

**TITLE****ALBUMIN-TARGETED ENDONUCLEASES, COMPOSITIONS, AND METHODS  
OF USE****CROSS-REFERENCE TO RELATED APPLICATIONS**

- 5           This application claims the benefit under 35 U.S.C. § 119(e) of U.S. Provisional Application No. 63/621,235, filed January 16, 2024, which is incorporated by reference herein in its entirety.

**REFERENCE TO SEQUENCE LISTING SUBMITTED ELECTRONICALLY**

- 10           This application contains a sequence listing, which is submitted electronically as a XML formatted sequence listing with a file name “240024WO01 Sequence Listing”, creation date of January 3, 2025, and having a size of 155,183 bytes. The sequence listing submitted electronically is part of the specification and is herein incorporated by reference in its entirety

15   **BACKGROUND**

**Technical Field**

The present disclosure relates to improved genome editing compositions. More particularly, the disclosure relates to engineered nucleases, compositions, and methods of using the same for editing the human Albumin (ALB) gene.

20           **Description of the Related Art**

The relatively recent surge in genome editing technologies has opened the possibility of directly targeting and modifying genomic sequences in almost any cell type including in eukaryotic cells *in vivo*. Such technologies include, but are not limited to, transcription

activator-like effector nucleases (TALENs), zinc-finger nucleases (ZFNs), clustered regularly interspaced short palindromic repeat (CRISPR)-Cas-associated nucleases, and homing endonucleases (HEs). Common to all of these editing techniques is that they create a breakpoint in the target nucleotide sequence, while the natural cellular repair mechanisms are left to re-ligate the nucleotide sequence either by non-homologous end-joining (NHEJ) or homology-directed repair (HDR). Moreover, one can determine genomic editing outcomes by providing to the edited cell recombinant polynucleotides, such as therapeutic transgene (*e.g.*, DNA encoding a therapeutic polypeptide/protein, factor, or antibody), for use in repair of the breakpoint. These advancements have raised the possibility of treating many types of diseases (genetic or otherwise), including, but not limited to blood disorders such as hemophilia and hemoglobinopathies, metabolic disorders, lysosomal storage diseases, and other mono or polygenic disorders. In addition, novel therapeutic approaches can be applied through the directed expression of therapeutic proteins such as antibodies or synthetic proteins yet to be devised.

However, while many of these gene editing technologies are highly specific to their target site, minimizing off-target editing by these gene-editing technologies while maintaining high editing efficiency remains a challenge. Minimizing off-target editing is of critical importance when developing a gene-editing therapeutic because such edits can lead to unintentional or even adverse permanent effects on the cell's genome. Naeem *et al.*, *Cells*. 2020 Jul; 9(7): 1608; and Guo *et al.*, *Front Bioeng Biotechnol*. 2023; 11: 1143157. Indeed, possible off-target editing was a topic of much discussion in the lead up to the very first CRISPR-based therapy approved by the FDA for sickle cell disease.

Therefore, there remains a need for highly specific and efficacious gene-editing enzymes for the treatment of disease.

## **BRIEF SUMMARY**

The present disclosure generally relates, in part, to compositions and kits comprising engineered homing endonucleases and megaTALs that cleave a target site in the human

Albumin (ALB) gene and methods of using the same, including gene editing and related methods of treatment.

In various embodiments, the present disclosure contemplates, in part, a polypeptide comprising an engineered homing endonuclease (HE) that cleaves a double strand DNA (dsDNA) target site in the human ALB gene.

In particular embodiments, the target site is within intron 1 of the ALB gene.

In certain embodiments, the target site comprises a nucleic acid sequence as set forth in any one of SEQ ID NOs: 32-34.

In particular embodiments, the target site comprises a nucleic acid sequence as set forth in SEQ ID NO: 34.

In particular embodiments, the engineered HE is an LAGLIDADG homing endonuclease (LHE) variant.

In certain embodiments, the engineered HE lacks the 1, 2, 3, 4, 5, 6, 7, or 8 N-terminal amino acids compared to a corresponding wild type or stabilized HE as set forth in SEQ ID NO: 1 or SEQ ID NO: 2.

In some embodiments, the engineered HE lacks the 4 N-terminal amino acids compared to a corresponding wild type or stabilized HE as set forth in SEQ ID NO: 1 or SEQ ID NO: 2.

In certain embodiments, the engineered HE lacks the 8 N-terminal amino acids compared to a corresponding wild type or stabilized HE as set forth in SEQ ID NO: 1 or SEQ ID NO: 2.

In particular embodiments, the engineered HE lacks the 1, 2, 3, 4, 5, or 6 C-terminal amino acids compared to a corresponding wild type or stabilized HE as set forth in SEQ ID NO: 1 or SEQ ID NO: 2.

In particular embodiments, the engineered HE lacks the C-terminal amino acid compared to a corresponding wild type or stabilized HE as set forth in SEQ ID NO: 1 or SEQ ID NO: 2.

In some embodiments, the engineered HE lacks the 2 C-terminal amino acids compared to a corresponding wild type or stabilized HE as set forth in SEQ ID NO: 1 or SEQ ID NO: 2.

In further embodiments, the engineered HE is a variant of an LHE selected from the group consisting of: I-OnuI, I-AabMI, I-AaeMI, I-AniI, I-ApaMI, I-CapIII, I-CapIV, I-CkaMI, I-CpaMI, I-CpaMII, I-CpaMIII, I-CpaMIV, I-CpaMV, I-CpaV, I-CraMI, I-EjeMI, I-GpeMI, I-GpiI, I-GzeMI, I-GzeMII, I-GzeMIII, I-HjeMI, I-LtrII, I-LtrI, I-LtrWI, I-MpeMI, I-MveMI, I-NcrII, I-NcrI, I-NcrMI, I-OheMI, I-OsoMI, I-OsoMII, I-OsoMIII, I-OsoMIV, I-PanMI, I-PanMII, I-PanMIII, I-PnoMI, I-ScuMI, I-SmaMI, I-SscMI, and I-Vdi141I.

In particular embodiments, the engineered HE is a variant of an LHE selected from the group consisting of: I-OnuI, I-CpaMI, I-HjeMI, I-PanMI, and SmaMI.

In particular embodiments, the engineered HE is an I-OnuI LHE variant.

In additional embodiments, the engineered HE comprises one or more amino acid substitutions in the DNA recognition interface at corresponding amino acid positions selected from the group consisting of: 24, 26, 28, 30, 31, 32, 34, 35, 36, 37, 38, 40, 42, 44, 46, 48, 68, 70, 71, 75, 76, 78, 80, 180, 182, 184, 186, 189, 190, 191, 192, 193, 197, 199, 201, 203, 223, 225, 229, 232, 234, 236, and 238 of an I-OnuI LHE amino acid sequence as set forth in any one of SEQ ID NOs: 1-3 and 5, or a biologically active fragment thereof.

In additional embodiments, the engineered HE comprises one or more amino acid substitutions in the DNA recognition interface at corresponding amino acid positions selected from the group consisting of: 20, 22, 24, 26, 27, 28, 30, 31, 32, 33, 34, 36, 38, 40, 42, 44, 64, 66, 67, 71, 72, 74, 76, 176, 178, 180, 182, 185, 186, 187, 188, 189, 193, 195, 197, 199, 219, 221, 225, 228, 230, 232, and 234 of an I-OnuI LHE amino acid sequence as set forth in SEQ ID NO: 4, or a biologically active fragment thereof.

In particular embodiments, the engineered HE comprises at least 5, at least 15, preferably at least 25, more preferably at least 35, or even more preferably at least 40, or 43 of the corresponding following amino acid substitutions: S24C, L26S, R28A, R28H, R28T, R28V, R30A, R30G, R30Q, N31K, N32A, N32F, N32L, N32G, N32C, N32M, K34G, S35W, S35R, S35G, S35K, S35C, S36T, V37T, G38R, S40N, S40T, S40S, S40G, S40V, S40L, E42K, G44V, G44M, Q46S, T48G, V68T, A70S, N71K, N75T, A76Q, S78R, K80T, C180S, F182M, N184I, I186A, K189R, S190T, K191G, L192A, G193R, Q197R, V199C, S201E, T203G, Y223H, K225S, K229L, F232S, W234F, D236I, and V238R of any one of SEQ ID NOs: 1-3 and 5, or a biologically active fragment thereof.



In particular embodiments, the engineered HE comprises at least 5, at least 15, preferably at least 25, more preferably at least 35, or even more preferably at least 40, or 43 of the corresponding following amino acid substitutions: S20C, L22S, R24A, R24H, R24T, R24V, R26A, R26G, R26Q, N27K, N28A, N28F, N28L, N28G, N28C, N28M, K30G, S31W, S31R, S31G, S31K, S31C, S32T, V33T, G34R, S36N, S36T, S36S, S36G, S36V, S36L, E38K, G40V, G40M, Q42S, T44G, V64T, A66S, N67K, N71T, A72Q, S74R, K76T, C176S, F178M, N180I, I182A, K185R, S186T, K187G, L188A, G189R, Q193R, V195C, S197E, T199G, Y219H, K221S, K225L, F228S, W230F, D232I, and V234R SEQ ID NO: 4, or a biologically active fragment thereof.

10 In certain embodiments, the engineered HE comprises the corresponding following amino acid substitutions: S24C, L26S, R28V, R30Q, N31K, N32A, K34G, S35W, S36T, V37T, G38R, S40N, E42K, G44V, Q46S, T48G, V68T, A70S, N71K, N75T, A76Q, S78R, K80T, C180S, F182M, N184I, I186A, K189R, S190T, K191G, L192A, G193R, Q197R, V199C, S201E, T203G, Y223H, K225S, K229L, F232S, W234F, D236I, and V238R of any one of SEQ ID NOs: 1-3 and 5, or a biologically active fragment thereof.

In certain embodiments, the engineered HE comprises the corresponding following amino acid substitutions: S20C, L22S, R24V, R26Q, N27K, N28A, K30G, S31W, S32T, V33T, G34R, S36N, E38K, G40V, Q42S, T44G, V64T, A66S, N67K, N71T, A72Q, S74R, K76T, C176S, F178M, N180I, I182A, K185R, S186T, K187G, L188A, G189R, Q193R, V195C, S197E, T199G, Y219H, K221S, K225L, F228S, W230F, D232I, and V234R of SEQ ID NO: 4, or a biologically active fragment thereof.

In particular embodiments, the engineered HE comprises the corresponding following amino acid substitutions: S24C, L26S, R28T, N31K, N32F, K34G, S35R, S36T, V37T, G38R, S40T, E42K, G44M, Q46S, T48G, V68T, A70S, N71K, N75T, A76Q, S78R, K80T, C180S, F182M, N184I, I186A, K189R, S190T, K191G, L192A, G193R, Q197R, V199C, S201E, T203G, Y223H, K225S, K229L, F232S, W234F, D236I, and V238R of any one of SEQ ID NOs: 1-3 and 5, or a biologically active fragment thereof.

In particular embodiments, the engineered HE comprises the corresponding following amino acid substitutions: S20C, L22S, R24T, N27K, N28F, K30G, S31R, S32T, V33T, G34R, S36T, E38K, G40M, Q42S, T44G, V64T, A66S, N67K, N71T, A72Q, S74R, K76T,

C176S, F178M, N180I, I182A, K185R, S186T, K187G, L188A, G189R, Q193R, V195C, S197E, T199G, Y219H, K221S, K225L, F228S, W230F, D232I, and V234R of SEQ ID NO: 4 or a biologically active fragment thereof.

In certain embodiments, the engineered HE comprises the corresponding following amino acid substitutions: S24C, L26S, R28A, R30G, N31K, N32L, K34G, S35G, S36T, V37T, G38R, E42K, G44V, Q46S, T48G, V68T, A70S, N71K, N75T, A76Q, S78R, K80T, C180S, F182M, N184I, I186A, K189R, S190T, K191G, L192A, G193R, Q197R, V199C, S201E, T203G, Y223H, K225S, K229L, F232S, W234F, D236I, and V238R of any one of SEQ ID NOs: 1-3 and 5, or a biologically active fragment thereof.

In certain embodiments, the engineered HE comprises the corresponding following amino acid substitutions: S20C, L22S, R24A, R26G, N27K, N28L, K30G, S31G, S32T, V33T, G34R, E38K, G40V, Q42S, T44G, V64T, A66S, N67K, N71T, A72Q, S74R, K76T, C176S, F178M, N180I, I182A, K185R, S186T, K187G, L188A, G189R, Q193R, V195C, S197E, T199G, Y219H, K221S, K225L, F228S, W230F, D232I, and V234R of SEQ ID NO: 4, or a biologically active fragment thereof.

In particular embodiments, the engineered HE comprises the corresponding following amino acid substitutions: S24C, L26S, R28H, R30A, N31K, N32F, K34G, S35K, S36T, V37T, G38R, S40G, E42K, G44V, Q46S, T48G, V68T, A70S, N71K, N75T, A76Q, S78R, K80T, C180S, F182M, N184I, I186A, K189R, S190T, K191G, L192A, G193R, Q197R, V199C, S201E, T203G, Y223H, K225S, K229L, F232S, W234F, D236I, and V238R of any one of SEQ ID NOs: 1-3 and 5, or a biologically active fragment thereof.

In particular embodiments, the engineered HE comprises the corresponding following amino acid substitutions: S20C, L22S, R24H, R26A, N27K, N28F, K30G, S31K, S32T, V33T, G34R, S36G, E38K, G40V, Q42S, T44G, V64T, A66S, N67K, N71T, A72Q, S74R, K76T, C176S, F178M, N180I, I182A, K185R, S186T, K187G, L188A, G189R, Q193R, V195C, S197E, T199G, Y219H, K221S, K225L, F228S, W230F, D232I, and V234R of SEQ ID NO: 4, or a biologically active fragment thereof.

In certain embodiments, the engineered HE comprises the corresponding following amino acid substitutions: S24C, L26S, R28V, R30A, N31K, N32G, K34G, S35G, S36T, V37T, G38R, S40T, E42K, G44V, Q46S, T48G, V68T, A70S, N71K, N75T, A76Q, S78R,

K80T, C180S, F182M, N184I, I186A, K189R, S190T, K191G, L192A, G193R, Q197R, V199C, S201E, T203G, Y223H, K225S, K229L, F232S, W234F, D236I, and V238R of any one of SEQ ID NOs: 1-3 and 5, or a biologically active fragment thereof.

In certain embodiments, the engineered HE comprises the corresponding following amino acid substitutions: S20C, L22S, R24V, R26A, N27K, N28G, K30G, S31G, S32T, V33T, G34R, S36T, E38K, G40V, Q42S, T44G, V64T, A66S, N67K, N71T, A72Q, S74R, K76T, C176S, F178M, N180I, I182A, K185R, S186T, K187G, L188A, G189R, Q193R, V195C, S197E, T199G, Y219H, K221S, K225L, F228S, W230F, D232I, and V234R of SEQ ID NO: 4, or a biologically active fragment thereof.

In some embodiments, the engineered HE comprises the corresponding following amino acid substitutions: S24C, L26S, R28A, R30G, N31K, N32C, K34G, S35C, S36T, V37T, G38R, E42K, G44V, Q46S, T48G, V68T, A70S, N71K, N75T, A76Q, S78R, K80T, C180S, F182M, N184I, I186A, K189R, S190T, K191G, L192A, G193R, Q197R, V199C, S201E, T203G, Y223H, K225S, K229L, F232S, W234F, D236I, and V238R of any one of SEQ ID NOs: 1-3 and 5, or a biologically active fragment thereof.

In some embodiments, the engineered HE comprises the corresponding following amino acid substitutions: S20C, L22S, R24A, R26G, N27K, N28C, K30G, S31C, S32T, V33T, G34R, E38K, G40V, Q42S, T44G, V64T, A66S, N67K, N71T, A72Q, S74R, K76T, C176S, F178M, N180I, I182A, K185R, S186T, K187G, L188A, G189R, Q193R, V195C, S197E, T199G, Y219H, K221S, K225L, F228S, W230F, D232I, and V234R of SEQ ID NO: 4, or a biologically active fragment thereof.

In certain embodiments, the engineered HE comprises the corresponding following amino acid substitutions: S24C, L26S, R28A, R30G, N31K, N32L, K34G, S35A, S36T, V37T, G38R, S40V, E42K, G44V, Q46S, T48G, V68T, A70S, N71K, N75T, A76Q, S78R, K80T, C180S, F182M, N184I, I186A, K189R, S190T, K191G, L192A, G193R, Q197R, V199C, S201E, T203G, Y223H, K225S, K229L, F232S, W234F, D236I, and V238R of any one of SEQ ID NOs: 1-3 and 5, or a biologically active fragment thereof.

In certain embodiments, the engineered HE comprises the corresponding following amino acid substitutions: S20C, L22S, R24A, R26G, N27K, N28L, K30G, S31A, S32T, V33T, G34R, S36V, E38K, G40V, Q42S, T44G, V64T, A66S, N67K, N71T, A72Q, S74R,

K76T, C176S, F178M, N180I, I182A, K185R, S186T, K187G, L188A, G189R, Q193R, V195C, S197E, T199G, Y219H, K221S, K225L, F228S, W230F, D232I, and V234R of SEQ ID NO: 4, or a biologically active fragment thereof.

In particular embodiments, the engineered HE comprises the following corresponding amino acid substitutions: S24C, L26S, R28H, R30G, N31K, N32M, K34G, S35R, S36T, V37T, G38R, S40L, E42K, G44V, Q46S, T48G, V68T, A70S, N71K, N75T, A76Q, S78R, K80T, C180S, F182M, N184I, I186A, K189R, S190T, K191G, L192A, G193R, Q197R, V199C, S201E, T203G, Y223H, K225S, K229L, F232S, W234F, D236I, and V238R of any one of SEQ ID NOs: 1-3 and 5, or a biologically active fragment thereof.

In particular embodiments, the engineered HE comprises the following corresponding amino acid substitutions: S20C, L22S, R24H, R26G, N27K, N28M, K30G, S31R, S32T, V33T, G34R, S36L, E38K, G40V, Q42S, T44G, V64T, A66S, N67K, N71T, A72Q, S74R, K76T, C176S, F178M, N180I, I182A, K185R, S186T, K187G, L188A, G189R, Q193R, V195C, S197E, T199G, Y219H, K224S, K225L, F228S, W230F, D232I, and V234R of any one of SEQ ID NO: 4, or a biologically active fragment thereof.

In additional embodiments, the engineered HE further comprises the following corresponding amino acid substitutions C115S, E121G, I125E, L138M, I153D, K156R, S159P, L160I, F168G, E178D, K207R, N246K, V261M, and L263H of SEQ ID NO: 1, or a biologically active fragment thereof.

In certain embodiments, the engineered HE further comprises: (a) a first DNA recognition interface comprising amino acid residues 20-46 of any one of SEQ ID NOs: 6-13, (b) a second DNA recognition interface comprising amino acid residues 64-78 of any one of SEQ ID NOs: 6-13, (c) a third DNA recognition interface comprising amino acid residues 176-199 of any one of SEQ ID NOs: 6-13, and (d) a fourth DNA recognition interface comprising amino acid residues 219-236 of any one of SEQ ID NOs: 6-13.

In particular embodiments, the engineered HE comprises an amino acid sequence that is at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to the amino acid sequence set forth in any one of SEQ ID NOs: 6-13, or a biologically active fragment thereof.

In particular embodiments, the engineered HE comprises the amino acid sequence set forth in SEQ ID NO: 6, or a biologically active fragment thereof.

In particular embodiments, the engineered HE comprises the amino acid sequence set forth in SEQ ID NO: 7, or a biologically active fragment thereof.

5 In particular embodiments, the engineered HE comprises the amino acid sequence set forth in SEQ ID NO: 8, or a biologically active fragment thereof.

In particular embodiments, the engineered HE comprises the amino acid sequence set forth in SEQ ID NO: 9, or a biologically active fragment thereof.

In particular embodiments, the engineered HE comprises the amino acid sequence set forth in SEQ ID NO: 10, or a biologically active fragment thereof.

In particular embodiments, the engineered HE comprises the amino acid sequence set forth in SEQ ID NO: 11, or a biologically active fragment thereof.

In particular embodiments, the engineered HE comprises the amino acid sequence set forth in SEQ ID NO: 12, or a biologically active fragment thereof.

15 In particular embodiments, the engineered HE comprises the amino acid sequence set forth in SEQ ID NO: 13, or a biologically active fragment thereof.

In further embodiments, the polypeptide comprises a nuclear localization signal (NLS).

In particular embodiments, the polypeptide comprises one or more NLS located N-terminal of the polypeptide (*e.g.*, N-terminal to the engineered HE or the megaTAL).

20 In further embodiments, the polypeptide further comprises a DNA binding domain. In certain embodiments, the DNA binding domain is selected from the group consisting of: a TALE DNA binding domain and a zinc finger DNA binding domain.

In certain embodiments, the TALE DNA binding domain comprises about 8.5 TALE repeat units to about 15.5 TALE repeat units.

25 In particular embodiments, the TALE DNA binding domain binds the polynucleotide sequence in the ALB gene.

In particular embodiments, the TALE DNA binding domain binds the polynucleotide sequence set forth in SEQ ID NO: 38.

In certain embodiments, the polypeptide binds and cleaves the polynucleotide sequence set forth in SEQ ID NO: 43.

In certain embodiments, the zinc finger DNA binding domain comprises 2, 3, 4, 5, 6, 7, or 8 zinc finger motifs.

In particular embodiments, the polypeptide comprises an amino acid sequence that is at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to the amino acid  
5 sequence set forth in any one of SEQ ID NOs: 14-21, or a biologically active fragment thereof.

In particular embodiments, the polypeptide comprises the amino acid sequence set forth in SEQ ID NO: 14, or a biologically active fragment thereof.

In particular embodiments, the polypeptide comprises the amino acid sequence set forth in SEQ ID NO: 15, or a biologically active fragment thereof.

10 In particular embodiments, the polypeptide comprises the amino acid sequence set forth in SEQ ID NO: 16, or a biologically active fragment thereof.

In certain embodiments, the polypeptide comprises the amino acid sequence set forth in SEQ ID NO: 17, or a biologically active fragment thereof.

In particular embodiments, the polypeptide comprises the amino acid sequence set forth in SEQ ID NO: 18, or a biologically active fragment thereof.

In particular embodiments, the polypeptide comprises the amino acid sequence set forth in SEQ ID NO: 19, or a biologically active fragment thereof.

In particular embodiments, the polypeptide comprises the amino acid sequence set forth in SEQ ID NO: 20, or a biologically active fragment thereof.

20 In particular embodiments, the polypeptide comprises the amino acid sequence set forth in SEQ ID NO: 21, or a biologically active fragment thereof.

In further embodiments, the polypeptide comprises a peptide linker and an end-processing enzyme or biologically active fragment thereof.

In further embodiments, the polypeptide comprises a viral self-cleaving 2A peptide and  
25 an end-processing enzyme or biologically active fragment thereof.

In certain embodiments, the polypeptide comprises Trex2 or a biologically active fragment thereof.

In various embodiments, the present disclosure contemplates, a polynucleotide encoding the polypeptide described herein.

In various embodiments, the present disclosure contemplates, a mRNA encoding the polypeptide described herein.

In certain embodiments, the mRNA comprises a nucleic acid sequence that is at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to the nucleic acid  
5 sequence set forth in any one of SEQ ID NOs: 22-29.

In particular embodiments, the mRNA comprises a nucleic acid sequence set forth in SEQ ID NO: 22.

In particular embodiments, the mRNA comprises a nucleic acid sequence set forth in SEQ ID NO: 108.

10 In particular embodiments, the mRNA comprises a nucleic acid sequence set forth in SEQ ID NO: 23.

In particular embodiments, the mRNA comprises a nucleic acid sequence set forth in SEQ ID NO: 24.

15 In particular embodiments, the mRNA comprises a nucleic acid sequence set forth in SEQ ID NO: 25.

In particular embodiments, the mRNA comprises a nucleic acid sequence set forth in SEQ ID NO: 26.

In particular embodiments, the mRNA comprises a nucleic acid sequence set forth in SEQ ID NO: 27.

20 In particular embodiments, the mRNA comprises a nucleic acid sequence set forth in SEQ ID NO: 28.

In particular embodiments, the mRNA comprises a nucleic acid sequence set forth in SEQ ID NO: 29.

25 In various embodiments, the present disclosure contemplates, a cDNA encoding the polypeptide described herein.

In various embodiments, the present disclosure contemplates, a vector comprising a polynucleotide encoding the polypeptide described herein.

In certain embodiments, the present disclosure contemplates, a cell comprising the polypeptide described herein.

In various embodiments, the present disclosure contemplates, a cell comprising a polynucleotide encoding the polypeptide described herein.

In various embodiments, the present disclosure contemplates, a cell comprising a vector described herein.

5 In particular embodiments, the present disclosure contemplates, a cell comprising one or more genome modifications introduced by a polypeptide described herein.

In further embodiments, the cell comprises a polynucleotide, mRNA, cDNA, or vector further comprises a heterologous polyadenylation signal.

In further embodiments, the cell further comprises a donor repair template comprising a  
10 polynucleotide encoding a therapeutic polypeptide or protein is integrated into the ALB gene at a DNA double stranded break site introduced by the polypeptide.

In certain embodiments, the therapeutic polypeptide is a therapeutic antihemophilic factor, antibody, protein, cytokine, chemokine, cytotoxin, cytokine receptor, hormone, or functional variants thereof.

15 In particular embodiments, the therapeutic polypeptide is Factor VII (FVII), Factor VIII (FVIII), Factor IX (FIX), Factor X (FX), Factor XI (FXI), or functional variants thereof.

In particular embodiments, the therapeutic polypeptide is FVIII, FIX, or functional variant thereof.

In particular embodiments, the therapeutic polypeptide is a FVIII or functional variant  
20 thereof.

In certain embodiments, the therapeutic polypeptide is a modified FVIII or functional variant thereof.

In certain embodiments, the therapeutic polypeptide comprises a B-domain variant FVIII (FVIII-BDV).

25 In certain embodiments, the therapeutic polypeptide comprises a modified FVIII comprising a shortened B domain.

In particular embodiments, the therapeutic polypeptide comprises a B domain deleted FVIII (FVIII-BDD).

In particular embodiments, the therapeutic polypeptide is a FIX or functional variant  
30 thereof.



In certain embodiments, the therapeutic polypeptide is a modified FIX or functional variant thereof.

In particular embodiments, the cell is a hepatocyte.

In particular embodiments, the cell comprises one or more modified ALB alleles.

5 In particular embodiments, the present disclosure contemplates, a plurality of cells comprising one or more cells contemplated herein.

In particular embodiments, the present disclosure contemplates, a composition comprising one or more cells contemplated herein.

10 In particular embodiments, the present disclosure contemplates, a composition comprising polynucleotides encoding the polypeptide contemplated herein.

In certain embodiments, the composition comprises a pharmaceutically acceptable carrier.

In particular embodiments, the pharmaceutically acceptable carrier is a lipid nanoparticle and wherein the polynucleotide is encapsulated in said lipid nanoparticle.

15 In various embodiments, the present disclosure contemplates, a method of editing a human ALB gene in a cell comprising: introducing a polynucleotide encoding a polypeptide contemplated herein into the cell, wherein expression of the polypeptide creates a double-strand break (DSB) at the target site.

20 In various embodiments, the present disclosure contemplates, a method of editing a human ALB gene in a cell comprising: introducing a polynucleotide encoding a contemplated herein and a donor repair template into the cell, wherein expression of the polypeptide creates a double-strand break (DSB) at the target site and the donor repair template is incorporated into the human ALB gene by homology directed repair (HDR) at the site of the DSB.

25 In various embodiments, the present disclosure contemplates, a method of treating a blood disorder, lysosomal storage disorder, or metabolic disorder, or condition associated therewith, comprising: introducing a polynucleotide encoding a polypeptide contemplated herein and a donor repair template into a cell, wherein expression of the polypeptide creates a double-strand break (DSB) at the target site and the donor repair template is incorporated into the human ALB gene by homology directed repair (HDR) at the site of the DSB.

In various embodiments, the present disclosure contemplates, a method of treating a blood disorder, or condition associated therewith, comprising: introducing a polynucleotide encoding the polypeptide contemplated herein and a donor repair template into a hepatocyte, wherein expression of the polypeptide creates a double-strand break (DSB) at the target site  
5 and the donor repair template is incorporated into the human ALB gene by homology directed repair (HDR) at the site of the DSB.

In various embodiments, the present disclosure contemplates, a method of treating a blood disorder, or condition associated therewith, comprising: introducing a polynucleotide encoding a polypeptide disclosed herein and a donor repair template into a hepatocyte, wherein  
10 expression of the polypeptide creates a double-strand break (DSB) at the target site and the donor repair template is incorporated into the human ALB gene by Non-homologous end-joining (NHEJ) at the site of the DSB.

In certain embodiments, the polynucleotide is an mRNA and comprises a nucleic acid sequence set forth in any one of SEQ ID NOs: 22-29.

15 In certain embodiments, the polynucleotide is a cDNA or a vector.

In particular embodiments, the polynucleotide, mRNA, cDNA, or vector further comprises a heterologous polyadenylation signal.

In particular embodiments, the polynucleotide, mRNA, cDNA, or vector further comprises a nuclear localization signal (NLS).

20 In certain embodiments, a polynucleotide encoding a 5'-3' exonuclease is introduced into the cell.

In particular embodiments, a polynucleotide encoding Trex2 or a biologically active fragment thereof is introduced into the cell.

In particular embodiments, a donor repair template comprising a polynucleotide  
25 encoding a therapeutic polypeptide or protein is integrated into the ALB gene at the DSB introduced by the polypeptide.

In particular embodiments, the therapeutic polypeptide is a therapeutic antihemophilic factor, antibody, protein, cytokine, chemokine, cytotoxin, cytokine receptor, hormone, or functional variants thereof.

In particular embodiments, the therapeutic polypeptide is a Factor VIII (FVIII), Factor IX (FIX), Factor X (FX), Factor XI (FXI), or functional variants thereof.

In particular embodiments, the therapeutic polypeptide is a FVIII or FIX or functional variant thereof.

5 In certain embodiments, the therapeutic polypeptide is a FVIII or functional variant thereof.

In particular embodiments, the therapeutic polypeptide is a modified FVIII, such as B-domain variant FVIII (FVIII-BDV).

10 In certain embodiments, the therapeutic polypeptide is a FIX or functional variant thereof.

In particular embodiments, the therapeutic polypeptide is a modified FIX.

In particular embodiments, the donor repair template further comprises a heterologous polyadenylation signal.

15 In particular embodiments, the donor repair template comprises a 5' homology arm homologous to a human ALB gene sequence 5' of the DSB and a 3' homology arm homologous to a human ALB gene sequence 3' of the DSB.

In certain embodiments, a viral vector is used to introduce the donor repair template into the cell.

20 In particular embodiments, a recombinant adeno-associated viral vector (rAAV) or a retrovirus is used to introduce the donor repair template into the cell.

In particular embodiments, the rAAV has one or more ITRs from an AAV serotype selected from the group consisting of: AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, and AAV10.

25 In particular embodiments, the rAAV comprises a capsid from an AAV serotype selected from the group consisting of: AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, and AAV10.

In particular embodiments, the retrovirus is a lentivirus.

In certain embodiments, the lentivirus is an integrase deficient lentivirus (IDLV).

In particular embodiments, the blood disorder is a Hemophilia or Hemoglobinopathy.

30 In certain embodiments, the blood disorder is Hemophilia

In particular embodiments, the is Hemophilia A, Hemophilia B, or Hemophilia C.

In certain embodiments, the Hemophilia is Hemophilia A.

In particular embodiments, the Hemophilia is Hemophilia B.

In particular embodiment, the Hemophilia is Hemophilia C.

- 5 In various embodiments, the present disclosure contemplates a polypeptide comprising an engineered I-OnuI homing endonuclease (HE) that binds and cleaves DNA at a target site within a double-stranded DNA (dsDNA) molecule, wherein the target site is within intron 1 of the human albumin (ALB) gene, wherein the I-OnuI HE cleaves both strands of the dsDNA molecule, and wherein the engineered I-OnuI HE comprises an amino acid sequence that is at  
10 least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to the amino acid sequence set forth in any one of SEQ ID NOs: 6-13.

In particular embodiments, the target site comprises a nucleic acid sequence as set forth in any one of SEQ ID NOs: 32-34.

- In certain embodiments, the target site comprises a nucleic acid sequence as set forth in  
15 SEQ ID NO: 34.

In particular embodiments, the engineered I-OnuI HE comprises the following amino acid substitutions in relation to the numbering of SEQ ID NO: 4:

- (a) S20C, L22S, R24V, R26Q, N27K, N28A, K30G, S31W, S32T, V33T, G34R, S36N, E38K, G40V, Q42S, T44G, V64T, A66S, N67K, N71T, A72Q, S74R, K76T, C176S,  
20 F178M, N180I, I182A, K185R, S186T, K187G, L188A, G189R, Q193R, V195C, S197E, T199G, Y219H, K221S, K225L, F228S, W230F, D232I, and V234R;
- (b) S20C, L22S, R24T, N27K, N28F, K30G, S31R, S32T, V33T, G34R, S36T, E38K, G40M, Q42S, T44G, V64T, A66S, N67K, N71T, A72Q, S74R, K76T, C176S, F178M, N180I, I182A, K185R, S186T, K187G, L188A, G189R, Q193R, V195C, S197E, T199G,  
25 Y219H, K221S, K225L, F228S, W230F, D232I, and V234R;
- (c) S20C, L22S, R24A, R26G, N27K, N28L, K30G, S31G, S32T, V33T, G34R, E38K, G40V, Q42S, T44G, V64T, A66S, N67K, N71T, A72Q, S74R, K76T, C176S, F178M, N180I, I182A, K185R, S186T, K187G, L188A, G189R, Q193R, V195C, S197E, T199G, Y219H, K221S, K225L, F228S, W230F, D232I, and V234R;

- (d) S20C, L22S, R24A, R26G, N27K, N28L, K30G, S31G, S32T, V33T, G34R, E38K, G40V, Q42S, T44G, V64T, A66S, N67K, N71T, A72Q, S74R, K76T, C176S, F178M, N180I, I182A, K185R, S186T, K187G, L188A, G189R, Q193R, V195C, S197E, T199G, Y219H, K221S, K225L, F228S, W230F, D232I, and V234R;
- 5 (e) S20C, L22S, R24V, R26A, N27K, N28G, K30G, S31G, S32T, V33T, G34R, S36T, E38K, G40V, Q42S, T44G, V64T, A66S, N67K, N71T, A72Q, S74R, K76T, C176S, F178M, N180I, I182A, K185R, S186T, K187G, L188A, G189R, Q193R, V195C, S197E, T199G, Y219H, K221S, K225L, F228S, W230F, D232I, and V234R;
- (f) S20C, L22S, R24A, R26G, N27K, N28C, K30G, S31C, S32T, V33T, G34R, E38K, G40V, Q42S, T44G, V64T, A66S, N67K, N71T, A72Q, S74R, K76T, C176S, F178M, N180I, I182A, K185R, S186T, K187G, L188A, G189R, Q193R, V195C, S197E, T199G, Y219H, K221S, K225L, F228S, W230F, D232I, and V234R;
- 10 (g) S20C, L22S, R24A, R26G, N27K, N28L, K30G, S31A, S32T, V33T, G34R, S36V, E38K, G40V, Q42S, T44G, V64T, A66S, N67K, N71T, A72Q, S74R, K76T, C176S, F178M, N180I, I182A, K185R, S186T, K187G, L188A, G189R, Q193R, V195C, S197E, T199G, Y219H, K221S, K225L, F228S, W230F, D232I, and V234R; or
- (h) S20C, L22S, R24H, R26G, N27K, N28M, K30G, S31R, S32T, V33T, G34R, S36L, E38K, G40V, Q42S, T44G, V64T, A66S, N67K, N71T, A72Q, S74R, K76T, C176S, F178M, N180I, I182A, K185R, S186T, K187G, L188A, G189R, Q193R, V195C, S197E, T199G, Y219H, K224S, K225L, F228S, W230F, D232I, and V234R.
- 20

In certain embodiments, the engineered HE comprises the amino acid sequence set forth in SEQ ID NO: 6.

In particular embodiments, the polypeptide further comprises one or more nuclear localization signal (NLS) and/or a TALE DNA binding domain.

- 25 In certain embodiments, the TALE DNA binding domain comprises about 8.5 TALE repeat units to about 15.5 TALE repeat units.

In certain embodiments, the NLS comprises a nucleic acid sequence set forth in SEQ ID NO: 86.

- In particular embodiments, the TALE DNA binding domain binds the polynucleotide sequence set forth in any one of SEQ ID NOs: 35-42.
- 30

In certain embodiments, the polypeptide binds and cleaves the polynucleotide sequence set forth in SEQ ID NO: 43.

In particular embodiments, the polypeptide comprises an amino acid sequence that is at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to the amino acid  
5 sequence set forth in any one of SEQ ID NOs: 14-21.

In certain embodiments, the polypeptide comprises the amino acid sequence set forth in SEQ ID NO: 14.

In particular embodiments, the polypeptide further comprising a peptide linker and an end-processing enzyme or biologically active fragment thereof.

10 In certain embodiments, the end-processing enzyme comprises Trex2.

In other embodiments, the end-processing enzyme comprises an amino acid sequence as set forth in SEQ ID NO: 44.

In various embodiments, the present disclosure contemplates a polynucleotide encoding the polypeptide contemplated herein.

15 In certain embodiments, the polynucleotide is an mRNA or a cDNA.

In particular embodiments, the polynucleotide is an mRNA, wherein the mRNA comprises a nucleic acid sequence that is at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to the nucleic acid sequence set forth in any one of SEQ ID NOs: 22-29.

In certain embodiments, the mRNA comprises a nucleic acid sequence set forth in SEQ  
20 ID NO: 22.

In certain embodiments, the mRNA comprises a nucleic acid sequence set forth in SEQ ID NO: 108.

In various embodiments, the present disclosure contemplates a pharmaceutical composition comprising a polynucleotide encoding the polypeptide contemplated herein and a  
25 physiologically acceptable carrier.

In particular embodiments, the polynucleotide is an mRNA comprising a nucleic acid sequence set forth in SEQ ID NO: 22.

In certain embodiments, the pharmaceutically acceptable carrier comprises a lipid nanoparticle and the polynucleotide is encapsulated in said lipid nanoparticle.

In particular embodiments, the pharmaceutical composition further comprising a donor repair template.

In certain embodiments, an AAV vector particle comprises the donor repair template.

In certain embodiments, the donor repair template comprises a FVIII transgene or a  
5 FIX transgene.

In various embodiments, the present disclosure contemplates a method of editing a human ALB gene in a cell comprising: introducing a polynucleotide encoding the polypeptide contemplated herein into the cell, wherein the polynucleotide is translated in the cell to produce a polypeptide, wherein the polypeptide creates a double-strand break (DSB) at the target site;  
10 optionally further comprising introducing a donor repair template into the cell and wherein the donor repair template is incorporated into the human ALB gene by homology directed repair (HDR) at the site of DSB.

In various embodiments, the present disclosure contemplates a method of treating a blood disorder, lysosomal storage disorder, or metabolic disorder, or condition associated  
15 therewith, comprising: introducing into the cell (a) a polynucleotide encoding the polypeptide contemplated herein, and (b) a donor repair template into a cell; wherein the polynucleotide is translated in the cell to produce a polypeptide, wherein the polypeptide creates a double-strand break (DSB) at the target site and the donor repair template is incorporated into the human ALB gene by homology directed repair (HDR) at the site of the DSB.

20 In certain embodiments, the cell is a hepatocyte and the donor repair template comprises a FVIII transgene or a FIX transgene.

In various embodiments, the present disclosure contemplates the pharmaceutical composition contemplated herein for use in editing a human ALB gene in a cell or for use in treating a blood disorder, lysosomal storage disorder, or metabolic disorder, or condition  
25 associated therewith, wherein the composition delivers the polynucleotide and the donor repair template into the cell, wherein the polynucleotide is translated in the cell to produce a polypeptide, wherein the polypeptide creates a double-strand break (DSB) at the target site, and wherein the donor repair template is incorporated into the human ALB gene by homology directed repair (HDR) at the site of the DSB.

In certain embodiments, the cell is a hepatocyte and the donor repair template is Factor VIII or Factor IX.

#### BRIEF DESCRIPTION OF SEVERAL VIEWS OF THE DRAWINGS

- FIG. 1** shows a schematic of a megaTAL targeting a portion of *ALB* Intron 1.
- 5 **FIG. 2A** shows percent indel for select enzymes targeting *ALB* intron 1 site S01.
- FIG. 2B** shows percent indel for select enzymes targeting *ALB* intron 1 site S03.
- FIG. 2C** shows percent indel for select enzymes targeting *ALB* intron 1 site S06.
- FIG. 3A** shows a schematic of various TALE arrays attached to an illustrative meganuclease and their respective target sites.
- 10 **FIG. 3B** shows percent indel for megaTALs having the indicated TALE array configuration, with and without Trex2.
- FIG. 4** shows off-target editing analysis of select engineered megaTALs.
- FIG. 5** shows percent indel for select enzymes targeting *ALB* intron 1.
- FIG. 6** shows percent indel and the calculated indel50s for select Gen1 megaTALs.
- 15 **FIG. 7** shows percent indel and the calculated indel50s for select Gen 1 and Gen2 megaTALs.
- FIG. 8** shows a graph depicting average percent indel at identified off-target sites in cells electroporated with the 327A10 Gen2 clone mRNA.
- FIG. 9** shows percent indel and the calculated indel50s for select Gen2 and Gen3
- 20 megaTALs.
- FIG. 10** shows average percent indel at identified off-target sites in cells electroporated with the indicated Gen3 clone mRNAs.
- FIG. 11A** shows percent indel for select Gen2, Gen3, and Gen4 megaTALs.
- FIG. 11B** shows the calculated indel50s for select Gen2, Gen3, and Gen4 megaTALs.
- 25 **FIGs. 12A and 12B** show average percent indel at identified off-target sites in cells electroporated with the indicated Gen4 clone mRNAs.
- FIG. 13** shows percent FVIII integration with increasing amounts of AAV-FVIII in primary human hepatocytes.



**FIG. 14A** shows megaTAL editing (% indel) in liver cells from animals administered AAV-FVIII and LNP encapsulated mRNA encoding an *ALB* targeted megaTAL.

**FIG. 14B** shows percent AAV integration in liver cells from animals administered AAV-FVIII and LNP encapsulated mRNA encoding an *ALB* targeted megaTAL.

5 **FIG. 14C** shows blood FVIII activity from animals administered AAV-FVIII and LNP encapsulated mRNA encoding an *ALB* targeted megaTAL.

**FIG. 15** shows an alignment of SEQ ID NOs: 1-6.

### BRIEF DESCRIPTION OF THE SEQUENCE IDENTIFIERS

**SEQ ID NO: 1** is an amino acid sequence of a wild type I-OnuI LAGLIDADG  
10 homing endonuclease (LHE).

**SEQ ID NO: 2** is an amino acid sequence of a stabilized I-OnuI LHE.

**SEQ ID NO: 3** is an amino acid sequence of a biologically active fragment of a stabilized I-OnuI LHE.

**SEQ ID NO: 4** is an amino acid sequence of a biologically active fragment of a  
15 stabilized I-OnuI LHE.

**SEQ ID NO: 5** is an amino acid sequence of a biologically active fragment of a stabilized I-OnuI LHE.

**SEQ ID NOs: 6-13 and 87-104** set forth amino acid sequences of engineered I-OnuI LHE variants reprogrammed to bind and cleave a target site in the human *ALB* gene.

20 **SEQ ID NOs: 14-21 and 105-107** set forth amino acid sequences of megaTALs engineered to bind and cleave a target site in the human *ALB* gene.

**SEQ ID NOs: 22-29** set forth mRNA sequences encoding *ALB* megaTALs.

**SEQ ID NO: 30** is the human *ALB* gene, Accession Number M12523.1.

**SEQ ID NO: 31** is intron 1 of the human *ALB* gene.

25 **SEQ ID NOs: 32-34** set forth HE target sites within intron 1 of human *ALB*.

**SEQ ID NOs: 35-42** set forth TALE array target sites.

**SEQ ID NO: 43** is megaTAL target site in intron 1 of human *ALB* gene.

**SEQ ID NO: 44** is an amino acid sequence encoding human Trex2.

**SEQ ID NOs: 45-55** set forth the amino acid sequences of various linkers.

SEQ ID NOs: 56-80 set forth the amino acid sequences of protease cleavage sites and self-cleaving polypeptide cleavage sites.

SEQ ID NO: 81 set forth the consensus Kozak sequence.

SEQ ID NOs: 82-85 set forth examples of polyA sequences.

5 SEQ ID NO: 86 set forth an example of NLS sequence.

SEQ ID NO: 108 set forth amino acid sequence of megaTAL engineered to bind and cleave a target site in the Human ALB gene and with a polyA sequence.

In the foregoing sequences, X or Xaa, if present, refers to any amino acid or the  
10 absence of an amino acid.

## DETAILED DESCRIPTION

### A. OVERVIEW

The present disclosure generally relates, in part, to compositions comprising highly specific and efficacious engineered homing endonucleases and megaTALs that cleave a target  
15 site in the human Albumin (ALB) gene and methods of using the same. In various embodiments, the engineered homing endonucleases bind and cleave an *ALB* intron 1 target site.

While almost any nucleic acid sequence could be targeted by select gene-editing enzymes (*e.g.*, site-specific nucleases), a threshold issue in developing a gene-editing  
20 therapeutic is selecting an appropriate target site for editing. Indeed, not all target sites are targeted with equal efficiency, nor do all target sites yield equally specific enzymes. Thus, there are many factors that contribute to an effect gene-editing enzyme and associated therapeutic treatment including, but not limited to, the enzyme type and modifications thereto (if needed), target site location, target site size, target site uniqueness, nucleic acid composition,  
25 chromatin state, and other epigenetic modifications. Moreover, while enzymes such as TALENs, Zinc Fingers (ZFs), and CRISPR/Cas9 can be designed to target specific sequences by mere selection of the appropriate linked domains (TALENs and ZFs) or providing an appropriate guide sequence (CRISPR/Cas9), engineered homing endonucleases/meganucleases

must be empirically designed, and at least to-date, the resulting sequence of which cannot be reasonably predicted.

Thus, the inventors surprisingly discovered a set of unique engineered homing endonucleases and megaTALs having high specificity for intron 1 of the ALB locus and are also highly efficacious. Such engineered homing endonucleases have important implications and usefulness (such as enabling insertion of coding sequences for therapeutic polypeptides/proteins, factors, or antibodies) for the treatment of a variety of diseases, including but not limited to, bleeding disorders (*e.g.*, hemophilia), hemoglobinopathies, metabolic disorders, lysosomal storage diseases and other mono or polygenic disorders.

In one aspect, a polypeptide comprising an engineered homing endonuclease that cleaves a selected double strand DNA (dsDNA) target site in the human albumin (ALB) gene is provided. In various embodiments, the target site is, found within, or part of, the ALB intron 1 (*e.g.*, SEQ ID NOs: 32-34). In various embodiments, the homing endonuclease is an engineered LAGLIDAGE homing endonuclease (LHE), *e.g.*, an engineered I-OnuI. In particular embodiments, the polypeptide further comprises a DNA binding domain (*e.g.*, TALE or Zinc Finger DNA-binding domain). In even more particular embodiments, the polypeptide comprises an amino acid sequence that is at least 90% identical to the amino acid sequence set forth in any one of SEQ ID NOs: 6-21.

In another aspect, a polynucleotide (*e.g.*, mRNA, cDNA, or vector) encoding any one of the polypeptides contemplated herein are provided. In particular embodiments, the polynucleotide comprises an mRNA nucleotide sequence that is at least 90% identical to the amino acid sequence set forth in any one of SEQ ID NOs: 22-29.

In another aspect, a cell (*e.g.*, hepatocyte) comprising any one of the gene-editing polypeptides or polynucleotides contemplated herein are provided. In various embodiments, the cell further comprises a therapeutic transgene or polynucleotide encoding a therapeutic transgene (*e.g.*, a therapeutic polypeptide). In some embodiments, the therapeutic transgene or therapeutic polypeptide is Factor VII (FVII), Factor VIII (FVIII), Factor IX (FIX), Factor X (FX), Factor XI (FXI), or functional variants thereof, such as FVIII-BDV.

In another aspect, a method of editing a human ALB gene in a cell is provided, comprising introducing a polynucleotide encoding a gene-editing polypeptide contemplated herein into the cell, wherein expression of the polypeptide creates a double-strand break (DSB) at a target site in a human ALB gene. In some embodiments, a donor repair template is also  
5 introduced into the cell and the donor repair template is incorporated in the human ALB gene at the site of the DSB.

In another aspect, a method of treating a disease or condition associated therewith is provided (*e.g.*, a hemophilia, hemoglobinopathy, metabolic disorder, or lysosomal storage disease) comprising: introducing a polynucleotide encoding a gene-editing polypeptide/a site-  
10 specific nuclease contemplated herein (*e.g.*, megaTAL) and a donor repair template into a cell (*e.g.*, hepatocyte), wherein expression of the polypeptide creates a DSB at a target site in a human ALB gene and the donor repair template, for example comprising a therapeutic transgene (*e.g.*, FVIII or FIX), is incorporated into the human ALB gene at the site of the DSB. In various embodiments, a viral vector (*e.g.*, AAV) is used to introduce the donor repair  
15 template into the cell. In particular embodiments, the hemophilia is hemophilia A or hemophilia B.

Techniques for recombinant (*i.e.*, engineered) DNA, peptide and oligonucleotide synthesis, immunoassays, tissue culture, transformation (*e.g.*, electroporation, lipofection), enzymatic reactions, purification and related techniques and procedures may be generally  
20 performed as described in various general and more specific references in microbiology, molecular biology, biochemistry, molecular genetics, cell biology, virology and immunology as cited and discussed throughout the present specification. *See, e.g.*, Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, 3d ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.; *Current Protocols in Molecular Biology* (John Wiley and Sons, updated  
25 July 2008); *Short Protocols in Molecular Biology: A Compendium of Methods from Current Protocols in Molecular Biology*, Greene Pub. Associates and Wiley-Interscience; Glover, *DNA Cloning: A Practical Approach*, vol. I & II (IRL Press, Oxford Univ. Press USA, 1985); *Current Protocols in Immunology* (Edited by: John E. Coligan, Ada M. Kruisbeek, David H. Margulies, Ethan M. Shevach, Warren Strober 2001 John Wiley & Sons, NY, NY); *Real-Time*  
30 *PCR: Current Technology and Applications*, Edited by Julie Logan, Kirstin Edwards and Nick

- Saunders, 2009, Caister Academic Press, Norfolk, UK; Anand, *Techniques for the Analysis of Complex Genomes*, (Academic Press, New York, 1992); Guthrie and Fink, *Guide to Yeast Genetics and Molecular Biology* (Academic Press, New York, 1991); *Oligonucleotide Synthesis* (N. Gait, Ed., 1984); *Nucleic Acid The Hybridization* (B. Hames & S. Higgins, Eds., 1985); *Transcription and Translation* (B. Hames & S. Higgins, Eds., 1984); *Animal Cell Culture* (R. Freshney, Ed., 1986); Perbal, *A Practical Guide to Molecular Cloning* (1984); *Next-Generation Genome Sequencing* (Janitz, 2008 Wiley-VCH); *PCR Protocols (Methods in Molecular Biology)* (Park, Ed., 3rd Edition, 2010 Humana Press); *Immobilized Cells And Enzymes* (IRL Press, 1986); the treatise, *Methods In Enzymology* (Academic Press, Inc., N.Y.);
- 10 *Gene Transfer Vectors For Mammalian Cells* (J. H. Miller and M. P. Calos eds., 1987, Cold Spring Harbor Laboratory); Harlow and Lane, *Antibodies*, (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1998); *Immunochemical Methods In Cell And Molecular Biology* (Mayer and Walker, eds., Academic Press, London, 1987); *Handbook Of Experimental Immunology*, Volumes I-IV (D. M. Weir and CC Blackwell, eds., 1986); Roitt,
- 15 *Essential Immunology*, 6th Edition, (Blackwell Scientific Publications, Oxford, 1988); *Current Protocols in Immunology* (Q. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach and W. Strober, eds., 1991); *Annual Review of Immunology*; as well as monographs in journals such as *Advances in Immunology*.

## B. DEFINITIONS

- 20 Prior to setting forth this disclosure in more detail, it may be helpful to an understanding thereof to provide definitions of certain terms to be used herein.

- Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by those of ordinary skill in the art to which the invention belongs. Although any methods and materials similar or equivalent to those described herein
- 25 can be used in the practice or testing of particular embodiments, preferred embodiments of compositions, methods and materials are described herein. For the purposes of the present disclosure, the following terms are defined below. Additional definitions are set forth throughout this disclosure.

The articles “a,” “an,” and “the” are used herein to refer to one or to more than one (i.e., to at least one, or to one or more) of the grammatical object of the article. By way of example, “an element” means one element or one or more elements.

The use of the alternative (*e.g.*, “or”) should be understood to mean either one, both, or  
5 any combination thereof of the alternatives.

The term “and/or” should be understood to mean either one, or both of the alternatives.

As used herein, the term “about” or “approximately” refers to a quantity, level, value, number, frequency, percentage, dimension, size, amount, weight or length that varies by as much as 15%, 10%, 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2% or 1% to a reference quantity, level,  
10 value, number, frequency, percentage, dimension, size, amount, weight or length. In one embodiment, the term “about” or “approximately” refers a range of quantity, level, value, number, frequency, percentage, dimension, size, amount, weight or length  $\pm 15\%$ ,  $\pm 10\%$ ,  $\pm 9\%$ ,  $\pm 8\%$ ,  $\pm 7\%$ ,  $\pm 6\%$ ,  $\pm 5\%$ ,  $\pm 4\%$ ,  $\pm 3\%$ ,  $\pm 2\%$ , or  $\pm 1\%$  about a reference quantity, level, value, number, frequency, percentage, dimension, size, amount, weight or length.

15 In one embodiment, a range, *e.g.*, 1 to 5, about 1 to 5, or about 1 to about 5, refers to each numerical value encompassed by the range. For example, in one non-limiting and merely illustrative embodiment, the range “1 to 5” is equivalent to the expression 1, 2, 3, 4, 5; or 1.0, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, or 5.0; or 1.0, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2.0, 2.1, 2.2, 2.3, 2.4, 2.5, 2.6, 2.7, 2.8, 2.9, 3.0, 3.1, 3.2, 3.3, 3.4, 3.5, 3.6, 3.7, 3.8, 3.9, 4.0, 4.1, 4.2, 4.3,  
20 4.4, 4.5, 4.6, 4.7, 4.8, 4.9, or 5.0.

As used herein, the term “substantially” refers to a quantity, level, value, number, frequency, percentage, dimension, size, amount, weight or length that is 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or higher compared to a reference quantity, level, value, number, frequency, percentage, dimension, size, amount, weight or length. In one  
25 embodiment, “substantially the same” refers to a quantity, level, value, number, frequency, percentage, dimension, size, amount, weight or length that produces an effect, *e.g.*, a physiological effect, that is approximately the same as a reference quantity, level, value, number, frequency, percentage, dimension, size, amount, weight or length.

Throughout this specification, unless the context requires otherwise, the words “comprise”, “comprises” and “comprising” will be understood to imply the inclusion of a stated step or element or group of steps or elements but not the exclusion of any other step or element or group of steps or elements. By “consisting of” is meant including, and limited to, whatever follows the phrase “consisting of.” Thus, the phrase “consisting of” indicates that the listed elements are required or mandatory, and that no other elements may be present. By “consisting essentially of” is meant including any elements listed after the phrase, and limited to other elements that do not interfere with or contribute to the activity or action specified in the disclosure for the listed elements. Thus, the phrase “consisting essentially of” indicates that the listed elements are required or mandatory, but that no other elements are present that materially affect the activity or action of the listed elements.

Reference throughout this specification to “one embodiment,” “an embodiment,” “a particular embodiment,” “a related embodiment,” “a certain embodiment,” “an additional embodiment,” or “a further embodiment” or combinations thereof means that a particular feature, structure or characteristic described in connection with the embodiment is included in at least one embodiment. Thus, the appearances of the foregoing phrases in various places throughout this specification are not necessarily all referring to the same embodiment. Furthermore, the particular features, structures, or characteristics may be combined in any suitable manner in one or more embodiments. It is also understood that the positive recitation of a feature in one embodiment, serves as a basis for excluding the feature in a particular embodiment.

The term “*ex vivo*” refers generally to activities that take place outside an organism, such as experimentation or measurements done in or on living tissue in an artificial environment outside the organism, preferably with minimum alteration of the natural conditions. In particular embodiments, “*ex vivo*” procedures involve living cells or tissues taken from an organism and cultured or modulated in a laboratory apparatus, usually under sterile conditions, and typically for a few hours or up to about 24 hours, but including up to 48 or 72 hours, depending on the circumstances. In certain embodiments, such tissues or cells can be collected and frozen, and later thawed for *ex vivo* treatment. Tissue culture experiments or procedures lasting longer than a few days using living cells or tissue are typically considered to

be “*in vitro*,” though in certain embodiments, this term can be used interchangeably with *ex vivo*.

The term “*in vivo*” refers generally to activities that take place inside an organism. In one embodiment, cellular genomes are engineered, edited, or modified *in vivo*.

5 By “enhance” or “promote” or “increase” or “expand” or “potentiate” refers generally to the ability of an engineered nuclease, genome editing composition, or genome edited cell contemplated herein to produce, elicit, or cause a greater response (*i.e.*, physiological response) compared to the response caused by either vehicle or control. A measurable response may include an increase in catalytic activity, binding affinity, binding site specificity, binding site  
10 selectivity, persistence, cytolytic activity, and/or an increase in proinflammatory cytokines, among others apparent from the understanding in the art and the description herein. An “increased” or “enhanced” amount is typically a “statistically significant” amount, and may include an increase that is 1.1, 1.2, 1.5, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 30 or more times (*e.g.*, 500, 1000 times) (including all integers and decimal points in between and above 1, *e.g.*, 1.5,  
15 1.6, 1.7, 1.8, *etc.*) the response produced by vehicle or control.

By “decrease” or “lower” or “lessen” or “reduce” or “abate” or “ablate” or “inhibit” or “dampen” refers generally to the ability of an engineered nuclease, genome editing composition, or genome edited cell contemplated herein to produce, elicit, or cause a lesser response (*i.e.*, physiological response) compared to the response caused by either vehicle or  
20 control. A measurable response may include a decrease in off-target binding affinity, off-target cleavage specificity, T cell exhaustion, and the like. A “decrease” or “reduced” amount is typically a “statistically significant” amount, and may include an decrease that is 1.1, 1.2, 1.5, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 30 or more times (*e.g.*, 500, 1000 times) (including all integers and decimal points in between and above 1, *e.g.*, 1.5, 1.6, 1.7, 1.8, *etc.*) the response  
25 (reference response) produced by vehicle, or control.

By “maintain,” or “preserve,” or “maintenance,” or “no change,” or “no substantial change,” or “no substantial decrease” refers generally to the ability of a engineered nuclease, genome editing composition, or genome edited cell contemplated herein to produce, elicit, or cause a substantially similar or comparable physiological response (*i.e.*, downstream effects) in



as compared to the response caused by either vehicle or control. A comparable response is one that is not significantly different or measurable different from the reference response.

