell.

187. The method of claim 186, wherein a polynucleotide encoding the CAR or the functional fragment thereof is inserted into the site specific cleavage generated by the Cas12 polypeptide.

188. The method of claim 187, wherein the Cas12 polypeptide is a Cas12b polypeptide.

189. A method for producing a modified immune cell with reduced immunogenicity and/or increased anti-neoplasia activity, the method comprising:

a) modifying a single target nucleobase in a first gene sequence or a regulatory element thereof in an immune cell; and

b) modifying a second gene sequence or a regulatory element thereof in the immune cell by inserting an exogenous functional chimeric antigen receptor (CAR) or a functional fragment thereof or an exogenous functional T cell receptor or a functional fragment thereof in the second gene;

wherein each of the first gene and the second gene is a immunogenic gene, a checkpoint inhibitor gene, or an immune response regulation gene,

thereby generating a modified immune cell with reduced immunogenicity and/or increased anti-neoplasia activity.

190. The method of claim 189, wherein step b) further comprises generating a site-specific cleavage in the second gene sequence with a nucleic acid programmable DNA binding protein (napDNAbp).

191. The method of claim 190, wherein the napDNAbp is a Cas12b.

192. The method of any one of claims 185-191, wherein expression of the first gene is reduced by at least 60% or wherein expression of the second gene is reduced by at least 60% as compared to a control cell of a same type without the modification.

193. The method of any one of claims 138-192, wherein the first gene is selected from the group consisting of CD3 epsilon, CD3 gamma, CD3 delta, CD4, TRAC, TRBC1, TRBC2, PDCD1, CD30, CD33, CD7, CD52, B2M, CD70, CIITA, CD2, and CD5.

194. The method of claim 193, wherein the first gene or the second gene is selected from the group consisting of TRAC, CIITA, CD2, CD5, CD7, and CD52.

195. The method of any one of claims 185-194, wherein the second gene is TRAC.

196. The method of any one of claims 185-195, wherein step a) further comprises modifying a single target nucleobase in two other gene sequences or regulatory elements thereof.

197. The method of any one of claims 185-195, wherein step a) further comprises modifying a single target nucleobase in three other gene sequences or regulatory elements thereof.

198. The method of any one of claims 185-195, wherein step a) further comprises modifying a single target nucleobase in four other gene sequences or regulatory elements thereof.

199. The method of any one of claims 185-195, wherein step a) further comprises modifying a single target nucleobase in five other gene sequences or regulatory elements thereof.

200. The method of any one of claims 185-195, wherein step a) further comprises modifying a single target nucleobase in six other gene sequences or regulatory elements thereof.

201. The method of any one of claims 185-195, wherein step a) further comprises modifying a single target nucleobase in seven other gene sequences or regulatory elements thereof.

202. The method of any one of claims 185-201, wherein the modifying in step a) comprises deaminating the single target nucleobase with a base editor comprising a deaminase and a nucleic acid programmable DNA binding protein (napDNAbp).

203. The method of claim 202, wherein the napDNAbp comprises a Cas9 nickase or nuclease dead Cas9.

204. The method of claim 203, wherein the deaminase is a cytidine deaminase and wherein the modification comprises conversion of a cytidine (C) to a thymine (T).

205. The method of claim 203, wherein the deaminase is an adenosine deaminase and wherein the modification comprises conversion of an adenine (A) to a guanine (G).

206. The method of any one of claims 185-205, wherein the modifying in a) comprises contacting the immune cell with a guide nucleic acid sequence.

207. The method of claim 206, wherein the guide nucleic acid sequence comprises a sequence selected from the group consisting of UUCGUAUCUGUAAAACCAAG, CCUACCUGUCACCAGGACCA, CUCUUACCUGUACCAUAACC, CACCUACCUAAGAACCAUCC, ACUCACGCUGGAUAGCCUCC, ACUCACCCAGCAUCCCCAGC, CACUCACCUUAGCCUGAGCA, and CACGCACCUGGACAGCUGAC.

208. The method of any one of claims 185-207, wherein the modifying in b) comprises contacting the immune cell with a guide nucleic acid sequence.