An “antibody” refers to a binding agent that is a polypeptide comprising at least a light chain or heavy chain immunoglobulin variable region which specifically recognizes and binds  
5 an epitope of a target antigen, such as a peptide, lipid, polysaccharide, or nucleic acid containing an antigenic determinant, such as those recognized by an immune cell. Antibodies include antigen binding fragments, e.g., Camel Ig (a camelid antibody or VHH fragment thereof), Ig NAR, Fab fragments, Fab' fragments, F(ab)'2 fragments, F(ab)'3 fragments, Fv, single chain Fv antibody (“scFv”), bis-scFv, (scFv)<sub>2</sub>, minibody, diabody, triabody, tetrabody,  
10 disulfide stabilized Fv protein (“dsFv”), and single-domain antibody (sdAb, Nanobody) or other antibody fragments thereof. The term also includes genetically engineered forms such as chimeric antibodies (for example, humanized murine antibodies), heteroconjugate antibodies (such as, bispecific antibodies) and antigen binding fragments thereof. See also, Pierce Catalog and Handbook, 1994-1995 (Pierce Chemical Co., Rockford, IL); Kuby, J., Immunology, 3rd  
15 Ed., W. H. Freeman & Co., New York, 1997.

The terms “specific binding affinity” or “specifically binds” or “specifically bound” or “specific binding” or “specifically targets” as used herein, describe binding of one molecule to another, *e.g.*, DNA binding domain of a polypeptide binding to DNA, at greater binding  
affinity than background binding. A binding domain “specifically binds” to a target site if it  
20 binds to or associates with a target site with an affinity or  $K_a$  (*i.e.*, an equilibrium association constant of a particular binding interaction with units of 1/M) of, for example, greater than or equal to about  $10^5 \text{ M}^{-1}$ . In certain embodiments, a binding domain binds to a target site with a  $K_a$  greater than or equal to about  $10^6 \text{ M}^{-1}$ ,  $10^7 \text{ M}^{-1}$ ,  $10^8 \text{ M}^{-1}$ ,  $10^9 \text{ M}^{-1}$ ,  $10^{10} \text{ M}^{-1}$ ,  $10^{11} \text{ M}^{-1}$ ,  $10^{12} \text{ M}^{-1}$ , or  $10^{13} \text{ M}^{-1}$ . “High affinity” binding domains refers to those binding domains with a  $K_a$  of at  
25 least  $10^7 \text{ M}^{-1}$ , at least  $10^8 \text{ M}^{-1}$ , at least  $10^9 \text{ M}^{-1}$ , at least  $10^{10} \text{ M}^{-1}$ , at least  $10^{11} \text{ M}^{-1}$ , at least  $10^{12} \text{ M}^{-1}$ , at least  $10^{13} \text{ M}^{-1}$ , or greater.

Alternatively, affinity may be defined in particular embodiments as an equilibrium dissociation constant ( $K_d$ ) of a particular binding interaction with units of M (*e.g.*,  $10^{-5} \text{ M}$  to  $10^{-13} \text{ M}$ , or less). Affinities of engineered nucleases comprising one or more DNA binding

domains for DNA target sites contemplated in particular embodiments can be readily determined using conventional techniques, *e.g.*, yeast cell surface display, or by binding association, or displacement assays using labeled ligands.

In one embodiment, the affinity of specific binding is about 2 times greater than  
5 background binding, about 5 times greater than background binding, about 10 times greater than background binding, about 20 times greater than background binding, about 50 times greater than background binding, about 100 times greater than background binding, or about 1000 times greater than background binding or more.

The terms “selectively binds” or “selectively bound” or “selectively binding” or  
10 “selectively targets” and describe preferential binding of one molecule to a target molecule (on-target binding) in the presence of a plurality of off-target molecules. In particular embodiments, an HE or megaTAL selectively binds an on-target DNA binding site about 5, 10, 15, 20, 25, 50, 100, or 1000 times more frequently than the HE or megaTAL binds an off-target DNA target binding site.

15 “On-target” refers to a target site sequence.

“Off-target” refers to a sequence similar to but not identical to a target site sequence.

A “target site” or “target sequence” is a chromosomal or extrachromosomal nucleic acid sequence that defines a portion of a nucleic acid to which a binding molecule will bind and/or cleave, provided sufficient conditions for binding and/or cleavage exist. When referring  
20 to a polynucleotide sequence or SEQ ID NO. that references only one strand of a target site or target sequence, it would be understood that the target site or target sequence bound and/or cleaved by an engineered nuclease is double-stranded and comprises the reference sequence and its complement. In a preferred embodiment, the target site is a sequence in a human ALB gene. In a further preferred embodiment, the target site is a sequence in intron 1 of the ALB  
25 gene.

“Recombination” refers to a process of exchange of genetic information between two polynucleotides, including but not limited to, donor or end capture by non-homologous end joining (NHEJ) and homologous recombination. For the purposes of this disclosure, “homologous recombination (HR)” refers to the specialized form of such exchange that takes

place, for example, during repair of double-strand breaks in cells via homology-directed repair (HDR) mechanisms. This process requires nucleotide sequence homology, uses a “donor” molecule as a template to repair a “target” molecule (*i.e.*, the one that experienced the double-strand break), and is variously known as “non-crossover gene conversion” or “short tract gene conversion,” because it leads to the transfer of genetic information from the donor to the target. Without wishing to be bound by any particular theory, such transfer can involve mismatch correction of heteroduplex DNA that forms between the broken target and the donor, and/or “synthesis-dependent strand annealing,” in which the donor is used to resynthesize genetic information that will become part of the target, and/or related processes. Such specialized HR often results in an alteration of the sequence of the target molecule such that part or all of the sequence of the donor polynucleotide is incorporated into the target polynucleotide. This is also referred to as targeted integration or genomic integration.

“NHEJ” or “non-homologous end joining” refers to the resolution of a double-strand break in the absence of a donor repair template homologous sequence. NHEJ can result in insertions and deletions at the site of the break. NHEJ is mediated by several sub-pathways, each of which has distinct mutational consequences. The classical NHEJ pathway (cNHEJ) requires the KU/DNA-PKcs/Lig4/XRCC4 complex, ligates ends back together with minimal processing and often leads to precise repair of the break. Alternative NHEJ pathways (altNHEJ) also are active in resolving dsDNA breaks, but these pathways are considerably more mutagenic and often result in imprecise repair of the break marked by insertions and deletions. While not wishing to be bound to any particular theory, it is contemplated that modification of dsDNA breaks by end-processing enzymes, such as, for example, exonucleases, *e.g.*, Trex2, may increase the likelihood of imprecise repair.

“Cleavage” refers to the breakage of the covalent backbone of a DNA molecule. Cleavage can be initiated by a variety of methods including, but not limited to, enzymatic or chemical hydrolysis of a phosphodiester bond. Both single-stranded cleavage and double-stranded cleavage are possible. Double-stranded cleavage can occur as a result of two distinct single-stranded cleavage events. DNA cleavage can result in the production of either blunt ends or staggered ends. In certain embodiments, polypeptides and engineered nucleases, *e.g.*, homing endonuclease variants, megaTALs, *etc.* contemplated herein are used for targeted

double-stranded DNA cleavage. Endonuclease cleavage recognition sites may be on either DNA strand.

An “exogenous” molecule is a molecule that is not normally present in a cell, but that is introduced into a cell by one or more genetic, biochemical or other methods. Exemplary  
5 exogenous molecules include, but are not limited to small organic molecules, protein, nucleic acid, carbohydrate, lipid, glycoprotein, lipoprotein, polysaccharide, any modified derivative of the above molecules, or any complex comprising one or more of the above molecules. Methods for the introduction of exogenous molecules into cells are known to those of skill in the art and include, but are not limited to, lipid-mediated transfer (*i.e.*, liposomes, including  
10 neutral and cationic lipids), lipid nanoparticles (also referred to as LNP particle or lipid particle), electroporation, direct injection, cell fusion, particle bombardment, biopolymer nanoparticle, calcium phosphate co-precipitation, DEAE-dextran-mediated transfer and viral vector-mediated transfer.

An “endogenous” molecule is one that is normally present in a particular cell at a  
15 particular developmental stage under particular environmental conditions. Additional endogenous molecules can include proteins.

A “gene,” refers to a DNA region encoding a gene product, as well as all DNA regions which regulate the production of the gene product, whether or not such regulatory sequences are adjacent to coding and/or transcribed sequences. A gene includes, but is not limited to,  
20 promoter sequences, enhancers, silencers, insulators, boundary elements, terminators, polyadenylation sequences, post-transcription response elements, translational regulatory sequences such as ribosome binding sites and internal ribosome entry sites, replication origins, matrix attachment sites, and locus control regions.

“Gene expression” refers to the conversion of the information, contained in a gene, into  
25 a gene product. A gene product can be the direct transcriptional product of a gene (*e.g.*, mRNA, tRNA, rRNA, antisense RNA, ribozyme, structural RNA or any other type of RNA) or a protein produced by translation of an mRNA. Gene products also include RNAs which are modified, by processes such as capping, polyadenylation, methylation, and editing, and

proteins modified by, for example, methylation, acetylation, phosphorylation, ubiquitination, ADP-ribosylation, myristilation, and glycosylation.

As used herein, the term “genetically engineered” or “genetically modified” refers to the chromosomal or extrachromosomal addition of extra genetic material in the form of DNA  
5 or RNA to the total genetic material in a cell. Genetic modifications may be targeted or non-targeted to a particular site in a cell’s genome. In one embodiment, genetic modification is site specific and may be referred to targeted integration. In one embodiment, genetic modification is not site specific.

As used herein, the term “genome editing” refers to the substitution, deletion, and/or  
10 introduction of genetic material at a target site in the cell’s genome, which restores, corrects, disrupts, and/or modifies expression of a gene or gene product. Genome editing contemplated in particular embodiments comprises introducing one or more engineered nucleases into a cell to generate DNA lesions at or proximal to a target site in the cell’s genome, optionally in the presence of a donor repair template or AAV vector used to introduce the donor repair template.

As used herein, the term “gene therapy” refers to the introduction of extra genetic  
15 material into the total genetic material in a cell that restores, corrects, or modifies expression of a gene or gene product, or for the purpose of expressing a therapeutic transgene (*e.g.*, therapeutic polypeptide). In particular embodiments, introduction of genetic material into the cell’s genome by genome editing that restores, corrects, disrupts, or  
20 modifies expression of a gene or gene product, or for the purpose of expressing a therapeutic transgene (*e.g.*, therapeutic polypeptide) is considered gene therapy.

As used herein, the term “functional variant” refers to a genomic variant protein (*e.g.*, mutated protein from wild-type protein) wherein the molecular function of the variant maintains at least a portion of its original wild-type protein function.

25 A “bleeding disorder” or “blood disorder” or “coagulopathy” refers to a disorder or disease in which the blood’s to form clots is impaired. Illustrative examples of bleeding disorders include, but are not limited to: Hemophilia A, Hemophilia B, Hemophilia C, von Willebrand disease, and the like.

A “lysosomal storage disease” or “LSD” refers to a disorder or disease resulting from inborn errors of metabolism characterized by the accumulation of substrates in excess in various organs' cells due to the defective functioning of lysosomes. Illustrative examples of lysosomal storage diseases include, but are not limited to, Gaucher's, Fabry's, Hunter's, 5 Hurler's, NeimannPick's, and the like. LSDs are considered a type of metabolic disorder.

An “metabolic disorder” refers to a disorder or disease in which there are genetic conditions that negatively affects the body's processing and distribution of macronutrients, such as proteins, fats, and carbohydrates. Typically, inherited metabolic disorders occur when a defective gene causes an enzyme deficiency.

10 As used herein, the terms “individual” and “subject” are often used interchangeably and refer to any animal that exhibits a symptom of a disorder that can be treated with the engineered nucleases, genome editing compositions, gene therapy vectors, genome editing vectors, genome edited cells, and methods contemplated elsewhere herein. Suitable subjects (e.g., patients) include laboratory animals (such as mouse, rat, rabbit, or guinea pig), farm 15 animals, and domestic animals or pets (such as a cat or dog). Non-human primates and, preferably, human subjects, are included. Typical subjects include human patients that have, have been diagnosed with, or are at risk of having a bleeding disorder (e.g., hemophilia), hemoglobinopathy, metabolic disorder, immune disorder, lysosomal storage disease or other mono or polygenic disorder.

20 As used herein, the term “patient” refers to a subject that has been diagnosed with a disease or disorder that can be treated with the engineered nucleases, genome editing compositions, gene therapy vectors, genome editing vectors, genome edited cells, and methods contemplated elsewhere herein.

25 As used herein “treatment” or “treating,” includes any beneficial or desirable effect on the symptoms or pathology of a disease or pathological condition, and may include even minimal reductions in one or more measurable markers of the disease or condition being treated. Treatment can optionally involve delaying of the progression of the disease or condition. “Treatment” does not necessarily indicate complete eradication or cure of the disease or condition, or associated symptoms thereof.

As used herein, “prevent,” and similar words such as “prevention,” “prevented,” “preventing” *etc.*, indicate an approach for preventing, inhibiting, or reducing the likelihood of the occurrence or recurrence of, a disease or condition. It also refers to delaying the onset or recurrence of a disease or condition or delaying the occurrence or recurrence of the symptoms of a disease or condition. As used herein, “prevention” and similar words also includes reducing the intensity, effect, symptoms and/or burden of a disease or condition prior to onset or recurrence of the disease or condition.

As used herein, the phrase “ameliorating at least one symptom of” refers to decreasing one or more symptoms of the disease or condition for which the subject is being treated. In particular embodiments, the disease or condition being treated is a bleeding disorder, wherein the one or more symptoms ameliorated include, but are not limited to, unexplained and/or excessive bleeding from cuts or injuries, many large or deep bruises, joint pain, joint swelling, joint tightness/stiffness, blood in urine or stool, nosebleeds, vomiting, difficulty walking, convulsions, and seizures.

As used herein, the term “amount” refers to “an amount effective” or “an effective amount” of an engineered nuclease, genome editing composition, or genome edited cell sufficient to achieve a beneficial or desired prophylactic or therapeutic result, including clinical results.

A “prophylactically effective amount” refers to an amount of an engineered nuclease, genome editing composition, or genome edited cell sufficient to achieve the desired prophylactic result. Typically, but not necessarily, since a prophylactic dose is used in subjects prior to or at an earlier stage of disease, the prophylactically effective amount is less than the therapeutically effective amount.

A “therapeutically effective amount” of an engineered nuclease, genome editing composition, or genome edited cell may vary according to factors such as the disease state, age, sex, and weight of the individual, and the ability to elicit a desired response in the individual. A therapeutically effective amount is also one in which any toxic or detrimental effects are outweighed by the therapeutically beneficial effects. The term “therapeutically effective amount” includes an amount that is effective to “treat” a subject (*e.g.*, a patient). When a

therapeutic amount is indicated, the precise amount of the compositions contemplated in particular embodiments, to be administered, can be determined by a physician in view of the specification and with consideration of individual differences in age, weight, tumor size, extent of infection or metastasis, and condition of the patient (subject).

5           “Physiologically acceptable carrier(s)” or “pharmaceutically acceptable carrier(s)” or “excipient(s)” are known in the art and include lipid nanoparticles. Lipid nanoparticles (also called LNP particle or lipid particle) are different types of compositions of nano-scale particles used as carriers containing essential phospholipids that encapsulate a component or composition.

10           As used herein, the term “transgene” or “therapeutic transgene” refers to an exogenous nucleic acid sequence that encodes a protein or functional nucleotide. The therapeutic transgene may encode for a non-natural or naturally occurring protein or polypeptide. As used herein, the terms “therapeutic polypeptide”, “therapeutic protein”, and “therapeutic polypeptide/protein” are used interchangeably and refer to a protein or polypeptide encoded by  
15 a therapeutic transgene useful for the treatment of a disease in a patient. In various embodiments, the therapeutic transgene or therapeutic polypeptide is a therapeutic antihemophilic factor (*e.g.*, FVIII or FIX), antibody, protein, enzyme (*e.g.* ALPL or other phosphatases) cytokine, chemokine, cytotoxin, cytokine receptor, hormone, or functional variant thereof.

20           In some embodiments, the therapeutic polypeptide is a chimeric antigen receptor (CAR), a chimeric costimulatory receptor (CCR), an  $\alpha\beta$  T cell receptor ( $\alpha\beta$ -TCR), a  $\gamma\delta$  T cell receptor ( $\gamma\delta$ -TCR), a dimerizing agent regulated immunoreceptor complex (DARIC), or switch receptor that specifically binds a target antigen. In some embodiments, the therapeutic transgene or therapeutic polypeptide is an exogenous costimulatory factor, immunomodulatory  
25 factor, agonist for a costimulatory factor, antagonist for an immunosuppressive factor, immune cell engager, or fusion protein. In some embodiments, the therapeutic transgene or therapeutic polypeptide is a costimulatory factor. In some embodiments, the therapeutic transgene or therapeutic protein is a cytokine. In some embodiments, the therapeutic transgene or



therapeutic polypeptide is a natural or modified extracellular or intracellular enzyme (e.g. phosphatase, lyase, or others).

Additional definitions are set forth throughout this disclosure.

### C. ENGINEERED NUCLEASES

5 Engineered nucleases contemplated in particular embodiments herein are suitable for genome editing a target site in the ALB gene and comprise one or more DNA cleavage domains (*e.g.*, one or more endonuclease and/or exonuclease domains), and optionally, one or more linkers contemplated herein. Beyond the intrinsic DNA-binding/cleaving capabilities of the engineered nucleases contemplated herein, in some embodiments, the engineered nucleases  
10 further comprise one or more additional DNA-binding domains. The terms “engineered nuclease,” “reprogrammed nuclease,” or “nuclease variant” are used interchangeably and refer to a nuclease comprising one or more DNA cleavage domains, wherein the nuclease has been designed and/or modified from a parental or naturally occurring nuclease, to bind and cleave a double-stranded DNA target sequence in an ALB gene. A “parental nuclease” or “parental  
15 enzyme” refers to the naturally occurring nuclease or nuclease variant from which the engineered nuclease is designed and/or modified from.

In particular embodiments, an engineered nuclease binds and cleaves a target sequence in intron 1 of an ALB gene, preferably at any one of SEQ ID NOs: 32-34 in intron 1 of a ALB gene, and more preferably at the sequence “TTAT” in SEQ ID NO: 34 in intron 1 of an ALB  
20 gene.

The engineered nuclease may be designed and/or modified from a naturally occurring nuclease or from a previous engineered nuclease. Engineered nucleases contemplated in particular embodiments may further comprise one or more additional functional domains, *e.g.*, an end-processing enzymatic domain of an end-processing enzyme that exhibits 5’-3’  
25 exonuclease, 5’-3’ alkaline exonuclease, 3’-5’ exonuclease (*e.g.*, Trex2), 5’ flap endonuclease, helicase, template-dependent DNA polymerases or template-independent DNA polymerase activity.

Illustrative examples of nuclease variants that bind and cleave a target sequence in the ALB gene include, but are not limited to engineered homing endonucleases (meganucleases) and megaTALs.

#### ***1. ENGINEERED HOMING ENDONUCLEASES (MEGANUCLEASES)***

5 In various embodiments, a homing endonuclease or meganuclease is reprogrammed to introduce a double-strand break (DSB) in a target site in a ALB gene. In particular embodiments, an engineered homing endonuclease introduces a DSB in intron 1 of a ALB gene, preferably at any one of SEQ ID NOs: 32-34 in intron 1 of an ALB gene, and more preferably at the sequence “TTAT” in SEQ ID NO: 34 in intron 1 of an ALB gene.

10 “Homing endonuclease” and “meganuclease” are used interchangeably and refer to naturally-occurring homing endonucleases that recognize 12-45 base-pair cleavage sites and are commonly grouped into five families based on sequence and structure motifs: LAGLIDADG, GIY-YIG, HNH, His-Cys box, and PD-(D/E)XK.

A “reference homing endonuclease” or “reference meganuclease” refers to a wild type  
15 homing endonuclease or a homing endonuclease found in nature. In one embodiment, a “reference homing endonuclease” refers to a wild type homing endonuclease that has been modified to increase basal activity (e.g., SEQ ID NOs: 2-5).

An “engineered homing endonuclease,” “engineered homing endonuclease variant,” “reprogrammed homing endonuclease,” “homing endonuclease variant,” “engineered  
20 meganuclease,” “reprogrammed meganuclease,” or “meganuclease variant” refers to a homing endonuclease comprising one or more DNA binding domains and one or more DNA cleavage domains, wherein the homing endonuclease has been designed and/or modified from a parental or naturally occurring homing endonuclease, to bind and cleave a DNA target sequence in a ALB gene. The engineered homing endonuclease may be designed and/or modified from a  
25 naturally occurring homing endonuclease or from another homing endonuclease variant. A “parental homing endonuclease” or “parental meganuclease” refers to the naturally occurring homing endonuclease or homing endonuclease variant from which the engineered homing endonuclease is designed and/or modified from.

Engineered homing endonucleases variants do not exist in nature and can be obtained by recombinant DNA technology or by random mutagenesis. Engineered homing endonucleases may be obtained by making one or more amino acid alterations, *e.g.*, mutating, substituting, adding, or deleting one or more amino acids, in a naturally occurring HE or HE variant. In particular embodiments, an engineered HE comprises one or more amino acid alterations to the DNA recognition interface.

Engineered homing endonucleases contemplated in particular embodiments may further comprise one or more linkers and/or additional functional domains, *e.g.*, an end-processing enzymatic domain of an end-processing enzyme that exhibits 5'-3' exonuclease, 5'-3' alkaline exonuclease, 3'-5' exonuclease (*e.g.*, Trex2), 5' flap endonuclease, helicase, template-dependent DNA polymerase or template-independent DNA polymerase activity. In particular embodiments, Engineered homing endonucleases are introduced into a T cell with an end-processing enzyme that exhibits 5'-3' exonuclease, 5'-3' alkaline exonuclease, 3'-5' exonuclease (*e.g.*, Trex2), 5' flap endonuclease, helicase, template-dependent DNA polymerase or template-independent DNA polymerase activity. The engineered homing endonuclease and 3' processing enzyme may be introduced separately, *e.g.*, in different vectors or separate mRNAs, or together, *e.g.*, as a fusion protein, or in a polycistronic construct separated by a viral self-cleaving peptide or an IRES element.

A "DNA recognition interface" refers to the HE amino acid residues that interact with nucleic acid target bases as well as those residues that are adjacent. For each HE, the DNA recognition interface comprises an extensive network of side chain-to-side chain and side chain-to-DNA contacts, most of which is necessarily unique to recognize a particular nucleic acid target sequence. Thus, the amino acid sequence of the DNA recognition interface corresponding to a particular nucleic acid sequence varies significantly and is a feature of any natural or engineered homing endonuclease variant. By way of non-limiting example, a HE variant contemplated in particular embodiments may be derived by constructing libraries of HE variants in which one or more amino acid residues localized in the DNA recognition interface of the natural HE (or a previously generated HE variant) are varied. The libraries may be screened for target cleavage activity against each predicted ALB target site using cleavage assays (*see e.g.*, Jarjour *et al.*, 2009. *Nuc. Acids Res.* **37**(20): 6871-6880).

LAGLIDADG homing endonucleases (LHE) are the most well studied family of homing endonucleases, are primarily encoded in archaea and in organellar DNA in green algae and fungi, and display the highest overall DNA recognition specificity. LHEs comprise one or two LAGLIDADG catalytic motifs per protein chain and function as homodimers or single chain monomers, respectively. Structural studies of LAGLIDADG proteins identified a highly conserved core structure (Stoddard 2005), characterized by an  $\alpha\beta\beta\alpha\beta\beta\alpha$  fold, with the LAGLIDADG motif belonging to the first helix of this fold. The highly efficient and specific cleavage of LHE's represent a protein scaffold to derive novel, highly specific endonucleases. However, engineering LHEs to bind and cleave a non-natural or non-canonical target site requires selection of the appropriate LHE scaffold, examination of the target locus, selection of putative target sites, and extensive alteration of the LHE to alter its DNA contact points and cleavage specificity, at up to two-thirds of the base-pair positions in a target site.

In one embodiment, LHEs from which reprogrammed LHEs or LHE variants may be designed include, but are not limited to I-OnuI, I-CreI, and I-SceI.

Illustrative examples of LHEs from which engineered homing endonucleases, reprogrammed LHEs, or LHE variants may be designed include, but are not limited to I-AabMI, I-AaeMI, I-AniI, I-ApaMI, I-CapIII, I-CapIV, I-CkaMI, I-CpaMI, I-CpaMII, I-CpaMIII, I-CpaMIV, I-CpaMV, I-CpaV, I-CraMI, I-EjeMI, I-GpeMI, I-GpiI, I-GzeMI, I-GzeMII, I-GzeMIII, I-HjeMI, I-LtrII, I-LtrI, I-LtrWI, I-MpeMI, I-MveMI, I-NcrII, I-NcrI, I-NcrMI, I-OheMI, I-OnuI, I-OsoMI, I-OsoMII, I-OsoMIII, I-OsoMIV, I-PanMI, I-PanMII, I-PanMIII, I-PanMIII, I-PnoMI, I-ScuMI, I-SmaMI, I-SscMI, and I-Vdi141I.

In one embodiment, the engineered homing endonuclease, reprogrammed LHE or LHE variant is selected from the group consisting of: an I-CpaMI variant, an I-HjeMI variant, an I-OnuI variant, an I-PanMI variant, and an I-SmaMI variant.

In one embodiment, the engineered homing endonuclease, reprogrammed LHE, or LHE variant is an engineered I-OnuI. *See e.g.*, SEQ ID NOs: 6-13.

In one embodiment, reprogrammed I-OnuI LHEs or I-OnuI variants targeting the ALB gene were generated from a natural I-OnuI or biologically active fragment thereof (SEQ ID NOs: 1-5). In a preferred embodiment, reprogrammed I-OnuI LHEs or I-OnuI variants targeting the human ALB gene were generated from an existing I-OnuI variant. In one

embodiment, reprogrammed I-OnuI LHEs were generated against a human ALB gene target site set forth in SEQ ID NO: 34.

In a particular embodiment, the reprogrammed I-OnuI LHE or I-OnuI variant that binds and cleaves a human ALB gene comprises one or more amino acid substitutions in the DNA recognition interface. In particular embodiments, the reprogrammed I-OnuI LHE that binds and cleaves a human ALB gene comprises at least 70%, at least 71%, at least 72%, at least 73%, at least 74%, at least 75%, at least 76%, at least 77%, at least 78%, at least 79%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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%, or at least 99% sequence identity with the DNA recognition interface of I-OnuI (Taekuchi *et al.* 2011. *Proc Natl Acad Sci U. S. A.* 2011 Aug 9; 108(32): 13077-13082) or comprises at least 70%, at least 71%, at least 72%, at least 73%, at least 74%, at least 75%, at least 76%, at least 77%, at least 78%, at least 79%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, 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%, or at least 99% sequence identity with the DNA recognition interface of an engineered I-OnuI LHE as set forth in any one of SEQ ID NOs: 6-13.

In one embodiment, the engineered I-OnuI LHE that binds and cleaves a human ALB gene comprises at least 70%, more preferably at least 80%, more preferably at least 85%, more preferably at least 90%, more preferably at least 95%, more preferably at least 97%, more preferably at least 99% sequence identity with the DNA recognition interface of I-OnuI (Taekuchi *et al.* 2011. *Proc Natl Acad Sci U. S. A.* 2011 Aug 9; 108(32): 13077-13082) or comprises at least 70%, more preferably at least 80%, more preferably at least 85%, more preferably at least 90%, more preferably at least 95%, more preferably at least 97%, more preferably at least 99% sequence identity with the DNA recognition interface of an engineered I-OnuI LHE as set forth in any one of SEQ ID NOs: 6-13.

In particular embodiments, the engineered I-OnuI LHE that binds and cleaves a human ALB gene comprises the DNA recognition interface of an engineered I-OnuI LHE as set forth in any one of SEQ ID NOs: 6-13.

In a particular embodiment, an engineered I-OnuI LHE that binds and cleaves a human ALB gene comprises one or more amino acid substitutions or modifications in the DNA recognition interface of an I-OnuI as set forth in any one of SEQ ID NOs: 1-13, biologically active fragments thereof, and/or further variants thereof.

- 5 In a particular embodiment, an engineered I-OnuI LHE that binds and cleaves a human ALB gene comprises one or more amino acid substitutions or modifications in the DNA recognition interface, particularly in the subdomains situated from positions 24 to 50, 68 to 82, 180 to 203 and 223 to 240 of I-OnuI (SEQ ID NOs: 1-3 and 5) or from positions 20 to 46, 64 to 78, 176 to 199 and 219 to 236 of I-OnuI as set forth in SEQ ID NO: 4 or comprises the DNA
- 10 recognition interface, particularly in the subdomains situated from positions 20 to 46, 64 to 78, 176 to 199 and 219 to 236 of an engineered I-OnuI as set forth in any one of SEQ ID NOs: 6-13, or biologically active fragments thereof.

- In particular embodiments, the engineered homing endonuclease comprises one or more amino acid substitutions in the DNA recognition interface at amino acid positions
- 15 selected from the group consisting of: 24, 26, 28, 30, 31, 32, 34, 35, 36, 37, 38, 40, 42, 44, 46, 48, 68, 70, 71, 75, 76, 78, 80, 180, 182, 184, 186, 189, 190, 191, 192, 193, 197, 199, 201, 203, 223, 225, 229, 232, 234, 236, and 238 of I-OnuI (SEQ ID NOs: 1-3 and 5) or at amino acid positions selected from the group consisting of: 20, 22, 24, 26, 27, 28, 30, 31, 32, 33, 34, 36, 38, 40, 42, 44, 64, 66, 67, 71, 72, 74, 76, 176, 178, 180, 182, 185, 186, 187, 188, 189, 193, 195,
- 20 197, 199, 219, 224, 225, 228, 230, 232, and 234 of I-OnuI as set forth in SEQ ID NO: 4.

- In particular embodiments, the engineered homing endonuclease comprises one or more amino acid residues in the DNA recognition interface at amino acid positions selected from the group consisting of: 20, 22, 24, 26, 27, 28, 30, 31, 32, 33, 34, 36, 38, 40, 42, 44, 64, 66, 67, 71, 72, 74, 76, 176, 178, 180, 182, 185, 186, 187, 188, 189, 193, 195, 197, 199, 219,
- 25 224, 225, 228, 230, 232, and 234 of an engineered I-OnuI as set forth in any one of SEQ ID NOs: 6-13.

In further particular embodiment, the engineered homing endonuclease comprises the amino acid residues at amino acid positions 20, 22, 24, 26, 27, 28, 30, 31, 32, 33, 34, 36, 38, 40, 42, 44, 64, 66, 67, 71, 72, 74, 76, 176, 178, 180, 182, 185, 186, 187, 188, 189, 193, 195,

197, 199, 219, 224, 225, 228, 230, 232, and 234 of an engineered I-OnuI as set forth in any one of SEQ ID NOs: 6-13.

In a particular embodiment, an I-OnuI LHE variant that binds and cleaves a human ALB gene comprises 5, 10, 15, 20, 25, 30, 35, 40 or more amino acid substitutions or  
 5 modifications in the DNA recognition interface, particularly in the subdomains situated from positions 24-50, 68-82, 180-203 and 223-240 of I-OnuI (SEQ ID NOs: 1-3 and 5) or from positions 20 to 46, 64 to 78, 176 to 199 and 219 to 236 of I-OnuI as set forth in SEQ ID NO: 4.

In a particular embodiment, an I-OnuI LHE variant that binds and cleaves a human ALB gene comprises 5, 10, 15, 20, 25, 30, 35, 40, 41, 42, or 43 amino acid substitutions or  
 10 modifications at amino acid positions selected from the group consisting of: 24, 26, 28, 30, 31, 32, 34, 35, 36, 37, 38, 40, 42, 44, 46, 48, 68, 70, 71, 75, 76, 78, 80, 180, 182, 184, 186, 189, 190, 191, 192, 193, 197, 199, 201, 203, 223, 225, 229, 232, 234, 236, and 238 of I-OnuI (SEQ ID NOs: 1-3 and 5) or at amino acid positions selected from the group consisting of: 20, 22, 24, 26, 27, 28, 30, 31, 32, 33, 34, 36, 38, 40, 42, 44, 64, 66, 67, 71, 72, 74, 76, 176, 178, 180,  
 15 182, 185, 186, 187, 188, 189, 193, 195, 197, 199, 219, 224, 225, 228, 230, 232, and 234 of I-OnuI as set forth in SEQ ID NO: 4.

In one embodiment, an engineered I-OnuI LHE that binds and cleaves a human ALB gene comprises one or more amino acid substitutions or modifications at additional positions situated anywhere within the entire I-OnuI sequence. The residues which may be substituted  
 20 and/or modified include but are not limited to amino acids that contact the nucleic acid target or that interact with the nucleic acid backbone or with the nucleotide bases, directly or via a water molecule. In one non-limiting example, an engineered I-OnuI LHE contemplated herein that binds and cleaves a human ALB gene comprises one or more substitutions and/or  
 25 modifications, preferably at least 5, preferably at least 10, preferably at least 15, preferably at least 20, more preferably at least 25, more preferably at least 30, even more preferably at least 35, or even more preferably at least 40 or more amino acid substitutions in at least one position selected from the position group consisting of positions: 24, 26, 28, 30, 31, 32, 34, 35, 36, 37, 38, 40, 42, 44, 46, 48, 68, 70, 71, 75, 76, 78, 80, 180, 182, 184, 186, 189, 190, 191, 192, 193, 197, 199, 201, 203, 223, 225, 229, 232, 234, 236, and 238 of I-OnuI (SEQ ID NOs: 1-3 and 5)  
 30 or in at least one position selected from the position group consisting of positions: 20, 22, 24,

26, 27, 28, 30, 31, 32, 33, 34, 36, 38, 40, 42, 44, 64, 66, 67, 71, 72, 74, 76, 176, 178, 180, 182, 185, 186, 187, 188, 189, 193, 195, 197, 199, 219, 224, 225, 228, 230, 232, and 234 of I-OnuI as set forth in SEQ ID NO: 4.

In certain embodiments, the engineered homing endonuclease cleaves a ALB intron 1 target site and comprises at least 5, at least 15, preferably at least 25, more preferably at least 35, or even more preferably at least 40 or more of the following amino acid substitutions: S24C, L26S, R28A, R28H, R28T, R28V, R30A, R30G, R30Q, N31K, N32A, N32F, N32L, N32G, N32C, N32M, K34G, S35W, S35R, S35G, S35K, S35C, S36T, V37T, G38R, S40N, S40T, S40S, S40G, S40V, S40L, E42K, G44V, G44M, Q46S, T48G, V68T, A70S, N71K, N75T, A76Q, S78R, K80T, C180S, F182M, N184I, I186A, K189R, S190T, K191G, L192A, G193R, Q197R, V199C, S201E, T203G, Y223H, K225S, K229L, F232S, W234F, D236I, and V238R of I-OnuI (SEQ ID NOs: 1-3 and 5) or a biologically active fragments thereof, and/or further variants thereof.

In particular embodiments, the engineered HE comprises at least 5, at least 15, preferably at least 25, more preferably at least 35, or even more preferably at least 40, or 43 of the corresponding following amino acid substitutions: S20C, L22S, R24A, R24H, R24T, R24V, R26A, R26G, R26Q, N27K, N28A, N28F, N28L, N28G, N28C, N28M, K30G, S31W, S31R, S31G, S31K, S31C, S32T, V33T, G34R, S36N, S36T, S36S, S36G, S36V, S36L, E38K, G40V, G40M, Q42S, T44G, V64T, A66S, N67K, N71T, A72Q, S74R, K76T, C176S, F178M, N180I, I182A, K185R, S186T, K187G, L188A, G189R, Q193R, V195C, S197E, T199G, Y219H, K221S, K225L, F228S, W230F, D232I, and V234R SEQ ID NO: 4, or a biologically active fragment thereof, and/or further variants thereof.

In some embodiments, the HE variant cleaves a ALB target site and comprises at least 5, at least 15, preferably at least 25, more preferably at least 35, or even more preferably at least 40 or more of the following amino acid substitutions: S24C, L26S, R28V, R30Q, N31K, N32A, K34G, S35W, S36T, V37T, G38R, S40N, E42K, G44V, Q46S, T48G, V68T, A70S, N71K, N75T, A76Q, S78R, K80T, C180S, F182M, N184I, I186A, K189R, S190T, K191G, L192A, G193R, Q197R, V199C, S201E, T203G, Y223H, K225S, K229L, F232S, W234F, D236I, and V238R of I-OnuI (SEQ ID NOs: 1-3 and 5) or a biologically active fragments thereof, and/or further variants thereof.



In certain embodiments, the HE variant cleaves a ALB target site and comprises at least 5, at least 15, preferably at least 25, more preferably at least 35, or even more preferably at least 40 or more of the following amino acid substitutions: S20C, L22S, R24V, R26Q, N27K, N28A, K30G, S31W, S32T, V33T, G34R, S36N, E38K, G40V, Q42S, T44G, V64T, A66S, N67K, N71T, A72Q, S74R, K76T, C176S, F178M, N180I, I182A, K185R, S186T, K187G, L188A, G189R, Q193R, V195C, S197E, T199G, Y219H, K221S, K225L, F228S, W230F, D232I, and V234R of SEQ ID NO: 4, or a biologically active fragment thereof, and/or further variants thereof.

In particular embodiments, the HE variant cleaves a ALB target site and comprises at least 5, at least 15, preferably at least 25, more preferably at least 35, or even more preferably at least 40 or more of the following amino acid substitutions: S24C, L26S, R28T, N31K, N32F, K34G, S35R, S36T, V37T, G38R, S40T, E42K, G44M, Q46S, T48G, V68T, A70S, N71K, N75T, A76Q, S78R, K80T, C180S, F182M, N184I, I186A, K189R, S190T, K191G, L192A, G193R, Q197R, V199C, S201E, T203G, Y223H, K225S, K229L, F232S, W234F, D236I, and V238R of I-OnuI (SEQ ID NOs: 1-3 and 5) or a biologically active fragments thereof, and/or further variants thereof.

In particular embodiments the HE variant cleaves a ALB target site and comprises at least 5, at least 15, preferably at least 25, more preferably at least 35, or even more preferably at least 40 or more of the following amino acid substitutions: S20C, L22S, R24T, N27K, N28F, K30G, S31R, S32T, V33T, G34R, S36T, E38K, G40M, Q42S, T44G, V64T, A66S, N67K, N71T, A72Q, S74R, K76T, C176S, F178M, N180I, I182A, K185R, S186T, K187G, L188A, G189R, Q193R, V195C, S197E, T199G, Y219H, K221S, K225L, F228S, W230F, D232I, and V234R of SEQ ID NO: 4 or a biologically active fragment thereof, and/or further variants thereof.

In particular embodiments, the HE variant cleaves a ALB target site and comprises at least 5, at least 15, preferably at least 25, more preferably at least 35, or even more preferably at least 40 or more of the following amino acid substitutions: S24C, L26S, R28A, R30G, N31K, N32L, K34G, S35G, S36T, V37T, G38R, E42K, G44V, Q46S, T48G, V68T, A70S, N71K, N75T, A76Q, S78R, K80T, C180S, F182M, N184I, I186A, K189R, S190T, K191G, L192A, G193R, Q197R, V199C, S201E, T203G, Y223H, K225S, K229L, F232S, W234F, D236I, and

V238R of I-OnuI (SEQ ID NOs: 1-3 and 5) or a biologically active fragments thereof, and/or further variants thereof.

In certain embodiments the HE variant cleaves a ALB target site and comprises at least 5, at least 15, preferably at least 25, more preferably at least 35, or even more preferably at least 40 or more of the following amino acid substitutions: S20C, L22S, R24A, R26G, N27K, N28L, K30G, S31G, S32T, V33T, G34R, E38K, G40V, Q42S, T44G, V64T, A66S, N67K, N71T, A72Q, S74R, K76T, C176S, F178M, N180I, I182A, K185R, S186T, K187G, L188A, G189R, Q193R, V195C, S197E, T199G, Y219H, K221S, K225L, F228S, W230F, D232I, and V234R of SEQ ID NO: 4 or a biologically active fragment thereof, and/or further variants thereof.

In particular embodiments, the HE variant cleaves a ALB target site and comprises at least 5, at least 15, preferably at least 25, more preferably at least 35, or even more preferably at least 40 or more of the following amino acid substitutions: S24C, L26S, R28H, R30A, N31K, N32F, K34G, S35K, S36T, V37T, G38R, S40G, E42K, G44V, Q46S, T48G, V68T, A70S, N71K, N75T, A76Q, S78R, K80T, C180S, F182M, N184I, I186A, K189R, S190T, K191G, L192A, G193R, Q197R, V199C, S201E, T203G, Y223H, K225S, K229L, F232S, W234F, D236I, and V238R of I-OnuI (SEQ ID NOs: 1-3 and 5) or a biologically active fragments thereof, and/or further variants thereof.

In particular embodiments, the HE variant cleaves a ALB target site and comprises at least 5, at least 15, preferably at least 25, more preferably at least 35, or even more preferably at least 40 or more of the following amino acid substitutions: S20C, L22S, R24H, R26A, N27K, N28F, K30G, S31K, S32T, V33T, G34R, S36G, E38K, G40V, Q42S, T44G, V64T, A66S, N67K, N71T, A72Q, S74R, K76T, C176S, F178M, N180I, I182A, K185R, S186T, K187G, L188A, G189R, Q193R, V195C, S197E, T199G, Y219H, K221S, K225L, F228S, W230F, D232I, and V234R of SEQ ID NO: 4 or a biologically active fragment thereof, and/or further variants thereof.

In particular embodiments, the HE variant cleaves a ALB target site and comprises at least 5, at least 15, preferably at least 25, more preferably at least 35, or even more preferably at least 40 or more of the following amino acid substitutions: S24C, L26S, R28V, R30A, N31K, N32G, K34G, S35G, S36T, V37T, G38R, S40T, E42K, G44V, Q46S, T48G, V68T, A70S,

N71K, N75T, A76Q, S78R, K80T, C180S, F182M, N184I, I186A, K189R, S190T, K191G, L192A, G193R, Q197R, V199C, S201E, T203G, Y223H, K225S, K229L, F232S, W234F, D236I, and V238R of I-OnuI (SEQ ID NOs: 1-3 and 5) or a biologically active fragments thereof, and/or further variants thereof.

- 5 In certain embodiments the HE variant cleaves a ALB target site and comprises at least 5, at least 15, preferably at least 25, more preferably at least 35, or even more preferably at least 40 or more of the following amino acid substitutions: S20C, L22S, R24V, R26A, N27K, N28G, K30G, S31G, S32T, V33T, G34R, S36T, E38K, G40V, Q42S, T44G, V64T, A66S, N67K, N71T, A72Q, S74R, K76T, C176S, F178M, N180I, I182A, K185R, S186T, K187G, 10 L188A, G189R, Q193R, V195C, S197E, T199G, Y219H, K221S, K225L, F228S, W230F, D232I, and V234R of SEQ ID NO: 4 or a biologically active fragment thereof, and/or further variants thereof.

- In particular embodiments, the HE variant cleaves a ALB target site and comprises at least 5, at least 15, preferably at least 25, more preferably at least 35, or even more preferably at least 15 least 40 or more of the following amino acid substitutions: S24C, L26S, R28A, R30G, N31K, 15 N32C, K34G, S35C, S36T, V37T, G38R, E42K, G44V, Q46S, T48G, V68T, A70S, N71K, N75T, A76Q, S78R, K80T, C180S, F182M, N184I, I186A, K189R, S190T, K191G, L192A, G193R, Q197R, V199C, S201E, T203G, Y223H, K225S, K229L, F232S, W234F, D236I, and V238R of I-OnuI (SEQ ID NOs: 1-3 and 5) or a biologically active fragments thereof, and/or 20 further variants thereof.

- In some embodiments, the HE variant cleaves a ALB target site and comprises at least 5, at least 15, preferably at least 25, more preferably at least 35, or even more preferably at least 40 or more of the following amino acid substitutions: S20C, L22S, R24A, R26G, N27K, N28C, K30G, S31C, S32T, V33T, G34R, E38K, G40V, Q42S, T44G, V64T, A66S, N67K, 25 N71T, A72Q, S74R, K76T, C176S, F178M, N180I, I182A, K185R, S186T, K187G, L188A, G189R, Q193R, V195C, S197E, T199G, Y219H, K221S, K225L, F228S, W230F, D232I, and V234R of SEQ ID NO: 4 or a biologically active fragment thereof, and/or further variants thereof.

- In particular embodiments, the HE variant cleaves a ALB target site and comprises at 30 least 5, at least 15, preferably at least 25, more preferably at least 35, or even more preferably at

least 40 or more of the following amino acid substitutions: S24C, L26S, R28A, R30G, N31K, N32L, K34G, S35A, S36T, V37T, G38R, S40V, E42K, G44V, Q46S, T48G, V68T, A70S, N71K, N75T, A76Q, S78R, K80T, C180S, F182M, N184I, I186A, K189R, S190T, K191G, L192A, G193R, Q197R, V199C, S201E, T203G, Y223H, K225S, K229L, F232S, W234F, D236I, and V238R of I-OnuI (SEQ ID NOs: 1-3 and 5) or a biologically active fragments thereof, and/or further variants thereof.

In certain embodiments, the HE variant cleaves a ALB target site and comprises at least 5, at least 15, preferably at least 25, more preferably at least 35, or even more preferably at least 40 or more of the following amino acid substitutions: S20C, L22S, R24A, R26G, N27K, N28L, K30G, S31A, S32T, V33T, G34R, S36V, E38K, G40V, Q42S, T44G, V64T, A66S, N67K, N71T, A72Q, S74R, K76T, C176S, F178M, N180I, I182A, K185R, S186T, K187G, L188A, G189R, Q193R, V195C, S197E, T199G, Y219H, K221S, K225L, F228S, W230F, D232I, and V234R of SEQ ID NO: 4 or a biologically active fragment thereof, and/or further variants thereof.

In particular embodiments, the HE variant cleaves a ALB target site and comprises at least 5, at least 15, preferably at least 25, more preferably at least 35, or even more preferably at least 40 or more of the following amino acid substitutions: S24C, L26S, R28H, R30G, N31K, N32M, K34G, S35R, S36T, V37T, G38R, S40L, E42K, G44V, Q46S, T48G, V68T, A70S, N71K, N75T, A76Q, S78R, K80T, C180S, F182M, N184I, I186A, K189R, S190T, K191G, L192A, G193R, Q197R, V199C, S201E, T203G, Y223H, K225S, K229L, F232S, W234F, D236I, and V238R of I-OnuI (SEQ ID NOs: 1-3 and 5) or a biologically active fragments thereof, and/or further variants thereof.

In particular embodiments, the HE variant cleaves a ALB target site and comprises at least 5, at least 15, preferably at least 25, more preferably at least 35, or even more preferably at least 40 or more of the following amino acid substitutions: S20C, L22S, R24H, R26G, N27K, N28M, K30G, S31R, S32T, V33T, G34R, S36L, E38K, G40V, Q42S, T44G, V64T, A66S, N67K, N71T, A72Q, S74R, K76T, C176S, F178M, N180I, I182A, K185R, S186T, K187G, L188A, G189R, Q193R, V195C, S197E, T199G, Y219H, K224S, K225L, F228S, W230F, D232I, and V234R of any one of SEQ ID NO: 4 or a biologically active fragment thereof, and/or further variants thereof.

The engineered HEs contemplated herein may further comprise scaffold mutations. In particular embodiments, the engineered HEs further comprise the following corresponding amino acid substitutions C115S, E121G, I125E, L138M, I153D, K156R, S159P, L160I, F168G, E178D, K207R, N246K, V261M, and L263H of SEQ ID NO: 1, or a biologically active fragment thereof.

In particular embodiments, the engineered HE further comprises: (a) a first DNA recognition interface comprising amino acid residues 20-46 of any one of SEQ ID NOs: 6-13, (b) a second DNA recognition interface comprising amino acid residues 64-78 of any one of SEQ ID NOs: 6-13, (c) a third DNA recognition interface comprising amino acid residues 176-199 of any one of SEQ ID NOs: 6-13, and (d) a fourth DNA recognition interface comprising amino acid residues 219-236 of any one of SEQ ID NOs: 6-13.

In certain embodiments, the engineered HE further comprising (a) a first DNA recognition interface comprising amino acid residues 20-46 of SEQ ID NO: 6, (b) a second DNA recognition interface comprising amino acid residues 64-78 of SEQ ID NO: 6, (c) a third DNA recognition interface comprising amino acid residues 176-199 of SEQ ID NO: 6, and (d) a fourth DNA recognition interface comprising amino acid residues 219-236 of SEQ ID NO: 6.

In certain embodiments, the engineered HE further comprising (a) a first DNA recognition interface comprising amino acid residues 20-46 of SEQ ID NO: 7, (b) a second DNA recognition interface comprising amino acid residues 64-78 of SEQ ID NO: 6, (c) a third DNA recognition interface comprising amino acid residues 176-199 of SEQ ID NO: 6, and (d) a fourth DNA recognition interface comprising amino acid residues 219-236 of SEQ ID NO: 6.

In certain embodiments, the engineered HE further comprising (a) a first DNA recognition interface comprising amino acid residues 20-46 of SEQ ID NO: 8, (b) a second DNA recognition interface comprising amino acid residues 64-78 of SEQ ID NO: 6, (c) a third DNA recognition interface comprising amino acid residues 176-199 of SEQ ID NO: 6, and (d) a fourth DNA recognition interface comprising amino acid residues 219-236 of SEQ ID NO: 6.

In certain embodiments, the engineered HE further comprising (a) a first DNA recognition interface comprising amino acid residues 20-46 of SEQ ID NO: 9, (b) a second DNA recognition interface comprising amino acid residues 64-78 of SEQ ID NO: 6, (c) a third

DNA recognition interface comprising amino acid residues 176-199 of SEQ ID NO: 6, and (d) a fourth DNA recognition interface comprising amino acid residues 219-236 of SEQ ID NO: 6.

In certain embodiments, the engineered HE further comprising (a) a first DNA recognition interface comprising amino acid residues 20-46 of SEQ ID NO: 10, (b) a second  
 5 DNA recognition interface comprising amino acid residues 64-78 of SEQ ID NO: 6, (c) a third DNA recognition interface comprising amino acid residues 176-199 of SEQ ID NO: 6, and (d) a fourth DNA recognition interface comprising amino acid residues 219-236 of SEQ ID NO: 6.

In certain embodiments, the engineered HE further comprising (a) a first DNA recognition interface comprising amino acid residues 20-46 of SEQ ID NO: 11, (b) a second  
 10 DNA recognition interface comprising amino acid residues 64-78 of SEQ ID NO: 6, (c) a third DNA recognition interface comprising amino acid residues 176-199 of SEQ ID NO: 6, and (d) a fourth DNA recognition interface comprising amino acid residues 219-236 of SEQ ID NO: 6.

In certain embodiments, the engineered HE further comprising (a) a first DNA recognition interface comprising amino acid residues 20-46 of SEQ ID NO: 12, (b) a second  
 15 DNA recognition interface comprising amino acid residues 64-78 of SEQ ID NO: 6, (c) a third DNA recognition interface comprising amino acid residues 176-199 of SEQ ID NO: 6, and (d) a fourth DNA recognition interface comprising amino acid residues 219-236 of SEQ ID NO: 6.

In certain embodiments, the engineered HE further comprising (a) a first DNA recognition interface comprising amino acid residues 20-46 of SEQ ID NO: 13, (b) a second  
 20 DNA recognition interface comprising amino acid residues 64-78 of SEQ ID NO: 6, (c) a third DNA recognition interface comprising amino acid residues 176-199 of SEQ ID NO: 6, and (d) a fourth DNA recognition interface comprising amino acid residues 219-236 of SEQ ID NO: 6.

In particular embodiments, an engineered I-OnuI that binds and cleaves a human ALB gene comprises an amino acid sequence that is at least 80%, preferably at least 85%, more  
 25 preferably at least 90%, or even more preferably at least 95% identical to the amino acid sequence set forth in any one of SEQ ID NOs: 6-13, or a biologically active fragment thereof.

In certain embodiments, an engineered I-OnuI that binds and cleaves a human ALB gene comprises an amino acid sequence that is at least 80%, preferably at least 85%, more preferably at least 90%, or even more preferably at least 95% identical to the amino acid  
 30 sequence set forth in SEQ ID NO: 6, or a biologically active fragment thereof.

In certain embodiments, an engineered I-OnuI that binds and cleaves a human ALB gene contemplated herein comprises an amino acid sequence that is at least 96%, at least 97%, at least 98%, or at least 99% identical to the amino acid sequence set forth in SEQ ID NO: 6, or a biologically active fragment thereof.

- 5           In certain embodiments, an engineered I-OnuI that binds and cleaves a human ALB gene comprises an amino acid sequence that is at least 80%, preferably at least 85%, more preferably at least 90%, or even more preferably at least 95% identical to the amino acid sequence set forth in SEQ ID NO: 7, or a biologically active fragment thereof.

- 10           In certain embodiments, an engineered I-OnuI that binds and cleaves a human ALB gene contemplated herein comprises an amino acid sequence that is at least 96%, at least 97%, at least 98%, or at least 99% identical to the amino acid sequence set forth in SEQ ID NO: 7, or a biologically active fragment thereof

- 15           In certain embodiments, an engineered I-OnuI that binds and cleaves a human ALB gene comprises an amino acid sequence that is at least 80%, preferably at least 85%, more preferably at least 90%, or even more preferably at least 95% identical to the amino acid sequence set forth in SEQ ID NO: 8, or a biologically active fragment thereof.

- 20           In certain embodiments, an engineered I-OnuI that binds and cleaves a human ALB gene contemplated herein comprises an amino acid sequence that is at least 96%, at least 97%, at least 98%, or at least 99% identical to the amino acid sequence set forth in SEQ ID NO: 8, or a biologically active fragment thereof

In certain embodiments, an engineered I-OnuI that binds and cleaves a human ALB gene comprises an amino acid sequence that is at least 80%, preferably at least 85%, more preferably at least 90%, or even more preferably at least 95% identical to the amino acid sequence set forth in SEQ ID NO: 9, or a biologically active fragment thereof.

- 25           In certain embodiments, an engineered I-OnuI that binds and cleaves a human ALB gene contemplated herein comprises an amino acid sequence that is at least 96%, at least 97%, at least 98%, or at least 99% identical to the amino acid sequence set forth in SEQ ID NO: 9, or a biologically active fragment thereof

- 30           In certain embodiments, an engineered I-OnuI that binds and cleaves a human ALB gene comprises an amino acid sequence that is at least 80%, preferably at least 85%, more

preferably at least 90%, or even more preferably at least 95% identical to the amino acid sequence set forth in SEQ ID NO: 10, or a biologically active fragment thereof.

In certain embodiments, an engineered I-OnuI that binds and cleaves a human ALB gene contemplated herein comprises an amino acid sequence that is at least 96%, at least 97%,  
5 at least 98%, or at least 99% identical to the amino acid sequence set forth in SEQ ID NO: 10, or a biologically active fragment thereof

In certain embodiments, an engineered I-OnuI that binds and cleaves a human ALB gene comprises an amino acid sequence that is at least 80%, preferably at least 85%, more preferably at least 90%, or even more preferably at least 95% identical to the amino acid  
10 sequence set forth in SEQ ID NO: 11, or a biologically active fragment thereof.

In certain embodiments, an engineered I-OnuI that binds and cleaves a human ALB gene contemplated herein comprises an amino acid sequence that is at least 96%, at least 97%, at least 98%, or at least 99% identical to the amino acid sequence set forth in SEQ ID NO: 11, or a biologically active fragment thereof

15 In certain embodiments, an engineered I-OnuI that binds and cleaves a human ALB gene comprises an amino acid sequence that is at least 80%, preferably at least 85%, more preferably at least 90%, or even more preferably at least 95% identical to the amino acid sequence set forth in SEQ ID NO: 12, or a biologically active fragment thereof.

In certain embodiments, an engineered I-OnuI that binds and cleaves a human ALB  
20 gene contemplated herein comprises an amino acid sequence that is at least 96%, at least 97%, at least 98%, or at least 99% identical to the amino acid sequence set forth in SEQ ID NO: 12, or a biologically active fragment thereof

In certain embodiments, an engineered I-OnuI that binds and cleaves a human ALB gene comprises an amino acid sequence that is at least 80%, preferably at least 85%, more  
25 preferably at least 90%, or even more preferably at least 95% identical to the amino acid sequence set forth in SEQ ID NO: 13, or a biologically active fragment thereof.

In certain embodiments, an engineered I-OnuI that binds and cleaves a human ALB gene contemplated herein comprises an amino acid sequence that is at least 96%, at least 97%, at least 98%, or at least 99% identical to the amino acid sequence set forth in SEQ ID NO: 13,  
30 or a biologically active fragment thereof



In particular embodiments, an I-OnuI LHE variant comprises an amino acid sequence set forth in any one of SEQ ID NOs: 6-13, or a biologically active fragment thereof.

In particular embodiments, an I-OnuI LHE variant comprises an amino acid sequence set forth in SEQ ID NO: 6, or a biologically active fragment thereof.

5 In particular embodiments, an I-OnuI LHE variant comprises an amino acid sequence set forth in SEQ ID NO: 7, or a biologically active fragment thereof.

In particular embodiments, an I-OnuI LHE variant comprises an amino acid sequence set forth in SEQ ID NO: 8, or a biologically active fragment thereof.

10 In particular embodiments, an I-OnuI LHE variant comprises an amino acid sequence set forth in SEQ ID NO: 9, or a biologically active fragment thereof.

In particular embodiments, an I-OnuI LHE variant comprises an amino acid sequence set forth in SEQ ID NO: 10, or a biologically active fragment thereof.

In particular embodiments, an I-OnuI LHE variant comprises an amino acid sequence set forth in SEQ ID NO: 11, or a biologically active fragment thereof.

15 In particular embodiments, an I-OnuI LHE variant comprises an amino acid sequence set forth in SEQ ID NO: 12, or a biologically active fragment thereof.

In particular embodiments, an I-OnuI LHE variant comprises an amino acid sequence set forth in SEQ ID NO: 13, or a biologically active fragment thereof.

## 2. *MEGATALS*

20 In various embodiments, a megaTAL comprising an engineered homing endonuclease is reprogrammed to introduce a double-strand break (DSB) in a target site in an ALB gene. In particular embodiments, a megaTAL introduces a DSB in intron 1 of an ALB gene, preferably at a target site as set forth in any one of SEQ ID NOs: 32-34 in intron 1 of an ALB gene, and more preferably at the sequence “TTAT” in SEQ ID NO: 34 in intron 1 of an ALB gene.

25 A “megaTAL” refers to a polypeptide comprising a TALE DNA binding domain and an engineered homing endonuclease that binds and cleaves a DNA target sequence in a ALB gene, and optionally comprises one or more linkers and/or additional functional domains, *e.g.*, an end-processing enzymatic domain of an end-processing enzyme that exhibits 5'-3' exonuclease, 5'-3' alkaline exonuclease, 3'-5' exonuclease (*e.g.*, Trex2), 5' flap endonuclease, 30 helicase or template-independent DNA polymerase activity.

In particular embodiments, a megaTAL can be introduced into a cell along with an end-processing enzyme that exhibits 5'-3' exonuclease, 5'-3' alkaline exonuclease, 3'-5' exonuclease (*e.g.*, Trex2), 5' flap endonuclease, helicase, template-dependent DNA polymerase, or template-independent DNA polymerase activity. The megaTAL and 3' processing enzyme may be introduced separately, *e.g.*, in different vectors or separate mRNAs, or together, *e.g.*, as a fusion protein, or in a polycistronic construct separated by a viral self-cleaving peptide or an IRES element.

A "TALE DNA binding domain" is the DNA binding portion of transcription activator-like effectors (TALE or TAL-effectors), which mimics plant transcriptional activators to manipulate the plant transcriptome (*see e.g.*, Kay *et al.*, 2007. *Science* 318:648-651). TALE DNA binding domains contemplated in particular embodiments are engineered *de novo* or from naturally occurring TALEs, *e.g.*, AvrBs3 from *Xanthomonas campestris* *pv. vesicatoria*, *Xanthomonas gardneri*, *Xanthomonas translucens*, *Xanthomonas axonopodis*, *Xanthomonas perforans*, *Xanthomonas alfalfa*, *Xanthomonas citri*, *Xanthomonas euvesicatoria*, and *Xanthomonas oryzae* and brg11 and hpx17 from *Ralstonia solanacearum*. Illustrative examples of TALE proteins for deriving and designing DNA binding domains are disclosed in U.S. Patent No. 9,017,967, and references cited therein, all of which are incorporated herein by reference in their entireties.

In particular embodiments, a megaTAL comprises a TALE DNA binding domain comprising one or more repeat units that are involved in binding of the TALE DNA binding domain to its corresponding target DNA sequence. A single "repeat unit" (also referred to as a "repeat") is typically 33-35 amino acids in length. Each TALE DNA binding domain repeat unit includes 1 or 2 DNA-binding residues making up the Repeat Variable Di-Residue (RVD), typically at positions 12 and/or 13 of the repeat. The natural (canonical) code for DNA recognition of these TALE DNA binding domains has been determined such that an HD sequence at positions 12 and 13 leads to a binding to cytosine (C), NG binds to T, NI to A, NN binds to G or A, and NG binds to T. In certain embodiments, non-canonical (atypical) RVDs are contemplated.

Illustrative examples of non-canonical RVDs suitable for use in particular megaTALs contemplated in particular embodiments include, but are not limited to HH, KH, NH, NK, NQ,

RH, RN, SS, NN, SN, KN for recognition of guanine (G); NI, KI, RI, HI, SI for recognition of adenine (A); NG, HG, KG, RG for recognition of thymine (T); RD, SD, HD, ND, KD, YG for recognition of cytosine (C); NV, HN for recognition of A or G; and H\*, HA, KA, N\*, NA, NC, NS, RA, S\* for recognition of A or T or G or C, wherein (\*) means that the amino acid at

5 position 13 is absent. Additional illustrative examples of RVDs suitable for use in particular megaTALs contemplated in particular embodiments further include those disclosed in U.S. Patent No. 8,614,092, which is incorporated herein by reference in its entirety.

In particular embodiments, a megaTAL contemplated herein comprises a TALE DNA binding domain comprising 3 to 30 repeat units. In certain embodiments, a megaTAL  
10 comprises 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 TALE DNA binding domain repeat units. In a preferred embodiment, a megaTAL contemplated herein comprises a TALE DNA binding domain comprising 5-15 repeat units, more preferably 7-15 repeat units, more preferably 9-15 repeat units, and more preferably 9, 10, 11, 12, 13, 14, or 15 repeat units.

15 In particular embodiments, a megaTAL contemplated herein comprises a TALE DNA binding domain comprising 3 to 30 repeat units and an additional single truncated TALE repeat unit comprising 20 amino acids located at the C-terminus of a set of TALE repeat units, *i.e.*, an additional C-terminal half-TALE DNA binding domain repeat unit (amino acids -20 to -1 of the C-cap disclosed elsewhere herein, *infra*). Thus, in particular embodiments, a megaTAL  
20 contemplated herein comprises a TALE DNA binding domain comprising 3.5 to 30.5 repeat units. In certain embodiments, a megaTAL comprises 3.5, 4.5, 5.5, 6.5, 7.5, 8.5, 9.5, 10.5, 11.5, 12.5, 13.5, 14.5, 15.5, 16.5, 17.5, 18.5, 19.5, 20.5, 21.5, 22.5, 23.5, 24.5, 25.5, 26.5, 27.5, 28.5, 29.5, or 30.5 TALE DNA binding domain repeat units. In a preferred embodiment, a megaTAL contemplated herein comprises a TALE DNA binding domain comprising 5.5-15.5  
25 repeat units, more preferably 7.5-15.5 repeat units, more preferably 9.5-15.5 repeat units, and more preferably 8.5, 9.5, 10.5, 11.5, 12.5, 13.5, 14.5, or 15.5 repeat units.

In particular embodiments, a megaTAL comprises a TAL effector architecture comprising an “N-terminal domain (NTD)” polypeptide, one or more TALE repeat domains/units, a “C-terminal domain (CTD)” polypeptide, and a homing endonuclease variant.  
30 In some embodiments, the NTD, TALE repeats, and/or CTD domains are from the same

species. In other embodiments, one or more of the NTD, TALE repeats, and/or CTD domains are from different species.

As used herein, the term “N-terminal domain (NTD)” polypeptide refers to the sequence that flanks the N-terminal portion or fragment of a naturally occurring TALE DNA binding domain. The NTD sequence, if present, may be of any length as long as the TALE DNA binding domain repeat units retain the ability to bind DNA. In particular embodiments, the NTD polypeptide comprises at least 120 to at least 140 or more amino acids N-terminal to the TALE DNA binding domain (0 is amino acid 1 of the most N-terminal repeat unit). In particular embodiments, the NTD polypeptide comprises at least about 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, or at least 140 amino acids N-terminal to the TALE DNA binding domain. In one embodiment, a megaTAL contemplated herein comprises an NTD polypeptide of at least about amino acids +1 to +122 to at least about +1 to +137 of a *Xanthomonas* TALE protein (0 is amino acid 1 of the most N-terminal repeat unit). In particular embodiments, the NTD polypeptide comprises at least about 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, or 137 amino acids N-terminal to the TALE DNA binding domain of a *Xanthomonas* TALE protein. In one embodiment, a megaTAL contemplated herein comprises an NTD polypeptide of at least amino acids +1 to +121 of a *Ralstonia* TALE protein (0 is amino acid 1 of the most N-terminal repeat unit). In particular embodiments, the NTD polypeptide comprises at least about 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, or 137 amino acids N-terminal to the TALE DNA binding domain of a *Ralstonia* TALE protein.

As used herein, the term “C-terminal domain (CTD)” polypeptide refers to the sequence that flanks the C-terminal portion or fragment of a naturally occurring TALE DNA binding domain. The CTD sequence, if present, may be of any length as long as the TALE DNA binding domain repeat units retain the ability to bind DNA. In particular embodiments, the CTD polypeptide comprises at least 20 to at least 85 or more amino acids C-terminal to the last full repeat of the TALE DNA binding domain (the first 20 amino acids are the half-repeat unit C-terminal to the last C-terminal full repeat unit). In particular embodiments, the CTD polypeptide comprises at least about 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, 51, 52, 53, 54, 55, 56, 57, 58, 59,

60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, or at least 85 amino acids C-terminal to the last full repeat of the TALE DNA binding domain. In one embodiment, a megaTAL contemplated herein comprises a CTD polypeptide of at least about amino acids -20 to -1 of a *Xanthomonas* TALE protein (-20 is amino acid 1 of a half-repeat unit C-terminal to the last C-terminal full repeat unit). In particular embodiments, the CTD polypeptide comprises at least about 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, or 1 amino acids C-terminal to the last full repeat of the TALE DNA binding domain of a *Xanthomonas* TALE protein. In one embodiment, a megaTAL contemplated herein comprises a CTD polypeptide of at least about amino acids -20 to -1 of a *Ralstonia* TALE protein (-20 is amino acid 1 of a half-repeat unit C-terminal to the last C-terminal full repeat unit). In particular embodiments, the CTD polypeptide comprises at least about 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, or 1 amino acids C-terminal to the last full repeat of the TALE DNA binding domain of a *Ralstonia* TALE protein.

In particular embodiments, a megaTAL contemplated herein, comprises a fusion polypeptide comprising a TALE DNA binding domain engineered to bind a target sequence, a homing endonuclease reprogrammed to bind and cleave a target sequence, and optionally an NTD and/or CTD polypeptide, optionally joined to each other with one or more linker polypeptides contemplated elsewhere herein. Without wishing to be bound by any particular theory, it is contemplated that a megaTAL comprising TALE DNA binding domain, and optionally an NTD and/or CTD polypeptide is fused to a linker polypeptide which is further fused to a homing endonuclease variant. Thus, the TALE DNA binding domain binds a DNA target sequence that is within about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, or 15 nucleotides away from the target sequence bound by the DNA binding domain of the homing endonuclease variant. In this way, the megaTALs contemplated herein, increase the specificity and efficiency of genome editing.

In one embodiment, a megaTAL comprises a homing endonuclease variant and a TALE DNA binding domain that binds a nucleotide sequence that is within about 2, 3, 4, 5, or 6 nucleotides upstream of the binding site of the reprogrammed homing endonuclease. In particular embodiments, a megaTAL comprises a homing endonuclease variant and a TALE

DNA binding domain that binds a nucleotide sequence that is 3 nucleotides upstream of the binding site of the reprogrammed homing endonuclease.

In one embodiment, a megaTAL comprises an engineered homing endonuclease variant and a TALE DNA binding domain that binds the nucleotide sequence set forth in any one of SEQ ID NOs: 35-42. In particular embodiments, a megaTAL comprises an engineered homing endonuclease variant and a TALE DNA binding domain that binds the nucleotide sequence set forth in SEQ ID NO: 38, which is 3 nucleotides upstream (*i.e.*, there are 2 nucleotides between the TALE binding site and the HE binding site) of the nucleotide sequence bound and cleaved by the engineered homing endonuclease (SEQ ID NO: 34). In particular  
 10   embodiments, the megaTAL target sequence is set forth in SEQ ID NO: 43.

In particular embodiments, a megaTAL contemplated herein, comprises one or more TALE DNA binding repeat units and an engineered LHE designed or reprogrammed from an LHE selected from the group consisting of: I-AabMI, I-AaeMI, I-AniI, I-ApaMI, I-CapIII, I-CapIV, I-CkaMI, I-CpaMI, I-CpaMII, I-CpaMIII, I-CpaMIV, I-CpaMV, I-CpaV, I-CraMI, I-EjeMI, I-GpeMI, I-GpiI, I-GzeMI, I-GzeMII, I-GzeMIII, I-HjeMI, I-LtrII, I-LtrI, I-LtrWI, I-MpeMI, I-MveMI, I-NcrII, I-NcrI, I-NcrMI, I-OheMI, I-OnuI, I-OsoMI, I-OsoMII, I-OsoMIII, I-OsoMIV, I-PanMI, I-PanMII, I-PanMIII, I-PnoMI, I-ScuMI, I-SmaMI, I-SscMI, I-Vdi141I and variants thereof, or preferably I-CpaMI, I-HjeMI, I-OnuI, I-PanMI, SmaMI and variants thereof, or more preferably I-OnuI and variants thereof.

In particular embodiments, a megaTAL contemplated herein, comprises an NTD, one or more TALE DNA binding repeat units, a CTD, and an engineered LHE is selected from the group consisting of: I-AabMI, I-AaeMI, I-AniI, I-ApaMI, I-CapIII, I-CapIV, I-CkaMI, I-CpaMI, I-CpaMII, I-CpaMIII, I-CpaMIV, I-CpaMV, I-CpaV, I-CraMI, I-EjeMI, I-GpeMI, I-GpiI, I-GzeMI, I-GzeMII, I-GzeMIII, I-HjeMI, I-LtrII, I-LtrI, I-LtrWI, I-MpeMI, I-MveMI, I-NcrII, I-NcrI, I-NcrMI, I-OheMI, I-OnuI, I-OsoMI, I-OsoMII, I-OsoMIII, I-OsoMIV, I-PanMI, I-PanMII, I-PanMIII, I-PnoMI, I-ScuMI, I-SmaMI, I-SscMI, I-Vdi141I and variants thereof, or preferably I-CpaMI, I-HjeMI, I-OnuI, I-PanMI, SmaMI and variants thereof, or more preferably I-OnuI and variants thereof.

In particular embodiments, a megaTAL contemplated herein, comprises an NTD, about  
 30   9.5 to about 15.5 TALE DNA binding repeat units, and an engineered LHE selected from the

group consisting of: I-AabMI, I-AaeMI, I-AniI, I-ApaMI, I-CapIII, I-CapIV, I-CkaMI, I-CpaMI, I-CpaMII, I-CpaMIII, I-CpaMIV, I-CpaMV, I-CpaV, I-CraMI, I-EjeMI, I-GpeMI, I-GpiI, I-GzeMI, I-GzeMII, I-GzeMIII, I-HjeMI, I-LtrII, I-LtrI, I-LtrWI, I-MpeMI, I-MveMI, I-NcrII, I-NcrI, I-NcrMI, I-OheMI, I-OnuI, I-OsoMI, I-OsoMII, I-OsoMIII, I-OsoMIV, I-PanMI, I-PanMII, I-PanMIII, I-PnoMI, I-ScuMI, I-SmaMI, I-SscMI, I-Vdi14II and variants thereof, or preferably I-CpaMI, I-HjeMI, I-OnuI, I-PanMI, SmaMI and variants thereof, or more preferably I-OnuI and variants thereof.

In particular embodiments, a megaTAL contemplated herein, comprises an NTD of about 122 amino acids to 137 amino acids, about 8.5, about 9.5, about 10.5, about 11.5, about 12.5, about 13.5, about 14.5, or about 15.5 binding repeat units, a CTD of about 20 amino acids to about 85 amino acids, and an engineered I-OnuI LHE. In particular embodiments, any one of, two of, or all of the NTD, DNA binding domain, and CTD can be designed from the same species or different species, in any suitable combination.

In further embodiments, a megaTAL contemplated herein further comprises a nuclear localization signal (NLS). An NLS is a sequence which provides subcellular localization for targeting to the nucleus. See, *e.g.*, Lu *et al.*, *Cell Commun Signal*, 2021 May 22;19(1):60 NLS(s). In certain embodiments, one or more NLS(s) are positioned at the 5' end of a polynucleotide encoding a megaTAL (*e.g.*, 5' to the engineered HE or megaTAL). In certain embodiments, one or more NLS(s) are position at the N-terminus of a megaTAL (*e.g.*, N-terminal to the engineered HE or the megaTAL). NLS sequences are well known in the art. An example of an NLS is EPPKRKKRKIGI (SEQ ID NO: 86).

In particular embodiments, a megaTAL contemplated herein comprises an amino acid sequence that is at least 80%, preferably at least 85%, more preferably at least 90%, or even more preferably at least 95% identical to the amino acid sequence set forth in any one of SEQ ID NOs: 14-21, or a biologically active fragment thereof.

In certain embodiments, a megaTAL contemplated herein comprises an amino acid sequence that is at least 96%, at least 97%, at least 98%, or at least 99% identical to the amino acid sequence set forth in any one of SEQ ID NOs: 14-21, or a biologically active fragment thereof.

In certain embodiments, a megaTAL contemplated herein comprises an amino acid sequence that is at least 80%, preferably at least 85%, more preferably at least 90%, or even more preferably at least 95% identical to the amino acid sequence set forth in SEQ ID NO: 14, or a biologically active fragment thereof.

- 5           In certain embodiments, a megaTAL contemplated herein comprises an amino acid sequence that is at least 96%, at least 97%, at least 98%, or at least 99% identical to the amino acid sequence set forth in SEQ ID NO: 14, or a biologically active fragment thereof.

- In certain embodiments, a megaTAL contemplated herein comprises an amino acid sequence that is at least 80%, preferably at least 85%, more preferably at least 90%, or even  
10   more preferably at least 95% identical to the amino acid sequence set forth in SEQ ID NO: 15, or a biologically active fragment thereof.

In certain embodiments, a megaTAL contemplated herein comprises an amino acid sequence that is at least 96%, at least 97%, at least 98%, or at least 99% identical to the amino acid sequence set forth in SEQ ID NO: 15, or a biologically active fragment thereof.

- 15           In certain embodiments, a megaTAL contemplated herein comprises an amino acid sequence that is at least 80%, preferably at least 85%, more preferably at least 90%, or even more preferably at least 95% identical to the amino acid sequence set forth in SEQ ID NO: 16, or a biologically active fragment thereof.

- In certain embodiments, a megaTAL contemplated herein comprises an amino acid  
20   sequence that is at least 96%, at least 97%, at least 98%, or at least 99% identical to the amino acid sequence set forth in SEQ ID NO: 16, or a biologically active fragment thereof.

- In certain embodiments, a megaTAL contemplated herein comprises an amino acid sequence that is at least 80%, preferably at least 85%, more preferably at least 90%, or even more preferably at least 95% identical to the amino acid sequence set forth in SEQ ID NO: 17,  
25   or a biologically active fragment thereof.

In certain embodiments, a megaTAL contemplated herein comprises an amino acid sequence that is at least 96%, at least 97%, at least 98%, or at least 99% identical to the amino acid sequence set forth in SEQ ID NO: 17, or a biologically active fragment thereof.

- In certain embodiments, a megaTAL contemplated herein comprises an amino acid  
30   sequence that is at least 80%, preferably at least 85%, more preferably at least 90%, or even



more preferably at least 95% identical to the amino acid sequence set forth in SEQ ID NO: 18, or a biologically active fragment thereof.

In certain embodiments, a megaTAL contemplated herein comprises an amino acid sequence that is at least 96%, at least 97%, at least 98%, or at least 99% identical to the amino acid sequence set forth in SEQ ID NO: 18, or a biologically active fragment thereof.

In certain embodiments, a megaTAL contemplated herein comprises an amino acid sequence that is at least 80%, preferably at least 85%, more preferably at least 90%, or even more preferably at least 95% identical to the amino acid sequence set forth in SEQ ID NO: 19, or a biologically active fragment thereof.

In certain embodiments, a megaTAL contemplated herein comprises an amino acid sequence that is at least 96%, at least 97%, at least 98%, or at least 99% identical to the amino acid sequence set forth in SEQ ID NO: 19, or a biologically active fragment thereof.

In certain embodiments, a megaTAL contemplated herein comprises an amino acid sequence that is at least 80%, preferably at least 85%, more preferably at least 90%, or even more preferably at least 95% identical to the amino acid sequence set forth in SEQ ID NO: 20, or a biologically active fragment thereof.

In certain embodiments, a megaTAL contemplated herein comprises an amino acid sequence that is at least 96%, at least 97%, at least 98%, or at least 99% identical to the amino acid sequence set forth in SEQ ID NO: 20, or a biologically active fragment thereof.

In certain embodiments, a megaTAL contemplated herein comprises an amino acid sequence that is at least 80%, preferably at least 85%, more preferably at least 90%, or even more preferably at least 95% identical to the amino acid sequence set forth in SEQ ID NO: 21, or a biologically active fragment thereof.

In certain embodiments, a megaTAL contemplated herein comprises an amino acid sequence that is at least 96%, at least 97%, at least 98%, or at least 99% identical to the amino acid sequence set forth in SEQ ID NO: 21, or a biologically active fragment thereof.

In particular embodiments, a megaTAL contemplated herein, comprises the amino acid sequence set forth in any one of SEQ ID NOs: 14-21 or a biologically active fragment thereof.

In particular embodiments, a megaTAL comprises an amino acid sequence set forth in SEQ ID NO: 14, or a biologically active fragment thereof.

In particular embodiments, a megaTAL comprises an amino acid sequence set forth in SEQ ID NO: 15, or a biologically active fragment thereof.

In particular embodiments, a megaTAL comprises an amino acid sequence set forth in SEQ ID NO: 16, or a biologically active fragment thereof.

5 In particular embodiments, a megaTAL comprises an amino acid sequence set forth in SEQ ID NO: 17, or a biologically active fragment thereof.

In particular embodiments, a megaTAL comprises an amino acid sequence set forth in SEQ ID NO: 18, or a biologically active fragment thereof.

In particular embodiments, a megaTAL comprises an amino acid sequence set forth in  
10 SEQ ID NO: 19, or a biologically active fragment thereof.

In particular embodiments, a megaTAL comprises an amino acid sequence set forth in SEQ ID NO: 21, or a biologically active fragment thereof.

In particular embodiments, a megaTAL comprises an amino acid sequence set forth in SEQ ID NO: 22, or a biologically active fragment thereof.

15 In particular embodiments, a megaTAL comprises an amino acid sequence set forth in SEQ ID NO: 108, or a biologically active fragment thereof.

In certain embodiments, a megaTAL comprises a TALE DNA binding domain and an I-OnuI LHE variant binds and cleaves the nucleotide sequence set forth in SEQ ID NO: 43.

### 3. *END-PROCESSING ENZYMES*

20 Genome editing compositions and methods contemplated in particular embodiments comprise editing cellular genomes using an engineered nuclease and one or more copies of an end-processing enzyme. In particular embodiments, a single polynucleotide encodes a homing endonuclease variant and an end-processing enzyme, separated by a linker, a self-cleaving peptide sequence, *e.g.*, 2A sequence, or by an IRES sequence. In particular embodiments,  
25 genome editing compositions comprise a polynucleotide encoding an engineered nuclease and a separate polynucleotide encoding an end-processing enzyme. In particular embodiments, genome editing compositions comprise a polynucleotide encoding an engineered homing endonuclease and an end-processing enzyme in a single fusion polypeptide. In one embodiment, a fusion polypeptide comprises a megaTAL and one or more copies of an end-  
30 processing enzyme, each separated by a self-cleaving peptide.

The term “end-processing enzyme” refers to an enzyme that modifies the exposed ends of a polynucleotide chain. The polynucleotide may be double-stranded DNA (dsDNA), single-stranded DNA (ssDNA), RNA, double-stranded hybrids of DNA and RNA, and synthetic DNA (for example, containing bases other than A, C, G, and T). An end-processing enzyme may modify exposed polynucleotide chain ends by adding one or more nucleotides, removing one or more nucleotides, removing or modifying a phosphate group and/or removing or modifying a hydroxyl group. An end-processing enzyme may modify ends at endonuclease cut sites or at ends generated by other chemical or mechanical means, such as shearing (for example by passing through fine-gauge needle, heating, sonicating, mini bead tumbling, and nebulizing), ionizing radiation, ultraviolet radiation, oxygen radicals, chemical hydrolysis and chemotherapy agents.

In particular embodiments, genome editing compositions and methods contemplated in particular embodiments comprise editing cellular genomes using an engineered homing endonuclease or megaTAL and a DNA end-processing enzyme.

The term “DNA end-processing enzyme” refers to an enzyme that modifies the exposed ends of DNA. A DNA end-processing enzyme may modify blunt ends or staggered ends (ends with 5' or 3' overhangs). A DNA end-processing enzyme may modify single stranded or double stranded DNA. A DNA end-processing enzyme may modify ends at endonuclease cut sites or at ends generated by other chemical or mechanical means, such as shearing (for example by passing through fine-gauge needle, heating, sonicating, mini bead tumbling, and nebulizing), ionizing radiation, ultraviolet radiation, oxygen radicals, chemical hydrolysis and chemotherapy agents. DNA end-processing enzyme may modify exposed DNA ends by adding one or more nucleotides, removing one or more nucleotides, removing or modifying a phosphate group and/or removing or modifying a hydroxyl group.

Illustrative examples of DNA end-processing enzymes suitable for use in particular embodiments contemplated herein include, but are not limited to: 5'-3' exonucleases, 5'-3' alkaline exonucleases, 3'-5' exonucleases, 5' flap endonucleases, helicases, phosphatases, hydrolases and template-independent DNA polymerases.

Additional illustrative examples of DNA end-processing enzymes suitable for use in particular embodiments contemplated herein include, but are not limited to, Trex2, Trex1,

- Trex1 without transmembrane domain, Apollo, Artemis, DNA2, Exo1, ExoT, ExoIII, Fen1, Fan1, MreII, Rad2, Rad9, TdT (terminal deoxynucleotidyl transferase), PNKP, RecE, RecJ, RecQ, Lambda exonuclease, Sox, Vaccinia DNA polymerase, exonuclease I, exonuclease III, exonuclease VII, NDK1, NDK5, NDK7, NDK8, WRN, T7-exonuclease Gene 6, avian
- 5 myeloblastosis virus integration protein (IN), Bloom, Antarctic Phosphatase, Alkaline Phosphatase, Poly nucleotide Kinase (PNK), ApeI, Mung Bean nuclease, Hex1, TTRAP (TDP2), Sgs1, Sae2, CUP, Pol mu, Pol lambda, MUS81, EME1, EME2, SLX1, SLX4 and UL-
- 12.

- In particular embodiments, genome editing compositions and methods for editing
- 10 cellular genomes contemplated herein comprise polypeptides comprising a homing endonuclease variant or megaTAL and an exonuclease. The term “exonuclease” refers to enzymes that cleave phosphodiester bonds at the end of a polynucleotide chain via a hydrolyzing reaction that breaks phosphodiester bonds at either the 3’ or 5’ end.

- Illustrative examples of exonucleases suitable for use in particular embodiments
- 15 contemplated herein include, but are not limited to: hExoI, Yeast ExoI, *E. coli* ExoI, hTREX2, mouse TREX2, rat TREX2, hTREX1, mouse TREX1, and rat TREX1.

In particular embodiments, the DNA end-processing enzyme is a 3’ or 5’ exonuclease, preferably Trex 1 or Trex2, more preferably Trex2, and even more preferably human or mouse Trex2.

## 20 D. TARGET SITES

- Engineered nucleases contemplated in particular embodiments can be designed to bind to any suitable target sequence and can have a novel binding specificity, compared to a naturally-occurring nuclease. In particular embodiments, the target site is a regulatory region of a gene including, but not limited to promoters, enhancers, repressor elements, and the like.
- 25 In particular embodiments, the target site is a coding or non-coding region of a gene or a splice site. In particular embodiments, the target site is a non-coding region of a gene (*e.g.*, intron 1 of ALB). In certain embodiments, engineered nucleases are designed to down-regulate or decrease expression of a gene. In particular embodiments, a nuclease variant and donor repair template can be designed to repair or delete a desired target sequence.

- In various embodiments, nuclease variants bind to and cleave a target sequence in a human albumin (ALB) gene. Human albumin (ALB) is also referred to as serum albumin (HGNC: 399 NCBI Gene: 213 Ensembl: ENSG00000163631 OMIM®: 103600 UniProtKB/Swiss-Prot: P02768). This gene encodes the most abundant protein in human blood and makes up about 50% of human plasma proteins. Albumin functions in the regulation of blood plasma colloid osmotic pressure and acts as a carrier protein for a wide range of endogenous molecules including hormones, fatty acids, and metabolites, as well as exogenous drugs, and exhibits an esterase-like activity with broad substrate specificity. The encoded preproprotein is proteolytically processed to generate the mature protein.
- 10 In particular embodiments, an engineered homing endonuclease or megaTAL introduces a double-strand break (DSB) in a target site in an ALB gene. In particular embodiments, an engineered homing endonuclease or megaTAL introduces a DSB in intron 1 of an ALB gene, preferably at any one of SEQ ID NOs: 32-34 in intron 1 of an ALB gene, and more preferably at the sequence “TTAT” in SEQ ID NO: 34 in intron 1 of a
- 15 CBLB gene.

In a preferred embodiment, an engineered homing endonuclease or megaTAL cleaves double-stranded DNA and introduces a DSB into the polynucleotide sequence set forth in SEQ ID NO: 34 or 43.

In a preferred embodiment, the ALB gene is a human ALB gene.

## 20 E. DONOR REPAIR TEMPLATES AND THERAPEUTIC POLYPEPTIDES

Engineered nucleases may be used to introduce a DSB in a target sequence; the DSB may be repaired through non-homologous end-joining (NHEJ) or homology directed repair (HDR) mechanisms in the presence of one or more donor repair templates.

- In various embodiments, the donor repair template comprises one or more
- 25 polynucleotides or transgenes encoding a therapeutic polypeptide/protein.

In particular embodiments, the donor repair template is used to insert a sequence into the genome, also referred to as targeted integration. In particular preferred embodiments, the donor repair template is used to repair or modify a sequence in the genome.

In various embodiments, a donor repair template is introduced into a cell, *e.g.*, a hepatocyte, by transducing the cell with an adeno-associated virus (AAV), retrovirus, *e.g.*, lentivirus, IDLV, *etc.*, herpes simplex virus, adenovirus, or vaccinia virus vector comprising the donor repair template.

- 5           In particular embodiments, the donor repair template comprises one or more homology arms that flank the DSB site. In particular embodiments, the donor repair template does not comprise one or more homology arms that flank the DSB site.

As used herein, the term “homology arms” refers to a nucleic acid sequence in a donor repair template that is identical, or nearly identical, to DNA sequence flanking the DNA break  
 10 introduced by the nuclease at a target site. In one embodiment, the donor repair template comprises a 5' homology arm that comprises a nucleic acid sequence that is identical or nearly identical to the DNA sequence 5' of the DNA break site. In one embodiment, the donor repair template comprises a 3' homology arm that comprises a nucleic acid sequence that is identical or nearly identical to the DNA sequence 3' of the DNA break site. In a preferred embodiment,  
 15 the donor repair template comprises a 5' homology arm and a 3' homology arm. The donor repair template may comprise homology to the genome sequence immediately adjacent to the DSB site, or homology to the genomic sequence within any number of base pairs from the DSB site. In particular embodiments, a pair of homology arms comprises a homology arm comprising a polynucleotide sequence that includes a target site for a double strand break with  
 20 a mutation in the target site to minimize re-cleavage of the target site. In one embodiment, the donor repair template comprises a nucleic acid sequence that is homologous to a genomic sequence or homology arm of about 5 bp, about 10 bp, about 25 bp, about 50 bp, about 100 bp, about 250 bp, about 500 bp, about 1000 bp, about 2500 bp, about 5000 bp, about 10000 bp or more, including any intervening length of homologous sequence.

- 25           Illustrative examples of suitable lengths of homology arms contemplated in particular embodiments, may be independently selected, and include but are not limited to: 5 bp, about 10 bp, about 25 bp, about 50 bp, about 100 bp, about 200 bp, about 300 bp, about 400 bp, about 500 bp, about 600 bp, about 700 bp, about 800 bp, about 900 bp, about 1000 bp, about 1100 bp, about 1200 bp, about 1300 bp, about 1400 bp, about 1500 bp, about 1600 bp, about  
 30 1700 bp, about 1800 bp, about 1900 bp, about 2000 bp, about 2100 bp, about 2200 bp, about

2300 bp, about 2400 bp, about 2500 bp, about 2600 bp, about 2700 bp, about 2800 bp, about 2900 bp, or about 3000 bp, or longer homology arms, including all intervening lengths of homology arms.

Additional illustrative examples of suitable homology arm lengths include, but are not limited to: about 100 bp to about 3000 bp, about 200 bp to about 3000 bp, about 300 bp to about 3000 bp, about 400 bp to about 3000 bp, about 500 bp to about 3000 bp, about 500 bp to about 2500 bp, about 500 bp to about 2000 bp, about 750 bp to about 2000 bp, about 750 bp to about 1500 bp, or about 1000 bp to about 1500 bp, including all intervening lengths of homology arms.

In a particular embodiment, the lengths of the 5' and 3' homology arms are independently selected from about 500 bp to about 1500 bp. In one embodiment, the 5' homology arm is about 1500 bp and the 3' homology arm is about 1000 bp. In one embodiment, the 5' homology arm is between about 200 bp to about 600 bp and the 3' homology arm is between about 200 bp to about 600 bp. In one embodiment, the 5' homology arm is about 200 bp and the 3' homology arm is about 200 bp. In one embodiment, the 5' homology arm is about 300 bp and the 3' homology arm is about 300 bp. In one embodiment, the 5' homology arm is about 400 bp and the 3' homology arm is about 400 bp. In one embodiment, the 5' homology arm is about 500 bp and the 3' homology arm is about 500 bp. In one embodiment, the 5' homology arm is about 600 bp and the 3' homology arm is about 600 bp.

Donor repair templates may comprise one or more expression cassettes in particular embodiments. In certain embodiments, donor repair templates may comprise one or more polynucleotides including, but not limited to promoters and/or enhancers, untranslated regions (UTRs), Kozak sequences, polyadenylation signals, additional restriction enzyme sites, multiple cloning sites, internal ribosomal entry sites (IRES), recombinase recognition sites (*e.g.*, LoxP, FRT, and Att sites), termination codons, transcriptional termination signals, polynucleotides encoding self-cleaving polypeptides, and epitope tags.

In various embodiments, the donor repair template comprises one or more polynucleotides encoding a therapeutic polypeptide/protein, wherein the donor repair template

does not comprise homology arms and expression of the one or more polynucleotides is governed by the endogenous ALB promoter.

In various embodiments, the donor repair template comprises a 5' homology arm, one or more polynucleotides encoding a therapeutic polypeptide/protein, and a 3' homology arm, wherein expression of the one or more polynucleotides is governed by the endogenous ALB promoter.

In various embodiments, the donor repair template comprises an RNA polymerase II promoter and one or more polynucleotides encoding therapeutic polypeptide/protein.

In various embodiments, the donor repair template comprises a 5' homology arm, an RNA polymerase II promoter, one or more polynucleotides encoding a therapeutic polypeptide/protein, and a 3' homology arm.

In various embodiments, the donor repair template further comprises a poly(A) signal.

In particular embodiments, the therapeutic polypeptide is a therapeutic antihemophilic factor, antibody, protein, enzyme, cytokine, chemokine, cytotoxin, cytokine receptor, engineered or chimeric antigen receptor, TCR, hormone, or functional variants thereof.

In various embodiments, any therapeutic transgene(s) or therapeutic polypeptide can be expressed using the polynucleotides described herein, including, but not limited to, therapeutic transgenes or therapeutic polypeptides encoding functional versions of proteins lacking or deficient in any genetic disease, including but not limited to, lysosomal storage disorders (*e.g.*, Gaucher's, Fabry's, Hunter's, Hurler's, Neimann-Pick's, Phenylketonuria (PKU) *etc.*), metabolic disorders, and/or blood disorders such as hemophilias and hemoglobinopathies, *etc.* See, *e.g.*, U.S. Publication No. 20140017212 and 20140093913; U.S. Pat. Nos. 9,255,250 and 9,175,280.

Non-limiting examples of therapeutic transgenes, proteins, and/or therapeutic polypeptides that may be expressed as described herein include fibrinogen, prothrombin, tissue factor, Factor V, Factor VII, Factor VIII, including FVIII-BDV, Factor IX, Factor X, Factor XI, Factor XII (Hageman factor), Factor XIII (fibrin-stabilizing factor), Factor XIIIa, von Willebrand factor, prekallikrein, high molecular weight kininogen (Fitzgerald



- factor), fibronectin, antithrombin III, heparin cofactor II, protein C, protein S, protein Z, protein Z-related protease inhibitor, plasminogen, alpha 2-antiplasmin, tissue plasminogen activator, urokinase, plasminogen activator inhibitor-1, plasminogen activator inhibitor-2, glucocerebrosidase (GBA),  $\alpha$ -galactosidase A (GLA), iduronate sulfatase (IDS),
- 5 iduronidase (IDUA), acid sphingomyelinase (SMPD1), MMAA, MMAB, MMACHC, MMADHC (C2orf25), MTRR, LMBRD1, MTR, propionyl-CoA carboxylase (PCC) (PCCA and/or PCCB subunits), a glucose-6-phosphate transporter (G6PT) protein or glucose-6-phosphatase (G6Pase), an LDL receptor (LDLR), ApoB, LDLRAP-1, a PCSK9, a mitochondrial protein such as NAGS (N-acetylglutamate synthetase), CPS1 (carbamoyl
- 10 phosphate synthetase I), and OTC (ornithine transcarbamylase), ASS (argininosuccinic acid synthetase), ASL (argininosuccinase acid lyase) and/or ARG1 (arginase), and/or a solute carrier family 25 (SLC25A13, an aspartate/glutamate carrier) protein, a UGT1A1 or UDP glucuronyltransferase polypeptide A1, a fumarylacetoacetate hydrolyase (FAH), an alanine-glyoxylate aminotransferase (AGXT) protein, a glyoxylate
- 15 reductase/hydroxypyruvate reductase (GRHPR) protein, a transthyretin gene (TTR) protein, an ATP7B protein, a phenylalanine hydroxylase (PAH) protein, a lipoprotein lyase (LPL) protein, an engineered nuclease, an engineered transcription factor, an antibody, a single chain variable fragment antibody (scFv, diabody, VHH, *etc.*), enzyme, cytokine, chemokine, cytotoxin, cytokine receptor, engineered or chimeric antigen receptor, TCR,
- 20 hormone, and/or functional variants thereof.

In various embodiments, the therapeutic transgene or therapeutic polypeptide is a Factor VIII (FVIII), Factor IX (FIX), Factor X (FX), Factor XI (FXI), Factor XII (FXII), Factor FXIII (FXIII) polypeptide, or functional variants thereof.

- In various embodiments, the therapeutic transgene or therapeutic polypeptide is a
- 25 Factor VIII (FVIII) or Factor IX (FIX) polypeptide, or functional variant thereof. In various embodiments, the therapeutic transgene or therapeutic polypeptide is a Factor VIII (FVIII) polypeptide, or functional variant thereof, *e.g.*, a modified FVIII.

A "modified FVIII" refers to a FVIII protein with a modified amino acid sequence compared to the native sequence. In particular embodiments, the modified FVIII

comprises an altered B domain, referred to as B-domain variant FVIII (FVIII-BDV). In some embodiments, the modified FVIII (e.g., FVIII-BDV) comprises a shortened B domain. In some embodiments, the modified FVIII comprises a deletion of the B domain (FVIII-BDD). As used herein “FVIII” or “Factor VIII” encompasses functional FVIII  
 5 protein, including modified FVIII, such as FVIII-BDV.

Therapeutic transgenes encoding, or therapeutic polypeptides comprising, functional FVIII or FIX proteins have been described. (See, e.g., U.S. Patent Nos. 6,936,243; 7,238,346 and 6,200,560; Shi *et al.* (2007) *J Thromb Haemost.* (2):352-61; Lee *et al.* (2004) *Pharm. Res.* 7:1229-1232; Graham *et al.* (2008) *Genet Vaccines Ther.* 3:6-9;  
 10 Manno *et al.* (2003) *Blood* 101(8): 2963-72; Manno *et al.* (2006) *Nature Medicine* 12(3):342-7; Nathwani *et al.* (2011) *Mol Ther* 19(5): 876-85; Nathwani *et al.* (2011); N Engl J Med. 365(25): 2357-65 and McIntosh *et al.* (2013) *Blood* 121(17):3335-44).

#### **F. POLYPEPTIDES**

Various polypeptides are contemplated herein, including, but not limited to,  
 15 engineered homing endonucleases, megaTALs, end-processing enzymes, fusion polypeptides, and therapeutic polypeptides, or one or more biologically active fragments or variants thereof. In preferred embodiments, a polypeptide comprises the amino acid sequence set forth in SEQ ID NOs: 1-21, 44-81, and 86-107. “Polypeptide,” “polypeptide fragment,” “peptide” and “protein” are used interchangeably, unless specified to the contrary, and according to  
 20 conventional meaning, *i.e.*, as a sequence of amino acids. In one embodiment, a “polypeptide” includes fusion polypeptides and other variants. Polypeptides can be prepared using any of a variety of well-known recombinant and/or synthetic techniques. Polypeptides are not limited to a specific length, *e.g.*, they may comprise a full-length protein sequence, a fragment of a full-length protein, or a fusion protein, and may include post-translational modifications of the  
 25 polypeptide, for example, glycosylations, acetylations, phosphorylations and the like, as well as other modifications known in the art, both naturally occurring and non-naturally occurring.

An “isolated protein,” “isolated peptide,” or “isolated polypeptide” and the like, as used herein, refer to *in vitro* synthesis, isolation, and/or purification of a peptide or polypeptide

molecule from a cellular environment, and from association with other components of the cell, *i.e.*, it is not significantly associated with *in vivo* substances.

Illustrative examples of polypeptides contemplated in particular embodiments include, but are not limited to engineered homing endonucleases, megaTALs, antihemophilic factors  
5 (e.g., FVIII or FIX), antibodies, BiTEs, recombinant proteins, enzymes, cytokines, chemokines, cytotoxins, cytokine receptors, engineered or chimeric antigen receptors, TCRs, zetakines, hormones, or functional variants thereof.

Polypeptides include “polypeptide variants.” Polypeptide variants may differ from a naturally occurring polypeptide in one or more amino acid substitutions, deletions, additions  
10 and/or insertions. Such variants may be naturally occurring or may be synthetically generated, for example, by modifying one or more amino acids of the above polypeptide sequences. For example, in particular embodiments, it may be desirable to improve the biological properties of a homing endonuclease, megaTAL or the like that binds and cleaves a target site in the human CBLB gene by introducing one or more substitutions, deletions, additions and/or insertions into  
15 the polypeptide. In particular embodiments, polypeptides include polypeptides having at least about 65%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% amino acid identity to any of the sequences contemplated herein, typically where the variant maintains at least one biological activity of the contemplated sequence.

20 Polypeptide variants include biologically active “polypeptide fragments.” Illustrative examples of biologically active polypeptide fragments include DNA binding domains, nuclease domains, and the like. As used herein, the term “biologically active fragment” or “minimal biologically active fragment” refers to a polypeptide fragment that retains at least 100%, at least 90%, at least 80%, at least 70%, at least 60%, at least 50%, at least 40%, at least  
25 30%, at least 20%, at least 10%, or at least 5% of the naturally occurring polypeptide activity. In preferred embodiments, the biological activity is binding affinity and/or cleavage activity for a target sequence. In certain embodiments, a polypeptide fragment can comprise an amino acid chain at least 5 to about 1700 amino acids long. It will be appreciated that in certain embodiments, fragments are at least 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, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700 or more amino acids long. In particular embodiments, a polypeptide comprises a

5 biologically active fragment of a homing endonuclease variant. In particular embodiments, the polypeptides set forth herein may comprise one or more amino acids denoted as “X” or “Xaa”. “X” or “Xaa” if present in an amino acid SEQ ID NO, refers to any amino acid or absence of an amino acid. One or more “X” (or “Xaa”) residues may be present at the N- and C-terminus of an amino acid sequence set forth in particular SEQ ID NOs contemplated herein. If the “X”

10 (or “Xaa”) amino acids are not present the remaining amino acid sequence set forth in a SEQ ID NO may be considered a biologically active fragment. For example, an amino acid sequence of positions 9-301 of SEQ ID NO: 5 is a biologically active fragment of SEQ ID NO: 2 (FIG. 15).

In particular embodiments, a polypeptide comprises a biologically active fragment of a

15 homing endonuclease variant, *e.g.*, SEQ ID NOs: 3-13, or a megaTAL (SEQ ID NOs: 14-21). The biologically active fragment may comprise an N-terminal truncation and/or C-terminal truncation. In a particular embodiment, a biologically active fragment lacks or comprises a deletion of the 1, 2, 3, 4, 5, 6, 7, or 8 N-terminal amino acids of a homing endonuclease variant compared to a corresponding wild type homing endonuclease sequence, more preferably a

20 deletion of the 4 N-terminal amino acids of a homing endonuclease variant compared to a corresponding wild type homing endonuclease sequence. In a particular embodiment, a biologically active fragment lacks or comprises a deletion of the 1, 2, 3, 4, or 5 C-terminal amino acids of a homing endonuclease variant compared to a corresponding wild type homing endonuclease sequence, more preferably a deletion of the 2 C-terminal amino acids of a

25 homing endonuclease variant compared to a corresponding wild type homing endonuclease sequence. In a particular preferred embodiment, a biologically active fragment lacks or comprises a deletion of the 4 N-terminal amino acids and 2 C-terminal amino acids of a homing endonuclease variant compared to a corresponding wild type homing endonuclease sequence.

In a particular embodiment, an I-OnuI variant comprises a deletion of 1, 2, 3, 4, 5, 6, 7, or 8 the following N-terminal amino acids: M, A, Y, M, S, R, R, E; and/or a deletion of the following 1 or 2 C-terminal amino acids: F, V, , respectively.

In a particular embodiment, an I-OnuI variant comprises a deletion or substitution of 1,  
5 2, 3, 4, 5, 6, 7, or 8 the following N-terminal amino acids: M, A, Y, M, S, R, R, E; and/or a deletion or substitution of the following 1 or 2 C-terminal amino acids: F, V, respectively.

As noted above, polypeptides may be altered in various ways including amino acid substitutions, deletions, truncations, and insertions. Methods for such manipulations are generally known in the art. For example, amino acid sequence variants of a reference  
10 polypeptide can be prepared by mutations in the DNA. Methods for mutagenesis and nucleotide sequence alterations are well known in the art. See, for example, Kunkel (1985, *Proc. Natl. Acad. Sci. USA*, 82: 488-492), Kunkel *et al.*, (1987, *Methods in Enzymol*, 154: 367-382), U.S. Pat. No. 4,873,192, Watson, J. D. *et al.*, (*Molecular Biology of the Gene*, Fourth Edition, Benjamin/Cummings, Menlo Park, Calif., 1987) and the references cited therein.  
15 Guidance as to appropriate amino acid substitutions that do not affect biological activity of the protein of interest may be found in the model of Dayhoff *et al.*, (1978) *Atlas of Protein Sequence and Structure* (*Natl. Biomed. Res. Found.*, Washington, D.C.).

In certain embodiments, a variant will contain one or more conservative substitutions. A “conservative substitution” is one in which an amino acid is substituted for another amino  
20 acid that has similar properties, such that one skilled in the art of peptide chemistry would expect the secondary structure and hydropathic nature of the polypeptide to be substantially unchanged. Modifications may be made in the structure of the polynucleotides and polypeptides contemplated in particular embodiments, polypeptides include polypeptides having at least about and still obtain a functional molecule that encodes a variant or derivative  
25 polypeptide with desirable characteristics. When it is desired to alter the amino acid sequence of a polypeptide to create an equivalent, or even an improved, variant polypeptide, one skilled in the art, for example, can change one or more of the codons of the encoding DNA sequence, *e.g.*, according to Table 1.

**TABLE 1- Amino Acid Codons**

| Amino Acids   | One letter code | Three letter code | Codons |     |     |     |     |
|---------------|-----------------|-------------------|--------|-----|-----|-----|-----|
| Alanine       | A               | Ala               | GCA    | GCC | GCG | GCU |     |
| Cysteine      | C               | Cys               | UGC    | UGU |     |     |     |
| Aspartic acid | D               | Asp               | GAC    | GAU |     |     |     |
| Glutamic acid | E               | Glu               | GAA    | GAG |     |     |     |
| Phenylalanine | F               | Phe               | UUC    | UUU |     |     |     |
| Glycine       | G               | Gly               | GGA    | GGC | GGG | GGU |     |
| Histidine     | H               | His               | CAC    | CAU |     |     |     |
| Isoleucine    | I               | Iso               | AUA    | AUC | AUU |     |     |
| Lysine        | K               | Lys               | AAA    | AAG |     |     |     |
| Leucine       | L               | Leu               | UUA    | UUG | CUA | CUC | CUG |
| Methionine    | M               | Met               | AUG    |     |     |     |     |
| Asparagine    | N               | Asn               | AAC    | AAU |     |     |     |
| Proline       | P               | Pro               | CCA    | CCC | CCG | CCU |     |
| Glutamine     | Q               | Gln               | CAA    | CAG |     |     |     |
| Arginine      | R               | Arg               | AGA    | AGG | CGA | CGC | CGG |
| Serine        | S               | Ser               | AGC    | AGU | UCA | UCC | UCG |
| Threonine     | T               | Thr               | ACA    | ACC | ACG | ACU |     |
| Valine        | V               | Val               | GUA    | GUC | GUG | GUU |     |
| Tryptophan    | W               | Trp               | UGG    |     |     |     |     |
| Tyrosine      | Y               | Tyr               | UAC    | UAU |     |     |     |

- Guidance in determining which amino acid residues can be substituted, inserted, or deleted without abolishing biological activity can be found using computer programs well known in the art, such as DNASTAR, GCG, DNA Strider, Geneious, Mac Vector, or Vector NTI software. Preferably, amino acid changes in the protein variants disclosed herein are conservative amino acid changes, *i.e.*, substitutions of similarly charged or uncharged amino acids. A conservative amino acid change involves substitution of one of a family of amino acids which are related in their side chains. Naturally occurring amino acids are generally divided into four families: acidic (aspartate, glutamate), basic (lysine, arginine, histidine), non-

polar (alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), and uncharged polar (glycine, asparagine, glutamine, cysteine, serine, threonine, tyrosine) amino acids. Phenylalanine, tryptophan, and tyrosine are sometimes classified jointly as aromatic amino acids. In a peptide or protein, suitable conservative substitutions of amino acids are known to those of skill in this art and generally can be made without altering a biological activity of a resulting molecule. Those of skill in this art recognize that, in general, single amino acid substitutions in non-essential regions of a polypeptide do not substantially alter biological activity (see, *e.g.*, Watson *et al. Molecular Biology of the Gene*, 4th Edition, 1987, The Benjamin/Cummings Pub. Co., p.224).

10           Additionally, engineered nuclease variants can be identified using known mutagenesis techniques, yeast display, fluorescence reporter assays, and/or next-gene sequencing or any combination thereof as described herein (see also *e.g.*, Jarjour *et al.*, 2009. *Nuc. Acids Res.* 37(20): 6871-6880; Certo *et al.*, *Nat Methods*. 2011 Jul 10;8(8):671-6; Certo *et al.*, *Nat Methods*. 2012 October ; 9(10): 973–975; WO2020/072059; and WO2007/123636.

15           In one embodiment, where expression of two or more polypeptides is desired, the polynucleotide sequences encoding them can be separated by an IRES sequence as disclosed elsewhere herein.

                  In further embodiments, a polypeptide contemplated herein further comprises a nuclear localization signal (NLS). NLS(s) is a sequence which provides subcellular localization for targeting to the nucleus. In certain embodiments, the NLS(s) is attached to the N-terminus of a polypeptide (*e.g.*, N-terminal to the engineered HE or the megaTAL).

20           Polypeptides contemplated in particular embodiments include fusion polypeptides. In particular embodiments, fusion polypeptides and polynucleotides encoding fusion polypeptides are provided. Fusion polypeptides and fusion proteins refer to a polypeptide having at least two, three, four, five, six, seven, eight, nine, or ten polypeptide segments.

25           In another embodiment, two or more polypeptides can be expressed as a fusion protein that comprises one or more self-cleaving polypeptide sequences as disclosed elsewhere herein.

In one embodiment, a fusion protein contemplated herein comprises one or more DNA binding domains and one or more nucleases, and one or more linker and/or self-cleaving polypeptides.

In one embodiment, a fusion protein contemplated herein comprises nuclease variant; a  
5 linker or self-cleaving peptide; and an end-processing enzyme including but not limited to a 5'-3' exonuclease, a 5'-3' alkaline exonuclease, and a 3'-5' exonuclease (*e.g.*, Trex2).

Fusion polypeptides can comprise one or more polypeptide domains or segments including, but are not limited to signal peptides, cell permeable peptide domains (CPP), DNA binding domains, nuclease domains, *etc.*, epitope tags (*e.g.*, maltose binding protein ("MBP"),  
10 glutathione S transferase (GST), HIS6, MYC, FLAG, V5, VSV-G, and HA), polypeptide linkers, and polypeptide cleavage signals. Fusion polypeptides are typically linked C-terminus to N-terminus, although they can also be linked C-terminus to C-terminus, N-terminus to N-terminus, or N-terminus to C-terminus. In particular embodiments, the polypeptides of the fusion protein can be in any order. Fusion polypeptides or fusion proteins can also include  
15 conservatively modified variants, polymorphic variants, alleles, mutants, subsequences, and interspecies homologs, so long as the desired activity of the fusion polypeptide is preserved. Fusion polypeptides may be produced by chemical synthetic methods or by chemical linkage between the two moieties or may generally be prepared using other standard techniques. Ligated DNA sequences comprising the fusion polypeptide are operably linked to suitable  
20 transcriptional or translational control elements as disclosed elsewhere herein.

Fusion polypeptides may optionally comprise a linker that can be used to link the one or more polypeptides or domains within a polypeptide. A peptide linker sequence may be employed to separate any two or more polypeptide components by a distance sufficient to ensure that each polypeptide folds into its appropriate secondary and tertiary structures so as to  
25 allow the polypeptide domains to exert their desired functions. Such a peptide linker sequence is incorporated into the fusion polypeptide using standard techniques in the art. Suitable peptide linker sequences may be chosen based on the following factors: (1) their ability to adopt a flexible extended conformation; (2) their inability to adopt a secondary structure that could interact with functional epitopes on the first and second polypeptides; and (3) the lack of



hydrophobic or charged residues that might react with the polypeptide functional epitopes. Preferred peptide linker sequences contain Gly, Asn, and Ser residues. Other near neutral amino acids, such as Thr and Ala may also be used in the linker sequence. Amino acid sequences which may be usefully employed as linkers include those disclosed in Maratea *et al.*,  
 5 *Gene* 40:39-46, 1985; Murphy *et al.*, *Proc. Natl. Acad. Sci. USA* 83:8258-8262, 1986; U.S. Patent No. 4,935,233 and U.S. Patent No. 4,751,180. Linker sequences are not required when a particular fusion polypeptide segment contains non-essential N-terminal amino acid regions that can be used to separate the functional domains and prevent steric interference. Preferred linkers are typically flexible amino acid sequences which are synthesized as part of a  
 10 recombinant fusion protein. Linker polypeptides can be between 1 and 200 amino acids in length, between 1 and 100 amino acids in length, or between 1 and 50 amino acids in length, including all integer values in between.

Exemplary linkers include, but are not limited to the following amino acid sequences: glycine polymers (G)<sub>n</sub>; glycine-serine polymers (G<sub>1-5</sub>S<sub>1-5</sub>)<sub>n</sub>, where n is an integer of at least  
 15 one, two, three, four, or five; glycine-alanine polymers; alanine-serine polymers; GGG (SEQ ID NO: 45); DGGGS (SEQ ID NO: 46); TGEKP (SEQ ID NO: 47) (*see e.g.*, Liu *et al.*, *PNAS* 5525-5530 (1997)); GGRR (SEQ ID NO: 48) (Pomerantz *et al.* 1995, *supra*); (GGGGS)<sub>n</sub> wherein n = 1, 2, 3, 4 or 5 (SEQ ID NO: 49) (Kim *et al.*, *PNAS* 93, 1156-1160 (1996.); EGKSSGSGSESKVD (SEQ ID NO: 50) (Chaudhary *et al.*, 1990, *Proc. Natl. Acad. Sci. U.S.A.*  
 20 87:1066-1070); KESGSVSSEQLAQFRSLD (SEQ ID NO: 51) (Bird *et al.*, 1988, *Science* 242:423-426), GGRRGGGS (SEQ ID NO: 52); LRQRDGERP (SEQ ID NO: 53); LRQKDGGGSERP (SEQ ID NO: 54); LRQKD(GGGS)<sub>2</sub>ERP (SEQ ID NO: 55). Alternatively, flexible linkers can be rationally designed using a computer program capable of modeling both DNA-binding sites and the peptides themselves (Desjarlais & Berg, *PNAS*  
 25 90:2256-2260 (1993), *PNAS* 91:11099-11103 (1994) or by phage display methods.

Fusion polypeptides may further comprise a polypeptide cleavage signal between each of the polypeptide domains described herein or between an endogenous open reading frame and a polypeptide encoded by a donor repair template. In addition, a polypeptide cleavage site can be put into any linker peptide sequence. Exemplary polypeptide cleavage signals include  
 30 polypeptide cleavage recognition sites such as protease cleavage sites, nuclease cleavage sites

(e.g., rare restriction enzyme recognition sites, self-cleaving ribozyme recognition sites), and self-cleaving viral oligopeptides (see deFelipe and Ryan, 2004. *Traffic*, 5(8); 616-26).

Suitable protease cleavages sites and self-cleaving peptides are known to the skilled person (see, e.g., in Ryan *et al.*, 1997. *J. Gener. Virol.* 78, 699-722; Scymczak *et al.* (2004) Nature Biotech. 5, 589-594). Exemplary protease cleavage sites include, but are not limited to the cleavage sites of potyvirus NIa proteases (e.g., tobacco etch virus protease), potyvirus HC proteases, potyvirus P1 (P35) proteases, byovirus NIa proteases, byovirus RNA-2-encoded proteases, aphthovirus L proteases, enterovirus 2A proteases, rhinovirus 2A proteases, picorna 3C proteases, comovirus 24K proteases, nepovirus 24K proteases, RTSV (rice tungro spherical virus) 3C-like protease, PYVF (parsnip yellow fleck virus) 3C-like protease, heparin, thrombin, factor Xa and enterokinase. Due to its high cleavage stringency, TEV (tobacco etch virus) protease cleavage sites are preferred in one embodiment, e.g., EXXYXQ(G/S) (SEQ ID NO: 56), for example, ENLYFQG (SEQ ID NO: 57) and ENLYFQS (SEQ ID NO: 58), wherein X represents any amino acid (cleavage by TEV occurs between Q and G or Q and S).

In certain embodiments, the self-cleaving polypeptide site comprises a 2A or 2A-like site, sequence or domain (Donnelly *et al.*, 2001. *J. Gen. Virol.* 82:1027-1041). In a particular embodiment, the viral 2A peptide is an aphthovirus 2A peptide, a potyvirus 2A peptide, or a cardiovirus 2A peptide.

In one embodiment, the viral 2A peptide is selected from the group consisting of: a foot-and-mouth disease virus (FMDV) (F2A) peptide, an equine rhinitis A virus (ERAV) (E2A) peptide, a Thosea asigna virus (TaV) (T2A) peptide, a porcine teschovirus-1 (PTV-1) (P2A) peptide, a Theilovirus 2A peptide, and an encephalomyocarditis virus 2A peptide.

Illustrative examples of 2A sites are provided in Table 2.

TABLE 2: Exemplary 2A sites include the following sequences:

|               |                        |
|---------------|------------------------|
| SEQ ID NO: 59 | GSGATNFSLLKQAGDVEENPGP |
| SEQ ID NO: 60 | ATNFSLLKQAGDVEENPGP    |
| SEQ ID NO: 61 | LLKQAGDVEENPGP         |
| SEQ ID NO: 62 | GSGEGRGSLLTCGDVEENPGP  |

|               |                                          |
|---------------|------------------------------------------|
| SEQ ID NO: 63 | EGRGSLLTCGDVEENPGP                       |
| SEQ ID NO: 64 | LLTCGDVEENPGP                            |
| SEQ ID NO: 65 | GSGQCTNYALLKLAGDVESNPGP                  |
| SEQ ID NO: 66 | QCTNYALLKLAGDVESNPGP                     |
| SEQ ID NO: 67 | LLKLAGDVESNPGP                           |
| SEQ ID NO: 68 | GSGVKQTLNFDLLKLAGDVESNPGP                |
| SEQ ID NO: 69 | VKQTLNFDLLKLAGDVESNPGP                   |
| SEQ ID NO: 70 | LLKLAGDVESNPGP                           |
| SEQ ID NO: 71 | LLNFDLLKLAGDVESNPGP                      |
| SEQ ID NO: 72 | TLNFDLLKLAGDVESNPGP                      |
| SEQ ID NO: 73 | LLKLAGDVESNPGP                           |
| SEQ ID NO: 74 | NFDLLKLAGDVESNPGP                        |
| SEQ ID NO: 75 | QLNFDLLKLAGDVESNPGP                      |
| SEQ ID NO: 76 | APVKQTLNFDLLKLAGDVESNPGP                 |
| SEQ ID NO: 77 | VTELLYRMKRAETYCPRPLLAIHPTEARHKQKIVAPVKQT |
| SEQ ID NO: 78 | LNFDLLKLAGDVESNPGP                       |
| SEQ ID NO: 79 | LLAIHPTEARHKQKIVAPVKQTLNFDLLKLAGDVESNPGP |
| SEQ ID NO: 80 | EARHKQKIVAPVKQTLNFDLLKLAGDVESNPGP        |

#### G. POLYNUCLEOTIDES

In particular embodiments, polynucleotides or therapeutic transgenes encoding one or more engineered homing endonucleases, megaTALs, end-processing enzymes, fusion

5 polypeptides, and therapeutic polypeptides, or one or more biologically active fragments or variants contemplated herein are provided. As used herein, the terms “polynucleotide” or “nucleic acid” refer to deoxyribonucleic acid (DNA), ribonucleic acid (RNA) and DNA/RNA hybrids. Polynucleotides may be single-stranded or double-stranded and either recombinant, synthetic, or isolated. Polynucleotides include, but are not limited to: pre-messenger RNA

10 (pre-mRNA), messenger RNA (mRNA), RNA, short interfering RNA (siRNA), short hairpin

RNA (shRNA), microRNA (miRNA), ribozymes, genomic RNA (gRNA), plus strand RNA (RNA(+)), minus strand RNA (RNA(-)), tracrRNA, crRNA, single guide RNA (sgRNA), synthetic RNA, synthetic mRNA, genomic DNA (gDNA), PCR amplified DNA, complementary DNA (cDNA), synthetic DNA, or recombinant DNA. In particular

5   embodiments, a polynucleotide is a polynucleotide fragment that encodes one or more biologically active polypeptide fragments or variants. Polynucleotides refer to a polymeric form of nucleotides of at least 5, at least 10, at least 15, at least 20, at least 25, at least 30, at least 40, at least 50, at least 100, at least 200, at least 300, at least 400, at least 500, at least 1000, at least 5000, at least 10000, or at least 15000 or more nucleotides in length, either

10   ribonucleotides or deoxyribonucleotides or a modified form of either type of nucleotide, as well as all intermediate lengths. It will be readily understood that “intermediate lengths,” in this context, means any length between the quoted values, such as 6, 7, 8, 9, *etc.*, 101, 102, 103, *etc.*; 151, 152, 153, *etc.*; 201, 202, 203, *etc.* In particular embodiments, polynucleotides or variants have at least or about 50%, 55%, 60%, 65%, 70%, 71%, 72%, 73%, 74%, 75%, 76%,

15   77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% sequence identity to a sequence contemplated herein, including any polynucleotide set forth in a SEQ ID NO herein.