209. The method of claim 208, wherein the guide nucleic acid sequence comprises a sequence selected from sequences in Table 1.

210. The method of any one of claims 185-209, wherein the modifying in a) comprises replacing the single target nucleobase with a different nucleobase by target-primed reverse transcription with a reverse transcriptase and an extended guide nucleic acid sequence, wherein the extended guide nucleic acid sequence comprises a reverse transcription template sequence, a reverse transcription primer binding site, or a combination thereof.

211. The method of any one of claims 185-210, wherein the modifying in a) and b) generates less than 1% indels in the immune cell.

212. The method of any one of claims 185-211, wherein the modifying in a) and b) generates less than 5% off target modification in the immune cell.

213. The method of any one of claims 185-211, wherein the modifying in a) and b) generate less than 5% non-target modification in the immune cell.

214. The method of any one of claims 185-213, wherein the immune cell is a human cell.

215. The method of claim 214, wherein the immune cell is a cytotoxic T cell, a regulatory T cell, a T helper cell, a dendritic cell, a B cell, or a NK cell.

216. The method of any one of claims 186-215, wherein the CAR specifically binds a marker associated with neoplasia.

217. The method of claim 216, wherein the CAR specifically binds CD7.

218. A modified immune cell with reduced immunogenicity and/or increased anti-neoplasia activity, wherein the modified immune cell comprises:

a) a single target nucleobase modification in a first gene sequence or a regulatory element thereof; and

b) a modification in a second gene sequence or a regulatory element thereof, wherein the modification is a Cas12 polypeptide generated site-specific cleavage;

wherein each of the first gene and the second gene is a immunogenic gene, a checkpoint inhibitor gene, or an immune response regulation gene.

219. The modified immune cell of claim 218, wherein the immune cell further comprises an exogenous functional chimeric antigen receptor (CAR) or a functional fragment thereof.

220. The modified immune cell of claim 219, wherein a polynucleotide encoding the CAR or the functional fragment thereof is inserted into the site specific cleavage generated by the Cas12 polypeptide.

221. A modified immune cell with reduced immunogenicity and/or increased anti-neoplasia activity, the modified immune cell comprising:

a) a single target nucleobase modification in a first gene sequence or a regulatory element thereof in an immune cell; and

b) a modification in a second gene sequence or a regulatory element thereof, wherein the modification is an insertion of an exogenous chimeric antigen receptor (CAR) or a functional fragment thereof or an exogenous T cell receptor or a functional fragment thereof;

wherein each of the first gene and the second gene is a immunogenic gene, a checkpoint inhibitor gene, or immune response regulation gene.

222. The modified immune cell of claim 221, wherein the modification in b) is generated by a site-specific cleavage with a Cas12b.

223. The modified immune cell of any one of claims 218-222, wherein expression of the first gene is reduced by at least 60% or wherein expression of the second gene is reduced by at least 60% as compared to a control cell of a same type without the modification.

224. The modified immune cell of any one of claims 218-223, wherein the first gene or the second gene is selected from the group consisting of CD3 epsilon, CD3 gamma, CD3 delta, CD4, TRAC, TRBC1, TRBC2, PDCD1, CD30, CD33, CD7, CD52, B2M, CD70, CIITA, CD2, and CD5.

225. The method of claim 193, wherein the first gene or the second gene is selected from the group consisting of TRAC, CD2, CD5, CD7, and CD52.

226. The modified immune cell of claim 225, wherein the second gene is TRAC.

227. The modified immune cell of any one of claims 218-226, wherein the immune cell further comprises modification in a single target nucleobase in two other gene sequences or regulatory elements thereof.

228. The modified immune cell of any one of claims 218-226, wherein the immune cell further comprises modification in a single target nucleobase in three other gene sequences or regulatory elements thereof.

229. The modified immune cell of any one of claims 218-226, wherein the immune cell further comprises modification in a single target nucleobase in four other gene sequences or regulatory elements thereof.

230. The modified immune cell of any one of claims 218-226, wherein the immune cell further comprises modification in a single target nucleobase in five other gene sequences or regulatory elements thereof.

231. The modified immune cell of any one of claims 218-226, wherein the immune cell further comprises modification in a single target nucleobase in six other gene sequences or regulatory elements thereof.