In particular embodiments, polynucleotides may be codon-optimized. As used herein, the term “codon-optimized” refers to substituting codons in a polynucleotide encoding a

20   polypeptide in order to increase the expression, stability and/or activity of the polypeptide. Factors that influence codon optimization include, but are not limited to one or more of: (i) variation of codon biases between two or more organisms or genes or synthetically constructed bias tables, (ii) variation in the degree of codon bias within an organism, gene, or set of genes, (iii) systematic variation of codons including context, (iv) variation of codons according to

25   their decoding tRNAs, (v) variation of codons according to GC %, either overall or in one position of the triplet, (vi) variation in degree of similarity to a reference sequence for example a naturally occurring sequence, (vii) variation in the codon frequency cutoff, (viii) structural properties of mRNAs transcribed from the DNA sequence, (ix) prior knowledge about the function of the DNA sequences upon which design of the codon substitution set is to be based,

(x) systematic variation of codon sets for each amino acid, and/or (xi) isolated removal of spurious translation initiation sites.

As used herein the term “nucleotide” refers to a heterocyclic nitrogenous base in N-glycosidic linkage with a phosphorylated sugar. Nucleotides are understood to include natural  
5 bases, and a wide variety of art-recognized modified bases. Such bases are generally located at the 1' position of a nucleotide sugar moiety. Nucleotides generally comprise a base, sugar and a phosphate group. In ribonucleic acid (RNA), the sugar is a ribose, and in deoxyribonucleic acid (DNA) the sugar is a deoxyribose, *i.e.*, a sugar lacking a hydroxyl group that is present in ribose. Exemplary natural nitrogenous bases include the purines, adenosine (A) and guanine  
10 (G), and the pyrimidines, cytosine (C) and thymine (T) (or in the context of RNA, uracil (U)). The C-1 atom of deoxyribose is bonded to N-1 of a pyrimidine or N-9 of a purine. Nucleotides are usually mono, di- or triphosphates. The nucleotides can be unmodified or modified at the sugar, phosphate and/or base moiety, (also referred to interchangeably as nucleotide analogs, nucleotide derivatives, modified nucleotides, non-natural nucleotides, and non-standard  
15 nucleotides; see for example, WO 92/07065 and WO 93/15187). Examples of modified nucleic acid bases are summarized by Limbach *et al.*, (1994, *Nucleic Acids Res.* 22, 2183-2196).

A nucleotide may also be regarded as a phosphate ester of a nucleoside, with esterification occurring on the hydroxyl group attached to C-5 of the sugar. As used herein, the  
20 term “nucleoside” refers to a heterocyclic nitrogenous base in N-glycosidic linkage with a sugar. Nucleosides are recognized in the art to include natural bases, and also to include well known modified bases. Such bases are generally located at the 1' position of a nucleoside sugar moiety. Nucleosides generally comprise a base and sugar group. The nucleosides can be unmodified or modified at the sugar, and/or base moiety, (also referred to interchangeably as  
25 nucleoside analogs, nucleoside derivatives, modified nucleosides, non-natural nucleosides, or non-standard nucleosides). As also noted above, examples of modified nucleic acid bases are summarized by Limbach *et al.*, (1994, *Nucleic Acids Res.* 22, 2183-2196).

Illustrative examples of polynucleotides include, but are not limited to polynucleotides encoding SEQ ID NOs: 1-21, 44-81, and 86-107 and polynucleotide sequences set forth in SEQ ID NOs: 22-43 and 82-85.

In various illustrative embodiments, polynucleotides contemplated herein include, but  
5 are not limited to polynucleotides encoding engineered homing endonucleases, megaTALs, end-processing enzymes, fusion polypeptides, therapeutic polypeptides, and expression vectors, viral vectors, and transfer plasmids comprising polynucleotides contemplated herein.

In further embodiments, polynucleotides contemplated herein further comprise encoding for a nuclear localization signal (NLS), including more than one NLS. NLS(s) is a  
10 sequence which provides subcellular localization for targeting to the nucleus. In certain embodiments, the polynucleotide encodes the NLS(s) to be at the N-terminus or N-terminal side of a polypeptide.

As used herein, the terms “polynucleotide variant”, “nucleic acid variant”, and “variant” and the like refer to polynucleotides displaying substantial sequence identity with a  
15 reference polynucleotide sequence or polynucleotides that hybridize with a reference sequence under stringent conditions that are defined hereinafter. These terms also encompass polynucleotides that are distinguished from a reference polynucleotide by the addition, deletion, substitution, or modification of at least one nucleotide. Accordingly, the terms “polynucleotide variant”, “nucleic acid variant”, and “variant” include polynucleotides in  
20 which one or more nucleotides have been added or deleted, or modified, or replaced with different nucleotides. In this regard, it is well understood in the art that certain alterations inclusive of mutations, additions, deletions and substitutions can be made to a reference polynucleotide whereby the altered polynucleotide retains the biological function or activity of the reference polynucleotide.

25 Polynucleotide variants include polynucleotide fragments that encode biologically active polypeptide fragments or variants. As used herein, the term “polynucleotide fragment” refers to a polynucleotide fragment at least 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, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 150, 200, 250, 300, 350, 400,

450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700 or more nucleotides in length that encodes a polypeptide variant that retains at least 100%, at least 90%, at least 80%, at least 70%, at least 60%, at least 50%, at least 40%, at least 30%, at least 20%, at least 10%, or at least 5% of the naturally occurring polypeptide activity.

- 5 Polynucleotide fragments refer to a polynucleotide that encodes a polypeptide that has an amino-terminal deletion, a carboxyl-terminal deletion, and/or an internal deletion or substitution of one or more amino acids of a naturally-occurring or recombinantly-produced polypeptide.

- In one embodiment, a polynucleotide comprises a nucleotide sequence that hybridizes  
 10 to a target nucleic acid sequence under stringent conditions. To hybridize under “stringent conditions” describes hybridization protocols in which nucleotide sequences at least 60% identical to each other remain hybridized. Generally, stringent conditions are selected to be about 5°C lower than the thermal melting point (T<sub>m</sub>) for the specific sequence at a defined ionic strength and pH. The T<sub>m</sub> is the temperature (under defined ionic strength, pH and  
 15 nucleic acid concentration) at which 50% of the probes complementary to the target sequence hybridize to the target sequence at equilibrium. Since the target sequences are generally present at excess, at T<sub>m</sub>, 50% of the probes are occupied at equilibrium.

- The recitations “sequence identity” or, for example, comprising a “sequence 50% identical to,” as used herein, refer to the extent that sequences are identical on a nucleotide-by-  
 20 nucleotide basis or an amino acid-by-amino acid basis over a window of comparison. Thus, a “percentage of sequence identity” may be calculated by comparing two optimally aligned sequences over the window of comparison, determining the number of positions at which the identical nucleic acid base (*e.g.*, A, T, C, G, I) or the identical amino acid residue (*e.g.*, Ala, Pro, Ser, Thr, Gly, Val, Leu, Ile, Phe, Tyr, Trp, Lys, Arg, His, Asp, Glu, Asn, Gln, Cys and  
 25 Met) occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison (*i.e.*, the window size), and multiplying the result by 100 to yield the percentage of sequence identity. Included are nucleotides and polypeptides having at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% sequence identity to any of the

reference sequences described herein, typically where the polypeptide variant maintains at least one biological activity of the reference polypeptide.

Terms used to describe sequence relationships between two or more polynucleotides or polypeptides include “reference sequence,” “comparison window,” “sequence identity,”

- 5 “percentage of sequence identity,” and “substantial identity”. Because two polynucleotides may each comprise (1) a sequence (*i.e.*, only a portion of the complete polynucleotide sequence) that is similar between the two polynucleotides, and (2) a sequence that is divergent between the two polynucleotides, sequence comparisons between two (or more) polynucleotides are typically performed by comparing sequences of the two polynucleotides over a “comparison window” to identify and compare local regions of sequence similarity. A “comparison window” refers to a conceptual segment of at least 6 contiguous positions, usually about 50 to about 100, more usually about 100 to about 150 in which a sequence is compared to a reference sequence of the same number of contiguous positions after the two sequences are optimally aligned. The comparison window may comprise additions or deletions (*i.e.*, gaps) of
- 15 about 20% or less as compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. Optimal alignment of sequences for aligning a comparison window may be conducted by computerized implementations of algorithms (GAP, BESTFIT, FASTA, and TFASTA in for example, DNASTAR, GCG, DNA Strider, Geneious, Mac Vector, or Vector NTI software) or by inspection and the best
- 20 alignment (*i.e.*, resulting in the highest percentage homology over the comparison window) generated by any of the various methods selected. Reference also may be made to the BLAST family of programs as for example disclosed by Altschul *et al.*, 1997, *Nucl. Acids Res.* 25:3389. A detailed discussion of sequence analysis can be found in Unit 19.3 of Ausubel *et al.*, *Current Protocols in Molecular Biology*, John Wiley & Sons Inc., 1994-1998, Chapter 15.

- 25 An “isolated polynucleotide,” as used herein, refers to a polynucleotide that has been purified from the sequences which flank it in a naturally-occurring state, *e.g.*, a DNA fragment that has been removed from the sequences that are normally adjacent to the fragment. In particular embodiments, an “isolated polynucleotide” refers to a complementary DNA (cDNA), a recombinant polynucleotide, a synthetic polynucleotide, or other polynucleotide that
- 30 does not exist in nature and that has been made by the hand of man.



In various embodiments, a polynucleotide comprises an mRNA encoding a polypeptide contemplated herein including, but not limited to, a homing endonuclease variant, a megaTAL, and an end-processing enzyme. In certain embodiments, the mRNA comprises a cap, one or more nucleotides, and a poly(A) tail.

5 In various embodiments, linearized DNA may be used in an *in vitro* transcription (IVT) system to generate mRNA for use in the methods describe herein. The IVT system typically comprises a transcription buffer, nucleotide triphosphates (NTPs), an RNase inhibitor and an RNA polymerase. Methods for *in vitro* transcription are known in the art. See, *e.g.*, Beckert *et al.*, Methods Mol Biol. 2011;703:29-41. Exemplary commercially supplied kits for IVT  
10 include, but are not limited to, HiScribe™ T7 Quick High Yield RNA Synthesis Kit (New England BioLabs™), MEGAscript® T7 Kit (ThermoFisher Scientific™), TranscriptAid T7 High Yield Transcription Kit (ThermoFisher Scientific™), Riboprobe® or RiboMAX™ RNA Production System (Promega™), AmpliScribe™ T7 Transcription kits (Lucigen®), and RNAMaxx™ (Agilent Technologies™).

15 Alternatively, IVT assays can be assembled and performed in-house by obtaining each component separately and using methods known in the art. The NTPs may be manufactured in house or purchase from commercial suppliers (*e.g.*, Trilink® and NewEngland BioLabs®). Any number of RNA polymerases or variants thereof may be used in the method described herein, and are readily available through commercial  
20 suppliers (*e.g.*, NewEngland BioLabs®, ThermoFisher Scientific™, and MilliporeSigma™). The polymerase may be selected from, but is not limited to, a phage RNA polymerase, *e.g.*, a T7 RNA polymerase, a T3 RNA polymerase, an SP6 RNA polymerase, and/or mutant polymerases such as, but not limited to, polymerases able to incorporate modified nucleic acids.

25 A typical *in vitro* transcription reaction includes the following: an RNA polymerase, *e.g.*, a T7 RNA polymerase; a DNA template; nucleotides (NTPs); MgCl<sub>2</sub>; and a buffer such as, *e.g.*, HEPES or Tris. IVT reactions can also include dithiothreitol (DTT) and/or spermidine, an RNase inhibitor, a pyrophosphatase, and/or EDTA. The *in*

*vitro* transcription reaction is allowed to proceed, for example, under constant mixing at 37° C for 4 hours.

In some embodiments, the mRNA described herein is capped. Capping RNA maximizes efficiency of expression in cells by increasing stability and reducing degradation. In some embodiments, the RNA molecules used in the methods are synthesized *in vitro* by incubating uncapped RNA in the presence a capping enzyme system. In some embodiments, the RNA is enzymatically capped at the 5' end after *in vitro* transcription. In some embodiments, the RNA is enzymatically capped at the 5' end co-transcriptionally.

As used herein, the terms “5' cap” or “5' cap structure” or “5' cap moiety” refer to a chemical modification, which has been incorporated at the 5' end of an mRNA. The 5' cap is involved in nuclear export, mRNA stability, and translation.

In particular embodiments, an mRNA contemplated herein comprises a 5' cap comprising a 5'-ppp-5'-triphosphate linkage between a terminal guanosine cap residue and the 5'-terminal transcribed sense nucleotide of the mRNA molecule. This 5'-guanylate cap may then be methylated to generate an N7-methyl-guanylate residue.

Illustrative examples of 5' cap suitable for use in particular embodiments of the mRNA polynucleotides contemplated herein include, but are not limited to: unmethylated 5' cap analogs, *e.g.*, G(5')ppp(5')G, G(5')ppp(5')C, G(5')ppp(5')A; methylated 5' cap analogs, *e.g.*, m<sup>7</sup>G(5')ppp(5')G, m<sup>7</sup>G(5')ppp(5')C, and m<sup>7</sup>G(5')ppp(5')A; dimethylated 5' cap analogs, *e.g.*, m<sup>2,7</sup>G(5')ppp(5')G, m<sup>2,7</sup>G(5')ppp(5')C, and m<sup>2,7</sup>G(5')ppp(5')A; trimethylated 5' cap analogs, *e.g.*, m<sup>2,2,7</sup>G(5')ppp(5')G, m<sup>2,2,7</sup>G(5')ppp(5')C, and m<sup>2,2,7</sup>G(5')ppp(5')A; dimethylated symmetrical 5' cap analogs, *e.g.*, m<sup>7</sup>G(5')pppm<sup>7</sup>(5')G, m<sup>7</sup>G(5')pppm<sup>7</sup>(5')C, and m<sup>7</sup>G(5')pppm<sup>7</sup>(5')A; and anti-reverse 5' cap analogs, *e.g.* Anti-Reverse Cap Analog (ARCA) cap, designated 3'O-Me-m<sup>7</sup>G(5')ppp(5')G, 2'O-Me-m<sup>7</sup>G(5')ppp(5')G, 2'O-Me-m<sup>7</sup>G(5')ppp(5')C, 2'O-Me-m<sup>7</sup>G(5')ppp(5')A, m<sup>7</sup>2'd(5')ppp(5')G, m<sup>7</sup>2'd(5')ppp(5')C, m<sup>7</sup>2'd(5')ppp(5')A, 3'O-Me-m<sup>7</sup>G(5')ppp(5')C, 3'O-Me-m<sup>7</sup>G(5')ppp(5')A, m<sup>7</sup>3'd(5')ppp(5')G, m<sup>7</sup>3'd(5')ppp(5')C, m<sup>7</sup>3'd(5')ppp(5')A

and their tetraphosphate derivatives) (*see, e.g., Jemielity et al., RNA, 9: 1108-1122 (2003)*).

In particular embodiments, mRNAs comprise a 5' cap that is a 7-methyl guanylate ("m<sup>7</sup>G") linked via a triphosphate bridge to the 5'-end of the first transcribed nucleotide,  
5 resulting in m<sup>7</sup>G(5')ppp(5')N, where N is any nucleoside.

In some embodiments, mRNAs comprise a 5' cap wherein the cap is a Cap0 structure (Cap0 structures lack a 2'-O-methyl residue of the ribose attached to bases 1 and 2), a Cap1 structure (Cap1 structures have a 2'-O-methyl residue at base 2), or a Cap2 structure (Cap2 structures have a 2'-O-methyl residue attached to both bases 2 and 3).

10 For example, the RNA can be enzymatically capped at the 5' end using Vaccinia guanylyltransferase, guanosine triphosphate and s-adenosyl-L-methionine to yield cap 0 structure. An inverted 7-methylguanosine cap is added via a 5' to 5' triphosphate bridge. Alternatively, use of a 2'-O-methyltransferase with Vaccinia guanylyltransferase yields the cap 1 structure where in addition to the cap 0 structure, the 2'OH group is methylated on  
15 the penultimate nucleotide. S-adenosyl-L-methionine (SAM) is a cofactor utilized as a methyl transfer reagent.

In one embodiment, an mRNA comprises a m<sup>7</sup>G(5')ppp(5')G cap.

In one embodiment, an mRNA comprises an ARCA cap or modified ARCA cap.

In various embodiments, the RNA is co-transcriptionally capped or enzymatically  
20 capped in a separate reaction. The 5' terminal caps may include endogenous caps or cap analogs. A 5' terminal cap may comprise a guanine analog. Useful guanine analogs include, but are not limited to, inosine, N1-methyl-guanosine, 2'fluoro-guanosine, 7-deaza-guanosine, 8-oxo-guanosine, 2-amino-guanosine, LNA-guanosine, and 2-azido-guanosine.

Further examples of 5' cap structures include glyceryl, inverted deoxy abasic  
25 residue (moiety), 4',5' methylene nucleotide, 1-(beta-D-erythrofuransyl) nucleotide, 4'-thio nucleotide, carbocyclic nucleotide, 1,5-anhydrohexitol nucleotide, L-nucleotides, alpha-nucleotide, modified base nucleotide, threo-pentofuransyl nucleotide, acyclic 3',4'-seco nucleotide, acyclic 3,4-dihydroxybutyl nucleotide, acyclic 3,5 dihydroxypentyl nucleotide, 3'-3'-inverted nucleotide moiety, 3'-3'-inverted abasic moiety, 3'-2'-inverted

nucleotide moiety, 3'-2'-inverted abasic moiety, 1,4-butanediol phosphate, 3'-phosphoramidate, hexylphosphate, aminohexyl phosphate, 3'-phosphate, 3'-phosphorothioate, phosphorodithioate, or bridging or non-bridging methylphosphonate moiety. Further modified 5'-CAP structures which may be used in the context of the present invention are CAP1 (methylation of the ribose of the adjacent nucleotide of m7GpppN), CAP2 (methylation of the ribose of the 2nd nucleotide downstream of the m7GpppN), CAP3 (methylation of the ribose of the 3rd nucleotide downstream of the m7GpppN), CAP4 (methylation of the ribose of the 4th nucleotide downstream of the m7GpppN), ARCA (anti-reverse CAP analogue, modified ARCA (*e.g.* phosphothioate modified ARCA), inosine, N1-methyl-guanosine, 2'-fluoro-guanosine, 7-deaza-guanosine, 8-oxo-guanosine, 2-amino-guanosine, LNA-guanosine, and 2-azido-guanosine.

In eukaryotes, at least three enzymatic activities are required to generate a functional cap 0 (RNA triphosphatase (TPase), RNA guanylyltransferase (GTase) and guanine-N7 methyltransferase (guanine-N7 MTase). For a cap 1 structure, an additional m7G-specific 2'O methyltransferase (2'O MTase) is required to methylate the +1 ribonucleotide at the 2'O position of the ribose. Eukaryote capping enzymes are known in the art (*Nucleic Acids Research*, Volume 44, Issue 16, 19 September 2016, Pages 7511–7526).

Viral RNA capping enzymes are also known in the art. In some instances, viral capping enzymes are known to couple enzymatic activities into multifunctional proteins. For example, Flavivirus, Dengue, West Nile, and Paramyxoviruses, couple the GTase and MTase activities into their RNA polymerase (RdRp). Alternatively, the Vaccinia virus capping enzyme and Bluetongue virus capping enzyme couple all the necessary enzymatic activities of RNA capping to generate cap 0 or cap 1. Thus, due to its simplicity and effectiveness, the Vaccinia virus capping enzyme Vaccinia guanylyltransferase is often a preferred capping enzyme, but not a requirement. Other viral capping enzymes known in the art include, but are not limited to, chlorella virus, alpha virus, rhabdovirus, and vesicular stomatitis virus capping enzymes. In some embodiments, the polyadenylated mRNA is capped at its 5' end using a Vaccinia guanylyltransferase, guanosine triphosphate, and S-adenosyl-L-methionine (SAM) to produce a cap 0 structure. In some embodiments,

the polyadenylated mRNA is capped at its 5' end using a Vaccinia guanylyltransferase, guanosine triphosphate, S-adenosyl-L-methionine (SAM), and an 2'-O-Methyltransferase to produce a cap 1 structure.

Capping methods and conditions are known in the art (see, *e.g.*, Wiley Interdiscip  
 5 Rev RNA, 2010 Jul-Aug; 1(1): 152–172; and *Nat Rev Microbiol.* 2011 Dec 5;10(1):51-65.  
 An exemplary capping reaction may include the following: S-adenosylmethionine chloride (SAM); RNase inhibitor; buffer (*e.g.*, NEB capping buffer); GTP; Vaccinia Enzyme; mRNA Cap 2'-O-Methyltransferase; and EDTA. The reaction is run under constant mixing at 37° C.

10 In particular embodiments, an mRNA contemplated herein comprises one or more modified nucleosides.

In one embodiment, an mRNA comprises one or more modified nucleosides selected from the group consisting of: pseudouridine, pyridin-4-one ribonucleoside, 5-aza-uridine, 2-thio-5-aza-uridine, 2-thiouridine, 4-thio-pseudouridine, 2-thio-pseudouridine, 5-  
 15 hydroxyuridine, 3-methyluridine, 5-carboxymethyl-uridine, 1-carboxymethyl-pseudouridine, 5-propynyl-uridine, 1-propynyl-pseudouridine, 5-taurinomethyluridine, 1-taurinomethyl-pseudouridine, 5-taurinomethyl-2-thio-uridine, 1-taurinomethyl-4-thio-uridine, 5-methyl-uridine, 1-methyl-pseudouridine, 4-thio-1-methyl-pseudouridine, 2-thio-1-methyl-pseudouridine, 1-methyl-1-deaza-pseudouridine, 2-thio-1-methyl-1-deaza-pseudouridine,  
 20 dihydrouridine, dihydropseudouridine, 2-thio-dihydrouridine, 2-thio-dihydropseudouridine, 2-methoxyuridine, 2-methoxy-4-thio-uridine, 4-methoxy-pseudouridine, 4-methoxy-2-thio-pseudouridine, 5-aza-cytidine, pseudoisocytidine, 3-methyl-cytidine, N4-acetylcytidine, 5-formylcytidine, N4-methylcytidine, 5-hydroxymethylcytidine, 1-methyl-pseudoisocytidine, pyrrolo-cytidine, pyrrolo-pseudoisocytidine, 2-thio-cytidine, 2-thio-5-methyl-cytidine, 4-thio-pseudoisocytidine, 4-thio-1-methyl-pseudoisocytidine, 4-thio-1-methyl-1-deaza-pseudoisocytidine, 1-methyl-1-deaza-pseudoisocytidine, zebularine, 5-aza-zebularine, 5-methyl-zebularine, 5-aza-2-thio-zebularine, 2-thio-zebularine, 2-methoxy-cytidine, 2-methoxy-5-methyl-cytidine, 4-methoxy-pseudoisocytidine, 4-methoxy-1-methyl-pseudoisocytidine, 2-aminopurine, 2,6-diaminopurine, 7-deaza-adenine, 7-deaza-8-aza-adenine, 7-deaza-2-

aminopurine, 7-deaza-8-aza-2-aminopurine, 7-deaza-2,6-diaminopurine, 7-deaza-8-aza-2,6-diaminopurine, 1-methyladenosine, N6-methyladenosine, N6-isopentenyladenosine, N6-(cis-hydroxyisopentenyl)adenosine, 2-methylthio-N6-(cis-hydroxyisopentenyl) adenosine, N6-glycylcarbamoyladenosine, N6-threonylcarbamoyladenosine, 2-methylthio-N6-threonyl  
 5 carbamoyladenosine, N6,N6-dimethyladenosine, 7-methyladenine, 2-methylthio-adenine, 2-methoxy-adenine, inosine, 1-methyl-inosine, wyosine, wybutosine, 7-deaza-guanosine, 7-deaza-8-aza-guanosine, 6-thio-guanosine, 6-thio-7-deaza-guanosine, 6-thio-7-deaza-8-aza-guanosine, 7-methyl-guanosine, 6-thio-7-methyl-guanosine, 7-methylinosine, 6-methoxy-guanosine, 1-methylguanosine, N2-methylguanosine, N2,N2-dimethylguanosine, 8-oxo-  
 10 guanosine, 7-methyl-8-oxo-guanosine, 1-methyl-6-thio-guanosine, N2-methyl-6-thio-guanosine, and N2,N2-dimethyl-6-thio-guanosine.

In one embodiment, an mRNA comprises one or more modified nucleosides selected from the group consisting of: pseudouridine, pyridin-4-one ribonucleoside, 5-aza-uridine, 2-thio-5-aza-uridine, 2-thiouridine, 4-thio-pseudouridine, 2-thio-pseudouridine, 5-  
 15 hydroxyuridine, 3-methyluridine, 5-carboxymethyl-uridine, 1-carboxymethyl-pseudouridine, 5-propynyl-uridine, 1-propynyl-pseudouridine, 5-taurinomethyluridine, 1-taurinomethyl-pseudouridine, 5-taurinomethyl-2-thio-uridine, 1-taurinomethyl-4-thio-uridine, 5-methyl-uridine, 1-methyl-pseudouridine, 4-thio-1-methyl-pseudouridine, 2-thio-1-methyl-pseudouridine, 1-methyl-1-deaza-pseudouridine, 2-thio-1-methyl-1-deaza-pseudouridine,  
 20 dihydrouridine, dihydropseudouridine, 2-thio-dihydrouridine, 2-thio-dihydropseudouridine, 2-methoxyuridine, 2-methoxy-4-thio-uridine, 4-methoxy-pseudouridine, and 4-methoxy-2-thio-pseudouridine.

In one embodiment, an mRNA comprises one or more modified nucleosides selected from the group consisting of: 5-aza-cytidine, pseudoisocytidine, 3-methyl-cytidine, N4-  
 25 acetylcytidine, 5-formylcytidine, N4-methylcytidine, 5-hydroxymethylcytidine, 1-methyl-pseudoisocytidine, pyrrolo-cytidine, pyrrolo-pseudoisocytidine, 2-thio-cytidine, 2-thio-5-methyl-cytidine, 4-thio-pseudoisocytidine, 4-thio-1-methyl-pseudoisocytidine, 4-thio-1-methyl-1-deaza-pseudoisocytidine, 1-methyl-1-deaza-pseudoisocytidine, zebularine, 5-aza-zebularine, 5-methyl-zebularine, 5-aza-2-thio-zebularine, 2-thio-zebularine, 2-methoxy-

cytidine, 2-methoxy-5-methyl-cytidine, 4-methoxy-pseudoisocytidine, and 4-methoxy-1-methyl-pseudoisocytidine.

In one embodiment, an mRNA comprises one or more modified nucleosides selected from the group consisting of: 2-aminopurine, 2,6-diaminopurine, 7-deaza-adenine, 7-deaza-8-  
 5 aza-adenine, 7-deaza-2-aminopurine, 7-deaza-8-aza-2-aminopurine, 7-deaza-2,6-diaminopurine, 7-deaza-8-aza-2,6-diaminopurine, 1-methyladenosine, N6-methyladenosine, N6-isopentenyladenosine, N6-(cis-hydroxyisopentenyl)adenosine, 2-methylthio-N6-(cis-hydroxyisopentenyl) adenosine, N6-glycinylicarbamoyl-adenosine, N6-threonylcarbamoyl-adenosine, 2-methylthio-N6-threonyl carbamoyl-adenosine, N6,N6-  
 10 dimethyladenosine, 7-methyladenine, 2-methylthio-adenine, and 2-methoxy-adenine.

In one embodiment, an mRNA comprises one or more modified nucleosides selected from the group consisting of: inosine, 1-methyl-inosine, wyosine, wybutosine, 7-deaza-guanosine, 7-deaza-8-aza-guanosine, 6-thio-guanosine, 6-thio-7-deaza-guanosine, 6-thio-7-deaza-8-aza-guanosine, 7-methyl-guanosine, 6-thio-7-methyl-guanosine, 7-methylinosine, 6-  
 15 methoxy-guanosine, 1-methyl-guanosine, N2-methyl-guanosine, N2,N2-dimethyl-guanosine, 8-oxo-guanosine, 7-methyl-8-oxo-guanosine, 1-methyl-6-thio-guanosine, N2-methyl-6-thio-guanosine, and N2,N2-dimethyl-6-thio-guanosine.

In one embodiment, an mRNA comprises one or more pseudouridines, one or more 5-methyl-cytosines, and/or one or more 5-methyl-cytidines.

20 In one embodiment, an mRNA comprises one or more pseudouridines.

In one embodiment, an mRNA comprises one or more 5-methyl-cytidines.

In one embodiment, an mRNA comprises one or more 5-methyl-cytosines.

In particular embodiments, an mRNA contemplated herein comprises a poly(A) tail to help protect the mRNA from exonuclease degradation, stabilize the mRNA, and facilitate  
 25 translation. In certain embodiments, an mRNA comprises a 3' poly(A) tail structure.

The poly(A) may be encoded into the DNA template or added after transcription. In particular embodiments, an RNA contemplated herein comprises a poly(A) tail to help protect the RNA from exonuclease degradation, stabilize the RNA, and facilitate translation. In certain

embodiments, an RNA comprises a 3' poly(A) tail structure. Methods for polyadenylating RNA are known in the art (PL Wigley *et al. Mol Cell Biol.* 1990 Apr; 10(4): 1705–1713; and Wakiyama *et al.*, *Biochimie.* 1997 Dec;79(12):781-5)

In particular embodiments, the length of the poly(A) tail is at least about 10, 25, 50, 75,  
 5 100, 150, 200, 250, 300, 350, 400, 450, or at least about 500 or more adenine nucleotides or  
 any intervening number of adenine nucleotides. In particular embodiments, the length of the  
 poly(A) tail is at least about 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137,  
 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156,  
 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175,  
 10 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194,  
 195, 196, 197, 198, 199, 200, 201, 202, 202, 203, 205, 206, 207, 208, 209, 210, 211, 212, 213,  
 214, 215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232,  
 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251,  
 252, 253, 254, 255, 256, 257, 258, 259, 260, 261, 262, 263, 264, 265, 266, 267, 268, 269, 270,  
 15 271, 272, 273, 274, or 275 or more adenine nucleotides.

In particular embodiments, the length of the poly(A) tail is about 10 to about 500  
 adenine nucleotides, about 50 to about 500 adenine nucleotides, about 100 to about 500  
 adenine nucleotides, about 150 to about 500 adenine nucleotides, about 200 to about 500  
 adenine nucleotides, about 250 to about 500 adenine nucleotides, about 300 to about 500  
 20 adenine nucleotides, about 50 to about 450 adenine nucleotides, about 50 to about 400 adenine  
 nucleotides, about 50 to about 350 adenine nucleotides, about 100 to about 500 adenine  
 nucleotides, about 100 to about 450 adenine nucleotides, about 100 to about 400 adenine  
 nucleotides, about 100 to about 350 adenine nucleotides, about 100 to about 300 adenine  
 nucleotides, about 150 to about 500 adenine nucleotides, about 150 to about 450 adenine  
 25 nucleotides, about 150 to about 400 adenine nucleotides, about 150 to about 350 adenine  
 nucleotides, about 150 to about 300 adenine nucleotides, about 150 to about 250 adenine  
 nucleotides, about 150 to about 200 adenine nucleotides, about 200 to about 500 adenine  
 nucleotides, about 200 to about 450 adenine nucleotides, about 200 to about 400 adenine  
 nucleotides, about 200 to about 350 adenine nucleotides, about 200 to about 300 adenine  
 30 nucleotides, about 250 to about 500 adenine nucleotides, about 250 to about 450 adenine



nucleotides, about 250 to about 400 adenine nucleotides, about 250 to about 350 adenine nucleotides, or about 250 to about 300 adenine nucleotides or any intervening range of adenine nucleotides.

Terms that describe the orientation of polynucleotides include: 5' (normally the end of the polynucleotide having a free phosphate group) and 3' (normally the end of the polynucleotide having a free hydroxyl (OH) group). Polynucleotide sequences can be annotated in the 5' to 3' orientation or the 3' to 5' orientation. For DNA and mRNA, the 5' to 3' strand is designated the "sense," "plus," or "coding" strand because its sequence is identical to the sequence of the pre-messenger (pre-mRNA) [except for uracil (U) in RNA, instead of thymine (T) in DNA]. For DNA and mRNA, the complementary 3' to 5' strand which is the strand transcribed by the RNA polymerase is designated as "template," "antisense," "minus," or "non-coding" strand. As used herein, the term "reverse orientation" refers to a 5' to 3' sequence written in the 3' to 5' orientation or a 3' to 5' sequence written in the 5' to 3' orientation.

The terms "complementary" and "complementarity" refer to polynucleotides (*i.e.*, a sequence of nucleotides) related by the base-pairing rules. For example, the complementary strand of the DNA sequence 5' A G T C A T G 3' is 3' T C A G T A C 5'. The latter sequence is often written as the reverse complement with the 5' end on the left and the 3' end on the right, 5' C A T G A C T 3'. A sequence that is equal to its reverse complement is said to be a palindromic sequence. Complementarity can be "partial," in which only some of the nucleic acid bases are matched according to the base pairing rules. Or, there can be "complete" or "total" complementarity between the nucleic acids.

The term "nucleic acid cassette" or "expression cassette" as used herein refers to genetic sequences within the vector which can express an RNA, and subsequently a polypeptide. In one embodiment, the nucleic acid cassette contains a gene(s)-of-interest, *e.g.*, a polynucleotide(s)-of-interest. In another embodiment, the nucleic acid cassette contains one or more expression control sequences, *e.g.*, a promoter, enhancer, poly(A) sequence, and a gene(s)-of-interest, *e.g.*, a polynucleotide(s)-of-interest. Vectors may comprise 1, 2, 3, 4, 5, 6, 7, 8, 9 or 10 or more nucleic acid cassettes. The nucleic acid cassette is positionally and

sequentially oriented within the vector such that the nucleic acid in the cassette can be transcribed into RNA, and when necessary, translated into a protein or a polypeptide, undergo appropriate post-translational modifications required for activity in the transformed cell, and be translocated to the appropriate compartment for biological activity by targeting to appropriate intracellular compartments or secretion into extracellular compartments. Preferably, in some embodiments, the cassette has its 3' and 5' ends adapted for ready insertion into a vector and/or genome, *e.g.*, it has restriction endonuclease sites at each end. In a preferred embodiment, the nucleic acid cassette contains the sequence of a therapeutic gene used to treat, prevent, or ameliorate a genetic disorder. The cassette can be removed and inserted into a plasmid or viral vector as a single unit.

Polynucleotides include polynucleotide(s)-of-interest. As used herein, the term "polynucleotide-of-interest" refers to a polynucleotide encoding a polypeptide or fusion polypeptide, such as an engineered HE, megaTAL, donor repair template (*e.g.*, comprising a therapeutic transgene), as contemplated herein or refers to a polynucleotide that serves as a template for the transcription of an inhibitory polynucleotide, as contemplated herein.

Moreover, it will be appreciated by those of ordinary skill in the art that, as a result of the degeneracy of the genetic code, there are many nucleotide sequences that may encode a polypeptide, or fragment of variant thereof, as contemplated herein. Some of these polynucleotides bear minimal homology to the nucleotide sequence of any native gene. Nonetheless, polynucleotides that vary due to differences in codon usage are specifically contemplated in particular embodiments, for example polynucleotides that are optimized for human and/or primate codon selection. In one embodiment, polynucleotides comprising particular allelic sequences are provided. Alleles are endogenous polynucleotide sequences that are altered as a result of one or more mutations, such as deletions, additions and/or substitutions of nucleotides.

In a certain embodiment, a polynucleotide-of-interest comprises a donor repair template. In a certain preferred embodiment, a polynucleotide-of-interest comprises a donor repair template that comprises a therapeutic transgene

In a certain embodiment, a polynucleotide-of-interest comprises an inhibitory polynucleotide including, but not limited to, an siRNA, an miRNA, an shRNA, a ribozyme or another inhibitory RNA.

In one embodiment, a donor repair template comprising an inhibitory RNA comprises  
5 one or more regulatory sequences, such as, for example, a strong constitutive pol III, *e.g.*, human or mouse U6 snRNA promoter, the human and mouse H1 RNA promoter, or the human tRNA-val promoter, or a strong constitutive pol II promoter, as described elsewhere herein.

The polynucleotides contemplated in particular embodiments, regardless of the length of the coding sequence itself, may be combined with other DNA sequences, such as promoters  
10 and/or enhancers, untranslated regions (UTRs), Kozak sequences, polyadenylation signals, additional restriction enzyme sites, multiple cloning sites, internal ribosomal entry sites (IRES), recombinase recognition sites (*e.g.*, LoxP, FRT, and Att sites), termination codons, transcriptional termination signals, post-transcription response elements, and polynucleotides encoding self-cleaving polypeptides, epitope tags, as disclosed elsewhere herein or as known in  
15 the art, such that their overall length may vary considerably. It is therefore contemplated in particular embodiments that a polynucleotide fragment of almost any length may be employed, with the total length preferably being limited by the ease of preparation and use in the intended recombinant DNA protocol.

Polynucleotides can be prepared, manipulated, expressed and/or delivered using any of  
20 a variety of well-established techniques known and available in the art. In order to express a desired polypeptide, a nucleotide sequence encoding the polypeptide, can be inserted into appropriate vector. A desired polypeptide can also be expressed by delivering an mRNA encoding the polypeptide into the cell.

Illustrative examples of vectors include, but are not limited to plasmid, autonomously  
25 replicating sequences, and transposable elements, *e.g.*, Sleeping Beauty, PiggyBac.

Additional illustrative examples of vectors include, without limitation, plasmids, phagemids, cosmids, artificial chromosomes such as yeast artificial chromosome (YAC), bacterial artificial chromosome (BAC), or P1-derived artificial chromosome (PAC), bacteriophages such as lambda phage or M13 phage, and animal viruses.

Illustrative examples of viruses useful as vectors include, without limitation, retrovirus (including lentivirus), adenovirus, adeno-associated virus, herpesvirus (*e.g.*, herpes simplex virus), poxvirus, baculovirus, papillomavirus, and papovavirus (*e.g.*, SV40).

Illustrative examples of expression vectors include, but are not limited to pCIneo  
5 vectors (Promega) for expression in mammalian cells; pLenti4/V5-DEST™, pLenti6/V5-DEST™, and pLenti6.2/V5-GW/lacZ (Invitrogen) for lentivirus-mediated gene transfer and expression in mammalian cells. In particular embodiments, coding sequences of polypeptides disclosed herein can be ligated into such expression vectors for the expression of the polypeptides in mammalian cells.

10 In particular embodiments, the vector is an episomal vector or a vector that is maintained extrachromosomally. As used herein, the term “episomal” refers to a vector that is able to replicate without integration into host’s chromosomal DNA and without gradual loss from a dividing host cell also meaning that said vector replicates extrachromosomally or episomally.

15 “Expression control sequences,” “control elements,” or “regulatory sequences” present in an expression vector are those non-translated regions of the vector—origin of replication, selection cassettes, promoters, enhancers, translation initiation signals (Shine Dalgarno sequence or Kozak sequence) introns, post-transcriptional regulatory elements, a polyadenylation sequence, 5’ and 3’ untranslated regions—which interact with host cellular  
20 proteins to carry out transcription and translation. Such elements may vary in their strength and specificity. Depending on the vector system and host utilized, any number of suitable transcription and translation elements, including ubiquitous promoters and inducible promoters may be used.

In particular embodiments, a polynucleotide comprises a vector, including but not  
25 limited to expression vectors and viral vectors. A vector may comprise one or more exogenous, endogenous, or heterologous control sequences such as promoters and/or enhancers. An “endogenous control sequence” is one which is naturally linked with a given gene in the genome. An “exogenous control sequence” is one which is placed in juxtaposition to a gene by means of genetic manipulation (*i.e.*, molecular biological techniques) such that

transcription of that gene is directed by the linked enhancer/promoter. A “heterologous control sequence” is an exogenous sequence that is from a different species than the cell being genetically manipulated. A “synthetic” control sequence may comprise elements of one more endogenous and/or exogenous sequences, and/or sequences determined *in vitro* or *in silico* that  
5 provide optimal promoter and/or enhancer activity for the particular therapy.

The term “promoter” as used herein refers to a recognition site of a polynucleotide (DNA or RNA) to which an RNA polymerase binds. An RNA polymerase initiates and transcribes polynucleotides operably linked to the promoter. In particular embodiments, promoters operative in mammalian cells comprise an AT-rich region located approximately 25  
10 to 30 bases upstream from the site where transcription is initiated and/or another sequence found 70 to 80 bases upstream from the start of transcription, a CNCAAT region where N may be any nucleotide.

The term “enhancer” refers to a segment of DNA which contains sequences capable of providing enhanced transcription and in some instances can function independent of their  
15 orientation relative to another control sequence. An enhancer can function cooperatively or additively with promoters and/or other enhancer elements. The term “promoter/enhancer” refers to a segment of DNA which contains sequences capable of providing both promoter and enhancer functions.

The term “operably linked”, refers to a juxtaposition wherein the components described  
20 are in a relationship permitting them to function in their intended manner. In one embodiment, the term refers to a functional linkage between a nucleic acid expression control sequence (such as a promoter, and/or enhancer) and a second polynucleotide sequence, *e.g.*, a polynucleotide-of-interest, wherein the expression control sequence directs transcription of the nucleic acid corresponding to the second sequence.

25 As used herein, the term “constitutive expression control sequence” refers to a promoter, enhancer, or promoter/enhancer that continually or continuously allows for transcription of an operably linked sequence. A constitutive expression control sequence may be a “ubiquitous” promoter, enhancer, or promoter/enhancer that allows expression in a wide variety of cell and tissue types or a “cell specific,” “cell type specific,” “cell lineage specific,”

or “tissue specific” promoter, enhancer, or promoter/enhancer that allows expression in a restricted variety of cell and tissue types, respectively.

Illustrative ubiquitous expression control sequences suitable for use in particular embodiments include, but are not limited to, a cytomegalovirus (CMV) immediate early promoter, a viral simian virus 40 (SV40) (*e.g.*, early or late), a Moloney murine leukemia virus (MoMLV) LTR promoter, a Rous sarcoma virus (RSV) LTR, a herpes simplex virus (HSV) (thymidine kinase) promoter, H5, P7.5, and P11 promoters from vaccinia virus, a short elongation factor 1-alpha (EF1a-short) promoter, a long elongation factor 1-alpha (EF1a-long) promoter, early growth response 1 (EGR1), ferritin H (FerH), ferritin L (FerL), Glyceraldehyde 3-phosphate dehydrogenase (GAPDH), eukaryotic translation initiation factor 4A1 (EIF4A1), heat shock 70kDa protein 5 (HSPA5), heat shock protein 90kDa beta, member 1 (HSP90B1), heat shock protein 70kDa (HSP70),  $\beta$ -kinesin ( $\beta$ -KIN), the human ROSA 26 locus (Irions *et al.*, *Nature Biotechnology* 25, 1477 - 1482 (2007)), a Ubiquitin C promoter (UBC), a phosphoglycerate kinase-1 (PGK) promoter, a cytomegalovirus enhancer/chicken  $\beta$ -actin (CAG) promoter, a  $\beta$ -actin promoter and a myeloproliferative sarcoma virus enhancer, negative control region deleted, dl587rev primer-binding site substituted (MND) promoter (Challita *et al.*, *J Virol.* 69(2):748-55 (1995)).

In a particular embodiment, it may be desirable to use a cell, cell type, cell lineage or tissue specific expression control sequence to achieve cell type specific, lineage specific, or tissue specific expression of a desired polynucleotide sequence (*e.g.*, to express a particular nucleic acid encoding a polypeptide in only a subset of cell types, cell lineages, or tissues or during specific stages of development).

As used herein, “conditional expression” may refer to any type of conditional expression including, but not limited to, inducible expression; repressible expression; expression in cells or tissues having a particular physiological, biological, or disease state, *etc.* This definition is not intended to exclude cell type or tissue specific expression. Certain embodiments provide conditional expression of a polynucleotide-of-interest, *e.g.*, expression is controlled by subjecting a cell, tissue, organism, *etc.*, to a treatment or condition that causes the

polynucleotide to be expressed or that causes an increase or decrease in expression of the polynucleotide encoded by the polynucleotide-of-interest.

Illustrative examples of inducible promoters/systems include, but are not limited to, steroid-inducible promoters such as promoters for genes encoding glucocorticoid or estrogen  
5 receptors (inducible by treatment with the corresponding hormone), metallothionein promoter (inducible by treatment with various heavy metals), MX-1 promoter (inducible by interferon), the “GeneSwitch” mifepristone-regulatable system (Sirin *et al.*, 2003, *Gene*, 323:67), the cumate inducible gene switch (WO 2002/088346), tetracycline-dependent regulatory systems, *etc.*

10 Conditional expression can also be achieved by using a site-specific DNA recombinase. According to certain embodiments, polynucleotides comprise at least one (typically two) site(s) for recombination mediated by a site-specific recombinase. As used herein, the terms “recombinase” or “site specific recombinase” include excisive or integrative proteins, enzymes, co-factors or associated proteins that are involved in recombination  
15 reactions involving one or more recombination sites (*e.g.*, two, three, four, five, six, seven, eight, nine, ten or more.), which may be wild-type proteins (*see* Landy, Current Opinion in Biotechnology 3:699-707 (1993)), or mutants, derivatives (*e.g.*, fusion proteins containing the recombination protein sequences or fragments thereof), fragments, and variants thereof. Illustrative examples of recombinases suitable for use in particular embodiments include, but  
20 are not limited to: Cre, Int, IHF, Xis, Flp, Fis, Hin, Gin,  $\Phi$ C31, Cin, Tn3 resolvase, TndX, XerC, XerD, TnpX, Hjc, Gin, SpCCE1, and ParA.

The polynucleotides may comprise one or more recombination sites for any of a wide variety of site-specific recombinases. It is to be understood that the target site for a site-specific recombinase is in addition to any site(s) required for integration of a vector, *e.g.*, a retroviral  
25 vector or lentiviral vector. As used herein, the terms “recombination sequence,” “recombination site,” or “site specific recombination site” refer to a particular nucleic acid sequence to which a recombinase recognizes and binds.

For example, one recombination site for Cre recombinase is loxP which is a 34 base pair sequence comprising two 13 base pair inverted repeats (serving as the recombinase

binding sites) flanking an 8 base pair core sequence (Sauer, B., *Current Opinion in Biotechnology* 5:521-527 (1994)). Other exemplary loxP sites include, but are not limited to: lox511 (Hoess *et al.*, 1996; Bethke and Sauer, 1997), lox5171 (Lee and Saito, 1998), lox2272 (Lee and Saito, 1998), m2 (Langer *et al.*, 2002), lox71 (Albert *et al.*, 1995), and lox66 (Albert  
5 *et al.*, 1995).

Suitable recognition sites for the FLP recombinase include, but are not limited to: FRT (McLeod, *et al.*, 1996), F<sub>1</sub>, F<sub>2</sub>, F<sub>3</sub> (Schlake and Bode, 1994), F<sub>4</sub>, F<sub>5</sub> (Schlake and Bode, 1994), FRT(LE) (Senecoff *et al.*, 1988), FRT(RE) (Senecoff *et al.*, 1988).

Other examples of recognition sequences are the attB, attP, attL, and attR sequences,  
10 which are recognized by the recombinase enzyme  $\lambda$  Integrase, *e.g.*, phi-c31. The  $\phi$ C31 SSR mediates recombination only between the heterotypic sites attB (34 bp in length) and attP (39 bp in length) (Groth *et al.*, 2000). attB and attP, named for the attachment sites for the phage integrase on the bacterial and phage genomes, respectively, both contain imperfect inverted repeats that are likely bound by  $\phi$ C31 homodimers (Groth *et al.*, 2000). The product sites, attL  
15 and attR, are effectively inert to further  $\phi$ C31-mediated recombination (Belteki *et al.*, 2003), making the reaction irreversible. For catalyzing insertions, it has been found that attB-bearing DNA inserts into a genomic attP site more readily than an attP site into a genomic attB site (Thyagarajan *et al.*, 2001; Belteki *et al.*, 2003). Thus, typical strategies position by  
20 homologous recombination an attP-bearing “docking site” into a defined locus, which is then partnered with an attB-bearing incoming sequence for insertion.

In one embodiment, a polynucleotide contemplated herein comprises a donor repair template polynucleotide flanked by a pair of recombinase recognition sites. In particular embodiments, the repair template polynucleotide is flanked by LoxP sites, FRT sites, or att sites.

25 In particular embodiments, polynucleotides contemplated herein, include one or more polynucleotides-of-interest that encode one or more polypeptides. In particular embodiments, to achieve efficient translation of each of the plurality of polypeptides, the polynucleotide sequences can be separated by one or more IRES sequences or polynucleotide sequences encoding self-cleaving polypeptides.



As used herein, an “internal ribosome entry site” or “IRES” refers to an element that promotes direct internal ribosome entry to the initiation codon, such as ATG, of a cistron (a protein encoding region), thereby leading to the cap-independent translation of the gene. *See, e.g., Jackson et al., 1990. Trends Biochem Sci 15(12):477-83 and Jackson and Kaminski.*

- 5 1995. *RNA* 1(10):985-1000. Examples of IRES generally employed by those of skill in the art include those described in U.S. Pat. No. 6,692,736. Further examples of “IRES” known in the art include, but are not limited to IRES obtainable from picornavirus (Jackson *et al.*, 1990) and IRES obtainable from viral or cellular mRNA sources, such as for example, immunoglobulin heavy-chain binding protein (BiP), the vascular endothelial growth factor (VEGF) (Huez *et al.*
- 10 1998. *Mol. Cell. Biol.* 18(11):6178-6190), the fibroblast growth factor 2 (FGF-2), and insulin-like growth factor (IGFII), the translational initiation factor eIF4G and yeast transcription factors TFIID and HAP4, the encephelomyocarditis virus (EMCV) which is commercially available from Novagen (Duke *et al.*, 1992. *J. Virol* 66(3):1602-9) and the VEGF IRES (Huez *et al.*, 1998. *Mol Cell Biol* 18(11):6178-90). IRES have also been reported in viral genomes of
- 15 Picornaviridae, Dicistroviridae and Flaviviridae species and in HCV, Friend murine leukemia virus (FrMLV) and Moloney murine leukemia virus (MoMLV).

In one embodiment, the IRES used in polynucleotides contemplated herein is an EMCV IRES.

- In particular embodiments, the polynucleotides comprise polynucleotides that have a
- 20 consensus Kozak sequence and that encode a desired polypeptide. As used herein, the term “Kozak sequence” refers to a short nucleotide sequence that greatly facilitates the initial binding of mRNA to the small subunit of the ribosome and increases translation. The consensus Kozak sequence is (GCC)RCCATGG (SEQ ID NO:81), where R is a purine (A or G) (Kozak, 1986. *Cell*. 44(2):283-92, and Kozak, 1987. *Nucleic Acids Res.* 15(20):8125-48).

- 25 Elements directing the efficient termination and polyadenylation of the heterologous nucleic acid transcripts increases heterologous gene expression. Transcription termination signals are generally found downstream of the polyadenylation signal. In particular embodiments, vectors comprise a polyadenylation sequence 3' of a polynucleotide encoding a polypeptide to be expressed. The term “polyA site” or “polyA sequence” as used herein

denotes a DNA sequence which directs both the termination and polyadenylation of the nascent RNA transcript by RNA polymerase II. Polyadenylation sequences can promote mRNA stability by addition of a polyA tail to the 3' end of the coding sequence and thus, contribute to increased translational efficiency. Cleavage and polyadenylation is directed by a poly(A) sequence in the RNA. The core poly(A) sequence for mammalian pre-mRNAs has two recognition elements flanking a cleavage-polyadenylation site. Typically, an almost invariant AAUAAA (SEQ ID NO: 82) hexamer lies 20-50 nucleotides upstream of a more variable element rich in U or GU residues. Cleavage of the nascent transcript occurs between these two elements and is coupled to the addition of up to 250 adenosines to the 5' cleavage product. In particular embodiments, the core poly(A) sequence is an ideal polyA sequence (e.g., AATAAA (SEQ ID NO: 83), ATTAAA (SEQ ID NO: 84), AGTAAA (SEQ ID NO: 85)). In particular embodiments, the poly(A) sequence is an SV40 polyA sequence, a bovine growth hormone polyA sequence (BGHPA), a rabbit  $\beta$ -globin polyA sequence (r $\beta$ gpA), variants thereof, or another suitable heterologous or endogenous polyA sequence known in the art. In certain embodiments the poly(A) sequence is represented by positions 2704-2864 of SEQ ID NO: 108.

In some embodiments, a polynucleotide or cell harboring the polynucleotide utilizes a suicide gene, including an inducible suicide gene to reduce the risk of direct toxicity and/or uncontrolled proliferation. In specific embodiments, the suicide gene is not immunogenic to the host harboring the polynucleotide or cell. A certain example of a suicide gene that may be used is caspase-9 or caspase-8 or cytosine deaminase. Caspase-9 can be activated using a specific chemical inducer of dimerization (CID).

In certain embodiments, polynucleotides comprise gene segments that cause the genetically modified cells contemplated herein to be susceptible to negative selection *in vivo*. "Negative selection" refers to an infused cell that can be eliminated as a result of a change in the *in vivo* condition of the individual. The negative selectable phenotype may result from the insertion of a gene that confers sensitivity to an administered agent, for example, a compound. Negative selection genes are known in the art, and include, but are not limited to: the Herpes simplex virus type I thymidine kinase (HSV-I TK) gene which confers ganciclovir sensitivity; the cellular hypoxanthine phosphoribosyltransferase (HPRT) gene, the cellular adenine phosphoribosyltransferase (APRT) gene, and bacterial cytosine deaminase.

In some embodiments, genetically modified cells comprise a polynucleotide further comprising a positive marker that enables the selection of cells of the negative selectable phenotype *in vitro*. The positive selectable marker may be a gene, which upon being introduced into the host cell, expresses a dominant phenotype permitting positive selection of  
5 cells carrying the gene. Genes of this type are known in the art, and include, but are not limited to hygromycin-B phosphotransferase gene (hph) which confers resistance to hygromycin B, the amino glycoside phosphotransferase gene (neo or aph) from Tn5 which codes for resistance to the antibiotic G418, the dihydrofolate reductase (DHFR) gene, the adenosine deaminase gene (ADA), and the multi-drug resistance (MDR) gene.

10 In one embodiment, the positive selectable marker and the negative selectable element are linked such that loss of the negative selectable element necessarily also is accompanied by loss of the positive selectable marker. In a particular embodiment, the positive and negative selectable markers are fused so that loss of one obligatorily leads to loss of the other. An example of a fused polynucleotide that yields as an expression product a polypeptide that  
15 confers both the desired positive and negative selection features described above is a hygromycin phosphotransferase thymidine kinase fusion gene (HyTK). Expression of this gene yields a polypeptide that confers hygromycin B resistance for positive selection *in vitro*, and ganciclovir sensitivity for negative selection *in vivo*. See also the publications of PCT US91/08442 and PCT/US94/05601, by S. D. Lupton, describing the use of bifunctional  
20 selectable fusion genes derived from fusing a dominant positive selectable markers with negative selectable markers.

Preferred positive selectable markers are derived from genes selected from the group consisting of hph, nco, and gpt, and preferred negative selectable markers are derived from genes selected from the group consisting of cytosine deaminase, HSV-I TK, VZV TK, HPRT,  
25 APRT and gpt. Exemplary bifunctional selectable fusion genes contemplated in particular embodiments include, but are not limited to genes wherein the positive selectable marker is derived from hph or neo, and the negative selectable marker is derived from cytosine deaminase or a TK gene or selectable marker.

In particular embodiments, polynucleotides encoding one or more nuclease variants, megaTALs, end-processing enzymes, or fusion polypeptides may be introduced into cells, *e.g.*, hepatocytes, by both non-viral and viral methods. In particular embodiments, delivery of one or more polynucleotides encoding nucleases and/or donor repair templates may be  
5 provided by the same method or by different methods, and/or by the same vector or by different vectors.

The term “vector” is used herein to refer to a nucleic acid molecule capable transferring or transporting another nucleic acid molecule. The transferred nucleic acid is generally linked to, *e.g.*, inserted into, the vector nucleic acid molecule. A vector may include sequences that  
10 direct autonomous replication in a cell, or may include sequences sufficient to allow integration into host cell DNA. In particular embodiments, non-viral vectors are used to deliver one or more polynucleotides contemplated herein to a T cell.

Illustrative examples of non-viral vectors include, but are not limited to plasmids (*e.g.*, DNA plasmids or RNA plasmids), transposons, cosmids, and bacterial artificial  
15 chromosomes.

Illustrative methods of non-viral delivery of polynucleotides contemplated in particular embodiments include, but are not limited to: electroporation, sonoporation, lipofection, microinjection, biolistics, virosomes, liposomes, lipid nanoparticles, immunoliposomes, nanoparticles, polycation or lipid:nucleic acid conjugates, naked DNA,  
20 artificial virions, DEAE-dextran-mediated transfer, gene gun, and heat-shock.

Illustrative examples of polynucleotide delivery systems suitable for use in particular embodiments contemplated in particular embodiments include, but are not limited to those provided by Amaxa Biosystems, Maxcyte, Inc., BTX Molecular Delivery Systems, and Copernicus Therapeutics Inc. Lipofection reagents are sold commercially (*e.g.*,  
25 Transfectam™ and Lipofectin™). Cationic and neutral lipids that are suitable for efficient receptor-recognition lipofection of polynucleotides have been described in the literature. See *e.g.*, Liu *et al.* (2003) *Gene Therapy*. 10:180–187; and Balazs *et al.* (2011) *Journal of Drug Delivery*. 2011:1-12. Antibody-targeted, bacterially derived, non-living nanocell-based delivery is also contemplated in particular embodiments.

In a certain embodiment, a polynucleotide encoding an engineered HE, a megaTALs, end-processing enzymes, or fusion polypeptides thereof is introduced into cells, *e.g.*, hepatocytes, by nanolipid particles. In a preferred embodiment, a mRNA encoding an engineered HE, an engineered megaTALs, end-processing enzymes, or fusion polypeptides thereof is introduced into hepatocytes, by lipid nanoparticles. The lipid nanoparticles comprising the mRNA encoding an engineered HE, an engineered megaTALs, end-processing enzymes, or fusion polypeptides thereof can be delivered *in vivo* by administration to an individual patient, typically by systemic administration (*e.g.*, intravenous, intraperitoneal, intramuscular, subdermal, or intracranial infusion), as described below.

In a certain embodiment, a polynucleotide encoding an engineered HE, a megaTALs, end-processing enzymes, or fusion polypeptides thereof or the polynucleotide of the donor repair template is introduced into cells, *e.g.*, hepatocytes, via self-replicating RNAs. As used herein, “self-replicating RNA” or “srRNA” is a type of RNA that can replicate itself. srRNA are derived from the genomes of positive strand RNA viruses. It would be understood by the skilled artisan how to employ srRNA with the polynucleotides contemplated herein. Illustrative examples of srRNA are disclosed in Lin *et al* Molecular Therapy: Nucleic Acids, Vol 23, June 2023: 650-666, and references cited therein, all of which are incorporated herein by reference in their entireties. The srRNA comprising the polynucleotides encoding an engineered HE, an engineered megaTALs, end-processing enzymes, or fusion polypeptides thereof or the polynucleotide of the donor repair template can be delivered *in vivo* by administration to an individual patient, typically by systemic administration (*e.g.*, intravenous, intraperitoneal, intramuscular, subdermal, or intracranial infusion), as described below.

Viral vectors comprising polynucleotides contemplated in particular embodiments, such as donor repair templates, can be delivered *in vivo* by administration to an individual patient, typically by systemic administration (*e.g.*, intravenous, intraperitoneal, intramuscular, subdermal, or intracranial infusion) or topical application, as described below. Alternatively, vectors can be delivered to cells *ex vivo*, such as cells explanted from an individual patient (*e.g.*, mobilized peripheral blood, lymphocytes, bone marrow aspirates, tissue biopsy, *etc.*) or

universal donor hematopoietic stem cells, followed by reimplantation of the cells into a patient.

In one embodiment, viral vectors comprising one or more polynucleotides encoding nuclease variants or megaTALs and/or comprising donor repair templates are administered  
5 directly to an organism for transduction of cells *in vivo*. Alternatively, naked DNA can be administered. Administration is by any of the routes normally used for introducing a molecule into ultimate contact with blood or tissue cells including, but not limited to, injection, infusion, topical application and electroporation. Suitable methods of administering such nucleic acids are available and well known to those of skill in the art, and,  
10 although more than one route can be used to administer a particular composition, a particular route can often provide a more immediate and more effective reaction than another route.

Illustrative examples of viral vector systems suitable for use in particular embodiments contemplated herein include, but are not limited to adeno-associated virus (AAV), retrovirus, herpes simplex virus, adenovirus, and vaccinia virus vectors.

In various embodiments, one or more polynucleotides encoding a nuclease variant or megaTAL and/or donor repair template are introduced into a cell, *e.g.*, a hepatocyte or hematopoietic cell, by transducing the cell with a recombinant adeno-associated virus (rAAV), comprising the one or more polynucleotides. In particular embodiments, a donor repair template comprising a therapeutic transgene is introduced into a hepatocyte by  
20 transducing the hepatocyte with a rAAV comprising the donor repair template.

AAV is a small (~26 nm) replication-defective, primarily episomal, non-enveloped virus. AAV can infect both dividing and non-dividing cells and may incorporate its genome into that of the host cell. Recombinant AAV (rAAV) are typically composed of, at a minimum, a therapeutic transgene and its regulatory sequences, and 5' and 3' AAV inverted  
25 terminal repeats (ITRs). The ITR sequences are about 145 bp in length. In particular embodiments, the rAAV comprises ITRs and/or capsid sequences isolated from AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, or AAV10.

In some embodiments, a chimeric rAAV is used the ITR sequences are isolated from one AAV serotype and the capsid sequences are isolated from a different AAV serotype. For

example, a rAAV with ITR sequences derived from AAV2 and capsid sequences derived from AAV6 is referred to as AAV2/AAV6. In particular embodiments, the rAAV vector may comprise ITRs from AAV2, and capsid proteins from any one of AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, or AAV10.

- 5           In some embodiments, engineering and selection methods can be applied to AAV capsids to make them more likely to transduce cells of interest.

Construction of rAAV vectors, production, and purification thereof have been disclosed, *e.g.*, in U.S. Patent Nos. 9,169,494; 9,169,492; 9,012,224; 8,889,641; 8,809,058; and 8,784,799, each of which is incorporated by reference herein, in its entirety.

- 10           In various embodiments, one or more polynucleotides encoding an engineered nuclease, a megaTAL, and/or donor repair template are introduced into a cell (*e.g.*, hepatocyte or hematopoietic cell) by transducing the cell with a retrovirus, *e.g.*, lentivirus, comprising the one or more polynucleotides.

- As used herein, the term “retrovirus” refers to an RNA virus that reverse transcribes  
15 its genomic RNA into a linear double-stranded DNA copy and subsequently covalently integrates its genomic DNA into a host genome. Illustrative retroviruses suitable for use in particular embodiments, include, but are not limited to: Moloney murine leukemia virus (M-MuLV), Moloney murine sarcoma virus (MoMSV), Harvey murine sarcoma virus (HaMuSV), murine mammary tumor virus (MuMTV), gibbon ape leukemia virus (GaLV),  
20 feline leukemia virus (FLV), spumavirus, Friend murine leukemia virus, Murine Stem Cell Virus (MSCV) and Rous Sarcoma Virus (RSV)) and lentivirus.

- As used herein, the term “lentivirus” refers to a group (or genus) of complex retroviruses. Illustrative lentiviruses include, but are not limited to: HIV (human immunodeficiency virus; including HIV type 1, and HIV 2); visna-maedi virus (VMV) virus;  
25 the caprine arthritis-encephalitis virus (CAEV); equine infectious anemia virus (EIAV); feline immunodeficiency virus (FIV); bovine immune deficiency virus (BIV); and simian immunodeficiency virus (SIV). In one embodiment, HIV based vector backbones (*i.e.*, HIV cis-acting sequence elements) are preferred.

In various embodiments, a lentiviral vector contemplated herein comprises one or more LTRs, and one or more, or all, of the following accessory elements: a cPPT/FLAP, a Psi ( $\Psi$ ) packaging signal, an export element, poly (A) sequences, and may optionally comprise a WPRE or HPRE, an insulator element, a selectable marker, and a cell suicide gene, as  
5 discussed elsewhere herein.

In particular embodiments, lentiviral vectors contemplated herein may be integrative or non-integrating or integration defective lentivirus. As used herein, the term “integration defective lentivirus” or “IDLV” refers to a lentivirus having an integrase that lacks the capacity to integrate the viral genome into the genome of the host cells. Integration-incompetent viral  
10 vectors have been described in patent application WO 2006/010834, which is herein incorporated by reference in its entirety.

Illustrative mutations in the HIV-1 pol gene suitable to reduce integrase activity include, but are not limited to: H12N, H12C, H16C, H16V, S81 R, D41A, K42A, H51A, Q53C, D55V, D64E, D64V, E69A, K71A, E85A, E87A, D116N, D116I, D116A, N120G,  
15 N120I, N120E, E152G, E152A, D35E, K156E, K156A, E157A, K159E, K159A, K160A, R166A, D167A, E170A, H171A, K173A, K186Q, K186T, K188T, E198A, R199c, R199T, R199A, D202A, K211A, Q214L, Q216L, Q221 L, W235F, W235E, K236S, K236A, K246A, G247W, D253A, R262A, R263A and K264H.

In one embodiment, the HIV-1 integrase deficient pol gene comprises a D64V, D116I,  
20 D116A, E152G, or E152A mutation; D64V, D116I, and E152G mutations; or D64V, D116A, and E152A mutations.

In one embodiment, the HIV-1 integrase deficient pol gene comprises a D64V mutation.

The term “long terminal repeat (LTR)” refers to domains of base pairs located at the  
25 ends of retroviral DNAs which, in their natural sequence context, are direct repeats and contain U3, R and U5 regions.

As used herein, the term “FLAP element” or “cPPT/FLAP” refers to a nucleic acid whose sequence includes the central polypurine tract and central termination sequences (cPPT and CTS) of a retrovirus, *e.g.*, HIV-1 or HIV-2. Suitable FLAP elements are described in U.S.



Pat. No. 6,682,907 and in Zennou, *et al.*, 2000, *Cell*, 101:173. In another embodiment, a lentiviral vector contains a FLAP element with one or more mutations in the cPPT and/or CTS elements. In yet another embodiment, a lentiviral vector comprises either a cPPT or CTS element. In yet another embodiment, a lentiviral vector does not comprise a cPPT or  
5 CTS element.

As used herein, the term “packaging signal” or “packaging sequence” refers to psi [ $\Psi$ ] sequences located within the retroviral genome which are required for insertion of the viral RNA into the viral capsid or particle, *see e.g.*, Clever *et al.*, 1995, *J. of Virology*, Vol. 69, No. 4; pp. 2101–2109.

10 The term “export element” refers to a cis-acting post-transcriptional regulatory element which regulates the transport of an RNA transcript from the nucleus to the cytoplasm of a cell. Examples of RNA export elements include, but are not limited to, the human immunodeficiency virus (HIV) rev response element (RRE) (*see e.g.*, Cullen *et al.*, 1991, *J. Virol.* 65: 1053; and Cullen *et al.*, 1991, *Cell* 58: 423), and the hepatitis B virus post-  
15 transcriptional regulatory element (HPRE).

In particular embodiments, expression of heterologous sequences in viral vectors is increased by incorporating posttranscriptional regulatory elements, efficient polyadenylation sites, and optionally, transcription termination signals into the vectors. A variety of posttranscriptional regulatory elements can increase expression of a heterologous nucleic acid  
20 at the protein, *e.g.*, woodchuck hepatitis virus posttranscriptional regulatory element (WPRE; Zufferey *et al.*, 1999, *J. Virol.*, 73:2886); the posttranscriptional regulatory element present in hepatitis B virus (HPRE) (Huang *et al.*, *Mol. Cell. Biol.*, 5:3864); and the like (Liu *et al.*, 1995, *Genes Dev.*, 9:1766).

Lentiviral vectors preferably contain several safety enhancements as a result of  
25 modifying the LTRs. “Self-inactivating” (SIN) vectors refers to replication-defective vectors, *e.g.*, in which the right (3') LTR enhancer-promoter region, known as the U3 region, has been modified (*e.g.*, by deletion or substitution) to prevent viral transcription beyond the first round of viral replication. An additional safety enhancement is provided by replacing the U3 region of the 5' LTR with a heterologous promoter to drive transcription of the viral genome during

production of viral particles. Examples of heterologous promoters which can be used include, for example, viral simian virus 40 (SV40) (*e.g.*, early or late), cytomegalovirus (CMV) (*e.g.*, immediate early), Moloney murine leukemia virus (MoMLV), Rous sarcoma virus (RSV), and herpes simplex virus (HSV) (thymidine kinase) promoters.

- 5           The terms “pseudotype” or “pseudotyping” as used herein, refer to a virus whose viral envelope proteins have been substituted with those of another virus possessing preferable characteristics. For example, HIV can be pseudotyped with vesicular stomatitis virus G-protein (VSV-G) envelope proteins, which allows HIV to infect a wider range of cells because HIV envelope proteins (encoded by the *env* gene) normally target the virus to
- 10   CD4<sup>+</sup> presenting cells.

In certain embodiments, lentiviral vectors are produced according to known methods. *See e.g.*, Kutner *et al.*, *BMC Biotechnol.* 2009;9:10. doi: 10.1186/1472-6750-9-10; Kutner *et al.* *Nat. Protoc.* 2009;4(4):495–505. doi: 10.1038/nprot.2009.22.

- According to certain specific embodiments contemplated herein, most or all of the
- 15   viral vector backbone sequences are derived from a lentivirus, *e.g.*, HIV-1. However, it is to be understood that many different sources of retroviral and/or lentiviral sequences can be used, or combined and numerous substitutions and alterations in certain of the lentiviral sequences may be accommodated without impairing the ability of a transfer vector to perform the functions described herein. Moreover, a variety of lentiviral vectors are known
- 20   in the art, *see* Naldini *et al.*, (1996a, 1996b, and 1998); Zufferey *et al.*, (1997); Dull *et al.*, 1998, U.S. Pat. Nos. 6,013,516; and 5,994,136, many of which may be adapted to produce a viral vector or transfer plasmid contemplated herein.

- In various embodiments, one or more polynucleotides encoding a nuclease variant or a megaTAL and/or donor repair template are introduced into a cell (*e.g.*, hepatocyte or
- 25   hematopoietic cell) by transducing the cell with an adenovirus (Ad) comprising the one or more polynucleotides.

Adenoviral based vectors are capable of very high transduction efficiency in many cell types and do not require cell division. With such vectors, high titer and high levels of expression have been obtained. This vector can be produced in large quantities in a relatively

simple system. Most adenovirus vectors are engineered such that a therapeutic transgene replaces the Ad E1a, E1b, and/or E3 genes; subsequently the replication defective vector is propagated in human 293 cells that supply deleted gene function in trans. Ad vectors can transduce multiple types of tissues *in vivo*, including non-dividing, differentiated cells such as those found in liver, kidney and muscle. Conventional Ad vectors have a large carrying capacity.

Generation and propagation of the current adenovirus vectors, which are replication deficient, may utilize a unique helper cell line, designated 293, which was transformed from human embryonic kidney cells by Ad5 DNA fragments and constitutively expresses E1 proteins (Graham *et al.*, 1977). Since the E3 region is dispensable from the adenovirus genome (Jones & Shenk, 1978), the current adenovirus vectors, with the help of 293 cells, carry foreign DNA in either the E1, the D3 or both regions (Graham & Prevec, 1991). Adenovirus vectors have been used in eukaryotic gene expression (Levrero *et al.*, 1991; Gomez-Foix *et al.*, 1992) and vaccine development (Grunhaus & Horwitz, 1992; Graham & Prevec, 1992). Studies in administering recombinant adenovirus to different tissues include trachea instillation (Rosenfeld *et al.*, 1991; Rosenfeld *et al.*, 1992), muscle injection (Ragot *et al.*, 1993), peripheral intravenous injections (Herz & Gerard, 1993) and stereotactic inoculation into the brain (Le Gal La Salle *et al.*, 1993). An example of the use of an Ad vector in a clinical trial involved polynucleotide therapy for antitumor immunization with intramuscular injection (Stermann *et al.*, *Hum. Gene Ther.* 7:1083-9 (1998)).

In various embodiments, one or more polynucleotides encoding nuclease variant, megaTAL, and/or donor repair template are introduced into a hepatocyte cell by transducing the cell with a herpes simplex virus, *e.g.*, HSV-1, HSV-2, comprising the one or more polynucleotides.

The mature HSV virion consists of an enveloped icosahedral capsid with a viral genome consisting of a linear double-stranded DNA molecule that is 152 kb. In one embodiment, the HSV based viral vector is deficient in one or more essential or non-essential HSV genes. In one embodiment, the HSV based viral vector is replication deficient. Most replication deficient HSV vectors contain a deletion to remove one or more intermediate-early,

early, or late HSV genes to prevent replication. For example, the HSV vector may be deficient in an immediate early gene selected from the group consisting of: ICP4, ICP22, ICP27, ICP47, and a combination thereof. Advantages of the HSV vector are its ability to enter a latent stage that can result in long-term DNA expression and its large viral DNA genome that can  
5 accommodate exogenous DNA inserts of up to 25 kb. HSV-based vectors are described in, for example, U.S. Pat. Nos. 5,837,532, 5,846,782, and 5,804,413, and International Patent Applications WO 91/02788, WO 96/04394, WO 98/15637, and WO 99/06583, each of which are incorporated by reference herein in its entirety.

#### H. GENOME EDITED CELLS

10 The genome edited cells contemplated herein may be edited *in vitro*, *ex vivo*, or *in vivo*. The genome edited cells contemplated in particular embodiments comprise one or more gene edits in an ALB gene and provide improved cell-based therapeutics for the prevention, treatment, or amelioration of at least one symptom, of a bleeding disorder, lysosomal storage disease, or metabolic disorder.

15 Genome edited cells contemplated in particular embodiments may be autologous/autogeneic (“self”) or non-autologous (“non-self,” *e.g.*, allogeneic, syngeneic or xenogeneic). “Autologous,” as used herein, refers to cells from the same subject. “Allogeneic,” as used herein, refers to cells of the same species that differ genetically to the cell in comparison. “Syngeneic,” as used herein, refers to cells of a different subject that are  
20 genetically identical to the cell in comparison. “Xenogeneic,” as used herein, refers to cells of a different species to the cell in comparison. In preferred embodiments, the cells are obtained from a mammalian subject. In a more preferred embodiment, the cells are obtained from a primate subject, optionally a non-human primate. In the most preferred embodiment, the cells are obtained from a human subject.

25 An “isolated cell” refers to a non-naturally occurring cell, *e.g.*, a cell that does not exist in nature, a modified cell, an engineered cell, *etc.*, that has been obtained from an *in vivo* tissue or organ and is substantially free of extracellular matrix.

As used herein, the term “population of cells” refers to a plurality of cells that may be made up of any number and/or combination of homogenous or heterogeneous cell types, as

described elsewhere herein. A population of cells may comprise about 10%, about 20%, about 30%, about 40%, about 50%, about 60%, about 70%, about 80%, about 90%, or about 100% of the target cell type to be edited. In certain embodiments, cells may be isolated or purified from a population of heterogeneous cells using methods known in the art.

5 Illustrative examples of cell types whose genome can be edited using the compositions and methods contemplated herein include, but are not limited to, cell lines, primary cells, stem cells, induced pluripotent stem cells (iPSCs), progenitor cells, and differentiated cells, and mixtures thereof.

In preferred embodiments, the genome editing compositions and methods are used to  
10 edit hepatocytes.

In other embodiments, the genome editing compositions and methods are used to edit hematopoietic cells, more preferably immune cells (*e.g.*, T cells).

In particular embodiments, a population of cells (*e.g.*, hepatocytes or hematopoietic cells) comprises an edited ALB gene, wherein the edit is a DSB repaired by NHEJ. In  
15 particular embodiments, the cell comprises an edited ALB gene, wherein the edit is a DSB repaired by NHEJ. In particular embodiments, the edit is an insertion or deletion (INDEL) of about 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 or more nucleotides in a coding sequence of the ALB gene, preferably in intron 1 of the ALB gene (SEQ ID NO: 31), more preferably at SEQ ID NO: 34 (or SEQ ID NO: 43) in intron 1 of the  
20 ALB gene. In particular embodiments, the edit comprises an insertion of a therapeutic transgene, such as FVIII or FIX, in a coding sequence of the ALB gene, preferably in intron 1 of the ALB gene (SEQ ID NO: 31), more preferably at SEQ ID NO: 34 (or SEQ ID NO: 43) in intron 1 of the ALB gene.

In particular embodiments, a population of cells (*e.g.*, hepatocytes or hematopoietic  
25 cells) comprises an edited ALB gene comprising a donor repair template incorporated at a DSB repaired by NHEJ.

In particular embodiments, a population of cells (*e.g.*, hepatocytes or hematopoietic cells) comprises an edited ALB gene comprising a donor repair template incorporated at a DSB repaired by HDR.

In certain particular embodiments, a population of hepatocytes cells comprises an edited ALB gene comprising a donor repair template comprising a therapeutic transgene, such as FVIII or FIX, incorporated at a DSB.

In various embodiments, a genome edited cell comprises an edit in the ALB gene and  
 5 further comprises a polynucleotide encoding an antihemophilic factor (*e.g.*, FVIII or FIX), antibody, BiTE, recombinant protein, enzyme, cytokine, chemokine, cytotoxin, cytokine receptor, engineered or chimeric antigen receptor, TCR, zetakine, hormone, or functional variant thereof. In particular embodiments, a genome edited hepatocyte cell comprises an edit in the ALB gene and further comprises a polynucleotide encoding FVIII or FIX.

## 10 I. COMPOSITIONS AND FORMULATIONS

The compositions contemplated in particular embodiments may comprise one or more polypeptides, polynucleotides, vectors comprising same, and genome editing compositions and genome edited cell compositions, as contemplated herein. The genome editing compositions and methods contemplated in particular embodiments are useful for editing a target site in the  
 15 human ALB gene in a cell or a population of cells (*e.g.*, hepatocytes or hematopoietic cells). In preferred embodiments, a genome editing composition is used to edit an ALB gene in a hepatocyte.

In various embodiments, the compositions contemplated herein comprise a nuclease variant, and optionally an end-processing enzyme, *e.g.*, a 3'-5' exonuclease (Trex2). The  
 20 nuclease variant may be in the form of an mRNA that is introduced into a cell via polynucleotide delivery methods disclosed *supra*, *e.g.*, electroporation, lipid nanoparticles, *etc.* In one embodiment, a composition comprising an mRNA encoding a homing endonuclease variant or megaTAL, and optionally a 3'-5' exonuclease, is introduced in a cell via polynucleotide delivery methods disclosed *supra*, *e.g.*, lipid nanoparticles. The composition  
 25 may be used to generate a genome edited cell or population of genome edited cells by NHEJ.

In various embodiments, the compositions contemplated herein comprise a donor repair template. The composition may be delivered to a cell that expresses or will express engineered nuclease or megaTAL, and optionally an end-processing enzyme. In one embodiment, the composition may be delivered to a cell that expresses or will express an

engineered homing endonuclease or megaTAL, and optionally a 3'-5' exonuclease.

Expression of the gene editing enzymes in the presence of the donor repair template can be used to generate a genome edited cell or population of genome edited cells by NHEJ or HDR.

In particular embodiments, the compositions contemplated herein comprise a  
5 population of cells, an engineered nuclease or a megaTAL, and optionally a 3'-5' exonuclease and further optionally, a donor repair template. In particular embodiments, the compositions contemplated herein comprise a population of cells, an engineered nuclease or a megaTAL, an end-processing enzyme, and optionally, a donor repair template. The engineered nuclease, megaTAL, and/or end-processing enzyme may be in the form of an mRNA that is introduced  
10 into the cell via polynucleotide delivery methods disclosed *supra*, e.g., lipid nanoparticles.

In particular embodiments, the compositions contemplated herein comprise a population of cells, an engineered homing endonuclease or megaTAL, and optionally, a donor repair template. In particular embodiments, the compositions contemplated herein comprise a population of cells, an engineered homing endonuclease or megaTAL, a 3'-5' exonuclease, and  
15 optionally, a donor repair template. The engineered homing endonuclease, megaTAL, and/or 3'-5' exonuclease may be in the form of an mRNA that is introduced into the cell via polynucleotide delivery methods disclosed *supra*, e.g., lipid nanoparticles or electroporation.

In particular embodiments, the population of cells comprise genetically modified hepatocytes.

20 Compositions include, but are not limited to pharmaceutical compositions. A "pharmaceutical composition" refers to a composition formulated in pharmaceutically-acceptable or physiologically-acceptable solutions for administration to a cell or an animal, either alone, or in combination with one or more other modalities of therapy. It will also be understood that, if desired, the compositions may be administered in combination with other  
25 agents as well, such as, e.g., cytokines, growth factors, hormones, small molecules, chemotherapeutics, pro-drugs, drugs, antibodies, or other various pharmaceutically-active agents. There is virtually no limit to other components that may also be included in the compositions, provided that the additional agents do not adversely affect the composition.

The phrase "pharmaceutically acceptable" is employed herein to refer to those  
30 compounds, materials, compositions, and/or dosage forms which are, within the scope of sound

medical judgment, suitable for use in contact with the tissues of human beings and animals without excessive toxicity, irritation, allergic response, or other problem or complication, commensurate with a reasonable benefit/risk ratio.

The term “pharmaceutically acceptable carrier” refers to a diluent, adjuvant, excipient, or vehicle with which the therapeutic cells are administered. Illustrative examples of pharmaceutical carriers can be sterile liquids, such as cell culture media, water and oils, including those of petroleum, animal, vegetable or synthetic origin, such as peanut oil, soybean oil, mineral oil, sesame oil and the like. Saline solutions and aqueous dextrose and glycerol solutions can also be employed as liquid carriers, particularly for injectable solutions. Suitable pharmaceutical excipients in particular embodiments, include starch, glucose, lactose, sucrose, gelatin, malt, rice, flour, chalk, silica gel, sodium stearate, glycerol monostearate, talc, sodium chloride, dried skim milk, glycerol, propylene, glycol, water, ethanol and the like. Except insofar as any conventional media or agent is incompatible with the active ingredient, its use in the therapeutic compositions is contemplated. Supplementary active ingredients can also be incorporated into the compositions.

In one embodiment, a composition comprising a pharmaceutically acceptable carrier is suitable for administration to a subject. In particular embodiments, a composition comprising a carrier is suitable for parenteral administration, *e.g.*, intravascular (intravenous or intraarterial), intraperitoneal or intramuscular administration. In particular embodiments, a composition comprising a pharmaceutically acceptable carrier is suitable for intraventricular, intraspinal, or intrathecal administration. Pharmaceutically acceptable carriers include sterile aqueous solutions, cell culture media, or dispersions. The use of such media and agents for pharmaceutically active substances is well known in the art. Except insofar as any conventional media or agent is incompatible with the transduced cells, use thereof in the pharmaceutical compositions is contemplated.

In particular embodiments, pharmaceutically acceptable carrier comprises lipid nanoparticles for encapsulation. In certain embodiments, polynucleotides, such as mRNA encoding a meganucleases/homing endonuclease, megaTAL, or polypeptide contemplated herein are encapsulated in lipid nanoparticles.



In particular embodiments, compositions contemplated herein comprise genetically modified T cells and a pharmaceutically acceptable carrier. A composition comprising a cell-based composition contemplated herein can be administered separately by enteral or parenteral administration methods or in combination with other suitable compounds to  
5 effect the desired treatment goals.

The pharmaceutically acceptable carrier must be of sufficiently high purity and of sufficiently low toxicity to render it suitable for administration to the human subject being treated. It further should maintain or increase the stability of the composition. The pharmaceutically acceptable carrier can be liquid or solid and is selected, with the planned  
10 manner of administration in mind, to provide for the desired bulk, consistency, *etc.*, when combined with other components of the composition. For example, the pharmaceutically acceptable carrier can be, without limitation, a binding agent (*e.g.*, pregelatinized maize starch, polyvinylpyrrolidone or hydroxypropyl methylcellulose, *etc.*), a filler (*e.g.*, lactose and other sugars, microcrystalline cellulose, pectin, gelatin, calcium sulfate, ethyl  
15 cellulose, polyacrylates, calcium hydrogen phosphate, *etc.*), a lubricant (*e.g.*, magnesium stearate, talc, silica, colloidal silicon dioxide, stearic acid, metallic stearates, hydrogenated vegetable oils, corn starch, polyethylene glycols, sodium benzoate, sodium acetate, *etc.*), a disintegrant (*e.g.*, starch, sodium starch glycolate, *etc.*), or a wetting agent (*e.g.*, sodium lauryl sulfate, *etc.*). Other suitable pharmaceutically acceptable carriers for the  
20 compositions contemplated herein include, but are not limited to, water, salt solutions, alcohols, polyethylene glycols, gelatins, amyloses, magnesium stearates, talcs, silicic acids, viscous paraffins, hydroxymethylcelluloses, polyvinylpyrrolidones and the like.

Such carrier solutions also can contain buffers, diluents and other suitable additives. The term “buffer” as used herein refers to a solution or liquid whose chemical  
25 makeup neutralizes acids or bases without a significant change in pH. Examples of buffers contemplated herein include, but are not limited to, Dulbecco’s phosphate buffered saline (PBS), Ringer’s solution, 5% dextrose in water (D5W), normal/physiologic saline (0.9% NaCl).

The pharmaceutically acceptable carriers may be present in amounts sufficient to  
30 maintain a pH of the composition of about 7. Alternatively, the composition has a pH in a

range from about 6.8 to about 7.4, *e.g.*, 6.8, 6.9, 7.0, 7.1, 7.2, 7.3, and 7.4. In still another embodiment, the composition has a pH of about 7.4.

Compositions contemplated herein may comprise a nontoxic pharmaceutically acceptable medium. The compositions may be a suspension. The term “suspension” as  
5 used herein refers to non-adherent conditions in which cells are not attached to a solid support. For example, cells maintained as a suspension may be stirred or agitated and are not adhered to a support, such as a culture dish.

In particular embodiments, compositions contemplated herein are formulated in a suspension, where the genome edited T cells are dispersed within an acceptable liquid  
10 medium or solution, *e.g.*, saline or serum-free medium, in an intravenous (IV) bag or the like. Acceptable diluents include, but are not limited to water, PlasmaLyte, Ringer’s solution, isotonic sodium chloride (saline) solution, serum-free cell culture medium, and medium suitable for cryogenic storage, *e.g.*, Cryostor® medium.

In certain embodiments, a pharmaceutically acceptable carrier is substantially free  
15 of natural proteins of human or animal origin, and suitable for storing a composition comprising a population of genome edited T cells. The therapeutic composition is intended to be administered into a human patient, and thus is substantially free of cell culture components such as bovine serum albumin, horse serum, and fetal bovine serum.

In some embodiments, compositions are formulated in a pharmaceutically  
20 acceptable cell culture medium. Such compositions are suitable for administration to human subjects. In particular embodiments, the pharmaceutically acceptable cell culture medium is a serum free medium.

Serum-free medium has several advantages over serum containing medium, including a simplified and better defined composition, a reduced degree of contaminants,  
25 elimination of a potential source of infectious agents, and lower cost. In various embodiments, the serum-free medium is animal-free, and may optionally be protein-free. Optionally, the medium may contain biopharmaceutically acceptable recombinant proteins. “Animal-free” medium refers to medium wherein the components are derived from non-animal sources. Recombinant proteins replace native animal proteins in animal-

free medium and the nutrients are obtained from synthetic, plant or microbial sources.

“Protein-free” medium, in contrast, is defined as substantially free of protein.

Illustrative examples of serum-free media used in particular compositions includes, but is not limited to QBSF-60 (Quality Biological, Inc.), StemPro-34 (Life Technologies),  
5 and X-VIVO 10.

In a preferred embodiment, the compositions comprising genome edited T cells are formulated in PlasmaLyte.

In various embodiments, compositions comprising genome edited T cells are formulated in a cryopreservation medium. For example, cryopreservation media with  
10 cryopreservation agents may be used to maintain a high cell viability outcome post-thaw. Illustrative examples of cryopreservation media used in particular compositions includes, but is not limited to, CryoStor CS10, CryoStor CS5, and CryoStor CS2.

In one embodiment, the compositions are formulated in a solution comprising 50:50 PlasmaLyte A to CryoStor CS10.

15 In particular embodiments, the composition is substantially free of mycoplasma, endotoxin, and microbial contamination. By “substantially free” with respect to endotoxin is meant that there is less endotoxin per dose of cells than is allowed by the FDA for a biologic, which is a total endotoxin of 5 EU/kg body weight per day, which for an average 70 kg person is 350 EU per total dose of cells. In particular embodiments, compositions  
20 comprising hematopoietic stem or progenitor cells transduced with a retroviral vector contemplated herein contain about 0.5 EU/mL to about 5.0 EU/mL, or about 0.5 EU/mL, 1.0 EU/mL, 1.5 EU/mL, 2.0 EU/mL, 2.5 EU/mL, 3.0 EU/mL, 3.5 EU/mL, 4.0 EU/mL, 4.5 EU/mL, or 5.0 EU/mL.

In certain embodiments, compositions and formulations suitable for the delivery of  
25 polynucleotides are contemplated including, but not limited to, one or more mRNAs encoding one or more reprogrammed nucleases or megaTALs, and optionally end-processing enzymes.

Exemplary formulations for *ex vivo* delivery may also include the use of various transfection agents known in the art, such as calcium phosphate, electroporation, heat  
30 shock and various liposome formulations (*i.e.*, lipid-mediated transfection). Liposomes,

as described in greater detail below, are lipid bilayers entrapping a fraction of aqueous fluid. DNA spontaneously associates to the external surface of cationic liposomes (by virtue of its charge) and these liposomes will interact with the cell membrane.

In particular embodiments, formulation of pharmaceutically-acceptable carrier solutions is well-known to those of skill in the art, as is the development of suitable dosing and treatment regimens for using the particular compositions described herein in a variety of treatment regimens, including *e.g.*, enteral and parenteral, *e.g.*, intravascular, intravenous, intrarterial, intraosseously, intraventricular, intracerebral, intracranial, intraspinal, intrathecal, and intramedullary administration and formulation. It would be understood by the skilled artisan that particular embodiments contemplated herein may comprise other formulations, such as those that are well known in the pharmaceutical art, and are described, for example, in *Remington: The Science and Practice of Pharmacy*, volume I and volume II. 22<sup>nd</sup> Edition. Edited by Loyd V. Allen Jr. Philadelphia, PA: Pharmaceutical Press; 2012, which is incorporated by reference herein, in its entirety.

#### 15 J. METHODS OF TREATMENT - GENE EDITING AND CELL THERAPIES

Provided herein are methods for editing a cell and/or preventing, treating, or ameliorating a disease or disorder, or ameliorating a disease condition or symptom associated therewith. Specifically, in one aspect, the engineered meganucleases/homing endonucleases, megaTALs, polypeptides, polynucleotides, compositions, cells, and associated methods of gene editing contemplated herein can be used in the prevention, treatment, and amelioration of a disease or disorder, or ameliorating a disease condition or symptom associated therewith.

In certain embodiments, the diseases or disorders include, but are not limited to, blood or bleeding disorders such as hemophilias and hemoglobinopathies, lysosomal storage disorders, metabolic disorders, or other mono or polygenic disorders or genetic diseases. See, *e.g.*, U.S. Patent Nos. 9,255,250; 9,877,988; 9,963,715; and 9,175,280.

In some embodiments, the diseases or disorders include, but are not limited to Hemophilia A, Hemophilia B, Hemophilia C, von Willebrand disease, galactosemia, Gaucher's disease, generalized gangliosidoses (GMI), Fabry disease, Tay-Sachs disease,

acid maltase deficiency, Pompe disease, Niemann-Pick disease, Hurler syndrome (MPS I), Hunter syndrome (MPS II), and urea cycle disorders.

In certain embodiments, the disease or disorder is a bleeding disorder. In various embodiments, the bleeding disorder is a hemophilia. In particular embodiments, the  
 5 hemophilia is Hemophilia A. In particular embodiments, the hemophilia is Hemophilia B. In particular embodiments, the hemophilia is Hemophilia C. In particular embodiments, the bleeding disorder is von Willebrand disease,

In one aspect, the disease or disorder is a lysosomal storage disorder or metabolic disorder. In particular embodiments, the disease or disorder is galactosemia. In particular  
 10 embodiments, the disease or disorder is Gaucher's disease. In particular embodiments, the disease or disorder is generalized gangliosidosis (GMI). In particular embodiments, the disease or disorder is Fabry disease. In particular embodiments, the disease or disorder is Tay-Sachs disease. In particular embodiments, the disease or disorder is acid maltase deficiency. In particular embodiments, the disease or disorder is Pompe disease. In  
 15 particular embodiments, the disease or disorder is Niemann-Pick disease. In particular embodiments, the disease or disorder is Hurler syndrome (MPS I). In particular embodiments, the disease or disorder is Hunter syndrome (MPS II). In particular embodiments, the disease or disorder is urea cycle disorder.

In still other embodiments, the diseases or disorders include, but are not limited to  
 20 achondroplasia, achromatopsia, adenosine deaminase deficiency, adrenoleukodystrophy, aicardi syndrome, alpha-I antitrypsin deficiency, alpha-thalassemia, androgen insensitivity syndrome, apert syndrome, arrhythmogenic right ventricular, dysplasia, ataxia telangiectasia, Barth syndrome, beta-thalassemia, blue rubber bleb nevus syndrome, Canavan disease, chronic granulomatous diseases (CGD), cri du chat syndrome, cystic fibrosis,  
 25 dercid's disease, ectodermal dysplasia, Fanconi anemia, fibrodysplasia ossificans progressiva, fragile X syndrome, hemochromatosis, the hemoglobin C mutation in the 6th codon of beta-globin (HbC), Huntington's disease, hypophosphatasia, Klinefelter syndrome, Krabbe's Disease, Langer-Giedion Syndrome, leukocyte adhesion deficiency, leukodystrophy, long QT syndrome, Marfan syndrome, Moebius syndrome, nail patella

- syndrome, nephrogenic diabetes insipidus, neurofibromatosis, osteogenesis imperfecta, porphyria, Prader-Willi syndrome, progeria, Proteus syndrome, retinoblastoma, Rett syndrome, Rubinstein-Taybi syndrome, Sanfilippo syndrome, severe combined immunodeficiency (SCID), Shwachman syndrome, sickle cell disease (sickle cell anemia),
- 5 Smith-Magenis syndrome, Stickler syndrome, Tay Sachs disease, Thrombocytopenia Absent Radius (TAR) syndrome, Treacher Collins syndrome, trisomy, tuberous sclerosis, von Hippel-Landau disease, Waardenburg syndrome, Williams syndrome, Wilson's disease, Wiskott-Aldrich syndrome, X-linked lymphoproliferative syndrome, and other acquired immunodeficiencies and hemoglobinopathies.
- 10 In still other embodiments, the diseases or disorders include cancers. In particular embodiments, the disease or disorder is a solid tumor or cancer including, but not limited to: adrenal cancer, adrenocortical carcinoma, anal cancer, appendix cancer, astrocytoma, atypical teratoid/rhabdoid tumor, basal cell carcinoma, bile duct cancer, bladder cancer, bone cancer, brain/CNS cancer, breast cancer, bronchial tumors, cardiac tumors, cervical
- 15 cancer, cholangiocarcinoma, chondrosarcoma, chordoma, colon cancer, colorectal cancer, craniopharyngioma, ductal carcinoma in situ (DCIS) endometrial cancer, ependymoma, esophageal cancer, esthesioneuroblastoma, Ewing's sarcoma, extracranial germ cell tumor, extragonadal germ cell tumor, eye cancer, fallopian tube cancer, fibrous histiosarcoma, fibrosarcoma, gallbladder cancer, gastric cancer, gastrointestinal carcinoid tumors,
- 20 gastrointestinal stromal tumor (GIST), germ cell tumors, glioma, glioblastoma, head and neck cancer, hemangioblastoma, hepatocellular cancer, hypopharyngeal cancer, intraocular melanoma, kaposi sarcoma, kidney cancer, laryngeal cancer, leiomyosarcoma, lip cancer, liposarcoma, liver cancer, lung cancer, non-small cell lung cancer, lung carcinoid tumor, malignant mesothelioma, medullary carcinoma, medulloblastoma, meningioma,
- 25 melanoma, Merkel cell carcinoma, midline tract carcinoma, mouth cancer, myxosarcoma, myelodysplastic syndrome, myeloproliferative neoplasms, nasal cavity and paranasal sinus cancer, nasopharyngeal cancer, neuroblastoma, oligodendroglioma, oral cancer, oral cavity cancer, oropharyngeal cancer, osteosarcoma, ovarian cancer, pancreatic cancer, pancreatic islet cell tumors, papillary carcinoma, paraganglioma, parathyroid cancer, penile cancer,
- 30 pharyngeal cancer, pheochromocytoma, pinealoma, pituitary tumor, pleuropulmonary

blastoma, primary peritoneal cancer, prostate cancer, rectal cancer, retinoblastoma, renal cell carcinoma, renal pelvis and ureter cancer, rhabdomyosarcoma, salivary gland cancer, sebaceous gland carcinoma, skin cancer, soft tissue sarcoma, squamous cell carcinoma, small cell lung cancer, small intestine cancer, stomach cancer, sweat gland carcinoma, synovioma, testicular cancer, throat cancer, thymus cancer, thyroid cancer, urethral cancer, uterine cancer, uterine sarcoma, vaginal cancer, vascular cancer, vulvar cancer, and Wilms Tumor.

In particular embodiments, genome edited cells contemplated herein are used in the treatment of solid tumors or cancers including, without limitation, liver cancer, pancreatic cancer, lung cancer, breast cancer, bladder cancer, brain cancer, bone cancer, thyroid cancer, kidney cancer, or skin cancer.

In particular embodiments, genome edited cells contemplated herein are used in the treatment of various cancers including but not limited to pancreatic, bladder, and lung.

In particular embodiments, genome edited cells contemplated herein are used in the treatment of liquid cancers or hematological cancers.

In particular embodiments, genome edited cells contemplated herein are used in the treatment of B-cell malignancies, including but not limited to: leukemias, lymphomas, and multiple myeloma.

In still other embodiments, the diseases or disorders include liquid cancers including, but not limited to: leukemias, lymphomas, and multiple myelomas, acute lymphocytic leukemia (ALL), acute myeloid leukemia (AML), myeloblastic, promyelocytic, myelomonocytic, monocytic, erythroleukemia, hairy cell leukemia (HCL), chronic lymphocytic leukemia (CLL), and chronic myeloid leukemia (CML), chronic myelomonocytic leukemia (CMML) and polycythemia vera, Hodgkin lymphoma, nodular lymphocyte-predominant Hodgkin lymphoma, Burkitt lymphoma, small lymphocytic lymphoma (SLL), diffuse large B-cell lymphoma, follicular lymphoma, immunoblastic large cell lymphoma, precursor B-lymphoblastic lymphoma, mantle cell lymphoma, marginal zone lymphoma, mycosis fungoides, anaplastic large cell lymphoma, Sézary syndrome, precursor T-lymphoblastic lymphoma, multiple myeloma, overt multiple myeloma, smoldering multiple

myeloma, plasma cell leukemia, non-secretory myeloma, IgD myeloma, osteosclerotic myeloma, solitary plasmacytoma of bone, and extramedullary plasmacytoma.

The edited cells useful for treatment of a disease or disorder described herein can be of any type, *e.g.*, those described above. In preferred embodiments, the cells are

5 hepatocytes.

In one aspect, methods of treating a patient suffering from a disease or disorder contemplated herein are provided and comprise administering to a patient an effective amount of any one of the engineered meganucleases/homing endonucleases, megaTALs, polypeptides, polynucleotides, compositions, or cells contemplated herein. In some

10 embodiments, the methods further comprise administering a therapeutic transgene contemplated herein.

In various embodiments, methods of treating a patient suffering from a disease or disorder contemplated herein are provided and comprise administering to a patient an effective amount of any one of the engineered meganucleases/homing endonucleases,

15 megaTALs, polypeptides, polynucleotides, compositions contemplated herein, and a vector comprising a therapeutic transgene. In some embodiments, the vector is a viral vector. In some embodiments, the viral vector is an LVV. In particular embodiments, the viral vector is an AVV (*e.g.*, rAVV).

In various embodiments, lipid nanoparticles are useful for encapsulating mRNA

20 encoding a meganucleases/homing endonuclease, megaTAL, or polypeptide contemplated herein. In some embodiments, the lipid nanoparticle encapsulated mRNA is administered to a subject or patient suffering from a disease or disorder contemplated herein. In particular embodiments, the lipid nanoparticle encapsulated mRNA and an AAV vector comprising a therapeutic transgene (*e.g.*, FVIII) is administered to a patient or subject

25 suffering from a disease or disorder contemplated herein.

The terms “lipid nanoparticle”, “LNP”, or “lipid particle” refer to a lipid formulation that can be used to deliver a nucleotide sequence (*e.g.*, mRNA encoding an engineered meganuclease/homing endonuclease or megaTAL and optionally a 3’-5’ exonuclease) to a target site of interest (such as hepatocytes).



In some embodiments, the lipid particle comprises a cationic lipid, a non-cationic lipid, and a mRNA sequence that is encapsulated by the lipid particle. Lipid particles of this type have been shown to efficiently deliver mRNA to the hepatocytes of the liver of rodents, primates and humans. The encapsulated mRNA undergoes a process of endosomal escape mediated by the ionizable nature of the cationic lipid. This delivers the mRNA into the cytoplasm where mRNA can be translated into the encoded protein. Lipid nanoparticles useful for the methods described herein are known in the art, see, e.g., WO2022/133344A1; Cullis and Hope, *Mol Ther.* 2017 Jul 5;25(7):1467-1475; Bottger *et al.*, *Adv Drug Deliv Rev.* 2020;154-155:79-101; and Hou *et al.*, *Nat Rev Mater.* 2021;6(12):1078-1094.

In another aspect, provided herein are compositions or kits or pharmaceutical compositions comprising: (a) an mRNA encoding a meganucleases/homing endonuclease, megaTAL, or polypeptide contemplated herein, and comprised in a lipid nanoparticle; and (b) an AAV vector comprising of a nucleotide sequence encoding a therapeutic transgene (e.g., FVIII) comprised in a AAV vector particle. These components (a) and (b), of the composition or pharmaceutical composition, can be administered together or separately to a cell or a patient. In various embodiments, the therapeutic transgene encodes a therapeutic polypeptide or factor useful for the treatment of hemophilia. In some embodiments, the hemophilia is Hemophilia A. In some embodiments, the hemophilia is Hemophilia B. In some embodiments, the factor is a FVIII or functional variant thereof (e.g., a modified FVIII). In some embodiments, the factor is a FIX or functional variant thereof (e.g., a modified FIX).

In various embodiments, the lipid nanoparticle encapsulated mRNA encoding an engineered meganucleases/homing endonuclease, megaTAL, or polypeptide contemplated herein are designed to edit the ALB gene (e.g., intron 1) of a hepatocyte.

In another aspect, cells that are edited by the compositions and methods contemplated herein and comprise an edited ALB gene provide improved drug products for use in the prevention, treatment, or amelioration of at least one symptom of a disease or disorder, including but not limited to a blood or bleeding disorder, lysosomal storage

disease, or metabolic disorder, cancer, GVHD, an infectious disease, an autoimmune disease, an inflammatory disease, or an immunodeficiency.

As used herein, the term “drug product” refers to genetically modified cells produced using the compositions and methods contemplated herein and refers to  
5 composition(s) or kit(s) used to produce genetically modified cells contemplated herein. In particular embodiments, the drug product is a composition or kit comprising (a) an mRNA encoding a meganucleases/homing endonuclease, megaTAL, or polypeptide contemplated herein, and comprised in a lipid nanoparticle; and (b) an AAV vector comprising of a nucleotide sequence encoding a therapeutic transgene (*e.g.*, FVIII or FIX) comprised in a  
10 AAV vector particle.

In particular embodiments, an effective amount of genome edited cells comprising an edited ALB gene are administered to a subject to prevent, treat, or ameliorate at least one symptom of a blood or bleeding disorder, lysosomal storage disease, metabolic disorder, cancer, GVHD, an infectious disease, an autoimmune disease, an inflammatory disease, or an  
15 immunodeficiency. In other embodiments, an effective amount of the composition(s) comprising (a) an mRNA encoding a meganucleases/homing endonuclease, megaTAL, or polypeptide contemplated herein, and comprised in a lipid nanoparticle; and (b) an AAV vector comprising of a nucleotide sequence encoding a therapeutic transgene (*e.g.*, FVIII or FIX) comprised in a AAV vector particle are administered to a subject to prevent, treat, or  
20 ameliorate at least one symptom of a blood or bleeding disorder.

In particular embodiments, a method of preventing, treating, or ameliorating at least one symptom of a disease or disorder comprises administering the subject an effective amount of genome edited cells comprising an edited ALB gene and a therapeutic transgene.

In particular embodiments, methods comprising administering a therapeutically  
25 effective amount of genome edited cells contemplated herein or a composition comprising the same, to a patient in need thereof, alone or in combination with one or more therapeutic agents, are provided.

In one embodiment, a method of treating a disease or disorder in a subject in need thereof comprises administering an effective amount, *e.g.*, therapeutically effective amount of a

composition comprising genome edited cells. The quantity and frequency of administration will be determined by such factors as the condition of the patient, and the type and severity of the patient's disease, although appropriate dosages may be determined by clinical trials.

In one illustrative embodiment, the effective amount of genome edited cells  
5 provided to a subject is at least  $2 \times 10^6$  cells/kg, at least  $3 \times 10^6$  cells/kg, at least  $4 \times 10^6$  cells/kg, at least  $5 \times 10^6$  cells/kg, at least  $6 \times 10^6$  cells/kg, at least  $7 \times 10^6$  cells/kg, at least  $8 \times 10^6$  cells/kg, at least  $9 \times 10^6$  cells/kg, or at least  $10 \times 10^6$  cells/kg, or more cells/kg, including all intervening doses of cells.

In another illustrative embodiment, the effective amount of genome edited cells  
10 provided to a subject is about  $2 \times 10^6$  cells/kg, about  $3 \times 10^6$  cells/kg, about  $4 \times 10^6$  cells/kg, about  $5 \times 10^6$  cells/kg, about  $6 \times 10^6$  cells/kg, about  $7 \times 10^6$  cells/kg, about  $8 \times 10^6$  cells/kg, about  $9 \times 10^6$  cells/kg, or about  $10 \times 10^6$  cells/kg, or more cells/kg, including all intervening doses of cells.

In another illustrative embodiment, the effective amount of genome edited cells  
15 provided to a subject is from about  $2 \times 10^6$  cells/kg to about  $10 \times 10^6$  cells/kg, about  $3 \times 10^6$  cells/kg to about  $10 \times 10^6$  cells/kg, about  $4 \times 10^6$  cells/kg to about  $10 \times 10^6$  cells/kg, about  $5 \times 10^6$  cells/kg to about  $10 \times 10^6$  cells/kg,  $2 \times 10^6$  cells/kg to about  $6 \times 10^6$  cells/kg,  $2 \times 10^6$  cells/kg to about  $7 \times 10^6$  cells/kg,  $2 \times 10^6$  cells/kg to about  $8 \times 10^6$  cells/kg,  $3 \times 10^6$  cells/kg to about  $6 \times 10^6$  cells/kg,  $3 \times 10^6$  cells/kg to about  $7 \times 10^6$  cells/kg,  $3 \times 10^6$  cells/kg to about  
20  $8 \times 10^6$  cells/kg,  $4 \times 10^6$  cells/kg to about  $6 \times 10^6$  cells/kg,  $4 \times 10^6$  cells/kg to about  $7 \times 10^6$  cells/kg,  $4 \times 10^6$  cells/kg to about  $8 \times 10^6$  cells/kg,  $5 \times 10^6$  cells/kg to about  $6 \times 10^6$  cells/kg,  $5 \times 10^6$  cells/kg to about  $7 \times 10^6$  cells/kg,  $5 \times 10^6$  cells/kg to about  $8 \times 10^6$  cells/kg, or  $6 \times 10^6$  cells/kg to about  $8 \times 10^6$  cells/kg, including all intervening doses of cells.

One of ordinary skill in the art would recognize that multiple administrations of the  
25 compositions contemplated in particular embodiments may be required to effect the desired therapy. For example, a composition may be administered 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 or more times over a span of 1 week, 2 weeks, 3 weeks, 1 month, 2 months, 3 months, 4 months, 5 months, 6 months, 1 year, 2 years, 5 years, 10 years, or more.

The administration of the compositions contemplated in particular embodiments may be carried out in any convenient manner, including by aerosol inhalation, injection, ingestion, transfusion, implantation or transplantation. In a preferred embodiment, compositions are administered parenterally. The phrases “parenteral administration” and “administered parenterally” as used herein refers to modes of administration other than enteral and topical administration, usually by injection, and includes, without limitation, intravascular, intravenous, intramuscular, intraarterial, intrathecal, intracapsular, intraorbital, intratumoral, intracardiac, intradermal, intraperitoneal, transtracheal, subcutaneous, subcuticular, intraarticular, subcapsular, subarachnoid, intraspinal and intrasternal injection and infusion. In one embodiment, the compositions contemplated herein are administered to a subject by direct injection into a tumor, lymph node, or site of infection.