232. The modified immune cell of any one of claims 218-226, wherein the immune cell further comprises modification in a single target nucleobase in seven other gene sequences or regulatory elements thereof.

233. The modified immune cell of claim any one of claims 218-232, wherein the modification in a) is generated by a base editor comprising a deaminase and a nucleic acid programmable DNA binding protein (napDNAbp).

234. The modified immune cell of claim 233, wherein the deaminase is a cytidine deaminase and wherein the modification comprises conversion of a cytidine (C) to a thymine (T).

235. The modified immune cell of claim 233, wherein the deaminase is an adenosine deaminase and wherein the modification comprises conversion of an adenine (A) to a guanine (G).

236. The modified immune cell of any one of claims 218-235, wherein the immune cell comprises less than 1% indels in the genome.

237. The modified immune cell of any one of claims 218-236, wherein the immune cell is a human cell.

238. The modified immune cell of any one of claims 218-237, wherein the immune cell is a cytotoxic T cell, a regulatory T cell, a T helper cell, a dendritic cell, a B cell, or a NK cell.

239. The modified immune cell of any one of claims 218-238, wherein the CAR specifically binds a marker associated with neoplasia.

240. The modified immune cell of claim 239, wherein the CAR specifically binds CD7.

241. The modified immune cell of any one of claims 218-240, wherein the modification in b) is an insertion in exon 1 in the TRAC gene sequence.

242. A population of modified immune cells, wherein a plurality of the population of immune cells comprises:

a) a single target nucleobase modification in a first gene sequence or a regulatory element thereof in an immune cell; and

b) a modification in a second gene sequence or a regulatory element thereof, wherein the modification is a Cas12 polypeptide generated site-specific cleavage;

wherein each of the first gene and the second gene is a immunogenic gene, a checkpoint inhibitor gene, or an immune response regulation gene,

and wherein the plurality of the population comprises an exogenous chimeric antigen receptor (CAR) or a functional fragment thereof.

243. The population of modified immune cells of claim 242, wherein a polynucleotide encoding the CAR or the functional fragment thereof is inserted into the site specific cleavage generated by the Cas12 polypeptide.

244. A population of modified immune cells, wherein a plurality of the population of immune cells comprises:

a) a single target nucleobase modification in a first gene sequence or a regulatory element thereof; and

b) a modification in a second gene sequence or a regulatory sequence thereof, wherein the modification is an insertion of an exogenous chimeric antigen receptor (CAR) or a functional fragment thereof or an exogenous T cell receptor or a functional fragment thereof;

wherein each of the first gene and the second gene is a immunogenic gene, a checkpoint inhibitor gene, or immune response regulation gene, and wherein the plurality of cells with the modification in a) or b) exhibit reduced immunogenicity and/or increased anti-neoplasia activity.

245. The population of modified immune cells of claim 244, wherein the modification in b) is generated by a site-specific cleavage with a Cas12b.

246. The population of modified immune cells of any one of claims 242-245, wherein expression of the first gene is reduced by at least 60% or wherein expression of the second gene is reduced by at least 60% in the plurality of cells with the modification in a) or b) as compared to plurality of control cells of a same type without the modification.

247. The population of modified immune cells of any one of claims 242-246, wherein the first gene or the second gene is selected from the group consisting of CD3 epsilon, CD3 gamma, CD3 delta, CD4, TRAC, TRBC1, TRBC2, PDCD1, CD30, CD33, CD7, CD52, B2M, CD70, CIITA, CD2, and CD5.

248. The population of modified immune cells of claim 247, wherein the first gene or the second gene is selected from the group consisting of TRAC, CIITA, CD2, CD5, , CD7, and CD52.

249. The population of modified immune cells of claim 248, wherein the first gene is TRAC, , CD7, or CD52.

250. The population of modified immune cells of claim 248, wherein the second gene is TRAC.

251. The population of modified immune cells of any one of claims 242-250, wherein the plurality of cells with the modification in a) or b) further comprises a modification in a single target nucleobase in two other gene sequences or regulatory elements thereof.

252. The population of modified immune cells of claim 251, wherein the plurality of cells with the modification in a) or b) further comprises a single target nucleobase in three, four, five, or six other gene sequences or regulatory elements thereof.

253. The population of modified immune cells of any one of claims 242-252, wherein the modification in a) is generated by a base editor comprising a deaminase and a nucleic acid programmable DNA binding protein (napDNAbp) to form a base editor.

254. The population of modified immune cells of claim 253, wherein the deaminase is a cytidine deaminase and wherein the modification comprises conversion of a cytidine (C) to a thymine (T).

255. The population of modified immune cells of claim 253, wherein the deaminase is an adenosine deaminase and wherein the modification comprises conversion of an adenine (A) to a guanine (G).

256. The population of modified immune cells of claim 254, wherein the base editor further comprises a uracil glycosylase inhibitor.

257. The population of modified immune cells of any one of claims 242-256, wherein at least 60% of the population of immune cells are viable.

258. The population of modified immune cells of any one of claims 242-256, wherein at least 60% of the population of immune cells expand at least 80% of expansion rate of a population of control cells of a same type without the modification.

259. The population of modified immune cells of any one of claims 242-258, wherein the immune cells are a human cells.

260. The population of modified immune cells of any one of claims 242-259, wherein the immune cells is are cytotoxic T cells, regulatory T cells, T helper cells, dendritic cells, B cells, or NK cells.

261. The population of modified immune cells of any one of claims 242-260, wherein the CAR specifically binds a marker associated with neoplasia.

262. The population of modified immune cells of claim 261, wherein the CAR specifically binds CD7.

263. The population of modified immune cells of any one of claims 242-262, wherein the modification in b) is an insertion in exon 1 in the TRAC gene sequence.

264. A method for producing a modified immune cell with increased anti-neoplasia activity, the method comprising: modifying a single target nucleobase in a Cbl Proto

Oncogene B (CBLB) gene sequence or a regulatory element thereof in an immune cell, wherein the modification reduces an activation threshold of the immune cell compared with an immune cell lacking the modification; thereby generating a modified immune cell with increased anti-neoplasia activity.

265. A composition comprising a modified immune cell with increased anti-neoplasia activity, wherein the modified immune cell comprises: a modification in a single target nucleobase in a Cbl Proto-Oncogene B (CBLB) gene sequence or a regulatory element thereof, wherein the modified immune cell exhibits a reduced activation threshold compared with a control immune cell of a same type without the modification.

266. A population of immune cells, wherein a plurality of the population of immune cells comprises: a modification in a single target nucleobase in a CBLB gene sequence or a regulatory element thereof, wherein the plurality of the population of the immune cells comprising the modification exhibit a reduced activation threshold compared with an control population of immune cells of a same type without the modification.

267. A method for producing a population of modified immune cells with increased anti-neoplasia activity, the method comprising: modifying a single target nucleobase in a Cbl Proto Oncogene B (CBLB) gene sequence or a regulatory element thereof in a population of immune cells, wherein at least 50% of the population of immune cells are modified to comprise the single target nucleobase modification.

268. A composition comprising at least four different guide nucleic acid sequences for base editing.

269. The composition of claim 268, further comprising a polynucleotide encoding a base editor polypeptide, wherein the base editor polypeptide comprises a nucleic acid programmable DNA binding protein (napDNAbp) and a deaminase.

270. The composition of claim 269, wherein the polynucleotide encoding the base editor is a mRNA sequence.

271. The composition of claim 269 or 270, wherein the deaminase is a cytidine deaminase or an adenosine deaminase.

272. The composition of claim 268, further comprising a base editor polypeptide, wherein the base editor polypeptide comprises a nucleic acid programmable DNA binding protein (napDNAbp) and a deaminase.

273. The composition of claim 272, wherein the deaminase is a cytidine deaminase or an adenosine deaminase.

274. The composition of claim 272 or 273, further comprising a lipid nanoparticle.

275. The composition of any one of claims 267-274, wherein the at least four guide nucleic acid sequences each hybridize with a gene sequence selected from the group consisting of CD2, CD3 epsilon, CD3 gamma, CD3 delta, CD4, CD5, CD7, CD30, CD33, CD52, CD70, and CIITA. In some embodiments, the at least 1, 2, 3, 4, 5, 6, 7, 8, or more genes or regulatory elements thereof are selected from CD2, CD3 epsilon, CD3 gamma, CD3 delta, CD4, CD5, CD7, CD30, CD33, CD52, CD70, and CIITA. In some embodiments, the at least 1, 2, 3, 4, 5, 6, 7, 8, or more genes or regulatory elements thereof comprise one or more genes selected from CD2, CD3 epsilon, CD3 gamma, CD3 delta, CD4, CD5, CD7, CD30, CD33, CD52, CD70, and CIITA. In some embodiments, the at least 1, 2, 3, 4, 5, 6, 7, 8, or more genes or regulatory elements thereof are selected from ACAT1, ACLY, ADORA2A, AXL, B2M , BATF, BCL2L11, BTLA, CAMK2D, cAMP, CASP8, Cblb, CCR5, CD2, CD3D, CD3E, CD3G, CD4, CD5, CD7, CD8A, CD33, CD38, CD52, CD70, CD82, CD86, CD96, CD123, CD160, CD244, CD276, CDK8, CDKN1B, Chi3l1, CIITA, CISH, CSF2CSK, CTLA-4, CUL3, Cyp11a1, DCK, DGKA, DGKZ, DHX37, ELOB(TCEB2), ENTPD1 (CD39), FADD, FAS, GATA3, IL6, IL6R, IL10, IL10RA, IRF4, IRF8, JUNB, Lag3, , LAIR-1 (CD305), LDHA, LIF, LYN, MAP4K4, MAPK14, MCJ, MEF2D, MGAT5, NR4A1, NR4A2, NR4A3, NT5E (CD73), ODC1, OTULINL (FAM105A), PAG1, PDCD1, PDIA3, PHD1 (EGLN2), PHD2 (EGLN1), PHD3 (EGLN3), PIK3CD, PIKFYVE, PPARa, PPARd, PRDMI1, PRKACA, PTEN, PTPN2, PTPN6, PTPN11, PVRIG (CD112R), RASA2, RFXANK, SELPG/PSGL1, SIGLEC15, SLA, SLAMF7, SOCS1, Spry1, Spry2, STK4, SUV39, H1TET2, TGFbRII, TIGIT, Tim-3, TMEM222, TNFAIP3, TNFRSF8 (CD30), TNFRSF10B, TOX, TOX2, , TRAC, TRBC1, TRBC2, UBASH3A, VHL, VISTA, XBP1, YAP1, and ZC3H12A.

276. The composition of any one of claims 267-274, wherein the at least four guide nucleic acid sequences each hybridize with a gene sequence selected from the group consisting of CD3epsilon, CD3 delta, CD3 gamma, TRAC, TRBC1, and TRBC2, CD2, CD5, CD7, CD52, CD70, and CIITA.

277. The composition of any one of claims 267-274, wherein the at least four guide nucleic acid sequences comprise a sequence selected from the group consisting of UUCGUAUCUGUAAAACCAAG, CCUACCUGUCACCAGGACCA,

CUCUUACCUGUACCAUAACC, CACCUACCUAAGAACCAUCC, ACUCACGCUGGAUAGCCUCC, ACUCACCCAGCAUCCCCAGC, CACUCACCUUAGCCUGAGCA, and CACGCACCUGGACAGCUGAC.

278. An immune cell comprising the composition of any one of claims 267-277, wherein the composition is introduced into the immune cell with electroporation.

279. An immune cell comprising the composition of any one of claims 267-277, wherein the composition is introduced into the immune cell with electroporation, nucleofection, viral transduction, or a combination thereof.

280. The modified immune cell of any one of claims 86 and 88-134 having increased growth or viability compared to a reference cell.

281. The modified immune cell of claim 280, wherein the reference cell is an immune cell modified with a Cas9 nuclease.

282. The population of modified immune cells of claim 87 and 135-179 having increased yield of modified immune cells compared to a reference population of cells.

283. The population of modified immune cells of claim 282, wherein the reference population is a population of immune cells modified with a Cas9 nuclease.