K. SEQUENCE LISTING

TABLE 3

| SEQ ID NO: | Description                   | Sequence                                                                                                                                                                                                                                                                                                                             |
|------------|-------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1          | wild type I-OnuI              | MAYMSRRESINPWILTGFADEGSFLLRIRNNNKSSVGYSTELGFQITLHNK<br>DKSILENIQSTWKGVIANS GDNVSLKVTRFEDLKVIIDHFEKYPLITQKL<br>GDYMLFKQAFVCMENKEHLKINGIKELVRIKAKLNWGLTDELKKAFFPEIISK<br>ERSLINKNIPNFKWLAGFTSGEGCFFVNLIKSKSKLGVQVQLVFSITQHIKD<br>KNLMNSLITYLGCGYIKEKNKSEFSWLD FVVTKFSDINDKIIPVFQENTLIG<br>VKLEDFEDWCKVAKLIEEKKHLTESGLDEIKKIKLNMNKG RVF |
| 2          | Stabilized full length I-OnuI | MAYMSRRESINPWILTGFADEGSFLLRIRNNNKSSVGYSTELGFQITLHNK<br>DKSILENIQSTWKGVIANS GDNVSLKVTRFEDLKVIIDHFEKYPLITQKL<br>GDYMLFKQAFVCMENKGLKENGIKELVRIKAKMNWGLTDELKKAFFPDISR<br>ERPIINKNIPNGKWLAGFTSGDGCFFVNLIKSKSKLGVQVQLVFSITQHIRD<br>KNLMNSLITYLGCGYIKEKNKSEFSWLD FVVTKFSDIKDKIIPVFQENTLIG<br>MKHEDFEDWCKVAKLIEEKKHLTESGLDEIKKIKLNMNKG RVF   |
| 3          | I-OnuI polypeptide            | XXXMSRRESINPWILTGFADEGSFLLRIRNNNKSSVGYSTELGFQITLHNK<br>DKSILENIQSTWKGVIANS GDNVSLKVTRFEDLKVIIDHFEKYPLITQKL<br>GDYMLFKQAFVCMENKGLKENGIKELVRIKAKMNWGLTDELKKAFFPDISR<br>ERPIINKNIPNGKWLAGFTSGDGCFFVNLIKSKSKLGVQVQLVFSITQHIRD<br>KNLMNSLITYLGCGYIKEKNKSEFSWLD FVVTKFSDIKDKIIPVFQENTLIG<br>MKHEDFEDWCKVAKLIEEKKHLTESGLDEIKKIKLNMNKG RVF   |
| 4          | I-OnuI polypeptide            | SRRESINPWILTGFADEGSFLLRIRNNNKSSVGYSTELGFQITLHNKDKSI<br>LENIQSTWKGVIANS GDNVSLKVTRFEDLKVIIDHFEKYPLITQKLGDYM<br>LFKQAFVCMENKGLKENGIKELVRIKAKMNWGLTDELKKAFFPDISRERPI<br>INKNIPNGKWLAGFTSGDGCFFVNLIKSKSKLGVQVQLVFSITQHIRDKNLM<br>NSLITYLGCGYIKEKNKSEFSWLD FVVTKFSDIKDKIIPVFQENTLIGMKHE<br>DFEDWCKVAKLIEEKKHLTESGLDEIKKIKLNMNKG RVF       |
| 5          | I-OnuI                        | XXXXXXXXSINPWILTGFADEGSFLLRIRNNNKSSVGYSTELGFQITLHNK                                                                                                                                                                                                                                                                                  |

|    |                                 |                                                                                                                                                                                                                                                                                                                                  |
|----|---------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|    | polypeptide                     | DKSILENIQSTWKVGVIANSGDNAVSLKVTRFEDLKVIIDHFEKYPLITQKL<br>GDYMLFKQAFSVMENKGHLKENGIKELVRIKAKMNWGLTDELKKAFFEDISR<br>ERPIINKNIPNGKWLAGFTSGDGCFFVNLIKSSKSLGVQVQLVFSITQHIRD<br>KNLMNSLITYLGCGYIKEKNKSEFSWLDVVTKFSDIKDKIIPVFQENTLIG<br>MKHEDFEDWCKVAKLIEEKKHLTESGLDEIKKIKLNMNKGXXX                                                       |
| 6  | Engineered HE<br>variant 399B05 | SRRESINPWILTGFAAEGCFSLVIQKANGWTRYNTKLVFSIGLHNKDKSI<br>LENIQSTWKVGTISKSGDTQVRLTVTRFEDLKVIIDHFEKYPLITQKLGDY<br>MLFKQAFSVMENKGHLKENGIKELVRIKAKMNWGLTDELKKAFFEDISRERPI<br>INKNIPNGKWLAGFTSGDGSFMVILAKSRTGARVQVRLCFEIGQHIRDKNLM<br>NSLITYLGCGHISEKNLSESSFLIFRVTKFSDIKDKIIPVFQENTLIGMKHE<br>DFEDWCKVAKLIEEKKHLTESGLDEIKKIKLNMNKG RVF   |
| 7  | Engineered HE<br>variant 399A08 | SRRESINPWILTGFAAEGCFSLTIRKFNGRTRYTTKLMFSIGLHNKDKSI<br>LENIQSTWKVGTISKSGDTQVRLTVTRFEDLKVIIDHFEKYPLITQKLGDY<br>MLFKQAFSVMENKGHLKENGIKELVRIKAKMNWGLTDELKKAFFEDISRERPI<br>INKNIPNGKWLAGFTSGDGSFMVILAKSRTGARVQVRLCFEIGQHIRDKNLM<br>NSLITYLGCGHISEKNLSESSFLIFRVTKFSDIKDKIIPVFQENTLIGMKHE<br>DFEDWCKVAKLIEEKKHLTESGLDEIKKIKLNMNKG RVF   |
| 8  | Engineered HE<br>variant 399A10 | SRRESINPWILTGFAAEGCFSLAIGKLNGGTTRYSTKLVFSIGLHNKDKSI<br>LENIQSTWKVGTISKSGDTQVRLTVTRFEDLKVIIDHFEKYPLITQKLGDY<br>MLFKQAFSVMENKGHLKENGIKELVRIKAKMNWGLTDELKKAFFEDISRERPI<br>INKNIPNGKWLAGFTSGDGSFMVILAKSRTGARVQVRLCFEIGQHIRDKNLM<br>NSLITYLGCGHISEKNLSESSFLIFRVTKFSDIKDKIIPVFQENTLIGMKHE<br>DFEDWCKVAKLIEEKKHLTESGLDEIKKIKLNMNKG RVF  |
| 9  | Engineered HE<br>variant 399C12 | SRRESINPWILTGFAAEGCFSLHIAKFNKGTRYGTIKLVFSIGLHNKDKSI<br>LENIQSTWKVGTISKSGDTQVRLTVTRFEDLKVIIDHFEKYPLITQKLGDY<br>MLFKQAFSVMENKGHLKENGIKELVRIKAKMNWGLTDELKKAFFEDISRERPI<br>INKNIPNGKWLAGFTSGDGSFMVILAKSRTGARVQVRLCFEIGQHIRDKNLM<br>NSLITYLGCGHISEKNLSESSFLIFRVTKFSDIKDKIIPVFQENTLIGMKHE<br>DFEDWCKVAKLIEEKKHLTESGLDEIKKIKLNMNKG RVF  |
| 10 | Engineered HE<br>variant 399D05 | SRRESINPWILTGFAAEGCFSLVIAKNGGGTTRYTTKLVFSIGLHNKDKSI<br>LENIQSTWKVGTISKSGDTQVRLTVTRFEDLKVIIDHFEKYPLITQKLGDY<br>MLFKQAFSVMENKGHLKENGIKELVRIKAKMNWGLTDELKKAFFEDISRERPI<br>INKNIPNGKWLAGFTSGDGSFMVILAKSRTGARVQVRLCFEIGQHIRDKNLM<br>NSLITYLGCGHISEKNLSESSFLIFRVTKFSDIKDKIIPVFQENTLIGMKHE<br>DFEDWCKVAKLIEEKKHLTESGLDEIKKIKLNMNKG RVF  |
| 11 | Engineered HE<br>variant 399D09 | SRRESINPWILTGFAAEGCFSLAIGKCNCGTTRYSTKLVFSIGLHNKDKSI<br>LENIQSTWKVGTISKSGDTQVRLTVTRFEDLKVIIDHFEKYPLITQKLGDY<br>MLFKQAFSVMENKGHLKENGIKELVRIKAKMNWGLTDELKKAFFEDISRERPI<br>INKNIPNGKWLAGFTSGDGSFMVILAKSRTGARVQVRLCFEIGQHIRDKNLM<br>NSLITYLGCGHISEKNLSESSFLIFRVTKFSDIKDKIIPVFQENTLIGMKHE<br>DFEDWCKVAKLIEEKKHLTESGLDEIKKIKLNMNKG RVF  |
| 12 | Engineered HE<br>variant 399D10 | SRRESINPWILTGFAAEGCFSLAIGKLNATTRYVTKLVFSIGLHNKDKSI<br>LENIQSTWKVGTISKSGDTQVRLTVTRFEDLKVIIDHFEKYPLITQKLGDY<br>MLFKQAFSVMENKGHLKENGIKELVRIKAKMNWGLTDELKKAFFEDISRERPI<br>INKNIPNGKWLAGFTSGDGSFMVILAKSRTGARVQVRLCFEIGQHIRDKNLM<br>NSLITYLGCGHISEKNLSESSFLIFRVTKFSDIKDKIIPVFQENTLIGMKHE<br>DFEDWCKVAKLIEEKKHLTESGLDEIKKIKLNMNKG RVF   |
| 13 | Engineered HE<br>variant 399H10 | SRRESINPWILTGFAAEGCFSLHIGKMNGRTRYLTCLKLVFSIGLHNKDKSI<br>LENIQSTWKVGTISKSGDTQVRLTVTRFEDLKVIIDHFEKYPLITQKLGDY<br>MLFKQAFSVMENKGHLKENGIKELVRIKAKMNWGLTDELKKAFFEDISRERPI<br>INKNIPNGKWLAGFTSGDGSFMVILAKSRTGARVQVRLCFEIGQHIRDKNLM<br>NSLITYLGCGHISEKNLSESSFLIFRVTKFSDIKDKIIPVFQENTLIGMKHE<br>DFEDWCKVAKLIEEKKHLTESGLDEIKKIKLNMNKG RVF |

|    |                |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
|----|----------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 14 | megaTAL 399B05 | VDLRTLGYSSQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPAALG<br>TVAVTYQHIITALPEATHEDIVGVGKQWSGARALEALLTDAGELRGPPQLD<br>TGQLVKIAKRGGVTAMEAVHASRNALTGAPLNLTDPQVVAIASNNGGKQALE<br>TVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPD<br>QVVAIASNIGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNIGGKQALETV<br>QRLLPVLCQDHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQDHGLTPDQV<br>VAIASNNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNIGGKQALETVQ<br>LLPVLCQDHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQDHGLTPDQVVA<br>IASNIGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRL<br>PVLCQDHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQDHGLTPDQVVAIA<br>SNGGKQALESIQAQLSRPDALAALTNDHLVALACLGGRPAMDVAVKGLPH<br>APELIRRVNRRIGERTSHRVAISRGGSSRRESINPWILTFADAEGCFSLV<br>IQKANGWTRYNTKLVSIGLHNKDKSILENIQSTWKVGTISKSGDTQVRLT<br>VTRFEDLKVIDHFEKYPLITQKLG DYMLFKQAFSVMENKGLKENGKIKELV<br>RIKAKMNWGLTDELKKAFFPDISRERPIINKNIPNGKWLAFGTSFGDSFMVI<br>LAKSRTGARVQVRLCFEIGQHIRDKNLMNSLITYLGCCHISEKNLSESSFLI<br>FRVTKFSDIKDKIIPVFQENTLIGMKHEDFEDWCKVAKLIEKKHLTESGLD<br>EIKKIKLNMNKG RVF  |
| 15 | megaTAL 399A08 | VDLRTLGYSSQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPAALG<br>TVAVTYQHIITALPEATHEDIVGVGKQWSGARALEALLTDAGELRGPPQLD<br>TGQLVKIAKRGGVTAMEAVHASRNALTGAPLNLTDPQVVAIASNNGGKQALE<br>TVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPD<br>QVVAIASNIGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNIGGKQALETV<br>QRLLPVLCQDHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQDHGLTPDQV<br>VAIASNNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNIGGKQALETVQ<br>LLPVLCQDHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQDHGLTPDQVVA<br>IASNIGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRL<br>PVLCQDHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQDHGLTPDQVVAIA<br>SNGGKQALESIQAQLSRPDALAALTNDHLVALACLGGRPAMDVAVKGLPH<br>APELIRRVNRRIGERTSHRVAISRGGSSRRESINPWILTFADAEGCFSLT<br>IRKFNGRTRYTTKLMFSIGLHNKDKSILENIQSTWKVGTISKSGDTQVRLT<br>VTRFEDLKVIDHFEKYPLITQKLG DYMLFKQAFSVMENKGLKENGKIKELV<br>RIKAKMNWGLTDELKKAFFPDISRERPIINKNIPNGKWLAFGTSFGDSFMVI<br>LAKSRTGARVQVRLCFEIGQHIRDKNLMNSLITYLGCCHISEKNLSESSFLI<br>FRVTKFSDIKDKIIPVFQENTLIGMKHEDFEDWCKVAKLIEKKHLTESGLD<br>EIKKIKLNMNKG RVF |
| 16 | megaTAL 399A10 | VDLRTLGYSSQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPAALG<br>TVAVTYQHIITALPEATHEDIVGVGKQWSGARALEALLTDAGELRGPPQLD<br>TGQLVKIAKRGGVTAMEAVHASRNALTGAPLNLTDPQVVAIASNNGGKQALE<br>TVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPD<br>QVVAIASNIGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNIGGKQALETV<br>QRLLPVLCQDHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQDHGLTPDQV<br>VAIASNNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNIGGKQALETVQ<br>LLPVLCQDHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQDHGLTPDQVVA<br>IASNIGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRL<br>PVLCQDHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQDHGLTPDQVVAIA<br>SNGGKQALESIQAQLSRPDALAALTNDHLVALACLGGRPAMDVAVKGLPH<br>APELIRRVNRRIGERTSHRVAISRGGSSRRESINPWILTFADAEGCFSLA<br>IGKLNGGTTRYSTKLVSIGLHNKDKSILENIQSTWKVGTISKSGDTQVRLT<br>VTRFEDLKVIDHFEKYPLITQKLG DYMLFKQAFSVMENKGLKENGKIKELV<br>RIKAKMNWGLTDELKKAFFPDISRERPIINKNIPNGKWLAFGTSFGDSFMVI<br>LAKSRTGARVQVRLCFEIGQHIRDKNLMNSLITYLGCCHISEKNLSESSFLI<br>FRVTKFSDIKDKIIPVFQENTLIGMKHEDFEDWCKVAKLIEKKHLTESGLD<br>EIKKIKLNMNKG RVF |
| 17 | megaTAL 399C12 | VDLRTLGYSSQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPAALG<br>TVAVTYQHIITALPEATHEDIVGVGKQWSGARALEALLTDAGELRGPPQLD<br>TGQLVKIAKRGGVTAMEAVHASRNALTGAPLNLTDPQVVAIASNNGGKQALE<br>TVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPD<br>QVVAIASNIGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNIGGKQALETV                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |

|    |                |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
|----|----------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|    |                | QRLLPVLCQDHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQDHGLTPDQV<br>VAIASNNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNIGGKQALETVQ<br>LLPVLCQDHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQDHGLTPDQVVA<br>IASNIGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRL<br>PVLCQDHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQDHGLTPDQVVAIA<br>SNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNIGGKQALETVQRLLP<br>APELIRRVNRRIGERTSHRVAISRVGSSRRESINPWILTGAFADAGCFS<br>IAKFNKGTRYGTCLVFSIGLHNKDKSILENIQSTWKVGTISKSGDTQVRLT<br>VTRFEDLKVIIDHFEKYPLITQKLGDMYLFKQAFSVMENKGHLKENGIKELV<br>RIKAKMNWGLTDELKKAFFPDISRERPIINKNIPNGKWLAFGFTSGDGSFMVI<br>LAKSRTGARVQVRLCFEIGQHIRDKNLMNSLITYLGCCHISEKNLSESSFLI<br>FRVTKFSDIKDKIIPVFQENTLIGMKHEDFEDWCKVAKLIEKKHLTESGLD<br>EIKKIKLNMNKG RVF                                                                                                                                                                                                                                                                                        |
| 18 | megaTAL 399D05 | VDLRTLGYSSQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPAALG<br>TVAVTYQHIITALPEATHEDIVGVGKQWSGARALEALLTDAGELRGPPQLD<br>TGQLVKIAKRGGVTAMEAVHASRNALTGAPLNLTDPQVVAIASNNGGKQALE<br>TVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPD<br>QVVAIASNIGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNIGGKQALETV<br>QRLLPVLCQDHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQDHGLTPDQV<br>VAIASNNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNIGGKQALETVQ<br>LLPVLCQDHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQDHGLTPDQVVA<br>IASNIGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRL<br>PVLCQDHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQDHGLTPDQVVAIA<br>SNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNIGGKQALETVQRLLP<br>APELIRRVNRRIGERTSHRVAISRVGSSRRESINPWILTGAFADAGCFS<br>IAKFNKGTRYGTCLVFSIGLHNKDKSILENIQSTWKVGTISKSGDTQVRLT<br>VTRFEDLKVIIDHFEKYPLITQKLGDMYLFKQAFSVMENKGHLKENGIKELV<br>RIKAKMNWGLTDELKKAFFPDISRERPIINKNIPNGKWLAFGFTSGDGSFMVI<br>LAKSRTGARVQVRLCFEIGQHIRDKNLMNSLITYLGCCHISEKNLSESSFLI<br>FRVTKFSDIKDKIIPVFQENTLIGMKHEDFEDWCKVAKLIEKKHLTESGLD<br>EIKKIKLNMNKG RVF |
| 19 | megaTAL 399D09 | VDLRTLGYSSQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPAALG<br>TVAVTYQHIITALPEATHEDIVGVGKQWSGARALEALLTDAGELRGPPQLD<br>TGQLVKIAKRGGVTAMEAVHASRNALTGAPLNLTDPQVVAIASNNGGKQALE<br>TVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPD<br>QVVAIASNIGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNIGGKQALETV<br>QRLLPVLCQDHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQDHGLTPDQV<br>VAIASNNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNIGGKQALETVQ<br>LLPVLCQDHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQDHGLTPDQVVA<br>IASNIGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRL<br>PVLCQDHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQDHGLTPDQVVAIA<br>SNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNIGGKQALETVQRLLP<br>APELIRRVNRRIGERTSHRVAISRVGSSRRESINPWILTGAFADAGCFS<br>IAKFNKGTRYGTCLVFSIGLHNKDKSILENIQSTWKVGTISKSGDTQVRLT<br>VTRFEDLKVIIDHFEKYPLITQKLGDMYLFKQAFSVMENKGHLKENGIKELV<br>RIKAKMNWGLTDELKKAFFPDISRERPIINKNIPNGKWLAFGFTSGDGSFMVI<br>LAKSRTGARVQVRLCFEIGQHIRDKNLMNSLITYLGCCHISEKNLSESSFLI<br>FRVTKFSDIKDKIIPVFQENTLIGMKHEDFEDWCKVAKLIEKKHLTESGLD<br>EIKKIKLNMNKG RVF |
| 20 | megaTAL 399D10 | VDLRTLGYSSQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPAALG<br>TVAVTYQHIITALPEATHEDIVGVGKQWSGARALEALLTDAGELRGPPQLD<br>TGQLVKIAKRGGVTAMEAVHASRNALTGAPLNLTDPQVVAIASNNGGKQALE<br>TVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPD<br>QVVAIASNIGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNIGGKQALETV<br>QRLLPVLCQDHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQDHGLTPDQV<br>VAIASNNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNIGGKQALETVQ<br>LLPVLCQDHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQDHGLTPDQVVA<br>IASNIGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRL<br>PVLCQDHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQDHGLTPDQVVAIA                                                                                                                                                                                                                                                                                                                                                                                                                       |

|    |                        |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
|----|------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|    |                        | SNGGGKQALESIVAQLSRPDPALAALTNDHLVALACLGRPAMDAVKKGLPH<br>APELIRRVNRRIGERTSHRVAISRVGSSRRRESINPWILTGFADAEFCFSLA<br>IGKLNATTRYVTKLVSIGLHNKDKSILENIQSTWKVGTISKSGDTQVRLT<br>VTRFEDLKVIIDHFEKYPLITQKLG DYMLFKQAFSVMENKGHLKENGIKELV<br>RIKAKMNWGLTDELKKAFFPDISRERPIINKNIPNGKWLAGFTSGDGSFMVI<br>LAKSRTGARVQVRLCFEIGQHIRDKNLMNSLITYLGCGHISEKNLSESSFLI<br>FRVTKFSDIKDKIIPVFQENTLIGMKHEDFEDWCKVAKLIEEKKHLTESGLD<br>EIKKIKLNMNKG RVF                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| 21 | megaTAL 399H10         | VDLRTLGYSSQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPAALG<br>TVAVTYQHIIITALPEATHEDIVGVGQWSGARALEALLTDAGELRGPPLQLD<br>TGQLVKIAKRGGVTAMEAVHASRNALTGAPLNLTDPQVVAIASNNGGKQALE<br>TVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPD<br>QVVAIASNIGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNIGGKQALETV<br>QRLLPVLCQDHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQDHGLTPDQV<br>VAIASNNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNIGGKQALETVQR<br>LLPVLCQDHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQDHGLTPDQVVA<br>IASNIGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLL<br>PVLCQDHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQDHGLTPDQVVAIA<br>SNGGGKQALESIVAQLSRPDPALAALTNDHLVALACLGRPAMDAVKKGLPH<br>APELIRRVNRRIGERTSHRVAISRVGSSRRRESINPWILTGFADAEFCFSLH<br>IGKMNGRTRYLTKLVSIGLHNKDKSILENIQSTWKVGTISKSGDTQVRLT<br>VTRFEDLKVIIDHFEKYPLITQKLG DYMLFKQAFSVMENKGHLKENGIKELV<br>RIKAKMNWGLTDELKKAFFPDISRERPIINKNIPNGKWLAGFTSGDGSFMVI<br>LAKSRTGARVQVRLCFEIGQHIRDKNLMNSLITYLGCGHISEKNLSESSFLI<br>FRVTKFSDIKDKIIPVFQENTLIGMKHEDFEDWCKVAKLIEEKKHLTESGLD<br>EIKKIKLNMNKG RVF                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| 22 | megaTAL 399B05<br>mRNA | GTGGACCTGAGAACCCTGGGCTACAGCCAGCAGCAGCAGGAGAAGATCAAAC<br>CCAAGGTGAGAAGCACCCGTGGCTCAGCACACGAGGCCCTGGTGGGCCACGG<br>CTTCACCCACGCCCACATCGTCGCCCTGAGCCAGCATCCCGCTGCCCTGGGA<br>ACCGTGGCCGTGACCTACCAGCACATCATCACAGCCCTGCCCGAAGCCACCC<br>ACGAGGACATCGTCGGAGTGGGCAAGCAGTGGAGCGGAGCCAGAGCCCTGGA<br>GGCCCTGCTGACCGACGCCGAGAGCTGAGAGGCCCTCCTCTGCAGCTGGAC<br>ACCGGACAGCTGGTGAAGATCGCCAAGAGAGCGGAGTGACCGCCATGGAGG<br>CCGTGCACGCCAGCAGAAACGCCCTGACCGGAGCTCCTCTGAACCTTACTCC<br>AGACCAAGTTGTCCGCATCGCCAGTAACAACGGAGGCAAGCAGGCTCTCGAA<br>ACCGTGCAGAGGTTGCTGCCTGTCTCTGCCAAGATCATGGGCTCACACCTG<br>ATCAGGTTGTGGCAATCGCCTCCAACAATGGCGGAAAGCAGGCCCTTGAAAC<br>CGTCCAACGGCTCCTCCCGTCTGTGTGTCAGGACCAGGCTCTTACACCTGAC<br>CAGGTTGTGGCCATTGCATCAAATATCGGGGGGAAGCAAGCTCTAGAAACCG<br>TGCAACGGTTGCTTCCAGTGCTTGTCAAGACCATGGACTGACTCCAGACCA<br>GGTTGTTGCCATCGCATCTAATATTGGAGGCAAACAGGCTCTAGAGACAGTA<br>CAACGCCTGCTCCCCGTGCTCTGCCAGGACCATGGGTTGACTCCAGATCAGG<br>TGTTGCTATCGCTAGTAATATCGGCGGAAAACAAGCCCTGGAGACGGTGCA<br>GCGCTTGCTCCCAAGTTCTGTGCCAGGATCATGGGCTGACACCCGATCAAGTG<br>GTTGCAATCGCCAGCAACAACGGTGGTAACAGGCTCTTGAGACTGTTCAAC<br>GGCTGTTGCCCGTGCTGTGCCAGGACCACGGACTTACCCAGATCAAGTTGT<br>GGCTATCGCTTCCAATATTGGTGCAAGCAAGCCCTCGAAACAGTGACAGCG<br>CTTCTACCCGTCTTTGTGTCAGGATCATGGACTCACACCAGACAGGTGGTGC<br>CTATTGCTAGCAACATCGGAGGGAAGCAGGCGTTGAAACTGTGCAAAGGCT<br>CTTGCCCGTGTTGTGTCAAGACCATGGTCTGACTCCCGATCAAGTCGTGGCT<br>ATTGCTTCAAACATTGGTGGAAGAACAGGCCCTTGGAACCGCTCCAAGACTGC<br>TTCCCGTGCTTTGCCAAGACCACGGCTTGACCCAGATCAGGTCGTGGCAAT<br>TGCAATCCAATGGAGGCGGTAAACAAGCTCTGGAGACAGTACAGAGACTCCTG<br>CCCGTCTCTGTCAAGATCATGGATTGACCCCTGACCAGGTCTGTCGAATTG<br>CCTCTAACATCGGCGGGAACAAGCTTTGGAGACTGTGCAGAGACTTCTCCC<br>AGTGCTATGTCAAGATCACGGCCTTACCCCGATCAGGTGGTGCGCATTGCA<br>AGCAATGGCGGTGGAAGCAAGCTCTCGAAAGCATCGTGGCCAGCTGAGCA<br>GACCCGACCTGCCCTGGCCGCTCTGACCAACGACCACTCGTGGCCCTGGC<br>CTGCCTGGGTGGCAGACCAGCCATGGACGCGTGAAGAAGGGCCTGCCTCAC<br>GCCCTGAGCTGATCAGAAGAGTGAACAGGAGAATCGGCGAGAGAACCAGCC |



|    |                        |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
|----|------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|    |                        | ACAGAGTGGCCATCTCTAGAGTGGGAGGAAGCTCCAGAAGAGAGAGCATCAA<br>TCCCTGGATCCTGACCGGCTTCGCCGACGCCGAGGGCTGCTTCAGCCTGGTG<br>ATCCAGAAGGCCAACGGCTGGACCACCAGATACAACACCAAGCTGGTGTTC<br>GCATCGGCCTGCACAACAAGGACAAGAGCATCCTGGAGAACATCCAGAGCAC<br>CTGGAAGGTGGGAACCATCAGCAAGTCTGGCGATACCCAAGTGAGGCTGACC<br>GTGACAAGATTTCGAGGACCTGAAGGTGATCATCGACCACTTCGAGAAGTACC<br>CTCTGATCACCAGAAAGCTGGGCGACTACATGCTGTTCAAGCAGGCCTTCTC<br>CGTGATGGAGAACAAGGGCCACCTGAAGGAGAACGGCATCAAGGAGCTGGTG<br>AGAATCAAGGCCAAGATGAACCTGGGGTCTGACCGACGAGCTGAAGAAGGCCT<br>TTCCCGAGGACATCAGCAGAGAGAGGCCTATCATCAACAAGAACATTCCCAA<br>CGGCAAGTGGCTGGCCGGATTACACAGCGGCACGGCAGCTTCATGGTGATC<br>CTGGCCAAGAGCAGAACCAGGAGCTAGAGTGCAGGTGAGGCTGTGCTTCGAGA<br>TCGGCCAGCACATTAGAGACAAGAACCCTGATGAACAGCTGATCACCTACCT<br>GGGCTGCGGACACATCAGCGAGAAGAACCCTCCGAATCCAGCTTCCTGATC<br>TTCAGAGTGACCAAGTTCTCCGACATCAAGGACAAAATCATTCCCGTGTTC<br>AGGAGAACACCCTGATCGGCATGAAGCACGAGGACTTCGAGGACTGGTGCAA<br>GGTGGCCAAGCTGATCGAAGAGAAGAAGCACCTGACCGAGAGCGGCCTGGAC<br>GAGATCAAGAAGATCAAGCTGAACATGAACAGGGCAGAGTGTTTTTGATAA                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| 23 | megaTAL 399A08<br>mRNA | GTGGACCTGAGAACCCTGGGCTACAGCCAGCAGCAGCAGGAGAAGATCAAAC<br>CCAAGGTGAGAAGCACCGTGGCTCAGCACCACGAGGCCCTGGTGGGCCACGG<br>CTTCACCCACGCCCACATCGTCGCCCTGAGCCAGCATCCCGCTGCCCTGGGA<br>ACCGTGGCCGTGACCTACCAGCACATCATCACAGCCCTGCCCGAAGCCACCC<br>ACGAGGACATCGTCGGAGTGGGCAAGCAGTGGAGCGGAGCCAGAGCCCTGGA<br>GGCCCTGCTGACCGACGCCGGAGAGCTGAGAGGCCCTCCTCTGCAGCTGGAC<br>ACCGGACAGCTGGTGAAGATCGCCAAGAGAGGCGGAGTGACCGCCATGGAGG<br>CCGTGCACGCCAGCAGAAACGCCCTGACCGGAGCTCCTCTGAACCTTACTCC<br>AGACCAAGTTGTGCGCATCGCCAGTAACAACGGAGGCCAAGCAGGCTCTCGAA<br>ACCGTGACAGAGTTGCTGCCTGTCTCTGCCAAGATCATGGGCTCACACCTG<br>ATCAGGTTGTGGCAATCGCCTCCAACAATGGCGGAAAGCAGGCCCTTGAAAC<br>CGTCCAACGGCTCCTTCCCGTCTGTGTGTCAGGACCACGGTCTTACACCTGAC<br>CAGGTTGTGGCCATTGCATCAAAATATCGGGGGGAAGCAAGCTCTAGAAACCG<br>TGCAACGGTTGCTTCCAGTGCTTTGTCAAGACCATGGACTCTCAGACAGTA<br>GGTTGTTGCCATCGCATCTAATATTGGAGGCCAAACAGGCTCTAGAGACAGTA<br>CAACGCCCTGCTCCCCGTGCTCTGCCAGGACCATGGGTTGACTCCAGATCAGG<br>TGGTTGCTATCGCTAGTAATATCGGCGGAAACAAGCCCTGGAGACGGTGCA<br>GGCTTGTGCTTCCAGTTCTGTGCCAGGATCATGGGCTGACACCCGATCAAGTG<br>GTTGCAATCGCCAGCAACAACGGTGGTAACAGGCTCTTGAGACTGTTCAAC<br>GGCTGTTGCCCGTGTGTGCCAGGACCACGGACTTACCCAGATCAAGTTGT<br>GGCTATCGCTTCCAATATTGGTGGCAAGCAAGCCCTCGAAACAGTGCAGCGG<br>CTTCTACCCGTCTTTGTGAGGATCATGGACTCACACCAGACCAGGTGGTTCG<br>CTATTGCTAGCAACATCGGAGGGAAGCAGGCGTTGGAACCTGTGCAAAAGGCT<br>CTTGCCCGTGTGTGTCAAGACCATGGTCTGACTCCCGATCAAGTCGTGGCT<br>ATTGCTTCAAACATTGGTGGAAACAGGCCCTTGGAACCCGTCCAAGACTGC<br>TTCCCGTGTCTTGCCAAGACCACGGCTTGACCCAGATCAGGTCGTGGCAAT<br>TGCATCCAATGGAGGCGGTAAACAAGCTCTGGAGACAGTACAGAGACTCCTG<br>CCCGTCTCTGTCAAGATCATGGATTGACCCCTGACCAGGTCTGTCGAATTG<br>CCTCTAACATCGGCGGAAACAAGCTTTGGAGACTGTGCAGAGACTTCTCCC<br>AGTGCTATGTCAAGATCACGGCCTTACCCCGATCAGGTGGTCGCCATTGCA<br>AGCAATGGCGGTGGAAGCAAGCTCTCGAAAGCATCGTGGCCCAGCTGAGCA<br>GACCCGACCCTGCCCTGGCCGCTCTGACCAACGACCACCTCGTGGCCCTGGC<br>CTGCCTGGGTGGCAGACCAGCCATGGACGCGCTGAAGAAGGGCCTGCCTCAC<br>GGCCCTGAGCTGATCAGAAGAGTGAACAGGAGAATCGGCGAGAGAACAGCC<br>ACAGAGTGGCCATCTCTAGAGTGGGAGGAAGCTCCAGAAGAGAGAGCATCAA<br>TCCCTGGATCCTGACCGGCTTCGCCGACGCCGAGGGCTGCTTCAGCCTGACG<br>ATCCGGAAGTTCAACGGCAGGACCACAGATACACCACCAAGCTGATGTTCA<br>GCATCGGCCTGCACAACAAGGACAAGAGCATCCTGGAGAACATCCAGAGCAC<br>CTGGAAGGTGGGAACCATCAGCAAGTCTGGCGATACCAAGTGAGGCTGACC<br>GTGACAAGATTTCGAGGACCTGAAGGTGATCATCGACCACTTCGAGAAGTACC<br>CTCTGATCACCAGAAAGCTGGGCGACTACATGCTGTTCAAGCAGGCCTTCTC<br>CGTGATGGAGAACAAGGGCCACCTGAAGGAGAACGGCATCAAGGAGCTGGTG<br>AGAATCAAGGCCAAGATGAACCTGGGGTCTGACCGACGAGCTGAAGAAGGCCT |

|    |                        |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
|----|------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|    |                        | <p>TTCCCGAGGACATCAGCAGAGAGAGGCCTATCATCAACAAGAACATTTCCCAA<br/> CGGCAAGTGGCTGGCCGGATTACCCAGCGGCGACGGCAGCTTCATGGTGATC<br/> CTGGCCAAGAGCAGAACCGGAGCTAGAGTGCAGGTGAGGCTGTGCTTCGAGA<br/> TCGGCCAGCACATTAGAGACAAGAACCTGATGAACAGCCTGATCACCTACCT<br/> GGGCTGCGGACACATCAGCGAGAAGAACCTCTCCGAATCCAGCTTCCTGATC<br/> TTCAGAGTGACCAAGTTCTCCGACATCAAGGACAAAATCATTCCCCTGTTCC<br/> AGGAGAACCCTGATCGGCATGAAGCACGAGGACTTCGAGGACTGGTGCAA<br/> GGTGGCCAAGCTGATCGAAGAGAAGAAGCACCTGACCAGAGCGGCCTGGAC<br/> GAGATCAAGAAGATCAAGCTGAACATGAACAAGGGCAGAGTGTTTTGATAA</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
| 24 | megaTAL 399A10<br>mRNA | <p>GTGGACCTGAGAACCCTGGGCTACAGCCAGCAGCAGCAGGAGAAGATCAAAC<br/> CCAAGGTGAGAAGCACCGTGGCTCAGCACCCAGAGGCCCTGGTGGGCCACGG<br/> CTTCACCCACGCCACATCGTCGCCCTGAGCCAGCATCCCGCTGCCCTGGGA<br/> ACCGTGGCCGTGACCTACCAGCACATCATCAGCCCTGCCCGAAGCCACCC<br/> ACGAGGACATCGTCGGAGTGGGCAAGCAGTGGAGCGGAGCCAGAGCCCTGGA<br/> GGCCCTGCTGACCGACCGCGAGAGCTGAGAGGCCCTCCTCTGCAGCTGGAC<br/> ACCGGACAGCTGGTGAAGATCGCCAAGAGAGGCGGAGTGACCGCCATGGAGG<br/> CCGTGCACGCCAGCAGAAACGCCCTGACCGGAGCTCCTCTGAACCTTACTCC<br/> AGACCAAGTTGTCGCCATCGCCAGTAACAACGGAGGCAAGCAGGCTCTCGAA<br/> ACCGTGCAGAGGTTGCTGCCTGTCTCTGCCAAGATCATGGGCTCACACCTG<br/> ATCAGGTTGTGGCAATCGCCTCCAACAATGGCGGAAAGCAGGCCCTTGAAAC<br/> CGTCCAACGGCTCCTTCCCGTCTGTGTGTCAGGACCAGGTCTTACACCTGAC<br/> CAGGTTGTGGCCATTGCATCAAATATCGGGGGGAAGCAAGCTCTAGAAACCG<br/> TGCAACGGTTGCTTCCAGTGCTTTGTCAAGACCATGGACTGACTCCAGACCA<br/> GGTTGTGTCATCGCATCTAATATTGGAGGCAACAGGCTCTAGAGACAGTA<br/> CAACGCCCTGCTCCCCGTGCTCTGCCAGGACCATGGTTGACTCCAGATCAGG<br/> TGGTTGCTATCGCTAGTAATATCGCGGAAAACAAGCCCTGGAGACGGTGCA<br/> GGCTTGTGCTCCAGTTCTGTGCCAGGATCATGGGCTGACACCCGATCAAGTG<br/> GTTGCAATCGCCAGCAACAACGGTGGTAAACAGGCTCTTGAGACTGTTCAAC<br/> GGCTGTTGCCCGTGCTGTGCCAGGACCACGGACTTACCCAGATCAAGTTGT<br/> GGCTATCGCTTCCAATATTGGTGGCAAGCAAGCCCTCGAAACAGTGCAGCGG<br/> CTTCTACCCGTCTTTGTGTCAGGATCATGGACTCACACAGACAGGTGGTGCG<br/> CTATTGCTAGCAACATCGGAGGGAAGCAGGCGTTGGAAGCTGTGCAAGGCT<br/> CTTGCCCGTGTTGTGTCAAGACCATGGTCTGACTCCCGATCAAGTCTGGCT<br/> ATTGCTTCAAACATTGGTGGAAAACAGGCCCTTGGAACCGTCCAAGAGACTGC<br/> TTCCCGTGCTTTGCCAAGACCACGGCTTGACCCAGATCAGGTCTGGCAAT<br/> TGCATCCAATGGAGGCGGTAAACAAGCTCTGGAGACATACAGAGACTCCTG<br/> CCCGTCTCTGTCAAGATCATGGATTGACCCCTGACCAGGTCTGTCGAATTG<br/> CCTCTAACATCGGCGGGAAACAAGCTTTGGAGACTGTGCAGAGACTTCTCCC<br/> AGTGCTATGTCAAGATCACGGCCTTACCCCGATCAGTGGTCTGCCATTGCA<br/> AGCAATGGCGGTGGAAGCAAGCTCTCGAAAGCATCGTGGCCAGCTGAGCA<br/> GACCCGACCTGCCCTGGCCGCTCTGACCAACGACACCTCTGTGGCCCTGGC<br/> CTGCCTGGGTGGCAGACCAGCCATGGACGCGGTGAAGAAGGGCCTGCCTCAC<br/> GCCCCTGAGCTGATCAGAAGAGTGAACAGGAGAATCGCGAGAGAAACAGCC<br/> ACAGAGTGGCCATCTCTAGAGTGGGAGGAAGCTCCAGAAGAGAGAGCATCAA<br/> TCCCTGGATCCTGACCGGCTTCGCCGACGCCGAGGGCTGCTTCAGCCTGGCG<br/> ATCGGGAAGCTGAACGGCGGCACCACAGATACTCGACCAAGCTGGTGTTCA<br/> GCATCGGCCTGCACAACAAGGACAAGAGCATCCTGGAGAACATCCAGAGCAC<br/> CTGGAAGGTGGGAACCATCAGCAAGTCTGGCGATACCCAAAGTGAAGCTGACC<br/> GTGACAAGATTGAGGACCTGAAGGTGATCATCGACCACTTCGAGAAGTACC<br/> CTCTGATCACCCAGAAGCTGGGCGACTACATGCTGTTCAAGCAGGCCCTCTC<br/> CGTGATGGAGAACAAGGGCCACCTGAAGGAGAACGGCATCAAGGAGCTGGTG<br/> AGAATCAAGGCCAAGATGAAGTGGGGTCTGACCGACGAGCTGAAGAAGGCCT<br/> TTCCCGAGGACATCAGCAGAGAGAGGCCTATCATCAACAAGAACATTTCCCAA<br/> CGGCAAGTGGCTGGCCGGATTACCCAGCGGCGACGGCAGCTTCATGGTGATC<br/> CTGGCCAAGAGCAGAACCGGAGCTAGAGTGCAGGTGAGGCTGTGCTTCGAGA<br/> TCGGCCAGCACATTAGAGACAAGAACCTGATGAACAGCCTGATCACCTACCT<br/> GGGCTGCGGACACATCAGCGAGAAGAACCTCTCCGAATCCAGCTTCCTGATC<br/> TTCAGAGTGACCAAGTTCTCCGACATCAAGGACAAAATCATTCCCCTGTTCC<br/> AGGAGAACCCTGATCGGCATGAAGCACGAGGACTTCGAGGACTGGTGCAA<br/> GGTGGCCAAGCTGATCGAAGAGAAGAAGCACCTGACCGAGAGCGGCCTGGAC<br/> GAGATCAAGAAGATCAAGCTGAACATGAACAAGGGCAGAGTGTTTTGATAA</p> |

|    |                     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
|----|---------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 25 | megaTAL 399C12 mRNA | GTGGACCTGAGAACCCTGGGCTACAGCCAGCAGCAGCAGGAGAAGATCAAAC<br>CCAAGGTGAGAAGCACCGTGGCTCAGCACCACGAGGCCCTGGTGGGCCACGG<br>CTTCACCCACGCCCACATCGTCGCCCTGAGCCAGCATCCCGCTGCCCTGGGA<br>ACCGTGGCCGTGACCTACCAGCACATCATCACAGCCCTGCCCGAAGCCACCC<br>ACGAGGACATCGTCGGAGTGGGCAAGCAGTGGAGCGGAGCCAGAGCCCTGGA<br>GGCCCTGCTGACCGACGCCGGAGAGCTGAGAGGCCCTCCTCTGCAGCTGGAC<br>ACCGGACAGCTGGTGAAGATCGCCAAGAGAGGCGGAGTGACCGCCATGGAGG<br>CCGTGCACGCCAGCAGAAACGCCCTGACCGGAGCTCCTCTGAACCTTACTCC<br>AGACCAAGTTGTCGCCATCGCCAGTAACAACGGAGGCAAGCAGGCTCTCGAA<br>ACCGTGCAGAGGTTGCTGCCTGTCTCTGCCAAGATCATGGGCTCACACCTG<br>ATCAGGTTGTGGCAATCGCCTCCAACAATGGCGGAAAGCAGGCCCTTGAAC<br>CGTCCAACGGCTCCTTCCCGTCTGTGTGTCAGGACACGGCTCTTACACCTGAC<br>CAGGTTGTGGCCATTGCATCAAATATCGGGGGGAAGCAAGCTCTAGAAACCG<br>TGCAACGGTTGCTTCCAGTGCTTTGTCAAGACCATGGACTGACTCCAGACCA<br>GGTGTGTGCCATCGCATCTAATATTGGAGGCAAAACAGGCTCTAGAGACAGTA<br>CAACGCCTGCTCCCCGTGCTCTGCCAGGACCATGGGTTGACTCCAGATCAGG<br>TGGTTGCTATCGCTAGTAATATCGGCGGAAACAAGCCCTGGAGACGGTGCA<br>GCGCTTGCTCCAGTTCTGTGCCAGGATCATGGGCTGACACCCGATCAAGTG<br>GTTGCAATCGCCAGCAACAACGGTGGTAAACAGGCTCTTGAGACTGTTCAAC<br>GGCTGTTGCCCGTGCTGTGCCAGGACCACGGACTTACCCAGATCAAGTTGT<br>GGCTATCGCTTCCAATATTGGTGGCAAGCAAGCCCTCGAAACAGTGCAGCGG<br>CTCTACCCGTCTTTGTGTCAGGATCATGGACTCACACAGACACAGGTGGTCG<br>CTATTGCTAGCAACATCGGAGGGAAGCAGGCGTTGGAACTGTGCAAAAGGCT<br>CTGCCCCGTGTTGTGTCAAGACCATGGTCTGACTCCCGATCAAGTCGTGGCT<br>ATTGCTTCAAACATTGGTGGAAAACAGGCCCTTGGAAACCGTCCAAGACTGC<br>TTCCCGTGCTTTGCCAAGACCACGGCTTGACCCAGATCAGGTCGTGGCAAT<br>TGCATCCAATGGAGGCGGTAAACAAGCTCTGGAGACAGTACAGAGACTCCTG<br>CCCGTCTCTGTCAAGATCATGGATTGACCCTGACCAGGTGCTCGCAATTG<br>CCTCTAACATCGGCGGGAACAAGCTTTGGAGACTGTGCAGAGACTTCTCCC<br>AGTGCTATGTCAAGATCACGGCCTTACCCCGATCAGGTGGTCGCCATTGCA<br>AGCAATGGCGGTGGAAGCAAGCTCTCGAAAGCATCGTGGCCAGCTGAGCA<br>GACCCGACCTGCCCCTGGCCGCTCTGACCAACGACCATCTGTCGGCCCTGGC<br>CTGCCTGGGTGGCAGACCAGCCATGGACGCCGTGAAGAAGGGCCTGCCTCAC<br>GCCCCGTAGCTGATCAGAAGAGTGAACAGGAGAATCGGCGAGAGAACAGCC<br>ACAGAGTGGCCATCTCTAGAGTGGGAGGAAGCTCCAGAAGAGAGAGCATCAA<br>TCCCTGGATCCTGACCGGCTTCGCCGACGCCGAGGGCTGCTTACAGCTGCAC<br>ATCGCGAAGTTCAACGGCAAGACCACAGATACGGGACCAAGCTGGTGTTC<br>GCATCGGCCTGCACAACAAGGACAAGAGCATCCTGGAGAATCCAGAGC<br>CTGGAAGGTGGGAACCATCAGCAAGTCTGGCGATACCCAAGTGAAGCTGACC<br>GTGACAAGATTGAGGACCTGAAGGTGATCATCGACCACTTCGAGAAGTACC<br>CTCTGATCACCAGAAAGCTGGGCGACTACATGCTGTTCAAGCAGGCCTTCTC<br>CGTGATGGAGAACAAGGGCCACCTGAAGGAGAAGCGCATCAAGGAGCTGGTG<br>AGAATCAAGGCCAAGATGAACTGGGGTCTGACCGACGAGCTGAAGAAGGCCT<br>TTCCCGAGGACATCAGCAGAGAGAGGCCCTATCATCAACAAGAATTCCTCAA<br>CGGCAAGTGGCTGGCCGATTACACAGCGCGACGGCAGCTTCATGGTGATC<br>CTGGCCAAGAGCAGAACCAGGAGCTAGAGTGCAGGTGAGGCTGTGCTTCGAGA<br>TCGGCCAGCACATTAGAGACAAGAACCTGATGAACAGCCTGATCACCTACCT<br>GGGCTGCGGACACATCAGCGAGAAGAACCTCTCCGAATCCAGCTTCCTGATC<br>TTCAGAGTGACCAAGTTCTCCGACATCAAGGACAAAATCATTCCCGTGTTC<br>AGGAGAACACCTGATCGGCATGAAGCACGAGGACTTCGAGGACTGGTGCAA<br>GGTGGCCAAGCTGATCGAAGAGAAGAAGCACCTGACCGAGAGCGGCCTGGAC<br>GAGATCAAGAAGATCAAGCTGAACATGAACAGGGCAGAGTGTTTGTATAA |
| 26 | megaTAL 399D05 mRNA | GTGGACCTGAGAACCCTGGGCTACAGCCAGCAGCAGCAGGAGAAGATCAAAC<br>CCAAGGTGAGAAGCACCGTGGCTCAGCACCACGAGGCCCTGGTGGGCCACGG<br>CTTCACCCACGCCCACATCGTCGCCCTGAGCCAGCATCCCGCTGCCCTGGGA<br>ACCGTGGCCGTGACCTACCAGCACATCATCACAGCCCTGCCCGAAGCCACCC<br>ACGAGGACATCGTCGGAGTGGGCAAGCAGTGGAGCGGAGCCAGAGCCCTGGA<br>GGCCCTGCTGACCGACGCCGGAGAGCTGAGAGGCCCTCCTCTGCAGCTGGAC<br>ACCGGACAGCTGGTGAAGATCGCCAAGAGAGGCGGAGTGACCGCATGGAGG<br>CCGTGCACGCCAGCAGAAACGCCCTGACCGGAGCTCCTCTGAACCTTACTCC                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |

|    |                        |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
|----|------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|    |                        | <p>AGACCAAGTTGTGCGCCATCGCCAGTAACAACGGAGGCAAGCAGGCTCTCGAA<br/>ACCGTGCAGAGGTTGCTGCCTGTCTCTGCCAAGATCATGGGCTCACACCTG<br/>ATCAGGTTGTGGCAATCGCCTCCAACAATGGCGGAAAGCAGGCCCTTGAAAC<br/>CGTCCAACGGCTCCTTCCCGTCTGTGTGTCAGGACCACGGTCTTACACCTGAC<br/>CAGGTTGTGGCCATTGCATCAAATATCGGGGGGAAGCAAGCTCTAGAAACCG<br/>TGCAACGGTTGCTTCCAGTGCTTTGTCAAGACCATGGACTGACTCCAGACCA<br/>GGTTGTTGCCATCGCATCTAATATTTGGAGGCAAACAGGCTCTAGAGACAGTA<br/>CAACGCCTGCTCCCCGTGCTCTGCCAGGACCATGGGTTGACTCCAGATCAGG<br/>TGGTTGCTATCGCTAGTAATATCGGCGGAAAACAAGCCCTGGAGACGGTGCA<br/>GCGCTTGCTCCAGTTCTGTGCCAGGATCATGGGCTGACACCCGATCAAGTG<br/>GTTGCAATCGCCAGCAACAACGGTGGTAACAGGCTCTTGAGACTGTTCAAC<br/>GGCTGTTGCCCGTGCTGTGCCAGGACCACGGACTTACCCAGATCAAGTTGT<br/>GGCTATCGCTTCCAATATTTGGTGGCAAGCAAGCCCTCGAAACAGTGCAGCGG<br/>CTTCTACCCGTCTTTTGTGAGGATCATGGACTCACACCAGACCAGGTGGTTCG<br/>CTATTGCTAGCAACATCGGAGGGAAGCAGGCGTTGGAAACTGTGCAAAAGGCT<br/>CTTGCCCGTGTTGTGTCAAGACCATGGTCTGACTCCCGATCAAGTCTGGCT<br/>ATTGCTTCAAACATTGGTGGAAAACAGGCCCTTGGAAACCGTCCAAGACTGC<br/>TTCCCGTGCTTTGCCAAGACCACGGCTTGACCCAGATCAGGTCTGGCAAT<br/>TGCAATCCAATGGAGGCGGTAAACAAGCTCTGGAGACAGTACAGAGACTCCTG<br/>CCCGTCTCTGTCAAGATCATGGATTGACCCCTGACCAGGTCTGTCGAATTG<br/>CCTCTAACATCGGCGGGAACAAGCTTTGGAGACTGTGCAGAGACTTCTCCC<br/>AGTGCTATGTCAAGATCACGGCCTTACCCCGATCAGGTGGTCCGCATTGCA<br/>AGCAATGGCGGTGGAAGCAAGCTCTCGAAAGCATCGTGGCCAGCTGAGCA<br/>GACCCGACCCTGCCCTGGCCGCTCTGACCAACGACCACCTCGTGGCCCTGGC<br/>CTGCCTGGGTGGCAGACCAGCCATGGACGCGCTGAAGAAGGGCCTGCCTCAC<br/>GCCCTGAGCTGATCAGAAGAGTGAACAGGAGAATCGGCGAGAGAACGAGCC<br/>ACAGAGTGGCCATCTCTAGAGTGGGAGGAAGCTCCAGAAGAGAGAGCATCAA<br/>TCCCTGGATCCTGACCGGCTTCGCCGACGCCGAGGGCTGCTTCAGCCTGGTG<br/>ATCGCCAAGGGGAACGGCGGCACCACCAGATACACCACCAAGCTGGTGTTC<br/>GCATCGGCCTGCACAACAAGGACAAGAGCATCCTGGAGAATCCAGAGCAC<br/>CTGGAAGGTGGGAACCATCAGCAAGTCTGGCGATACCCAAAGTGAAGCTGACC<br/>GTGACAAAGATTGAGGACCTGAAGGTGATCATCGACCACTTCGAGAAGTACC<br/>CTCTGATCACCCAGAAGCTGGGCGACTACATGCTGTTCAAGCAGGCTTCTC<br/>CGTGATGGAGAACAAGGGCCACCTGAAGGAGAACGGCATCAAGGAGCTGGTG<br/>AGAATCAAGGCCAAGATGAACCTGGGGTCTGACCGACGAGCTGAAGAAGGCCT<br/>TTCCCGAGGACATCAGCAGAGAGAGGCCTATCATCAACAAGAATTCCTCAA<br/>CGCAAGTGGCTGGCCGGATTACACAGCGGCGACGGCAGCTTCATGGTGATC<br/>CTGGCCAAGAGCAGAACCGGAGCTAGAGTGCAGGTGAGGCTGTGCTTCGAGA<br/>TCGGCCAGCACATTAGAGACAAGAACCTGATGAACAGCCTGATCACTACCT<br/>GGGCTGCGGACACATCAGCGAGAAGAACCCTCTCCGAATCCAGTCTCTGATC<br/>TTCAGAGTGACCAAGTTCTCCGACATCAAGGACAAAATCATTCCCCTGTTCC<br/>AGGAGAACACCCTGATCGGCATGAAGCACGAGGACTTCGAGGACTGGTGCAA<br/>GGTGGCCAAGCTGATCGAAGAGAAGAAGCACCTGACCAGAGCGGCCTGGAC<br/>GAGATCAAGAAGATCAAGCTGAACATGAACAAGGGCAGAGTGTTTTGATAA</p> |
| 27 | megaTAL 399D09<br>mRNA | <p>GTGGACCTGAGAACCCTGGGCTACAGCCAGCAGCAGCAGGAGAAGATCAAAC<br/>CCAAGGTGAGAAGCACCGTGGCTCAGCACCACGAGGCCCTGGTGGGCCACGG<br/>CTTCACCCACGCCCACATCGTCGCCCTGAGCCAGCATCCCGCTGCCCTGGGA<br/>ACCGTGGCCGTGACCTACCAGCACATCATCACAGCCCTGCCCGAAGCCACCC<br/>ACGAGGACATCGTCGGAGTGGGCAAGCAGTGGAGCGGAGCCAGAGCCCTGGA<br/>GGCCCTGCTGACCGACGCCGGAGAGCTGAGAGGCCCTCCTCTGCAGCTGGAC<br/>ACCGGACAGCTGGTGAAGATCGCCAAGAGAGGCGGAGTGACCGCATGGAGG<br/>CCGTGCACGCCAGCAGAAACGCCCTGACCGGAGCTCCTCTGAACCTTACTCC<br/>AGACCAAGTTGTGCGCATCGCCAGTAACAACGGAGGCAAGCAGGCTCTCGAA<br/>ACCGTGCAGAGGTTGCTGCCTGTCTCTGCCAAGATCATGGGCTCACACCTG<br/>ATCAGGTTGTGGCAATCGCCTCCAACAATGGCGGAAAGCAGGCCCTTGAAAC<br/>CGTCCAACGGCTCCTTCCCGTCTGTGTGTCAGGACCACGGTCTTACACCTGAC<br/>CAGGTTGTGGCCATTGCATCAAATATCGGGGGGAAGCAAGCTCTAGAAACCG<br/>TGCAACGGTTGCTTCCAGTGCTTTGTCAAGACCATGGACTGACTCCAGACCA<br/>GGTTGTTGCCATCGCATCTAATATTTGGAGGCAAACAGGCTCTAGAGACAGTA<br/>CAACGCCTGCTCCCCGTGCTCTGCCAGGACCATGGGTTGACTCCAGATCAGG<br/>TGGTTGCTATCGCTAGTAATATCGGCGGAAAACAAGCCCTGGAGACGGTGCA</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |

|    |                     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
|----|---------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|    |                     | <p>             GCGCTTGCTCCAGTTCTGTGCCAGGATCATGGGCTGACACCCGATCAAGTG<br/>             GTTGCAATCGCCAGCAACAACGGTGGTAAACAGGCTCTTGAGACTGTTCAAC<br/>             GGCTGTTGCCCCGTGCTGTGCCAGGACCACGGACTTACCCAGATCAAGTTGT<br/>             GGCTATCGCTTCCAATATTGGTGGCAAGCAAGCCCTCGAAACAGTGCAGCGG<br/>             CTTCTACCCGTCTTTGTGTCAGGATCATGGACTCACACCAGACCAGGTGGTCG<br/>             CTATTGCTAGCAACATCGGAGGGGAAGCAGGCGTTGGAACTGTGCAAAGGCT<br/>             CTGCCCCGTGTTGTGTCAAGACCATGGTCTGACTCCCGATCAAGTCGTGGCT<br/>             ATTGCTTCAAACATTGGTGGAAAACAGGCCCTTGGAAACCGTCCAAGACTGC<br/>             TTCCCGTGCTTTGCCAAGACCACGGCTTGACCCAGATCAGGTCGTGGCAAT<br/>             TGCATCCAATGGAGGCGGTAAACAAGCTCTGGAGACAGTACAGAGACTCCTG<br/>             CCCGTCTCTGTCAAGATCATGGATTGACCCCTGACCAGGTCGTGCAATTG<br/>             CCTCTAACATCGGCGGGAAACAAGCTTTGGAGACTGTGCAGAGACTTCTCC<br/>             AGTGCTATGTCAAGATCACGGCTTACCCCGATCAGGTGGTCGCCATTGCA<br/>             AGCAATGGCGGTGGAAAGCAAGCTCTCGAAAGCATCGTGGCCAGCTGAGCA<br/>             GACCCGACCTGCCCCTGGCCGCTCTGACCAACGACCACCTCGTGGCCCTGGC<br/>             CTGGCTGGGTGGCAGACCAGCCATGGACGCCGTGAAGAAGGGCCCTGCTCAC<br/>             GCCCTGAGCTGATCAGAAGAGTGAACAGGAGAATCGGCAGAGAACCCAGCC<br/>             ACAGAGTGGCCATCTCTAGAGTGGGAGGAAGCTCCAGAGAGAGAGCATCAA<br/>             TCCCTGGATCCTGACCGGCTTCGCCGACGCCGAGGGCTCTTCAAGCTGGCC<br/>             ATCGGCAAGTGAACCGCTGCACCACAGATACAGCACAAGCTGGTGTTC<br/>             GCATCGGCCTGCACAACAAGGACAAGAGCATCCTGGAGAACATCCAGAGCAC<br/>             CTGGAAGGTGGGAACCATCAGCAAGTCTGGCGATACCCAAGTGAGGCTGACC<br/>             GTGACAAGATTTCGAGGACCTGAAGGTGATCATCGACCACTTCGAGAAGTACC<br/>             CTCTGATCACCCAGAAGCTGGGCGACTACATGCTGTTCAAGCAGGCCTTCTC<br/>             CGTGATGGAGAACAAGGGCCACCTGAAGGAGAACGGCATCAAGGAGCTGGTG<br/>             AGAATCAAGGCCAAGATGAAGTGGGTCTGACCGACGAGCTGAAGAAGGCCT<br/>             TTCCCGAGGACATCAGCAGAGAGAGGCCTATCATCAACAAGAATTTCCAA<br/>             CGGCAAGTGGCTGGCCGGATTACACAGCGGCGACGGCAGCTTCATGGTGATC<br/>             CTGGCCAAGAGCAGAACCCGAGCTAGAGTGCAGGTGAGGCTGTGCTTCGAGA<br/>             TCGGCCAGCACATTAGAGACAAGAACCTGATGAACAGCTGATCACTACCT<br/>             GGGCTGCGGACACATCAGCGAGAAGAACCCTCTCCGAATCCAGCTTCTGATC<br/>             TTCAGAGTGACCAAGTTCTCCGACATCAAGGACAAAATCATTCCGTGTTCC<br/>             AGGAGAACACCCTGATCGGCATGAAGCACGAGGACTTCGAGGACTGGTGCAA<br/>             GGTGGCCAAGCTGATCGAAGAGAAGAAGCACCTGACCGAGAGCGGCCTGGAC<br/>             GAGATCAAGAAGATCAAGCTGAACATGAACAAGGGCAGAGTGTTTTTGATAA           </p> |
| 28 | megaTAL 399D10 mRNA | <p>             GTGGACCTGAGAACCCTGGGCTACAGCCAGCAGCAGCAGGAGAAGATCAAAC<br/>             CCAAGGTGAGAAGCACCGTGGCTCAGCACCACGAGGCCCTGGTGGGCCACGG<br/>             CTTACCCACGCCCACATCGTCGCCCTGAGCCAGCATCCCGCTGCCCTGGGA<br/>             ACCGTGGCCGTGACCTACCAGCACATCATCACAGCCCTGCCCGAAGCCACCC<br/>             ACAGGACATCGTCGGAGTGGGCAAGCAGTGGAGCGGAGCCAGAGCCCTGGA<br/>             GGCCCTGCTGACCGACGCCGAGAGCTGAGAGGCCCTCCTCTGCAGCTGGAC<br/>             ACCGGACAGCTGGTGAAGATCGCCAAGAGAGGCGGAGTGACCGCCATGGAGG<br/>             CCGTGCACGCCAGCAGAAACGCCCTGACCGGAGCTCCTCTGAACCTTACTCC<br/>             AGACCAAGTTGTCGCCATCGCCAGTAACAACGGAGGCAAGCAGGCTCTCGAA<br/>             ACCGTGCAGAGGTTGCTGCCTGTCTCTGCCAAGATCATGGGCTCACACCTG<br/>             ATCAGGTTGTGGCAATCGCCTCCAACAATGGCGGAAAGCAGGCCCTTGAAC<br/>             CGTCCAACGGCTCCTTCCCGTCTGTGTGTCAGGACCACGGTCTTACACCTGAC<br/>             CAGGTTGTGGCCATTGCATCAAATATCGGGGGGAAGCAAGCTCTAGAAACCG<br/>             TGCAACGGTTGCTTCCAGTGCTTTGTCAAGACCATGGACTGACTCCAGACCA<br/>             GGTGTTGCCATCGCATCTAATATTGGAGGCAACAGGCTCTAGAGACAGTA<br/>             CAACGCCCTGCTCCCGTGCTCTGCCAGGACCATGGGTTGACTCCAGATCAGG<br/>             TGGTTGCTATCGCTAGTAATATCGGCGGAAACAAGCCCTGGAGACGGTGCA<br/>             GCGCTTGCTCCAGTTCTGTGCCAGGATCATGGGCTGACACCCGATCAAGTG<br/>             GTTGCAATCGCCAGCAACAACGGTGGTAAACAGGCTCTTGAGACTGTTCAAC<br/>             GGCTGTTGCCCCGTGCTGTGCCAGGACCACGGACTTACCCAGATCAAGTTGT<br/>             GGCTATCGCTTCCAATATTGGTGGCAAGCAAGCCCTCGAAACAGTGCAGCGG<br/>             CTTCTACCCGTCTTTGTGTCAGGATCATGGACTCACACCAGACCAGGTGGTCG<br/>             CTATTGCTAGCAACATCGGAGGGGAAGCAGGCGTTGGAACTGTGCAAAGGCT<br/>             CTGCCCCGTGTTGTGTCAAGACCATGGTCTGACTCCCGATCAAGTCGTGGCT<br/>             ATTGCTTCAAACATTGGTGGAAAACAGGCCCTTGGAAACCGTCCAAGACTGC<br/>             TTCCCGTGCTTTGCCAAGACCACGGCTTGACCCAGATCAGGTCGTGGCAAT           </p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |

|    |                        |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
|----|------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|    |                        | <p>TGCATCCAATGGAGGCGGTAAACAAGCTCTGGAGACAGTACAGAGACTCCTG<br/>         CCGTCTCTGTCAAGATCATGGATTGACCCCTGACCAGGTCGTCGCAATTG<br/>         CCTCTAACATCGGCGGGAACAAGCTTTGGAGACTGTGCAGAGACTTCTCCC<br/>         AGTGCTATGTCAAGATCACGGCCTTACCCCGATCAGGTGGTCGCCATTGCA<br/>         AGCAATGGCGGTGGAAAGCAAGCTCTCGAAAGCATCGTGGCCAGCTGAGCA<br/>         GACCCGACCCCTGCCCTGGCCGCTCTGACCACGACCACCTCGTGGCCCTGGC<br/>         CTGCCTGGGTGGCAGACCAGCCATGGACGCCGTGAAGAGGGGCTGCCTCAC<br/>         GCCCCTGAGCTGATCAGAAGAGTGAACAGGAGAATCGGCGAGAGAACCAGCC<br/>         ACAGAGTGGCCATCTCTAGAGTGGGAGGAAGCTCCAGAAGAGAGAGCATCAA<br/>         TCCCTGGATCCTGACCGGCTTCGCCGACGCCGAGGGTCTCTCAGCCTGGCC<br/>         ATCGGCAAGTTGAACGGCGCCACCACCAGATACGTACCAAGCTGGTGTTC<br/>         GCATCGGCTGCACAACAAGGACAAGAGCATCCTGGAGAATCCAGAGCAC<br/>         CTGGAAGGTGGGAACCATCAGCAAGTCTGGCGATACCCAAGTGAAGCTGACC<br/>         GTGACAAGATTGAGGACCTGAAGGTGATCATCGACCACTTCGAGAAGTACC<br/>         CTCTGATCACCCAGAAGCTGGGCGACTACATGCTGTTCAAGCAGGCCTTCTC<br/>         CGTGATGGAGAACAAGGGCCACCTGAAGGAGAACGGCATCAAGGAGCTGGTG<br/>         AGAATCAAGGCCAAGATGAACGGGGTCTGACCGACGAGCTGAAGAAGGCCT<br/>         TTCCCGAGGACATCAGCAGAGAGAGGCCTATCATCAACAAGAATTTCCCAA<br/>         CGGCAAGTGGCTGGCCGGATTACACAGCGGCGACGGCAGCTTCAGGTGATC<br/>         CTGGCCAAGAGCAGAACCGGAGCTAGAGTGCAGGTGAGGCTGTGCTTCGAGA<br/>         TCGGCCAGCACATTAGAGACAAGAACCTGATGAACAGCTGATCAGCTACCT<br/>         GGGCTGCGGACACATCAGCGAGAAGAACCTTCTCCGAATCCAGCTTCTCTGATC<br/>         TTCAGAGTGACCAAGTTCTCCGACATCAAGGACAAAATCATTCCCGTGTTC<br/>         AGGAGAACACCTGATCGGCATGAAGCACGAGGACTTCGAGGACTGGTGCAA<br/>         GGTGGCCAAGCTGATCGAAGAGAAGAAGCACCTGACCAGAGCGGCTGGAC<br/>         GAGATCAAGAAGATCAAGCTGAACATGAACAAGGGCAGAGTGTTTTGATAA</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
| 29 | megaTAL 399H10<br>mRNA | <p>GTGGACCTGAGAACCCTGGGCTACAGCCAGCAGCAGGAGAGAAGATCAAAC<br/>         CCAAGGTGAGAAGCACCGTGGCTCAGCACCACGAGGCCCTGGTGGGCCACGG<br/>         CTTCAACCCACGCCACATCGTCGCCCTGAGCCAGCATCCCGCTGCCCTGGGA<br/>         ACCGTGGCCGTGACCTACCAGCACATCATCACAGCCCTGCCCGAAGCCACCC<br/>         ACGAGGACATCGTCGGAGTGGGCAAGCAGTGGAGCGGAGCCAGAGCCCTGGA<br/>         GGCCCTGCTGACCGACGCCGAGAGCTGAGAGGCCCTCCTCTGCAAGCTGGAC<br/>         ACCGGACAGCTGGTGAAGATCGCCAAGAGAGGCGGAGTGACCGCCATGGAGG<br/>         CCGTGCACGCCAGCAGAAACGCCCTGACCGGAGCTCCTCTGAACCTTACTCC<br/>         AGACCAAGTTGTCGCCATCGCCAGTAACAACGGAGGCAAGCAGGCTCTCGAA<br/>         ACCGTGCAGAGGTTGCTGCCTGTCTCTGCCAAGATCATGGGCTCACACTG<br/>         ATCAGGTTGTGGCAATCGCCTCCAACAATGGCGGAAAGCAGGCCCTTGAAAC<br/>         CGTCCAACGGCTCCTTCCCGTCTGTGTGTCAGGACCACGGTCTTACACCTGAC<br/>         CAGGTTGTGGCCATTGCATCAAATATCGGGGGGAAGCAAGCTCTAGAAACCG<br/>         TGCAACGGTTGCTTCCAGTGCTTTGTCAAGACCATGGACTGACTCCAGACCA<br/>         GGTTGTTGCCATCGCATCTAATATTGGAGGCAAACAGGCTCTAGAGACAGTA<br/>         CAACGCCCTGCTCCCGTGCTCTGCCAGGACCATGGGTTGACTCCAGATCAGG<br/>         TGGTTGCTATCGCTAGTAATATCGGCGGAAAACAAGCCCTGGAGACGGTGCA<br/>         GCGCTTGCTCCAGTTCTGTGCCAGGATCATGGGCTGACACCCGATCAAGTG<br/>         GTTGCAATCGCCAGCAACAACGGTGGTAAACAGGCTCTTGAGACTGTTCAAC<br/>         GGCTGTTGCCCGTGCTGTGCCAGGACCACGGACTTACCCAGATCAAGTTGT<br/>         GGCTATCGCTTCCAATATTGGTGGCAAGCAAGCCCTCGAAACAGTGCAGCGG<br/>         CTTCTACCCGTCTTTGTGTCAGGATCATGGACTCACACCAGACCAGGTGGTCG<br/>         CTATTGCTAGCAACATCGGAGGGAAGCAGGCGTTGGAACCTGTGCAAAAGGCT<br/>         CTTGCCCCTGTGTGTCAAGACCATGGTCTGACTCCCGATCAAGTCGTGGCT<br/>         ATTGCTTCAAACATTGGTGGAAAACAGGCTTGGAAACCGTCCAAGACTGC<br/>         TTCCCGTGCTTTGCCAAGACCACGGCTTGACCCAGATCAGGTCTGGCAAT<br/>         TGCATCCAATGGAGGCGGTAAACAAGCTCTGGAGACAGTACAGAGACTCCTG<br/>         CCCGTCTCTGTCAAGATCATGGATTGACCCCTGACCAGGTCTGTCGAATTG<br/>         CCTCTAACATCGGCGGGAACAAGCTTTGGAGACTGTGCAGAGACTTCTCCC<br/>         AGTGCTATGTCAAGATCACGGCCTTACCCCGATCAGGTGGTCGCCATTGCA<br/>         AGCAATGGCGGTGGAAAGCAAGCTCTCGAAAGCATCGTGGCCAGCTGAGCA<br/>         GACCCGACCCCTGCCCTGGCCGCTCTGACCAACGACCACCTCGTGGCCCTGGC<br/>         CTGCCTGGGTGGCAGACCAGCCATGGACGCCGTGAAGAGGGGCTGCCTCAC<br/>         GCCCCTGAGCTGATCAGAAGAGTGAACAGGAGAATCGGCGAGAGAACCAGCC<br/>         ACAGAGTGGCCATCTCTAGAGTGGGAGGAAGCTCCAGAAGAGAGAGCATCAA</p> |

|    |                                                     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
|----|-----------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|    |                                                     | <p>TCCCTGGATCCTGACCGGCTTCGCCGACGCCGAGGGCTGCTTCAGCCTGCAC<br/> ATCGGGAAGATGAACGGCCGGACCACAGATACCTGACCAAGCTGGTGTTC<br/> GCATCGGCTGCACAACAAGGACAAGAGCATCCTGGAGAATCCAGAGCAC<br/> CTGGAAGGTGGGAACCATCAGCAAGTCTGGCGATACCCAAGTGAGGCTGACC<br/> GTGACAAGATTCGAGGACCTGAAGGTGATCATCGACCACTTCGAGAAGTACC<br/> CTCTGATCAGCCAGAGCTGGGCGACTACATGCTGTTCAAGCAGGCTTCTC<br/> CGTGATGGAGAACAAGGGCCACCTGAAGGAGAACGGCATCAAGGAGCTGGTG<br/> AGAATCAAGGCCAAGATGAACTGGGGTCTGACCGACGAGCTGAAGAAGGCCT<br/> TTCCCGAGGACATCAGCAGAGAGAGGCCTATCATCAACAAGAATTCCCAA<br/> CGGCAAGTGGCTGGCCGGATTACCAGCGGCGACGGCAGCTTCATGGTGATC<br/> CTGGCCAAGAGCAGAACCGGAGCTAGAGTGCAGGTGAGGCTGTGCTTCGAGA<br/> TCGGCCAGCACATTAGAGACAAGAACCTGATGAACAGCCTGATCACCTACCT<br/> GGGCTGCGGACATCAGCGAGAAGAACCCTCCGAATCCAGCTTCTGATC<br/> TTCAGAGTGACCAAGTTCTCCGACATCAAGGACAAAATCATTCCCCTGTTCC<br/> AGGAGAACACCTGATCGGCATGAAGCACGAGGACTTCGAGGACTGGTGCA<br/> GGTGCCCAAGCTGATCGAAGAGAAGAAGCACCTGACCGAGAGCGGCTGGAC<br/> GAGATCAAGAAGATCAAGCTGAACATGAACAAGGGCAGAGTGTTTTGATAA</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
| 30 | ALB Gene DNA<br>sequence -<br>accession<br>M12523.1 | <p>CCTTTCCAGGACTTCTACAAGGAAAAAGCTAGAGTTGGTTACTGACTTCT<br/> AATAAATAATGCCTACAATTTCTAGGAAGTTAAAAGTTGACATAATTTATCC<br/> AAGAAAGAAATATTTTCTTAACCTAGAATAGTTTCTTTTCTTTTCAGATG<br/> TAGGTTTTTCTGGCTTTAGAAAAAATGCTTGTTTTCTTCAATGGAAAAATAG<br/> GCACACTTGTTTTATGCTGTTCTGTCATCTGTAGTCAGAAAGACAAGTCTGGTAT<br/> TTCCTTTTCCAGGACTCCCTTGAGTCATTAAAAAAAATCTTCTATCTATCTAT<br/> GTATCTATCATCCATCTAGCTTTGATTTTTTCTCTCTGTGCTTTATTAGT<br/> TAATTAGTACCCATTTCTGAAGAAGAAATAACATAAGATTATAGAAAATAAT<br/> TTCTTTTCAATTGTAAGACTGAATAGAAAAAATTTCTTTTCAATTATAGACTGA<br/> GTAGAAAAAATAATACTTTGTTAGTCTCTGTGCCTCTATGTGCCATGAGGAA<br/> ATTTGACTACTGGTTTTGACTGACTGAGTTATTTAATTAAGTAAAAAATACTG<br/> GCTTAGTACTAATTATTTGTTCTGTAGTATCAGAGAAAGTTGTTCTTCTACT<br/> GGTTGAGCTCAGTAGTCTTTCATATTCTGAGCAAAAGGGCAGAGGTAGGATA<br/> GCTTTTCTGAGGTAGAGATAAGAACCCTGGGTAGGGGAAGGAAATTTATGAA<br/> ATATTTAAAAAATTAATCTTCTTCTGCTTTGTTTTAGACATAATGTAAAT<br/> TTATTTTGAATTTAAAGCAACATAAAAGACATGTGATTTTTCTACTTATT<br/> GAAAGAGAGAAAGGAAAAAATATGAAACAGGGATGGAAGAATCCTATGCC<br/> TGGTGAAGGTCAAGGGTTCTCATAACCTACAGAGAATTTGGGGTCAGCCTGT<br/> CCTATTGTATATTATGGCAAGATAATCATCATCTCATTTGGGTCCATTTTC<br/> CTCTCCATCTCTGCTTAACTGAAGATCCCATGAGATATACTCACACTGAATC<br/> TAAATAGCCTATCTCAGGGCTTGAATCACATGTGGGCACAGCAGGAATGGG<br/> AACATGGAATTTCTAAGTCCATCTTACTTGTATTGTTGCTATGTCTTTTT<br/> CTTAGTTTGCATCTGAGGCAACATCAGCTTTTTTCCAGACAGAATGGCTTTGGA<br/> ATAGTAAAAAAGACACAGAAGCCCTAAAAATATGTATGTATGTATATGTGTGT<br/> GTGCATGCGTGAGTACTTGTGTGTAAATTTTTTCAATTATCTATAGGTAAAAGC<br/> ACACTTGGAATTAGCAATAGATGCAATTTGGGACTTAACTCTTTTCAATATGT<br/> CTTATTTCTAAGCAAAGTATTTAGTTTGGTTAGTAATTACTAAACACTGAGA<br/> ACTAAATTGCAAAACACCAAGAATAAAATGTTCAAGTGGGAAATTACAGTTA<br/> AATACCATGGTAATGAATAAAAGGTACAAATCGTTTAAACTCTTATGTAAAA<br/> TTTGATAAGATGTTTTACACAACTTTAATACATTGACAAGGTCTTGTGGAGA<br/> AAACAGTTCCAGATGGTAAATATACACAAGGGATTTAGTCAAACAAATTTTTT<br/> GGCAAGAATATTATGAATTTTGTAAATCGTTTGGCAGCCAATGAAATACAAAG<br/> ATGAGTCTAGTTAATAATCTACAATTATTGGTTAAAGAAGTATATTAGTGCT<br/> AATTTCCCTCCGTTTGTCTAGCTTTTCTCTCTGTCAACCCACACGCCTT<br/> TGGCACAATGAAGTGGGTAACCTTTATTTCCCTTCTTTTCTCTTAGCTCG<br/> GCTTATTTCCAGGGGTGTGTTTCTGTCGAGATGCACGTAAGAAATCCATTTTCT<br/> TATTGTTCAACTTTTATCTATTTTCCAGTAAAAATAAAGTTTTAGTAAACT<br/> CTGCATCTTTAAAGAATATTTTGGCATTATTTCTAAAAATGGCATAGTATT<br/> TTGTATTTGTGAAGTCTTACAAGTTATCTTATTAATAAAATTCAAACATCC<br/> TAGGTAAAAAAGGTCAGAATTGTTTAGTGACTGTAATTTTCTTTT<br/> GCGCACTAAGGAAAGTCAAAGTAACTTAGAGTGACTGAAACTTCACAGAAT<br/> AGGGTTGAAGATTGAATTCATAACTATCCCAAAGACCTATCCATTGCATAT<br/> GCTTTATTTAAAAACCACAAAACCTGTGCTGTTGATCTCATAAATAGAAGCTT<br/> GTATTTATATTTATTTTCAATTTTAGTCTGTCTTCTGGTTGCTGTTGATAGA</p> |

|  |  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
|--|--|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|  |  | CACTAAAAGAGTATTAGATATTATCTAAGTTTGAATATAAGGCTATAAATAT<br>TTAATAATTTTAAAAATAGTATTCTTGGTAATTGAATTATTCTTCTGTTTAA<br>AGGCAGAAGAAATAATTGAACATCATCCTGAGTTTTTCTGTAGGAATCAGAG<br>CCCAATATTTTGAAACAAATGCATAATCTAAGTCAAATGGAAAGAAATATAA<br>AAAGTAACATTATTACTTCTTGTCTTCTCAGTATTTAAACAATCCTTTTTTT<br>TCTTCCCTTGCCAGACAAGAGTGAGGTTGCTCATCGGTTTAAAGATTGGG<br>AGAAGAAAATTTCAAAGCCTTGTAAGTTAAATATTGATGAATCAAATTTAA<br>TGTTTCTAATAGTGTTGTTTATTATTCTAAAGTGCTTATATTTCTTGTGCT<br>CAGGGTTTCAAGTTCTAAAACAGTGCTGCCTCGTAGAGTTTCTGCGTTGAGG<br>AAGATATTCTGTATCTGGGCTATCCAATAAGGTAGTCACTGGTCACATGGCT<br>ATTGAGTACTTCAAATATGACAAGTGCAACTGAGAAACAAAACCTTAAATTG<br>TATTTAATTGTAGTTAATTTGAATGTATATAGTCACATGTGGCTAATGGCTA<br>CTGTATTGGACAGTACAGCTCTGGAACTTGCTTGGTGAAAGGACTTTAATA<br>TAGGTTTCTTTGGTGGCTTACCCACTAAATCTTCTTTACATAGCAAGCATT<br>CCTGTGCTTAGTTGGGAATATTTAATTTTTTTTTTTTTTAAAGACAGGGTCT<br>CGCTCTGTGCGCCAGGCTGGAGTGCAAGTGCGCAATCTCGGCTCACTGCAAA<br>CTCCGCTCCCGGGTTCACGCCATTCTCCTGCCTCAGCCTCCCGAGTAGCTGG<br>GACTACAGGCGCCCGCATCACGCGCGGCTAATCTTTGTATTTTTTAGTAGA<br>GATGGGGTTTACCCTGTGCCAGGATGGTCTCAATCTCCTGACATCGTGATC<br>TGCCACCTCGGCCTCCCAAAGTGCTGGGATTACAGGAGTGAGTCACCGCGC<br>CCGGCCTATTTAAATGTTTTTAAATCTAGTAAAAAATGAGAAAATGTTTTT<br>TTAAAGTCTACCTAATCCTACAGGCTAATTAAAGACGTGTGTGGGGATCAG<br>GTGCGGTGGTTACACCTGTAATCCCAGCACTTTGGAAGGCTGATCGAGAG<br>GATTGCTTGAGCCAGGAGTACAAGACCAGCCTGGGCAAGTCTCTTTAAAAA<br>AAACAAAACAAAACAAAACAAAATAGGCATGGTGGCACATGCCTGTAGTC<br>CTAGCTACTTAGGAGGCTGACGTAGGAGGATCGTTTGGACCTGAGAGGTCAA<br>GGCTACAGTGAGCCATGATTGTGCCACTGCACTCCAGCCTGGGTGACAGAGT<br>GAGACTCTGTCTAAAAAAGAAAAGGAAATCTGTGGGGTTTGTTTTAGTTTT<br>TAAGTAATCTAAGGACTTTAAAAATGCCTAGTCTTGACAATTAGATCTATT<br>TGGCATACAATTTGCTTGCTTAATCTATGTGTGTGCATAGATCTACTGACAC<br>ACGCATACATATAACATTAGGGAACCTACCATTCTCTTTGCGTAGGAAGCCA<br>CATATGCCTATCTAGGCCTCAGATCATACCTGATATGAATAGGCTTTCTGGA<br>TAATGGTGAAGAAGATGTATAAAAGATAGAACCATACCCATACATGATTG<br>TTCTCTAGCGTAGCAACCTGTTACATATTAAGTTTTATTATACTACATTTT<br>TCTACATCCTTTGTTTCAGGGTGTGATTGCCTTTGCTCAGTATCTTCAGCA<br>GTGTCCATTTGAAGATCATGTAAATTAGTGAATGAAGTAACTGAATTTGCA<br>AAAACATGTGTGCTGATGAGTCAGCTGAAAATTTGTGACAAATCACTTTGTAA<br>GTACATTCTAATTGTGGAGATTCTTCTTCTGTTTGAAGTAATCCCAAGCAT<br>TTCAAAGGAATTTTTTTTAAAGTTTCTCAATTATTATTAAAGTGCTCTGATT<br>GTAAGAAACACTAAAAAGTTGCTCATAGACTGATAAGCCATTGTTCTTTTG<br>TGATAGAGATGCTTTAGCTATGTCCACAGTTTTAAATCATTCTTTATTGA<br>GACCAAAACAACAGTCATGGTGTATTTAAATGGCAATTTGTCAATTATAAA<br>CACCTCTTTTTAAATTTGAGGTTTGGTTTTCTTTTGTAAGGCTAATAGGG<br>ATATGATAGCATGTATTTATTTATTTATTTATCTTATTTTATTATAGTAAGA<br>ACCCTTAACATGAGATCTACCCTGTTATATTTTAAAGTGACAATCCATTAT<br>TGTTAACTACGGGTACACTGTTGTATAGCTTACTCATCTTGCTGTATTA<br>CTTTGTGCCATTGATTAGTAACCCCTCGTTTCGTCCCTCCCCAGCCACTGG<br>CAACCAGCATTATACTCTTTGATTCTATGAGTTTGACTACTTTAGTACCTT<br>ATATAAGTGGTATTATGTACTGTTTATCTTTTATGACTGACTTATTTCCCT<br>TAGCATAGTGCAATCAAAGTCCAACCATGTTGTTGCCATTGCAGAATTTCC<br>TTCTTTTCAAGGCTGAATAATATTCAGTGATGTGTGTACCATTTTCTT<br>TATCCATTAAATTTGTTGATTGATAGACATTAGGTTGGTTTTCTACATCTTG<br>ACTATCATGAATAGTGTTGCAATGAACACAGGAGAGCTACTATCTCTTAGAG<br>ATGATATCATGGTTTTTATCATCAGAAAACCCCACTGATTTCTATGCTAAT<br>TTTGTTACCTGGGTGGAATAATAGTACAGCTATATATTCCTCATTTTAGATA<br>TCTTTGTATTTCTACATACAATAAAAAAGCAGAGTACTTAGTCATGTTGAAG<br>AACTTTAACTTTTAGTATTTCCAGATCAATCTTCAAAACAAGGACAGGTTT<br>ATCTTTCTCTCACCCTCAATCTATATATACCTCTTGTGGGCAAGGCCAGTT<br>TTTATCACTGGAGCCTTTCCCTTTTTTATTATGTACCTCTCCCTCACAGCAG<br>AGTCAGGACTTTAACTTTACACAATACTATGGCTCTACATATGAAATCTTAA<br>AAATACATAAAAAATTAATAAATTTCTGTCTAGAGTAGTATATTTTCCCTGGGG<br>TTACGATTACTTTCATAATAAAAAATTAGAGATAAGGAAAGGACTCATTATT |
|--|--|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|



|  |  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
|--|--|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|  |  | <p>GGAAAGTGATTTTAGGTAACATTTCTGGAAGAAAAATGTCTATATCTTAATA<br/>GTCACTTAATATATGATGGATTGTGTACTCCTCAGTTTTCAATGGCATATA<br/>CTAAAACATGGCCCTCTAAAAAGGGGGCAAAATGAAATGAGAACTCTCTGAA<br/>TGTTTTTCTCCCCTAGGTGAATTCACCTGCTGCTTAGAAGCTTATTTCTCT<br/>TGATTTCTGTTATAATGATTGCTCTTACCCTTTAGTTTTAAGTTTCAAAATA<br/>GGAGTCATATAAATTTCCCTTAAAGCTATTGACTGTCCTTTTGCTGTTTTTA<br/>TTCACCATGAGTTATAGTGTGACAGTTAATTCCTATGAAAATTATATAGAGA<br/>TGGTTAAATCATCAGAACTGTAAACCTCGATTGGGAGGGGAAGCGGATTTT<br/>TAAATGATTTCTGACCAAGCTTAACCAGTATATTAATCCTTTGTACTGTT<br/>CTTTGGCTATAAAGAAAAAGGTACTGTCCAGCAACTGAAACCTGCTTTCTT<br/>CCATTTAGCATACCCTTTTTGGAGACAAATTATGCACAGTTGCAACTCTTCG<br/>TGAAACCTATGGTGAATGGCTGACTGCTGTGCAAAACAAGAACCTGAGAGA<br/>AATGAATGCTTCTTGCAACACAAAGATGCAACCCAAACCTCCCCGATTGG<br/>TGAGACCAGAGGTTGATGTGATGTGCACTGCTTTTCATGACAATGAAGAGAC<br/>ATTTTTGAAAAAGTAAGTAATCAGATGTTTATAGTTCAAAATTA AAAAGCAT<br/>GGAGTAACCTCCATAGGCCAACACTCTATAAAAATTACCAATACAAAAATATT<br/>TTCAACATTAAGACTTGAAGTTTTGTATGATGATTTTTTAAGAAGTAGT<br/>ATTTGATACCACAAAATTTCTACACAGCAAAAAATATGATCAAAGATATTTTG<br/>AAGTTTATTGAAACAGGATACAATCTTCTGAAAAATTTAAGATAGACAAAT<br/>TATTTAATGTATTACGAAGATATGTATATATGGTTGTTATAATTGATTTCTG<br/>TTTAGTCAGCAACATTATATTGCCAAAATTTAACCATTATGCACACACACA<br/>CACACACACACACTTAACCTTTTTTCCACATACTTAAAGAATGACAGAG<br/>ACAAGACCATCATGTGCAAATTGAGCTTAATGGTTAATTAGATATCTTTGG<br/>AATTTGGAGGTTCTGGGGAGAATGTCGATTACAATTATTTCTGTAATATTGT<br/>CTGCTATAGAAAAGTGACTGTTTTCTTTTTCAAAATTTAGATACTTATATG<br/>AAATTGCCAGAAGACATCCTTACTTTTATGCCCGGAACCTCCTTTCTTTGCT<br/>TAAAAGGTATAAAGCTGCTTTTACAGAATGTTGCCAAGCTGCTGATAAAGCT<br/>GCCTGCCTGTTGCCAAGGTATTATGCAAAAGAATAGAAAAAAGAGTT CAT<br/>TATCCAACCTGATTTTGTCCATTTTGTGGCTAGATTTAGGGAACCTGAGTGT<br/>CTGATACAACTTTCCGACATGGTCAAAAAAGCCTTCCCTTTTATCTGTCTTG<br/>AAAACTTTTCATCTTTGAAGGCCTACACTCTCGTTTCTCTTTTAAAGATTTG<br/>CCAATGATGATCTGTGAGAGTAATCACTGTGCAATGTGTTAAAGATTTTAC<br/>CACTTTTTTATGGTGGTGATCACTATAGTGAAATACTGAAACTGTTTGTCAA<br/>ATTGCACAGCAAGGGGACACAGTTCTTGTTTATCTTTTCATGATAATTTTTTA<br/>GTAGGGAGGGAATTCAAAGTAGAGAATTTTACTGCATCTAGATGCTGAGTT<br/>CATGCATTCATTCCATAAATATATATTATGGAATGCTTTATTTCTTTCTG<br/>AGGAGTTTACTGATGTTGGTGGAGGAGAGACTGAAATGAATTATACACAAAA<br/>TTTAAAAATTAGCAAAATTCAGCCCTGGGATATTAGCGTACTCTTCTCT<br/>GACTTTTCTCCACTTTTAAAGGCTCTTTTTCTGCAATGTTTCCAGTTGGT<br/>TTCTAACTACATAGGGAATTCGCTGTGACCAGAATGATCGAATGATCTTTC<br/>CTTTTCTTAGAGAGCAAAATCATTATTGCTAAAGGGAGTACTTGGGAATTT<br/>AGGCATAAATATATGCTTCAAAATTTAATTTGGCACAGTCTCATCTGAGCTT<br/>ATGGAGGGGTGTTTCATGTAGAAATTTTCTTCTAATTTTCATCAAATTATTC<br/>CTTTTGTAGCTCGATGAACCTTCGGGATGAAGGGAAGGCTTCGCTGCCCCAA<br/>CAGAGACTCAAGTGTGCCAGTCTCCAAAAATTTGGAGAAAGAGCTTCAAAG<br/>CATGGTAAATACTTTTAAACATAGTTGGCATCTTTATAACGATGTAATGAT<br/>AATGCTTCAGTGACAAATTGTACATTTTTATGTATTTTGCAAAGTGCTGTCA<br/>AATACATTTCTTGGTGTCTAACAGGTAGAATCTAATAGAGGTAAAAATC<br/>AGAATATCAATGACAATTTGACATATTTTTAATCTTTTCTTTCTAAATAG<br/>TTGAATAATTTAGAGGACGCTGTCTTTTTGTCTTAAAAAAGGGACAGATA<br/>TTTAAGTTCTATTTATTTATAAAATCTTGGACTCTTATCTAATGGTTTCAAT<br/>ATTTTATAGAGCTGTAGGCATGGTTCTTTATTTAATTTTTTAAAGTTATTT<br/>TTAATTTTTGTGGATACAGAGTAGGTATACATATTTACGGGGTATATGAGAT<br/>ATTTTGATATAAGTATACAACATATATAATCCCTTTATTTAATTTTATCTTC<br/>CCCCAATGATCTAAACTATTTGCTTGTCTTTTATGTCTTATAGTTAAAT<br/>TCAGTCACCAACTAAGTTGAAGTTACTTCTTATTTTTGCATAGCTCCAGCTC<br/>TGATCTTCATCTCATGTTTTTGCCTGAGCTCTGTTTTCATATTACTTAGTT<br/>GGTCTGGGAGCATACTTTAATAGCCGAGTCAAGAAAAATACAGCTGCCCC<br/>GTCACCCACACTCCTCACCTGCTAGTCAACAGCAAAATCAACACAACAGGAAA<br/>TAAAATGAAAAATAAGACATTATGCATGCTCTCTAGAACTGTCAATTGAA<br/>CTGTATTTGCTCATCATTCCTACCATCTACACCACCAAAATCAACCAATTT<br/>ATGAAAAAAAACAGCCCAACATAAAATTTATACACAGATAAACAGGCTATG</p> |
|--|--|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

|  |  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
|--|--|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|  |  | <p>ATTGGTTTTGGGAAAGAAGTCACCTTTACCTGATTTAGGCAACTGTGAAATG<br/>ACTAGAGAATGAAGAAAATTAGACGTTTACATCTTGTGCATAGAGTTTGAAGA<br/>TAGTGCTGGATCTTTCTTTTATAAGTAAGATCAATAAAAACTCCCTCATTC<br/>TGTAAGAAGTTATGATTTCTTTTCTAAGAGACCTTTAGAAAGTCAGAAAAAATG<br/>TGTTTCAATTGAGAAAAAGATAACTGGAGTTTGTGTAGTACTTCCAGATT<br/>ATAAAATGCTTTTGTATGTATTATCTAATTAAATCCTCAAAACTTCTTCAAT<br/>TTAGCATGTTGTGCATGACACTGCAGAGGCTGAAGCTCAGAGACGCTGAGCCC<br/>TCTGCTAACCAAGTCTACTGCTAACCAAGTGATAAAGCCAGAGCTGGAAGTCA<br/>CATCTGGACTCCAAACCTGATGCTTCTCAGCCTGTTGCCCTTTTAGAGTTT<br/>CTTTTAAATTTCTGCTTTTATGACTTGTCTAGATTTCTACCTACCACACACAC<br/>TCTTAAATGGATAATTCTGCCCTAAGGATAAGTGATTACCATTGTTGTTT<br/>CTAGAACTAATGAATTTTAAAAATTATTTCTGTATGTCATTTTGAATTTT<br/>CTTATGAGAAATAGTATTTGCCCTAGTGTTTTCATATAAAATATCGCATGATA<br/>ATACCATTTTGATTGGCGATTTTCTTTTAGGGCAGTAGCTCGCCTGAGCCA<br/>GAGATTTCCCAAAGCTGAGTTTGCAGAAGTTTCCAAGTTAGTGACAGATCTT<br/>ACCAAAGTCCACACGGAATGCTGCCATGGAGATCTGCTTGAATGTCTGATG<br/>ACAGGGTAAAGAGTCTGTCGATATGCTTTTTGGTAGCTTGCATGCTCAAGTTG<br/>GTAGAATGGATGCGTTTGGTATCATTTGGTGATAGCTGACAGTGGGTGAGAT<br/>TGTCTTCTGTGCTTTCTGCTCTGCTATCTTCAATCTTTCCCTGTTGGTG<br/>GTGGTACCTTTCTGTTTTAACCTGCTATAAATTACCAGATAAAACCCATTCA<br/>CTGATTTGTAACCTCTTTGAGTCATGCTCTAACTGTAATGAAGGCTTAAAC<br/>TGAAGTAGAACAGTTACAAGTTTTACTTGGCAGAACATCTTGCAAGGTAGA<br/>TGCTAAGAAGATTTTTTTTTCTTTTTTTAAGACAGAGTTTCGCTCTTGT<br/>CCCAGGCTGGGGTGCAATGGTGTGATCTTGGCTCAGCGCAACCTCTGCCTCC<br/>TGGGTTCAAGTGATTTTCATGCCCTCAGCCTCCCAAGTAGCTGGGATTACAGG<br/>CATGCGCCACACACCTGGCTAATTTTGTATTTTAGTAGAGCGGGGTTT<br/>ACCATATTGTCCAGACTGGTCTCGAACTCCTGACCTCAGGTGATCCACCCGC<br/>CTTGGCCTCCCAAAGTGCTGGGATTACAGGCATGAGCCACCTTGCCAGCCT<br/>AAGAAGATTTTTGAGGGAGGTAGGTGGACTTGGAGAAGGTCACTACTTGAA<br/>GAGATTTTTGGAAATGATGTATTTTCTTCTCTATATTCCTTCCCTTAATTA<br/>ACTCTGTTTGTAGATGTGCAAAATATTTGGAATGATATCTTTTTCTCAAAA<br/>CTTATAATATTTCTTCTCCCTTTCTTCAAGATTAACTTATGGGCAAAATA<br/>CTAGAATCCTAATCTCTCATGGCACTTTCTGGAAAAATTTAAGGCGGTTATTT<br/>TATATATGTAAGCAGGCGCTATGACTATGATCTTGACTCATTTTTTCAAAAAT<br/>CTTCTATATTTTATTTAGTTATTTGGTTTTCAAAAGGCCTGCACTTAATTTTG<br/>GGGGATTATTTGAAAAACAGCATTGAGTTTTAATGAAAAAACTTAAATGC<br/>CCTAACAGTAGAAACATAAAATTAATAAATAACTGAGCTGAGCCTGCTAC<br/>TGATTAGTCTATTTTAATTAAGTGGGAATGTTTTGTAGTCTTATCTACATC<br/>TCCAGGTTTAGGAGCAACAGAGTATGTTCATAGAAGGAATATGTGTATGGT<br/>CTTAGAATAACAATGAACATGTTCTGCCAACTTAATAAAGGTCTGAGGAGAAA<br/>GTGTAGCAATGTCAATTCGTGTTGAACAATTTCCACCAACTTACTTATAGGC<br/>GGACCTTGCCAAGTATATCTGTGAAAATCAAGATTTCGATCTCCAGTAAACTG<br/>AAGGAATGCTGTGAAAAACCTCTGTTGGAAAAATCCCCTGCACTTGCCGAAG<br/>TGAAAAATGATGAGATGCCCTGCTGACTTGCCCTTCAATAGCTGCTGATTTTGT<br/>TGAAAGTAAGGATGTTTGCAAAAATATGCTGAGGCAAGGATGTCTTCTG<br/>GGCATGTAAGTAGATAAGAAATTAATCTTTTATAGCTTTGGCATGACCTCAC<br/>AACTTAGGAGGATAGCCTAGGCTTTTCTGTGGAGTTGTACAATTTCCCTGC<br/>TGCCAGAAATGTTTCTTCTATCCTTCCCTTTCCAGGCTTTAACAATTTTTGA<br/>AATAGTTAATTAGTTGAATACATTTGTATATAAATAATACATGTTACGGCAA<br/>AGCTCAACATTCCTTACTCCTTAGGGGATTTCTGAAAAACGCTAGAAAC<br/>ATTTTGTGTATATATAAATTATGTATACTTCAGTCATTCATTCCAAGTGAT<br/>TTCTTGAACATCTATAATATATGTGTGACTATGTATGCTGTCTATCTA<br/>ACTAATCTAATCTAATCTAGTCTATCTATCTAATCTATGCAATGATAGCAAA<br/>GAAGTATAAAAAGAAATATAGAGTCTGACACAGGTGCTTATATTTGGTGAA<br/>AAGACCAGAAGTTCAGTATAATGGCAATATGGTAGGCAACTCAATTACAAAA<br/>TAAATGTTTACGTATTGTGAGAAGTTGTGGTGATAAACTGCATTTTGTGT<br/>TGGATTATGATAATGCACTAAATAATTTTCTAAAAATATGTACCCTACAA<br/>GATTTCACTCATAACAGAGAGAAAGAGAAATATTTAAGAACATATCTCTGCC<br/>CATCTATTTATCAGAATCCTTTTGAAGTGTAGTTTAAATCAAACAAAATGTT<br/>AATAAAAAATAACAAGTATCATTCATCAAAGACTTCATATGTGCCAAGCAGTG<br/>TGTGCTTTGTGTAGATTATGTCATATAGTTCTCATAATCCACCTTCCGAGAC<br/>AGATACTATTTATTTTTTGAACAGAGTTTTACTCTTGTGCCAGGCTGGA</p> |
|--|--|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

|  |  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
|--|--|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|  |  | <p>GTGCAATGGTGCCATCTCGGCTCACCACAACCTTCGCCTCCCAGGTTCAAGC<br/>GATTCTCCTGCCTCAGCCTCCTGGGATTACAGGCATGCACCACCATGCCTGG<br/>CTAATTTTGTATTTTGTAGTAGAGATGGGGTTTCACCATGTTGGTCAGACTGG<br/>TCTCAAACCTCCTGACCTCTGGTGATATGCCTGCCTCAGCCTCCTAAAGTGCT<br/>GGGATTACAGGCATGAGCCACTGTGCCCAGCCGACAGATACTATTATTATTT<br/>CCATTCTACCGAGAAGGAGACTAAGGCTCTGATCATTTAAATAAGTTGCCTA<br/>AGGTGATGCAGTGATATAAGTAGCAGAGCTAGGAATTGAGCCTTGTTAACTT<br/>TAACTCTGGACCCCAAGTCTCTAGCTACTAAGCTTTACTGCATGGGGTTTAG<br/>TCAAATTAAGACTTTTGGAAATATGAGTTACTTTTGGAGATTAGCTTTGTGATA<br/>TTTTTTGTGCTCATTTGTCCAACAAAGTCTATTTTATTTTCATCTTAATTAG<br/>GTTTTTGTATGAATATGCAAGAAGGCATCCTGATTACTCTGTGCTGCTGCTG<br/>CTGAGACTTGCCAAGACATATGAAACCACTCTAGAGAAGTGCTGTGCCGCTG<br/>CAGATCCTCATGAATGCTATGCCAAAGTGGTAGGTTTATTTGTTGGAAAAAAA<br/>TGTAGTTCTTTGACTGATGATTCGAATAATGAGAAAGAAAAATAATGCAAGA<br/>ATGTAANAATGATATACAGTGCAATTTAGATCTTTCTTGAGATGGTTTCAAT<br/>TCTGGAATCTTAAACATGAAAGAAAAAGTAGCCTTAGAATGATTAAACAAAT<br/>TTAGACTAGTTAGAATAGAAAGATCTGAATAGAGCAATCTCTAAAAATTTT<br/>GATCTTTTTTTCTCTTTTTCACATCCTGAGAACAAAAAAAATTAATTTTA<br/>AATGTTAATTAGAAGATATTTAACTTAGATGTAAGTGAGTTAACTGATTC<br/>CAGGATTAATCAAGTACTAGAATTAGTATCTTATGGCAAAATTAAGAACCTA<br/>TCCCTTTAGAATATTTTCAAATCTTTTGGAGGATGTTTAGGAATAGTTTTAC<br/>AAGAAATTAAGTTAGGAGAGGAAATCTGTTCTGGAGGATTTTAGGGTTCCC<br/>ACTAGCATATGTAATGGTTTCTGAACTATTCAGAATCAGAGAAAACTCATTT<br/>TTCCTGCTTTCAAGAAGCTACTGTATGCCAGGCACCATGCACAAACAATGAC<br/>CAACGTAANAATCTCTCATTTTGGAGAGCCTGGAATCTAACTGGAAGGTGAA<br/>CTAATAATAATAATATGTACAATCATAGCCATCATTTATTAACCTTTTATTA<br/>TATGCAAGGCACTGTTAATTTTCATTAGCTTACCTGGTTTACAGAGCAGCTC<br/>TATGAGATGAGTGCCATCTTTGCCCTATTTTAGGGATAAGGATTCGAAAT<br/>GTGGAGATGGTAAGTAAATTCACAACCTGAAGAATGAGTTACATGACTTGG<br/>CTCAAATACTGGTCATTTGAACTCCAGAGCCTGAATATTTCTAACCCTTACA<br/>TGATGCAAGCTCACCAATAAATAGTTCGAATGTATTGTGACAGAGCGGCAT<br/>TGATATTCATCTATTCTATGTGGCTTTGAGTAGGAAGAAAGGATATCATTT<br/>CTGACCAGAGGGGTGAAAAACAACCTGCATCTGATCCTGAGGCATTAATCTA<br/>TTAACACAATTTCTTTTATGTTTCAGTTTCGATGAATTTAAACCTCTTGTGGAA<br/>GAGCCTCAGAATTTAATCAAACAAAATTTGTGAGCTTTTTGAGCAGCTTGGAG<br/>AGTACAAATCCAGAATGCGTAAGTAATTTTTATTGACTGATTTTTTTTATC<br/>AATTTGTAATTTATTTAAGACTTAATATATGAGCCACCTAGCATAGAATCTTT<br/>AAGAATGAAAATACATTCATATTTCTAATCACTCTTTGTCAAGAAAGATAG<br/>GAGAGGAGAGATAAAATAGTTGATGGGGTGGAGAGGTCTATATTTGAATGTA<br/>GTCTAAAAATTTGTTCTTTAAGATTGGAAGTATGTAGGCTGGGAGGGTAAAT<br/>ACCAATCTTGGTATATCAGAACCTGAGCATGTCCCTTGAAGGTTAAGAAATA<br/>GTTAATGGGCAATAGAGCATGGCAATATTTGTAGAGCAGCAAGTAGTAGG<br/>CCTTGAATAGATGTCGCTCAAAAAGTAATATGTAAGCTGAACACAAAATGT<br/>AACAAATGAATTTAGATACATATTTGAATATTAATTCAGGTTGTTTGGGAG<br/>ATGCACCTAGTCTTTGATGGTTAAACCTTTCCCTCCATAGAAGAGACAGAGA<br/>CAGAATGGCTTGCTGGACTAATGTCCCAATTCATAGAGTCTTATCTACGAA<br/>GGTTAAAAACAAGAAGAGACATATTATACAGTAGATATTTATTGTGTGGCTC<br/>ATACACATGGTGCTCTTCTGATTATGGATTTTAGAGATAATAACAGTGAACA<br/>AGACATAGTTTCTTTCTCGAGTAGATTAAAGTCATACATTGACTTTTAATG<br/>GTGACTGGCATTCTTAATACATGATTATTATATATTAGGTACCATGTCAGAT<br/>TAATTATAATACTTTACTATTTTAAATTTAACCTTGAACATATCCCTATTGA<br/>GTCAGATATATTTCTTCCATTTTCTACTTGATCTTTCAAGTTTAGCATAT<br/>GCTGATACATATGAAGCTCTCTCCAGGTTTATTGAAAGAAGAAATTAATAA<br/>ATTTATTAATGTCACTGAATTAGGCAACTCACTTTCCCAAGATTATGCAAGT<br/>GGTACAGGTGGAACCTCAAAGCCAAGTTAACTAGTTGTTTCAAGGAGATGTTT<br/>TCTACCTCCACTAACCCACTACTCTGCAGATGGAGATAATATGATGAATGG<br/>AACATAGCAACATCTTAGTTGATTCGGCCAAAGTGTTCTCTGTTTTATCTAC<br/>TATGTTAGACAGTTTCTTGCTTGCTGAAAACACATGACTCTTTTTTTTCAG<br/>GCTATTAGTTTCGTTACACCAAGAAAGTACCCCAAGTGTCAACTCCAACCTCT<br/>GTAGAGGTCTCAAGAAACCTAGGAAAAGTGGGCAGCAATGTTGTAACATC<br/>CTGAAGCAAAAAGAAATGCCCTGTGCAGAAGACTATGTGAGTCTTTAAAAAAA<br/>TATAATAAATTAATAATGAAAAATTTTACCTTTAGATATTGATAATGCTAG</p> |
|--|--|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

|  |  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
|--|--|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|  |  | CTTTCATAAGCAGAAGGAAGTAATGTGTGTGTGTGCATGTTTGTGTGCATGT<br>GTGTGTGCATGCACGTGTGTGTATGTGTGATATTGGCAGTCAAGGCCCCGAG<br>GATGATAATTTTTTTTTTTTTTTTTGAGACGGAGTCTCGCTTTGTTGTCCAGG<br>CTGGAGTGCAGTGGTGCCATCTCGGCTCACTGCAACCTCCGCTCCCAAGTT<br>CAAGCCATTCTCCTGCCTCAGCCTCCCAAGTAGCTGGGACTACAGGTGCATG<br>CCACCATGCCTGGCTAATTTTTTTGTATTTTGTAGTAGAAAATTTTCAGCTTCA<br>CCTCTTTTGAATTTCTGCTCTCCTGCCTGTTCTTTAGCTATCCGTGGTCTCTG<br>AACCAGTTATGTGTGTGTCATGAGAAAACGCCAGTAAGTGACAGAGTCACCA<br>AATGCTGCACAGAATCCTTGGTGAACAGGCGACCATGCTTTTCAGCTCTGGA<br>AGTCGATGAAACATACGTTCCCAAAGAGTTTAATGCTGAAACATTACCTTC<br>CATGCAGATATATGCACACTTTCTGAGAAGGAGAGACAAATCAAGAAACAAA<br>CGTGAGGAGTATTTCACTACTGCATGTGTTTGTAGTCTTGATAGCAAGAACT<br>GTCAATTCAAGCTAGCAACTTTTTCTGAAAGTAGTGATATATTTCTTAGAG<br>GAAAGTATTGGAGTGTGCCCTTATTATGCTGATAAGAGTACCCAGAATAAAA<br>ATGAATAACTTTTTAAAGACAAAATCCTCTGTTATAATATTGCTAAAATAT<br>TCAGAGTAATATTGTGGATTAAAGCCACAAATAGAATAACATGTTAGACCATA<br>TTCAGTAGAAAAAGATGAACAATTAAGTATAAATTTGTGCACATGGCAAAT<br>TAGTTAATGGGAACCATAGGAGAAATTTATTCTAGATGTAAATAATATTTT<br>AAGTTTGCCTATGGTGGCCCCACACATGAGACAAACCCCAAGATGTGACT<br>TTTGAGAAATGAGACTTGGATAAAAAACATGTAGAAATGCAAGCCTGAAGCT<br>CAACTCCCTATTGCTATCACAGGGGTATAAATTGCATAAAATTTAGCTATAG<br>AAAGTTGCTGTCTCTTGTGGGCTGTAATCATCGTCTAGGCTTAAGAGTA<br>ATATTGCAAAACCTGTCATGCCACACAAATCTCCTCGGCATTGTTGTCT<br>TTGCAGATGTGAGTGAAGAGAACAGCAGCTCCCATGAGTTTGGATAGCCT<br>TATTTTCTATAGCCTCCCCACTGAAGGGAGCAAAGTTTAAAGAACAAATATA<br>AAGTTTCTCATCTTTATAGATGAGAAAAATTTAAATAAAGTCCAAGATAAT<br>TAAATTTTAAAGGATCATTTTTTAGCTCTTTAATAGCAATAAACTCAATATG<br>ACATAATATGGCACTTCCAAAATCTGAATAATATATAATTGCAATGACATAC<br>TTCTTTTCAGAGATTTACTGAAAAGAAATTTGTTGACACTACATAACGTGAT<br>GAGTGGTTTATACTGATTGTTTCAGTTGGTCTTCCCACCAACTCCATGAAAG<br>TGGATTTTATTATCCTCATCATGCAGATGAGAATATTGAGACTTATAGCGGT<br>ATGCCTGGCCCAAGTACTCAGAGTTGCCTGGCTCCAAGATTATATCTTAA<br>ATGATGGGACTACCCTTACTCTCTCCATTTTCTATACCTGATGTAATGT<br>TTTTTCTGTTTTTTTTTTTTCTTTTCCATTCAAACCTCAGTGCCTTGTGGA<br>GCTCGTGAAACACAAGCCCAAGGCAACAAAAGAGCAACTGAAAGCTGTTATG<br>GATGATTTTCGCAGCTTTTGTAGAGAAGTGTGCAAGGCTGACGATAAGGAGA<br>CCTGCTTTGCCGAGGAGGTACTACAGTTCTCTTCATTTTAATATGTCCAGTA<br>TTCATTTTTCATGTTTGGTTAGGCTAGGGCTTAGGGATTATATATCAAAG<br>GAGGCTTTGTACATGTGGGACAGGGATCTTATTTTACAAACAATTGTCTTAC<br>AAAATGAATAAAACAGCACTTTGTTTTTATCTCTGCTCTATTGTGCCATAC<br>TGTTGAATGTTTATAATGCATGTTCTGTTTCCAAATTTGTGATGCTTATGAA<br>TATTAATAGGAATATTGTAAAGGCTGAAATATTTTGATCATGAAATCAAAA<br>CATTAAATTTATTTAAACATTTACTTGAAATGTGGTGGTTTGTGATTTAGTTG<br>ATTTTATAGGCTAGTGGGAGAATTTACATTCAAATGTCTAAATCACTTAAAA<br>TTTCCCTTTATGGCCTGACAGTAACTTTTTTTTATTTCATTTGGGGACAAC<br>TGTCCTGAGCTTCCATCCAGAGATTATAGTAGTAAATTTGTAATTAAGGAT<br>ATGATGCACGTGAAATCACTTTGCAATCATCAATAGCTTCATAAATGTTAAT<br>TTTGTATCCTAATAGTAATGCTAATATTTTCTAACATCTGTCTGCTTTG<br>TGTTCAGGGTAAAAAATTTGTTGCTGCAAGTCAAGCTGCCTTAGGCTTATAA<br>CATCACATTTAAAAGCATCTCAGGTAACATATATTTGAAATTTTTAAAAAG<br>TAACTATAATAGTTATTATTTAAATAGCAAAGATTGACCATTTCCAAAGAGCC<br>ATATAGACCAGCACCAGCACTATTTCTAACTATTTATGTATGTAATATTA<br>GCTTTTAAATTTCTCAAAATAGTTGCTGAGTTGGGAACCACTATTATTTCTA<br>TTTTGTAGATGAGAAAATGAAGATAAACATCAAAGCATAGATTAAAGTAATTT<br>TCCAAAGGGTCAAAATTCAAATTGAAACCAAGGTTTCAGTGTGCCCATTG<br>TCCTGTTCTGACTTATATGATGCGGTACACAGAGCCATCCAAGTAAGTGATG<br>GCTCAGCAGTGAATACTCTGGGAATTAGGCTGAACCACATGAAAGAGTGCT<br>TTATAGGGCAAAAACAGTTGAATATCAGTGATTTACATGGTTCAACCTAAT<br>AGTTCAACTCATCCTTTCCATTGGAGAATATGATGGATCTACCTTCTGTGAA<br>CTTTATAGTGAAGAATCTGCTATTACATTTCCAATTTGTCAACATGCTGAGC<br>TTAATAGGACTTATCTTCTTATGACAACATTTATTGGTGTGTCCCTTGCC<br>TAGCCCAACAGAAGAATTGAGCAGCCGTAAGTCTAGGACAGGCTTAAATTTGT |
|--|--|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

|    |                            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
|----|----------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|    |                            | TTTCACTGGTGTAATTCGAGAAAGATGATCTAAGTAATTTGGCATTTATTT<br>TAATAGGTTTGAAAAACACATGCCATTTTACAAATAAGACTTATATTTGTCC<br>TTTTGTTTTTCAGCCTACCATGAGAATAAGAGAAAGAAAATGAAGATCAAAA<br>GCTTATTCATCTGTTTTTCTTTTTTCGTTGGTGTAAGCCAACACCCTGTCTA<br>AAAAACATAAAATTTCTTTAATCATTTTGCCTCTTTTCTCTGTGCTTCAATTA<br>ATAAAAAATGGAAAGAATCTAATAGAGTGGTACAGCACTGTTATTTTCAAA<br>GATGTGTTGCTATCCTGAAAATTCGTAGGTTCTGTGGAAGTTCAGTGTTTC<br>TCTCTTATTCACCTTCGGTAGAGGATTTCTAGTTCCTGTGGGCTAATTTAA<br>TAAATCATTAATACTCTTCTAAGTTATGGATTATAAACATTCAAAATAATAT<br>TTTGACATTATGATAATTCTGAATAAAAGAACAAAACCATGGTATAGGTAA<br>GGAATATAAACATGGCTTTTACCTTAGAAAAACAATTCTAAATTCATAT<br>GGAATCAAAAAGAGCCTGCAG                                                                                                               |
| 31 | ALB intron 1               | GTAAAGAAATCCATTTTCTATTGTTCAACTTTTATTCTATTTTCCAGTAAA<br>ATAAAGTTTTAGTAACTCTGCATCTTTAAGAATTATTTTGGCATTTATTT<br>CTAAAATGGCATAGTATTTTGTATTTGTGAAGTCTTACAAGGTTATCTTATT<br>AATAAAATTCAAACATCCTAGGTAAAAAAGGTCAGAATTGTTTAG<br>TGACTGTAATTTTCTTTGCGCACTAAGGAAAGTGCAAAGTAACTTAGAGTG<br>ACTGAACTTCACAGAATAGGGTTGAAGATTGAATTCATAACTATCCCAAAG<br>ACCTATCCATTGCACTATGCTTTATTTAAAAACCAAAACCTGTGCTGTTG<br>ATCTCATAAATAGAACTGTATTTATATTTATTTTCATTTTAGTCTGTCTTC<br>TTGGTTGCTGTTGATAGACACTAAAAGAGTATTAGATATATCTAAGTTTGA<br>ATATAAGGCTATAAATATTTAATAATTTTAAAAATAGTATCTTGGTAATTG<br>AATTATTCTTCTGTTTAAAGGCAGAAGAAATAATTGAACATCATCCTGAGTT<br>TTTCTGTAGGAATCAGAGCCCAATATTTTGAACAAATGCATAATCTAAGTC<br>AAATGGAAGAAATATAAAAAGTAACATTATTACTTCTGTTTTCTTCAGTA<br>TTTAACAATCCTTTTTTTTCTTCCCTTGCCAG |
| 32 | HE Target site<br>S01      | TGGGCTCTGATTCCTACAGAAA                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
| 33 | HE Target site<br>S03      | CAGAGCCCAATATTTTGAACA                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| 34 | HE Target site<br>S06      | AAAGTAACATTATTACTTCTTG                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
| 35 | TALE target<br>site 8.5+5  | TGGAAAGAAA                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
| 36 | TALE target<br>site 9.5+4  | TGGAAAGAAAT                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| 37 | TALE target<br>site 10.5+3 | TGGAAAGAAATA                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| 38 | TALE target<br>site 11.5+2 | TGGAAAGAAATAT                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| 39 | TALE target<br>site 13.5+5 | TCAAATGGAAAGAAA                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| 40 | TALE target<br>site 12.5+6 | TCAAATGGAAAGAA                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| 41 | TALE target<br>site 11.5+7 | TCAAATGGAAAGA                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| 42 | TALE target<br>site 10.5+8 | TCAAATGGAAAG                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| 43 | megaTAL S06                | TGGAAAGAAATATAAAAAGTAACATTATTACTTCTTG                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |

|    |                                        |                                                                                                                                                                                                                                                                                                                                   |
|----|----------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|    | target site                            |                                                                                                                                                                                                                                                                                                                                   |
| 44 | Human Trex2                            | MGSCRPPKKRKVSEAPRAETFFVFLDLEATGLPSVEPEIAELSLFAVHRSSL<br>ENPEHDESGALVLPRLDKLTLCMCPPERPFTAKASEITGLSSEGLARCRKAG<br>FDGAVVRTLQAFLSRQAGPICLVAHNGFDYDFPLLCAELRRLGARLPDRDVC<br>LDTLPALRGLDRAHSHGTRARGRQGYSLGSLFHRYFRAEPSAAHSAEGDVHT<br>LLLI FLHRAAELLAWADEQARGWAHIEPMYLPDDPSLEA                                                          |
| 87 | Engineered HE<br>variant 206A06        | SRRESINPWTLTGFADAEGCFWLVRNQLNQPGLRYRTRLGFSIGLHNKDKSI<br>LENIQSTWKVGTISNSGDTAVRLTVTRFEDLKVIIDHFEKYPLITQKLGDYK<br>LFKQAFSVMENKEHLKENGIKELVRIKAKMNWGLTDELKKAFFEVISRERPL<br>INKNIPNLKWLAGFTSGDGYFGVILAKTRKPSRVQVRLYFEIGQHIKDKNLM<br>NSLITYLGCGGIYERNLSETSFLTRVTKFSDIKDKIIPVFQENTLIGVKLE<br>DFEDWCKVAKLIEEKKHLTESGLDEIKKIKLNMNKGKRVF   |
| 88 | Engineered HE<br>variant 206B07        | SRRESINPWTLTGFADAEGCFWLVIKNVNGHTTRYRTRLGFSIGLHNKDKSI<br>LENIQSTWKVGTISNSGDTAVRLTVTRFEDLKVIIDHFEKYPLITQKLGDYK<br>LFKQAFSVMENKEHLKENGIKELVRIKAKMNWGLTDELKKAFFEVISRERPL<br>INKNIPNGKWLAGFTSGDGSFMVILEKTRKPSRVQVRLSFEIGQHIRDKNLM<br>NSLITYLGCGHISEKNLSEPSFLIFRVTKFSDIKDKIIPVFQENTLIGVKLE<br>DFEDWCKVAKLIEEKKHLTESGLDEIKKIKLNMNKGKRVF  |
| 89 | Engineered HE<br>variant 206C09        | SRRESINPWTLTGFADAEGCFWLGIKNVNGHTTRYRTRLGFSIGLHNKDKSI<br>LENIQSTWKVGTISNSGDTAVRLTVTRFEDLKVIIDHFEKYPLITQKLGDYK<br>LFKQAFSVMENKEHLKENGIKELVRIKAKMNWGLTDELKKAFFEVISRERPL<br>INKNIPNLKWLAGFTSGDGSFMVILGKSRTGARVQVRLYFEIGQHIKDKNLM<br>NSLITYLGCGGIYERNLSETSFLTRVTKFSDIKDKIIPVFQENTLIGVKLE<br>DFEDWCKVAKLIEEKKHLTESGLDEIKKIKLNMNKGKRVF   |
| 90 | Engineered HE<br>variant 206G08        | SRRESINPWTLTGFADAEGCFWLVRNNGHTTRYRTRLGFSIGLHNKDKSI<br>LENIQSTWKVGTISNSGDTAVRLTVTRFEDLKVIIDHFEKYPLITQKLGDYK<br>LFKQAFSVMENKEHLKENGIKELVRIKAKMNWGLTDELKKAFFEVISRERPL<br>INKNIPNGKWLAGFTSGDGSFMVILAKTRKPSRVQVRLRFEIGQHIRDKNLM<br>NSLITYLGCGHISEKNFSETSFLTFGVTKFSDIKDKIIPVFQENTLIGVKLE<br>DFEDWCKVAKLIEEKKHLTESGLDEIKKIKLNMNKGKRVF    |
| 91 | Engineered HE<br>variant 206G09        | SRRESINPWILTGFADAEGCFWLVRNVNGHTTRYRTRLGFSIGLHNKDKSI<br>LENIQSTWKVGTISNSGDTAVRLTVTRFEDLKVIIDHFEKYPLITQKLGDYK<br>LFKQAFSVMENKEHLKENGIKELVRIKAKMNWGLTDELKKAFFEVISRERPL<br>INKNIPNGKWLAGFTSGDGSFMVILAKSRTGARVQVRLCFFEIGQHIRDKNLM<br>NSLITYLGCGHISEKNLSEPSFLIFRVTKFSDIKDKIIPVFQENTLIGMKLE<br>DFEDWCKVAKLIEEKKHLTESGLDEIKKIKLNMNKGKRVF  |
| 92 | Engineered HE<br>variant 206H06        | SRRESINPWTLTGFADAEGCFWLVIKNLNRGNSRYRTRLGFSIGLHNKDKSI<br>LENIQSTWKVGTISNSGDTAVRLTVTRFEDLKVIIDHFEKYPLITQKLGDYK<br>LFKQAFSVMENKEHLKENGIKELVRIKAKMNWGLTDELKKAFFEVISRERPL<br>INKNIPNGKWLAGFTSGDGSFMVILAKTRKPSRVQVRLCFFEIGQHIRDKNLM<br>NSLITYLGCGHISEKNLSEPSFLIFRVTKFSDIKDKIIPVFQENTLIGVKLE<br>DFEDWCKVAKLIEEKKHLTESGLDEIKKIKLNMNKGKRVF |
| 93 | Engineered HE<br>variant 214E11        | SRRESINPWTLTGFADAEGCFWLVIKNVNGHTTRYRTRLGFSIGLHNKDKSI<br>LENIQSTWKVGVITNSGDRYVRLTVSRFEDLKVIIDHFEKYPLITQKLGDYK<br>LFKQAFSVMENKEHLKENGIKELVRIKAKMNWGLTDELKKAFFPENISKERPL<br>INKNIPNGKWLAGFTSGDGTFFVVLAKSRTGARVQVRLSFEIGQHIRDKNLM<br>NSLITYLGCGHISEKNFSEPSFLIFRVTKFSDIKDKIIPVFQENTLIGMKLE<br>DFEDWCKVAKLIEEKKHLTESGLDEIKKIKLNMNKGKRVF |
| 94 | Engineered HE<br>variant 264B08        | SRRESINPWILTGFADAEGCFWLVRKVNGHTTRYRTRLGFSIGLHNKDKSI<br>LENIQSTWKVGTISNSGDTAVRLTVTRFEDLKVIIDHFEKYPLITQKLGDYK<br>LFKQAFSVMENKEHLKENGIKELVRIKAKMNWGLTDELKKAFFEVISRERPL<br>INKNIPNGKWLAGFTSGDGSFMVILAKSRTGARVQVRLCFFEIGQHIRDKNLM<br>NSLITYLGCGHISEKNLSEPSFLIFRVTKFSDIKDKIIPVFQENTLIGMKLE<br>DFEDWCKVAKLIEEKKHLTESGLDEIKKIKLNMNKGKRVF  |
| 95 | Engineered HE<br>variant<br>264B08+B09 | SRRESINPWILTGFADAEGCFWLVRKVNGHTTRYRTRLGFSIGLHNKDKSI<br>LENIQSTWKVGTISNSGDTAVRLTVTRFEDLKVIIDHFEKYPLITQKLGDYK<br>LFKQAFSVMENKHLKENGIKELVRIKAKMNWGLTDELKKAFFEVISRERPL<br>INKNIPNGKWLAGFTSGDGSFMVILAKSRTGARVQVRLCFFEIGQHIRDKNLM<br>NSLITYLGCGHISEKNLSEPSFLIFRVTKFSDIKDKIIPVFQENTLIGKKLE<br>DFEDWCKVAKLIEEKKHLTESGLDEIKKIKLNMNKGKRVF   |

|     |                                  |                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
|-----|----------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 96  | Engineered HE variant 264F08     | SRRESINPWILTGFADEAGCFWLVRNVNGHTTRYRTRLGFSIGLHYKDKSI<br>LENIQSTWKVGTISNSGDTAVRLTVTRFEDLKVIIDHFEKYPLITQKLG DYK<br>LFKQAFSVMENKGHLKENGIKELVRIKAKMNWGLTDELKKAPEVISRERPL<br>INKNIPNGKWLAGFTSGDGSFMVILAKSRTGARVQVRLCFEIGQHIRDKNLM<br>NSLITYLGCGHISEKNLSESSFLIFRVTKFSDIKDKIIPVFQENTLIGMKLE<br>DFEDWCKVAKLIEEKKHLTESGLDEIKKIKLNMNKG RVF                                                                                                                             |
| 97  | Engineered HE variant 327A10     | SRRESINPWILTGFADEAGCFSLVRNVNGHTTRYRTRLVFSIGLHYKDKSI<br>LENIQSTWKVGTISNSGDTQVRLTVTRFEDLKVIIDHFEKYPLITQKLG DYK<br>LFKQAFSVMENKGHLKENGIKELVRIKAKMNWGLTDELKKAPEVISRERPL<br>INKNIPNGKWLAGFTSGDGSFMVILAKSRTGARVQVRLCFEIGQHIRDKNLM<br>NSLITYLGCGHISEKNLSESSFLIFRVTKFSDIKDKIIPVFQENTLIGMKLE<br>DFEDWCKVAKLIEEKKHLTESGLDEIKKIKLNMNKG RVF                                                                                                                             |
| 98  | Engineered HE variant 327G01     | SRRESINPWILTGFADEAGCFVLVRNVNGHTTRYRTRLGFSISLHYKDKSI<br>LENIQSTWKIGTISNSGDTQVRLQVTRFEDLKVIIDHFEKYPLITQKLG DYK<br>LFKQAFSVMENKGHLKENGIKELVRIKAKMNWGLTDELKKAPEVISRERPL<br>INKNIPNGKWLAGFTSGDGSFMVILAKSRTGARVQVRLCFEIGQHIRDKNLM<br>NSLITYLGCGHISEKNLSESSFLIFRVTKFSDIKDKIIPVFQENTLIGMKLE<br>DFEDWCKVAKLIEEKKHLTESGLDEIKKIKLNMNKG RVF                                                                                                                             |
| 99  | Engineered HE variant 327H05     | SRRESINPWILTGFADEAGCFVLVRNVNGHTTRYRTRLFSIGLHYKDKSI<br>LENIQSTWKVGTISNSGDTGVRLQVTRFEDLKVIIDHFEKYPLITQKLG DYK<br>LFKQAFSVMENKGHLKENGIKELVRIKAKMNWGLTDELKKAPEVISRERPL<br>INKNIPNGKWLAGFTSGDGSFMVILAKSRTGARVQVRLCFEIGQHIRDKNLM<br>NSLITYLGCGHISEKNLSESSFLIFRVTKFSDIKDKIIPVFQENTLIGMKLE<br>DFEDWCKVAKLIEEKKHLTESGLDEIKKIKLNMNKG RVF                                                                                                                              |
| 100 | Engineered HE variant 327H05V26S | SRRESINPWILTGFADEAGCFSLVRNVNGHTTRYRTRLFSIGLHYKDKSI<br>LENIQSTWKVGTISNSGDTGVRLQVTRFEDLKVIIDHFEKYPLITQKLG DYK<br>LFKQAFSVMENKGHLKENGIKELVRIKAKMNWGLTDELKKAPEVISRERPL<br>INKNIPNGKWLAGFTSGDGSFMVILAKSRTGARVQVRLCFEIGQHIRDKNLM<br>NSLITYLGCGHISEKNLSESSFLIFRVTKFSDIKDKIIPVFQENTLIGMKLE<br>DFEDWCKVAKLIEEKKHLTESGLDEIKKIKLNMNKG RVF                                                                                                                              |
| 101 | Engineered HE variant 326H04     | SRRESINPWILTGFADEAGCFVLVRNVNGHTTRYRTRLFSIGLHYKDKSI<br>LENIQSTWKVGTISNSGDTQVRLTVTRFEDLKVIIDHFEKYPLITQKLG DYK<br>LFKQAFSVMENKGHLKENGIKELVRIKAKMNWGLTDELKKAPEVISRERPL<br>INKNIPNGKWLAGFTSGDGSFMVILAKSRTGARVQVRLCFEIGQHIRDKNLM<br>NSLITYLGCGHISEKNLSESSFLIFRVTKFSDIKDKIIPVFQENTLIGMKLE<br>DFEDWCKVAKLIEEKKHLTESGLDEIKKIKLNMNKG RVF                                                                                                                              |
| 102 | Engineered HE variant 370P03     | SRRESINPWILTGFADEAGCFSLVIRKVNNGHTTRYRTRLVFSIGLHNKDKSI<br>LENIQSTWKVGTISNSGDTQVRLTVTRFEDLKVIIDHFEKYPLITQKLG DYK<br>LFKQAFSVMENKGHLKENGIKELVRIKAKMNWGLTDELKKAPEVISRERPI<br>INKNIPNGKWLAGFTSGDGSFMVILAKSRTGARVQVRLCFEIGQHIRDKNLM<br>NSLITYLGCGHISEKNLSESSFLIFRVTKFSDIKDKIIPVFQENTLIGMKHE<br>DFEDWCKVAKLIEEKKHLTESGLDEIKKIKLNMNKG RVF                                                                                                                           |
| 103 | Engineered HE variant 370P07     | SRRESINPWTLTGFADEAGCFSLVIRSVNGHTTRYRTRLVFSIGLHNKDKSI<br>LENIQSTWKVGTISNSGDTQVRLTVTRFEDLKVIIDHFEKYPLITQKLG DYK<br>LFKQAFSVMENKEHLKENGIKELVRIKAKMNWGLTDELKKAPEVISRERPI<br>INKNIPNGKWLAGFTSGDGSFMVILAKSRTGARVQVRLCFEIGQHIRDKNLM<br>NSLIPYLGCGHISEKNLSESSFLIFRVTKFSDIKDKIIPVFQENTLIGMKHE<br>DFEDWCKVAKLIEEKKHLTESGLDEIKKIKLNMNKG RVF                                                                                                                            |
| 104 | Engineered HE variant 378N03     | SRRESINPWTLTGFADEAGCFVLVIRKVNNGHTTRYRTRLFSIGLHNKDKSI<br>LENIQSTWKVGTISNSGDTGVRLQVTRFEDLKVIIDHFEKYPLITQKLG DYK<br>LFKQAFSVMENKEHLKENGIKELVRIKAKMNWGLTDELKKAPEVISRERPL<br>INKNIPNGKWLAGFTSGDGSFMVILAKSRTGARVQVRLCFEIGQHIRDKNLM<br>NSLIPYLGCGHISEKNLSESSFLIFRVTKFSDINDKIIPVFQENTLIGMKHE<br>DFEDWCKVAKLIEEKKHLTESGLDEIKKIKLNMNKG RVF                                                                                                                            |
| 105 | megaTAL 206G09                   | VDLRTLGYSQQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPAALG<br>TVAVTYQHIITALPEATHEDIVGVGQWSGARALEALLTDAGELRGPPLQLD<br>TGQLVKIAKRGGVTAMEAVHASRNALTGAPLNLTPDQVVAIASNNGGKQALE<br>TVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPD<br>QVVAIASNIGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNIGGKQALETV<br>QRLLPVLCQDHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQDHGLTPDQV<br>VAIASNNGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNIGGKQALETVQV<br>LLPVLCQDHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQDHGLTPDQVVA |

[illegible]



|  |  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
|--|--|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|  |  | <p>             TGGTTGCTATCGCTAGTAATATCGGCGGAAAAACAAGCCCTGGAGACGGTGCA<br/>             GCGCTTGCTCCCAGTTCTGTGCCAGGATCATGGGCTGACACCCGATCAAGTG<br/>             GTTGCAATCGCCAGCAACAACGGTGGTAAACAGGCTCTTGAGACTGTTCAAC<br/>             GGCTGTTGCCCGTGCTGTGCCAGGACCACGGACTTACCCAGATCAAGTTGT<br/>             GGCTATCGCTTCCAATATTGGTGGCAAGCAAGCCCTCGAAACAGTGCAGCGG<br/>             CTTCTACCCGTCTTTGTGAGGATCATGGACTCACACCAGACAGGTGGTTCG<br/>             CTATTGCTAGCAACATCGGAGGGAAGCAGGCGTTGGAACTGTGCAAGGCT<br/>             CTGCCCCGTGTTGTGTCAAGACCATGGTCTGACTCCCGATCAAGTCGTGGCT<br/>             ATTGCTTCAAACATTGGTGGAAAAACAGGCCTTGGAAACCGTCCAAAGACTGC<br/>             TTCCCGTGCTTTGCCAAGACCACGGCTTGACCCAGATCAGGTCTGTGGCAAT<br/>             TGCATCCAATGGAGGCGGTAAACAAGCTCTGGAGACAGTACAGAGACTCCTG<br/>             CCCGTCTCTGTCAAGATCATGGATTGACCCCTGACCAGGTCTGTCGCAATTG<br/>             CCTCTAACATCGGCGGGAAAAACAAGCTTTGGAGACTGTGCAGAGACTTCTCCC<br/>             AGTGCTATGTCAAGATCACGGCTTACCCCGATCAGGTGGTCGCCATTGCA<br/>             AGCAATGGCGGTGGAAAGCAAGCTCTCGAAAGCATCGTGGCCCAGCTGAGCA<br/>             GACCCGACCCTGCCCTGGCCGCTCTGACCAACGACCACTCGTGGCCCTGGC<br/>             CTGCCTGGGTGGCAGACCAGCCATGGACGCCGTGAAGAAGGGCCTGCCTCAC<br/>             GCCCCGTGAGCTGATCAGAAGAGTGAACAGGAGAATCGGCAGAGAACAGCC<br/>             ACAGAGTGGCCATCTCTAGAGTGGGAGGAAGCTCCAGAAGAGAGCATCAA<br/>             TCCCTGGATCCTGACCGGCTTCGCCGACGCCGAGGGCTGCTTCAGCCTGGTG<br/>             ATCCAGAAGGCCAACGGCTGGACCACCAGATACAACCCAAGCTGGTGTTC<br/>             GCATCGGCTGCACAACAAGGACAAGAGCATCCTGGAGAATCCAGAGCAC<br/>             CTGGAAGGTGGGAACCATCAGCAAGTCTGGCGATACCCAAGTGAGGCTGACC<br/>             GTGACAAGATTGAGGACCTGAAGGTGATCATCGACCACTTCGAGAAGTACC<br/>             CTCTGATCAGCCAGAAGCTGGGCGACTACATGCTGTTCAAGCAGGCTTCTC<br/>             CGTGATGGAGAACAAAGGGCCACCTGAAGGAGAACGGCATCAAGGAGCTGGTG<br/>             AGAATCAAGGCCAAGATGAAGTGGGGTCTGACCGACGAGCTGAAGAAGGCCT<br/>             TTCCCGAGGACATCAGCAGAGAGAGGCCTATCATCAACAAGAATTCCTCAA<br/>             CGGCAAGTGGCTGGCCGATTACCAGCGGCGACGGCAGCTTCATGGTGATC<br/>             CTGGCCAAGAGCAGAACCAGGAGCTAGAGTGAGGTGAGGCTGTGCTTCGAGA<br/>             TCGGCCAGCACATTAGAGACAAGAACCTGATGAACAGCTGATCACCTACCT<br/>             GGGCTGCGGACACATCAGCGAGAAGAACCCTCCGAATCCAGCTTCCTGATC<br/>             TTCAGAGTGACCAAGTTCTCCGACATCAAGGACAAAATCATTCCCGTGTTC<br/>             AGGAGAACACCCTGATCGGCATGAAGCACGAGGACTTCGAGGACTGGTGCAA<br/>             GGTGGCCAAGCTGATCGAAGAGAAGAAGCACCTGACCGAGAGCGGCCTGGAC<br/>             GAGATCAAGAAGATCAAGCTGAACATGAACAAGGGCAGAGTGTTTTGATAAC<br/>             GTACGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA<br/>             AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA<br/>             AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA<br/>             AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA<br/>             AAAAA           </p> |
|--|--|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

**L. LIST OF EMBODIMENTS**

The invention is further described by the following non-limiting embodiments.

Embodiment 1: A polypeptide comprising an engineered I-OnuI homing endonuclease  
5 (HE) that binds and cleaves DNA at a target site within a double-stranded DNA (dsDNA)  
molecule, wherein the target site is within intron 1 of the human albumin (ALB) gene, wherein  
the I-OnuI HE cleaves both strands of the dsDNA molecule, and wherein the engineered I-  
OnuI HE comprises an amino acid sequence that is at least 90%, 91%, 92%, 93%, 94%, 95%,  
96%, 97%, 98%, or 99% identical to the amino acid sequence set forth in any one of SEQ ID  
10 NOs: 6-13.

Embodiment 2: The polypeptide according to embodiment 1, wherein the target site  
comprises a nucleic acid sequence as set forth in any one of SEQ ID NOs: 32-34, preferably  
the target site comprises a nucleic acid sequence as set forth in SEQ ID NO: 34.

15

Embodiment 3: The polypeptide according to embodiment 1 or 2, wherein the  
engineered I-OnuI HE comprises the following amino acid substitutions in relation to the  
numbering of SEQ ID NO: 4:

(a) S20C, L22S, R24V, R26Q, N27K, N28A, K30G, S31W, S32T, V33T, G34R,  
20 S36N, E38K, G40V, Q42S, T44G, V64T, A66S, N67K, N71T, A72Q, S74R, K76T, C176S,  
F178M, N180I, I182A, K185R, S186T, K187G, L188A, G189R, Q193R, V195C, S197E,  
T199G, Y219H, K221S, K225L, F228S, W230F, D232I, and V234R;

(b) S20C, L22S, R24T, N27K, N28F, K30G, S31R, S32T, V33T, G34R, S36T, E38K,  
G40M, Q42S, T44G, V64T, A66S, N67K, N71T, A72Q, S74R, K76T, C176S, F178M,  
25 N180I, I182A, K185R, S186T, K187G, L188A, G189R, Q193R, V195C, S197E, T199G,  
Y219H, K221S, K225L, F228S, W230F, D232I, and V234R;

(c) S20C, L22S, R24A, R26G, N27K, N28L, K30G, S31G, S32T, V33T, G34R,  
E38K, G40V, Q42S, T44G, V64T, A66S, N67K, N71T, A72Q, S74R, K76T, C176S, F178M,  
N180I, I182A, K185R, S186T, K187G, L188A, G189R, Q193R, V195C, S197E, T199G,  
30 Y219H, K221S, K225L, F228S, W230F, D232I, and V234R;

- (d) S20C, L22S, R24A, R26G, N27K, N28L, K30G, S31G, S32T, V33T, G34R, E38K, G40V, Q42S, T44G, V64T, A66S, N67K, N71T, A72Q, S74R, K76T, C176S, F178M, N180I, I182A, K185R, S186T, K187G, L188A, G189R, Q193R, V195C, S197E, T199G, Y219H, K221S, K225L, F228S, W230F, D232I, and V234R;
- 5 (e) S20C, L22S, R24V, R26A, N27K, N28G, K30G, S31G, S32T, V33T, G34R, S36T, E38K, G40V, Q42S, T44G, V64T, A66S, N67K, N71T, A72Q, S74R, K76T, C176S, F178M, N180I, I182A, K185R, S186T, K187G, L188A, G189R, Q193R, V195C, S197E, T199G, Y219H, K221S, K225L, F228S, W230F, D232I, and V234R;
- (f) S20C, L22S, R24A, R26G, N27K, N28C, K30G, S31C, S32T, V33T, G34R, 10 E38K, G40V, Q42S, T44G, V64T, A66S, N67K, N71T, A72Q, S74R, K76T, C176S, F178M, N180I, I182A, K185R, S186T, K187G, L188A, G189R, Q193R, V195C, S197E, T199G, Y219H, K221S, K225L, F228S, W230F, D232I, and V234R;
- (g) S20C, L22S, R24A, R26G, N27K, N28L, K30G, S31A, S32T, V33T, G34R, S36V, E38K, G40V, Q42S, T44G, V64T, A66S, N67K, N71T, A72Q, S74R, K76T, C176S, 15 F178M, N180I, I182A, K185R, S186T, K187G, L188A, G189R, Q193R, V195C, S197E, T199G, Y219H, K221S, K225L, F228S, W230F, D232I, and V234R; or
- (h) S20C, L22S, R24H, R26G, N27K, N28M, K30G, S31R, S32T, V33T, G34R, S36L, E38K, G40V, Q42S, T44G, V64T, A66S, N67K, N71T, A72Q, S74R, K76T, C176S, F178M, N180I, I182A, K185R, S186T, K187G, L188A, G189R, Q193R, V195C, S197E, 20 T199G, Y219H, K224S, K225L, F228S, W230F, D232I, and V234R.

Embodiment 4: The polypeptide according to any one of embodiments 1-3, wherein the engineered HE comprises the amino acid sequence set forth in SEQ ID NO: 6.

- 25 Embodiment 5: The polypeptide according to any one of embodiments 1-4, further comprising a nuclear localization signal (NLS) and/or a TALE DNA binding domain, preferably the TALE DNA binding domain comprises about 8.5 TALE repeat units to about 15.5 TALE repeat units and/or preferably the NLS comprises a nucleic acid sequence set forth in SEQ ID NO: 86.

30

Embodiment 6: The polypeptide according to embodiment 5, wherein the TALE DNA binding domain binds the polynucleotide sequence set forth in any one of SEQ ID NOs: 35-42, preferably the polypeptide binds and cleaves the polynucleotide sequence set forth in SEQ ID NO: 43.

5

Embodiment 7: The polypeptide according to any one of embodiments 1-6, wherein the polypeptide comprises an amino acid sequence that is at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to the amino acid sequence set forth in any one of SEQ ID NOs: 14-21, preferably the polypeptide comprises the amino acid sequence set forth in SEQ ID NO: 14.

10

Embodiment 8: The polypeptide according to any one of embodiments 1-7, further comprising a peptide linker and an end-processing enzyme or biologically active fragment thereof, preferably wherein the end-processing enzyme comprises Trex2, more preferably, the end-processing enzyme comprises an amino acid sequence as set forth in SEQ ID NO: 44.

15

Embodiment 9: A polynucleotide encoding the polypeptide according to any one of embodiments 1-8, preferably wherein the polynucleotide is an mRNA or a cDNA.

Embodiment 10: The polynucleotide according to embodiment 9, wherein the polynucleotide is an mRNA, wherein the mRNA comprises a nucleic acid sequence that is at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to the nucleic acid sequence set forth in any one of SEQ ID NOs: 22-29 and 108, preferably wherein the mRNA comprises a nucleic acid sequence set forth in SEQ ID NO: 22.

20

25

Embodiment 11: A pharmaceutical composition comprising a polynucleotide encoding the polypeptide according to any embodiment 1-8 or the polynucleotide according to embodiment 9 or 10, and a physiologically acceptable carrier, preferably the polynucleotide is an mRNA comprising a nucleic acid sequence set forth in SEQ ID NO:

22, and preferably wherein the pharmaceutically acceptable carrier comprises a lipid nanoparticle and the polynucleotide is encapsulated in said lipid nanoparticle.

Embodiment 12: The pharmaceutical composition according to embodiment 11  
5 further comprising a donor repair template, preferably an AAV vector particle comprises the donor repair template, and more preferably the donor repair template comprises a FVIII transgene or a FIX transgene.

Embodiment 13: A method of editing a human ALB gene in a cell comprising:  
10 introducing a polynucleotide encoding the polypeptide according to any embodiment 1-8 or the polynucleotide according to embodiment 9 or 10 into the cell, wherein the polynucleotide is translated in the cell to produce a polypeptide, wherein the polypeptide creates a double-strand break (DSB) at the target site; optionally further comprising introducing a donor repair template into the cell and wherein the donor repair template is  
15 incorporated into the human ALB gene by homology directed repair (HDR) at the site of DSB.

Embodiment 14: A method of treating a blood disorder, lysosomal storage disorder, or metabolic disorder, or condition associated therewith, comprising: introducing into the  
20 cell (a) a polynucleotide encoding the polypeptide according to any embodiment 1-8 or the polynucleotide according to embodiment 9 or 10, and (b) a donor repair template into a cell;

wherein the polynucleotide is translated in the cell to produce a polypeptide, wherein the polypeptide creates a double-strand break (DSB) at the target site and the  
25 donor repair template is incorporated into the human ALB gene by homology directed repair (HDR) at the site of the DSB,

preferably wherein the cell is a hepatocyte and the donor repair template comprises a FVIII transgene or a FIX transgene.

Embodiment 15: The pharmaceutical composition according to embodiment 12 for use in editing a human ALB gene in a cell or for use in treating a blood disorder, lysosomal storage disorder, or metabolic disorder, or condition associated therewith,

wherein the composition delivers the polynucleotide and the donor repair  
5 template into the cell,

wherein the polynucleotide is translated in the cell to produce a polypeptide,  
wherein the polypeptide creates a double-strand break (DSB) at the target site, and

wherein the donor repair template is incorporated into the human ALB gene  
by homology directed repair (HDR) at the site of the DSB,

10 preferably wherein the cell is a hepatocyte and the donor repair template is FVIII or FIX.

Embodiment 16: A polypeptide comprising an engineered homing endonuclease (HE) that cleaves a selected double strand DNA (dsDNA) target site in the human albumin  
15 (ALB) gene.

Embodiment 17: The polypeptide according to embodiment 16, wherein the dsDNA target site is within intron 1 of the ALB gene.

Embodiment 18: The polypeptide according to embodiment 16 or 17, wherein the dsDNA target site comprises a nucleic acid sequence as set forth in any one of SEQ ID  
20 NOs: 32-34.

Embodiment 19: The polypeptide according to embodiment 16 or 17, wherein the dsDNA target site comprises a nucleic acid sequence as set forth in SEQ ID NO: 34.

Embodiment 20: The polypeptide according to any one of embodiments 16-19, wherein the engineered HE is an LAGLIDADG homing endonuclease (LHE) variant.

25 Embodiment 21: The polypeptide according to any one of embodiments 16-20, wherein the engineered HE lacks the 1, 2, 3, 4, 5, 6, 7, or 8 N-terminal amino acids compared to a corresponding wild type or stabilized HE as set forth in SEQ ID NO: 1 or SEQ ID NO: 2.

Embodiment 22: The polypeptide according to any one of embodiments 16-21, wherein the engineered HE lacks the 4 N-terminal amino acids compared to a corresponding wild type or stabilized HE as set forth in SEQ ID NO: 1 or SEQ ID NO: 2.

Embodiment 23: The polypeptide according to any one of embodiments 16-22, wherein the engineered HE lacks the 8 N-terminal amino acids compared to a corresponding wild type or stabilized HE as set forth in SEQ ID NO: 1 or SEQ ID NO: 2.

Embodiment 24: The polypeptide according to any one of embodiments 16-23, wherein the engineered HE lacks the 1, 2, 3, 4, 5, or 6 C-terminal amino acids compared to a corresponding wild type or stabilized HE as set forth in SEQ ID NO: 1 or SEQ ID NO: 2.

Embodiment 25: The polypeptide according to any one of embodiments 16-24, wherein the engineered HE lacks the C-terminal amino acid compared to a corresponding wild type or stabilized HE as set forth in SEQ ID NO: 1 or SEQ ID NO: 2.

Embodiment 26: The polypeptide according to any one of embodiments 16-25, wherein the engineered HE lacks the 2 C-terminal amino acids compared to a corresponding wild type or stabilized HE as set forth in SEQ ID NO: 1 or SEQ ID NO: 2.

Embodiment 27: The polypeptide according to any one of embodiments 16-26, wherein the engineered HE is a variant of an LHE selected from the group consisting of: I-OnuI, I-AabMI, I-AaeMI, I-AniI, I-ApaMI, I-CapIII, I-CapIV, I-CkaMI, I-CpaMI, I-CpaMII, I-CpaMIII, I-CpaMIV, I-CpaMV, I-CpaV, I-CraMI, I-EjeMI, I-GpeMI, I-GpiI, I-GzeMI, I-GzeMII, I-GzeMIII, I-HjeMI, I-LtrII, I-LtrI, I-LtrWI, I-MpeMI, I-MveMI, I-NcrII, I-NcrI, I-NcrMI, I-OheMI, I-OsoMI, I-OsoMII, I-OsoMIII, I-OsoMIV, I-PanMI, I-PanMII, I-PanMIII, I-PnoMI, I-ScuMI, I-SmaMI, I-SscMI, and I-Vdi141I.

Embodiment 28: The polypeptide according to any one of embodiments 16-27, wherein the engineered HE is a variant of an LHE selected from the group consisting of: I-OnuI, I-CpaMI, I-HjeMI, I-PanMI, and SmaMI.

Embodiment 29: The polypeptide according to any one of embodiments 16-28, wherein the engineered HE is an I-OnuI LHE variant.

Embodiment 30: The polypeptide according to any one of embodiments 16-29, wherein the engineered HE comprises one or more amino acid substitutions in the DNA

recognition interface at corresponding amino acid positions selected from the group consisting of: 24, 26, 28, 30, 31, 32, 34, 35, 36, 37, 38, 40, 42, 44, 46, 48, 68, 70, 71, 75, 76, 78, 80, 180, 182, 184, 186, 189, 190, 191, 192, 193, 197, 199, 201, 203, 223, 225, 229, 232, 234, 236, and 238 of an I-OnuI LHE amino acid sequence as set forth in any one of  
 5 SEQ ID NOs: 1-3 and 5, or a biologically active fragment thereof.

Embodiment 31: The polypeptide according to any one of embodiments 16-30, wherein the engineered HE comprises at least 5, at least 15, preferably at least 25, more preferably at least 35, or even more preferably at least 40, or 43 of the corresponding following amino acid substitutions: S24C, L26S, R28A, R28H, R28T, R28V, R30A,  
 10 R30G, R30Q, N31K, N32A, N32F, N32L, N32G, N32C, N32M, K34G, S35W, S35R, S35G, S35K, S35C, S36T, V37T, G38R, S40N, S40T, S40S, S40G, S40V, S40L, E42K, G44V, G44M, Q46S, T48G, V68T, A70S, N71K, N75T, A76Q, S78R, K80T, C180S, F182M, N184I, I186A, K189R, S190T, K191G, L192A, G193R, Q197R, V199C, S201E, T203G, Y223H, K225S, K229L, F232S, W234F, D236I, and V238R of any one of SEQ ID  
 15 NOs: 1-3 and 5, or a biologically active fragment thereof

Embodiment 32: The polypeptide according to any one of embodiments 16-31, wherein the engineered HE comprises the corresponding following amino acid substitutions: S24C, L26S, R28V, R30Q, N31K, N32A, K34G, S35W, S36T, V37T, G38R, S40N, E42K, G44V, Q46S, T48G, V68T, A70S, N71K, N75T, A76Q, S78R, K80T,  
 20 C180S, F182M, N184I, I186A, K189R, S190T, K191G, L192A, G193R, Q197R, V199C, S201E, T203G, Y223H, K225S, K229L, F232S, W234F, D236I, and V238R of any one of SEQ ID NOs: 1-3 and 5, or a biologically active fragment thereof

Embodiment 33: The polypeptide according to any one of embodiments 16-31, wherein the engineered HE comprises the corresponding following amino acid  
 25 substitutions: S24C, L26S, R28T, N31K, N32F, K34G, S35R, S36T, V37T, G38R, S40T, E42K, G44M, Q46S, T48G, V68T, A70S, N71K, N75T, A76Q, S78R, K80T, C180S, F182M, N184I, I186A, K189R, S190T, K191G, L192A, G193R, Q197R, V199C, S201E, T203G, Y223H, K225S, K229L, F232S, W234F, D236I, and V238R of any one of SEQ ID NOs: 1-3 and 5, or a biologically active fragment thereof.



Embodiment 34: The polypeptide according to any one of embodiments 16-31, wherein the engineered HE comprises the corresponding following amino acid substitutions: S24C, L26S, R28A, R30G, N31K, N32L, K34G, S35G, S36T, V37T, G38R, E42K, G44V, Q46S, T48G, V68T, A70S, N71K, N75T, A76Q, S78R, K80T, C180S, 5 F182M, N184I, I186A, K189R, S190T, K191G, L192A, G193R, Q197R, V199C, S201E, T203G, Y223H, K225S, K229L, F232S, W234F, D236I, and V238R of any one of SEQ ID NOs: 1-3 and 5, or a biologically active fragment thereof.

Embodiment 35: The polypeptide according to any one of embodiments 16-31, wherein the engineered HE comprises the corresponding following amino acid 10 substitutions: S24C, L26S, R28H, R30A, N31K, N32F, K34G, S35K, S36T, V37T, G38R, S40G, E42K, G44V, Q46S, T48G, V68T, A70S, N71K, N75T, A76Q, S78R, K80T, C180S, F182M, N184I, I186A, K189R, S190T, K191G, L192A, G193R, Q197R, V199C, S201E, T203G, Y223H, K225S, K229L, F232S, W234F, D236I, and V238R of any one of SEQ ID NOs: 1-3 and 5, or a biologically active fragment thereof.

Embodiment 36: The polypeptide according to any one of embodiments 16-31, wherein the engineered HE comprises the corresponding following amino acid substitutions: S24C, L26S, R28V, R30A, N31K, N32G, K34G, S35G, S36T, V37T, 15 G38R, S40T, E42K, G44V, Q46S, T48G, V68T, A70S, N71K, N75T, A76Q, S78R, K80T, C180S, F182M, N184I, I186A, K189R, S190T, K191G, L192A, G193R, Q197R, V199C, S201E, T203G, Y223H, K225S, K229L, F232S, W234F, D236I, and V238R of any one of SEQ ID NOs: 1-3 and 5, or a biologically active fragment thereof.

Embodiment 37: The polypeptide according to any one of embodiments 16-31, wherein the engineered HE comprises the corresponding following amino acid substitutions: S24C, L26S, R28A, R30G, N31K, N32C, K34G, S35C, S36T, V37T, G38R, 25 E42K, G44V, Q46S, T48G, V68T, A70S, N71K, N75T, A76Q, S78R, K80T, C180S, F182M, N184I, I186A, K189R, S190T, K191G, L192A, G193R, Q197R, V199C, S201E, T203G, Y223H, K225S, K229L, F232S, W234F, D236I, and V238R of any one of SEQ ID NOs: 1-3 and 5, or a biologically active fragment thereof.

Embodiment 38: The polypeptide according to any one of embodiments 16-31, wherein the engineered HE comprises the corresponding following amino acid substitutions: S24C, L26S, R28A, R30G, N31K, N32L, K34G, S35A, S36T, V37T, G38R, S40V, E42K, G44V, Q46S, T48G, V68T, A70S, N71K, N75T, A76Q, S78R, K80T, 5 C180S, F182M, N184I, I186A, K189R, S190T, K191G, L192A, G193R, Q197R, V199C, S201E, T203G, Y223H, K225S, K229L, F232S, W234F, D236I, and V238R of any one of SEQ ID NOs: 1-3 and 5, or a biologically active fragment thereof.

Embodiment 39: The polypeptide according to any one of embodiments 16-31, wherein the engineered HE comprises following corresponding amino acid substitutions: 10 S24C, L26S, R28H, R30G, N31K, N32M, K34G, S35R, S36T, V37T, G38R, S40L, E42K, G44V, Q46S, T48G, V68T, A70S, N71K, N75T, A76Q, S78R, K80T, C180S, F182M, N184I, I186A, K189R, S190T, K191G, L192A, G193R, Q197R, V199C, S201E, T203G, Y223H, K225S, K229L, F232S, W234F, D236I, and V238R of any one of SEQ ID NOs: 1-3 and 5, or a biologically active fragment thereof.

15 Embodiment 40: The polypeptide according to any one of embodiments 16-39, wherein the engineered HE further comprises the following corresponding amino acid substitutions C115S, E121G, I125E, L138M, I153D, K156R, S159P, L160I, F168G, E178D, K207R, N246K, V261M, and L263H of SEQ ID NO: 1, or a biologically active fragment thereof.

20 Embodiment 41: The polypeptide according to any one of embodiments 16-40, wherein the engineered HE further comprises: (a) a first DNA recognition interface comprising amino acid residues 20-46 of any one of SEQ ID NOs: 6-13, (b) a second DNA recognition interface comprising amino acid residues 64-78 of any one of SEQ ID NOs: 6-13, (c) a third DNA recognition interface comprising amino acid residues 176-199 of any 25 one of SEQ ID NOs: 6-13, and (d) a fourth DNA recognition interface comprising amino acid residues 219-236 of any one of SEQ ID NOs: 6-13.

Embodiment 42: The polypeptide according to any one of embodiments 16-41, wherein the engineered HE comprises an amino acid sequence that is at least 90%, 91%,

92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to the amino acid sequence set forth in any one of SEQ ID NOs: 6-13, or a biologically active fragment thereof.

Embodiment 43: The polypeptide according to any one of embodiments 16-42, wherein the engineered HE comprises an amino acid sequence that is at least 90% identical  
5 to the amino acid sequence set forth in any one of SEQ ID NOs: 6-13, or a biologically active fragment thereof.

Embodiment 44: The polypeptide according to any one of embodiments 16-42, wherein the engineered HE comprises an amino acid sequence that is at least 93% identical  
10 to the amino acid sequence set forth in any one of SEQ ID NOs: 6-13, or a biologically active fragment thereof.

Embodiment 45: The polypeptide according to any one of embodiments 16-42, wherein the engineered HE comprises an amino acid sequence that is at least 95% identical  
to the amino acid sequence set forth in any one of SEQ ID NOs: 6-13, or a biologically active fragment thereof.

15 Embodiment 46: The polypeptide according to any one of embodiments 16-42, wherein the engineered HE comprises an amino acid sequence that is at least 96% identical to the amino acid sequence set forth in any one of SEQ ID NOs: 6-13, or a biologically active fragment thereof.

Embodiment 47: The polypeptide according to any one of embodiments 16-42,  
20 wherein the engineered HE comprises an amino acid sequence that is at least 97% identical to the amino acid sequence set forth in any one of SEQ ID NOs: 6-13, or a biologically active fragment thereof.

Embodiment 48: The polypeptide according to any one of embodiments 16-42, wherein the engineered HE comprises an amino acid sequence that is at least 98% identical  
25 to the amino acid sequence set forth in any one of SEQ ID NOs: 6-13, or a biologically active fragment thereof.

Embodiment 49: The polypeptide according to any one of embodiments 16-42, wherein the engineered HE comprises an amino acid sequence that is at least 99% identical

to the amino acid sequence set forth in any one of SEQ ID NOs: 6-13, or a biologically active fragment thereof.

Embodiment 50: The polypeptide according to any one of embodiments 16-42, wherein the engineered HE comprises the amino acid sequence set forth in SEQ ID NO: 6,  
5 or a biologically active fragment thereof.

Embodiment 51: The polypeptide according to any one of embodiments 16-42, wherein the engineered HE comprises the amino acid sequence set forth in SEQ ID NO: 7, or a biologically active fragment thereof.

Embodiment 52: The polypeptide according to any one of embodiments 16-42,  
10 wherein the engineered HE comprises the amino acid sequence set forth in SEQ ID NO: 8, or a biologically active fragment thereof.

Embodiment 53: The polypeptide according to any one of embodiments 16-42, wherein the engineered HE comprises the amino acid sequence set forth in SEQ ID NO: 9, or a biologically active fragment thereof.

Embodiment 54: The polypeptide according to any one of embodiments 16-42, wherein the engineered HE comprises the amino acid sequence set forth in SEQ ID NO:  
15 10, or a biologically active fragment thereof.

Embodiment 55: The polypeptide according to any one of embodiments 16-42, wherein the engineered HE comprises the amino acid sequence set forth in SEQ ID NO:  
20 11, or a biologically active fragment thereof.

Embodiment 56: The polypeptide according to any one of embodiments 16-42, wherein the engineered HE comprises the amino acid sequence set forth in SEQ ID NO:  
12, or a biologically active fragment thereof.

Embodiment 57: The polypeptide according to any one of embodiments 16-42,  
25 wherein the engineered HE comprises the amino acid sequence set forth in SEQ ID NO: 13, or a biologically active fragment thereof.

Embodiment 58: The polypeptide according to any one of embodiments 16-57, further comprising a nuclear localization signal (NLS).

Embodiment 59: The polypeptide according to embodiment 58, wherein the NLS is located N-terminal to the engineered HE.

Embodiment 60: The polypeptide according to any one of embodiments 16-59, further comprising a DNA binding domain

5        Embodiment 61: The polypeptide according to embodiment 60, wherein the DNA binding domain is selected from the group consisting of: a TALE DNA binding domain and a zinc finger DNA binding domain.

Embodiment 62: The polypeptide according to embodiment 61, wherein the DNA binding domain is a TALE DNA binding domain.

10        Embodiment 63: The polypeptide according to embodiment 61 or 62, wherein the TALE DNA binding domain comprises about 8.5 TALE repeat units to about 15.5 TALE repeat units.

Embodiment 64: The polypeptide according to any one of embodiments 61-63, wherein the TALE DNA binding domain comprises about 11.5 TALE repeat units.

15        Embodiment 65: The polypeptide according to any one of embodiments 61-63, wherein the TALE DNA binding domain comprises 10.5 TALE repeat units.

Embodiment 66: The polypeptide according to any one of embodiments 61-65, wherein the TALE DNA binding domain binds the polynucleotide sequence set forth in any one of SEQ ID NOs: 35-42

20        Embodiment 67: The polypeptide according to any one of embodiments 61-65, wherein the TALE DNA binding domain binds the polynucleotide sequence set forth in SEQ ID NO: 38.

Embodiment 68: The polypeptide according to any one of embodiments 16-67, wherein the polypeptide binds and cleaves the polynucleotide sequence set forth in SEQ ID  
25    NO: 43.

Embodiment 69: The polypeptide according to embodiment 61, wherein the DNA binding domain is a zinc finger DNA binding domain.

Embodiment 70: The polypeptide according to embodiment 69, wherein the zinc finger DNA binding domain comprises 2, 3, 4, 5, 6, 7, or 8 zinc finger motifs.

Embodiment 71: The polypeptide according to any one of embodiments 16-70, wherein the polypeptide comprises an amino acid sequence that is at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to the amino acid sequence set forth in any one of SEQ ID NOs: 14-21, or a biologically active fragment thereof.

5           Embodiment 72: The polypeptide according to any one of embodiments 16-71, wherein the polypeptide comprises an amino acid sequence set forth in any one of SEQ ID NOs: 14-21, or a biologically active fragment thereof.

          Embodiment 73: The polypeptide according to embodiment 72, wherein the polypeptide comprises the amino acid sequence set forth in SEQ ID NO: 14, or a  
10   biologically active fragment thereof.

          Embodiment 74: The polypeptide according to embodiment 72, wherein the polypeptide comprises the amino acid sequence set forth in SEQ ID NO: 15, or a biologically active fragment thereof.

          Embodiment 75: The polypeptide according to embodiment 72, wherein the  
15   polypeptide comprises the amino acid sequence set forth in SEQ ID NO: 16, or a biologically active fragment thereof.

          Embodiment 76: The polypeptide according to embodiment 72, wherein the polypeptide comprises the amino acid sequence set forth in SEQ ID NO: 17, or a biologically active fragment thereof.

20           Embodiment 77: The polypeptide according to embodiment 72, wherein the polypeptide comprises the amino acid sequence set forth in SEQ ID NO: 18, or a biologically active fragment thereof.

          Embodiment 78: The polypeptide according to embodiment 72, wherein the polypeptide comprises the amino acid sequence set forth in SEQ ID NO: 19, or a  
25   biologically active fragment thereof.

          Embodiment 79: The polypeptide according to embodiment 72, wherein the polypeptide comprises the amino acid sequence set forth in SEQ ID NO: 20, or a biologically active fragment thereof.

Embodiment 80: The polypeptide according to embodiment 72, wherein the polypeptide comprises the amino acid sequence set forth in SEQ ID NO: 21, or a biologically active fragment thereof.

Embodiment 81: The polypeptide according to any one of embodiments 16-80,  
5 further comprising a peptide linker and an end-processing enzyme or biologically active fragment thereof.

Embodiment 82: The polypeptide according to any one of embodiments 16-81, further comprising a viral self-cleaving 2A peptide and an end-processing enzyme or biologically active fragment thereof.

10 Embodiment 83: The polypeptide according to embodiment 81 or 82, wherein the end-processing enzyme or biologically active fragment thereof has 5'-3' exonuclease, 5'-3' alkaline exonuclease, 3'-5' exonuclease, 5' flap endonuclease, helicase, TdT, or template-independent DNA polymerase activity.

Embodiment 84: The polypeptide according to any one of embodiments 81-83,  
15 wherein the end-processing enzyme comprises Trex2 or a biologically active fragment thereof.

Embodiment 85: The polypeptide according to embodiment 84, wherein the end-processing enzyme comprises an amino acid sequence as set forth in SEQ ID NO: 44.

Embodiment 86: A polynucleotide encoding the polypeptide according to any one  
20 of embodiments 16-85.

Embodiment 87: An mRNA encoding the polypeptide according to any one of embodiments 16-85.

Embodiment 88: The mRNA according to embodiment 87, wherein the mRNA comprises a nucleic acid sequence that is at least 90%, 91%, 92%, 93%, 94%, 95%, 96%,  
25 97%, 98%, or 99% identical to the nucleic acid sequence set forth in any one of SEQ ID NOs: 22-29.

Embodiment 89: The mRNA according to embodiment 87, wherein the mRNA comprises a nucleic acid sequence set forth in any one of SEQ ID NOs: 22-29.

Embodiment 90: A cDNA encoding the polypeptide according to any one of embodiments 16-85.

Embodiment 91: A vector comprising a polynucleotide encoding the polypeptide according to any one of embodiments 16-85.

5        Embodiment 92: A cell comprising the polypeptide according to any one of embodiments 16-85.

Embodiment 93: A cell comprising a polynucleotide encoding the polypeptide according to any one of embodiments 16-85.

Embodiment 94: A cell comprising the polynucleotide according to embodiment  
10    86.

Embodiment 95: A cell comprising the mRNA of any one of embodiments 87-89.

Embodiment 96: A cell comprising the cDNA of embodiment 90.

Embodiment 97: A cell comprising the vector of embodiment 91.

Embodiment 98: A cell comprising one or more genome modifications introduced  
15    by the polypeptide according to any one of embodiments 16-85.

Embodiment 99: The cell according to any one of embodiments 92-97, wherein the polynucleotide, mRNA, cDNA, or vector further comprises a heterologous polyadenylation signal.

Embodiment 100: The cell according to any one of embodiments 92-99, wherein a  
20    donor repair template comprising a polynucleotide encoding a therapeutic polypeptide or protein is integrated into the ALB gene at a DNA double stranded break site introduced by the polypeptide according to any one of embodiments 16-85.

Embodiment 101: The cell according to embodiment 100, wherein the therapeutic polypeptide is a therapeutic antihemophilic factor, antibody, protein, cytokine, chemokine,  
25    cytotoxin, cytokine receptor, hormone, or functional variants thereof.

Embodiment 102: The cell according to embodiment 100 or 101, wherein the therapeutic polypeptide is Factor VII (FVII), Factor VIII (FVIII), Factor IX (FIX), Factor X (FX), Factor XI (FXI), or functional variants thereof.



Embodiment 103: The cell according to any one of embodiments 100-102, wherein the therapeutic polypeptide is FVIII, FIX, or functional variant thereof

Embodiment 104: The cell according to any one of embodiments 100-103, wherein the therapeutic polypeptide is a FVIII or functional variant thereof.

5        Embodiment 105: The cell according to any one of embodiments 100-104, wherein the therapeutic polypeptide is a modified FVIII or functional variant thereof.

Embodiment 106: The cell according to any one of embodiments 100-105, wherein the therapeutic polypeptide comprises a modified FVIII comprising a shortened B domain.

10        Embodiment 107: The cell according to any one of embodiments 100-105, wherein the therapeutic polypeptide comprises a B domain deleted FVIII (FVIII-BDD).

Embodiment 108: The cell according to any one of embodiments 92-107, wherein the cell is a hepatocyte.

Embodiment 109: The cell according to any one of embodiments 92-108, wherein the cell comprises one or more modified ALB alleles.

15        Embodiment 110: A plurality of cells comprising one or more cells of any one of embodiments 92-109.

Embodiment 111: A composition comprising one or more cells according to any one of embodiments 92-110.

20        Embodiment 112: A composition comprising one or more cells according to any one of embodiments 92-110 and a physiologically acceptable carrier.

Embodiment 113: A composition comprising polynucleotides encoding the polypeptide of any one of embodiments 16-85.

Embodiment 114: A composition comprising polynucleotides encoding the polypeptide of any one of embodiments 16-85 and a physiologically acceptable carrier.

25        Embodiment 115: The composition according to embodiment 114, wherein the pharmaceutically acceptable carrier is a lipid nanoparticle and wherein the polynucleotide is encapsulated in said lipid nanoparticle.

Embodiment 116: A method of editing a human ALB gene in a cell comprising: introducing a polynucleotide encoding the polypeptide of any one of embodiments 16-85

into the cell, wherein expression of the polypeptide creates a double-strand break (DSB) at the target site.

Embodiment 117: A method of editing a human ALB gene in a cell comprising: introducing a polynucleotide encoding the polypeptide of any one of embodiments 16-85  
5 and a donor repair template into the cell, wherein expression of the polypeptide creates a double-strand break (DSB) at the target site and the donor repair template is incorporated into the human ALB gene by homology directed repair (HDR) at the site of the DSB.

Embodiment 118: A method of treating a blood disorder, lysosomal storage disorder, or metabolic disorder, or condition associated therewith, comprising: introducing  
10 a polynucleotide encoding the polypeptide of any one of embodiments 16-85 and a donor repair template into a cell, wherein expression of the polypeptide creates a double-strand break (DSB) at the target site and the donor repair template is incorporated into the human ALB gene by homology directed repair (HDR) at the site of the DSB.

Embodiment 119: The method according to any one of embodiments 116-118,  
15 wherein the cell is a hepatocyte.

Embodiment 120: A method of treating a blood disorder, or condition associated therewith, comprising: introducing a polynucleotide encoding the polypeptide of any one of embodiments 16-85 and a donor repair template into a hepatocyte, wherein expression of the polypeptide creates a double-strand break (DSB) at the target site and the donor repair  
20 template is incorporated into the human ALB gene by homology directed repair (HDR) at the site of the DSB.

Embodiment 121: A method of treating a blood disorder, or condition associated therewith, comprising: introducing a polynucleotide encoding the polypeptide of any one of embodiments 16-85 and a donor repair template into a hepatocyte, wherein expression of the polypeptide creates a double-strand break (DSB) at the target site and the donor repair  
25 template is incorporated into the human ALB gene by Non-homologous end-joining (NHEJ) at the site of the DSB.

Embodiment 122: The method according to any one of embodiments 116-121, wherein the polynucleotide is an mRNA.

Embodiment 123: The method according to embodiment 122, wherein the mRNA comprises a nucleic acid sequence set forth in any one of SEQ ID NOs: 14-21.

Embodiment 124: The method according to any one of embodiments 116-121, wherein the polynucleotide is an cDNA.

5        Embodiment 125: The method according to any one of embodiments 116-121, wherein the polynucleotide is a vector.

Embodiment 126: The method according to any one of embodiments 116-125, wherein the polynucleotide, mRNA, cDNA, or vector further comprises a heterologous polyadenylation signal.

10        Embodiment 127: The method according to any one of embodiments 116-126, wherein the polynucleotide, mRNA, cDNA, or vector further comprises a nuclear localization signal (NLS).

Embodiment 128: The method according to any one of embodiments 116-127, wherein a polynucleotide encoding a 5'-3' exonuclease is introduced into the cell.

15        Embodiment 129: The method according to any one of embodiments 116-128, wherein a polynucleotide encoding Trex2 or a biologically active fragment thereof is introduced into the cell.

Embodiment 130: The method according to any one of embodiments 125-129, wherein a donor repair template comprising a polynucleotide encoding a therapeutic polypeptide or protein is integrated into the ALB gene at the DSB introduced by the polypeptide according to any one of embodiments 16-85.

Embodiment 131: The method according to embodiment 130, wherein the therapeutic polypeptide is a therapeutic antihemophilic factor, antibody, protein, cytokine, chemokine, cytotoxin, cytokine receptor, hormone, or functional variants thereof.

25        Embodiment 132: The method according to embodiment 130 or 131, wherein the therapeutic polypeptide is a Factor VIII (FVIII), Factor IX (FIX), Factor X (FX), Factor XI (FXI), or functional variants thereof.

Embodiment 133: The method according to any one of embodiments 130-132, wherein the therapeutic polypeptide is a FVIII or FIX or functional variant thereof.

Embodiment 134: The method according to any one of embodiments 130-133, wherein the therapeutic polypeptide is a FVIII or functional variant thereof.

Embodiment 135: The method according to any one of embodiments 130-134, wherein the therapeutic polypeptide is a modified FVIII.

5        Embodiment 136: The method according to any one of embodiments 117-135, wherein the donor repair template further comprises a heterologous polyadenylation signal.

Embodiment 137: The method according to any one of embodiments 117-136, wherein the donor repair template comprises a 5' homology arm homologous to a human ALB gene sequence 5' of the DSB and a 3' homology arm homologous to a human ALB  
10    gene sequence 3' of the DSB.

Embodiment 138: The method according to any one of embodiments 116-137, wherein a viral vector is used to introduce the donor repair template into the cell.

Embodiment 139: The method according to embodiment 138, wherein the viral vector is a recombinant adeno-associated viral vector (rAAV) or a retrovirus.

15        Embodiment 140: The method according to embodiment 139, wherein the viral vector is a recombinant adeno-associated viral vector (rAAV).

Embodiment 141: The method according to embodiment 140, wherein the rAAV has one or more ITRs from an AAV serotype selected from the group consisting of: AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, and AAV10.

20        Embodiment 142: The method according to any one of embodiments 139-141, wherein the rAAV comprises a capsid from an AAV serotype selected from the group consisting of: AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, and AAV10.

Embodiment 143: The method according to embodiment 139, wherein the  
25    retrovirus is a lentivirus.

Embodiment 144: The method according to embodiment 143, wherein the lentivirus is an integrase deficient lentivirus (IDLV).

Embodiment 145: The method according to any one of embodiments 118-144, wherein the blood disorder is a Hemophilia or Hemoglobinopathy.

Embodiment 146: The method according to any one of embodiments 118-145, wherein the blood disorder is a Hemophilia.

Embodiment 147: The method according to embodiment 145 or 146, wherein the Hemophilia is Hemophilia A, Hemophilia B, or Hemophilia C.

5        Embodiment 148: The method according to any one of embodiments 145-147, wherein the Hemophilia is Hemophilia A.

Embodiment 149: The method according to any one of embodiments 145-147, wherein the Hemophilia is Hemophilia B.

Embodiment 150: The method according to any one of embodiments 145-147,  
10        wherein the Hemophilia is Hemophilia C.

All publications, patent applications, and issued patents cited in this specification are herein incorporated by reference as if each individual publication, patent application, or issued  
15        patent were specifically and individually indicated to be incorporated by reference.

Although the foregoing embodiments have been described in some detail by way of illustration and example for purposes of clarity of understanding, it will be readily apparent to one of ordinary skill in the art in light of the teachings contemplated herein that certain changes and modifications may be made thereto without departing from the spirit or scope of the  
20        appended claims. The following examples are provided by way of illustration only and not by way of limitation. Those of skill in the art will readily recognize a variety of noncritical parameters that could be changed or modified to yield essentially similar results.

## EXAMPLES

EXAMPLE 1TARGET SITE SELECTION AND REPROGRAMING OF MEGANUCLEASES TO DISRUPT THE  
ALBUMIN GENE

5

The sequence of human *ALB* intron 1 was screened to identify amenable regions to meganuclease engineering, specifically putative targets which would accommodate previously defined restrictions in the central 4 base-pair (bp) region of the target. Of the 687 possible 22 bp sequence target sites (22mers) in the human *ALB* intron 1, 186 met the central 4 sequence requirements and were considered target candidates. The 186 candidate sites were then scored for engineering probability using a database of homing endonuclease cleavage data. The top-ranking sites were manually assessed for uniqueness and presence of undesirable nucleotide sequences. In total, 12 candidate target sites were chosen for targeting.

Next, the meganuclease I-OnuI was reprogrammed/engineered to target the selected sites in intron 1 of the human *ALB* gene (**FIG. 1**) as described in Jarjour *et al.*, 25 2009. *Nuc. Acids Res.* 37(20): 6871-6880 and patent publication WO2020072059. In brief, variant libraries against all 12 candidate target sites were constructed spanning either the N- or C-terminal I-OnuI DNA recognition domain and transformed into *S. cerevisiae* with each library containing  $\sim 10^7$  to  $10^8$  unique transformants. The resulting libraries were displayed on the surface of yeast and screened by flow cytometry for cleavage activity against target sites comprising the corresponding domains "half-sites".

Yeast displaying the reprogrammed N- and C-terminal domains of I-OnuI were purified and the plasmid DNA was extracted. PCR reactions were performed to amplify the reprogrammed domains which were subsequently fused and transformed into *S. cerevisiae* to create a library of reprogrammed domain combinations. Out of the twelve initial candidate target sites in the hAlb intronic sequence, only libraries directed against 3 candidate target sites (S01, S03 and S06; SEQ ID NOs: 32-34) yielded yeast populations capable of cleaving their respective 22 bp on-targets.

To identify the most active variants from each library, the Lead Identification mRNA Assay (LIMA) was run. In brief, mRNA was made from unique variants from each library (S01 target site library = 56 clones; S03 target site library = 70 clones; S06 target site library = 88 clones) and electroporated into K562 cells at a single dose of 200ng/uL with the addition of 125ng/uL TREX2 mRNA. gDNA was harvested 72 hours post electroporation and used as template for next generation sequencing to elucidate indel (insertion and deletion) frequencies at the respective target sites for the reprogrammed meganucleases (**FIGs 2A - 2C**). Given the totality of the data, we chose to move forward with enzymes targeting site S06 (SEQ ID NO: 34) (**FIG. 2C**, SEQ ID NOs: 87-93). Next, several TALE arrays of different lengths were fused to the selected meganucleases targeting S06 to create megaTALs and tested in the LIMA assay to identify the optimal TALE for future megaTAL construction (**FIG. 3A**). The selected TALE array has 11.5 TALE repeats and targets the 11.5+2 sequence as shown in **FIG. 3B** (SEQ ID NO: 38).

MegaTALs for the most active meganucleases identified in **FIG. 2C** were constructed and assessed for off-target editing. The number of putative off-targets discovered and found to meet statistical significance (FDR corrected p-value < 0.05) for the selected megaTALs are shown in **FIG. 4**. Variant 206G09 (SEQ ID NO: 105) was found to have the fewest putative off-targets and was selected as the megaTAL for additional engineering efforts to improve both on-target cleavage activity and specificity.

In an effort to increase the on-target activity of the 206G09 megaTAL, error-prone PCR was performed using its DNA sequence as a template and the resulting library was re-cloned into the yeast display vector. The library was displayed on the surface of yeast and screened by flow cytometry for cleavage activity and binding affinity against the on-target site as described above. LIMA assays were performed as previously described in this Example, but in the absence of TREX2 mRNA, to identify Gen1 megaTAL sequences with improved on-target activity (**FIG. 5**). For reference, the parent megaTAL (206G09) (SEQ ID NO: 105) was also run in the assay (**FIG. 5**; last bar and dotted line). Off-target discovery was not performed on the Gen1 clones.

To better understand the on-target cleavage potency, a full mRNA titration for the top Gen1 clones (264B08, 264B08+B09, and 264F08) (SEQ ID NOs: 94-96) was electroporated into K562s and assessed for on-target indel formation using next generation sequencing (**FIG. 6**). Indel50s, the mRNA concentration where 50% indels are observed, were calculated and reported in the table associated with **FIG. 6**. Clone 264F08 was selected for an additional round of engineering to improve on-target cleavage activity and specificity.

## EXAMPLE 2

### OPTIMIZING MEGANUCLEASE TARGET SITE ACTIVITY AND SPECIFICITY

To improve specificity and on-target activity of the Gen1 enzymes, a fourth round of engineering was undertaken using the enzyme 264F08 (SEQ ID NO: 96) as the parent sequence. Using the off-target data generated from the enzymes in Example 1, six regions within the megaTAL were selected to introduce amino acid variations. Using NNS degenerate codons where N indicates any nucleotide and S indicates either a guanine or cytosine and assembly PCR, six libraries were constructed to cover these regions and cloned into the yeast-display vector system. The resulting yeast-display library was screened by flow cytometry for cleavage activity against the on-target site as well as for loss of cleavage activity against previously identified off-targets. As in Example 1, 93 unique clones identified in the yeast library screening were subjected to the LIMA assay (Gen2 clones). In this iteration of the LIMA assay, both activity against the on-target and select off-targets were monitored, allowing for selection of the most active and most specific enzymes.

Similar to Example 1, a full mRNA titration for the top Gen2 clones identified in the LIMA assay (327A10, 327G01, 327H05, 327H05.V26S and 326H04) (SEQ ID NOs: 97-101) were electroporated into K562s and assessed for on-target indel formation using next generation sequencing. Indel50s were calculated and reported in the table associated with **FIG. 7**.

Next, we examined the off-target profiles of the Gen2 clones. For off-target discovery, we utilized both experimental and computational approaches. Our experimental approach involved performing an oligo-capture assay (as in Example 1) on all enzymes in an unbiased



manner to identify putative off-targets. Our *in silico*, computational approach looked for sites in the human genome that were either perfect matches to the TALE-array target site (SEQ ID NO:38) or had equal to or less than 4 base pair mismatches to the meganuclease target site (SEQ ID NO: 34). To verify if the putative off-targets from the oligo-capture assay and *in*  
5 *silico* predictions were true off-targets we harvested gDNA from K562 cells that were electroporated with either megaTAL mRNA or buffer and performed a multiplexed PCR reaction (rhAMPseq: IDT) across all putative off-target sites, sequenced via next generation sequencing and monitored indel rates for treated samples compared to control samples (**FIG. 8**). Those sites that had an FDR corrected p-value <0.05 were considered verified off targets  
10 (black circles) all other sites were considered unverified (grey circles).

### EXAMPLE 3

#### FURTHER OPTIMIZING MEGANUCLEASE ACTIVITY

Another round of engineering was undertaken with Gen2 enzyme 327A10 (SEQ ID NO: 97) as the parent enzyme. In an effort to improve on-target cleavage activity, amino acid  
15 positions were identified in the meganuclease to randomize, as in Example 2. Using oligonucleotides with selected degenerate codons at positions in the scaffold of the homing endonuclease known to affect protein stability, a library of variants was constructed and displayed on the surface of yeast. The resulting yeast-display library was screened by flow cytometry for cleavage activity and binding affinity against the on-target site. As in Example  
20 2, 296 unique clones identified in the yeast library screening (Gen3 clones) were subjected to the LIMA assay, allowing for selection of the most active and most specific enzymes.

A full mRNA titration for the top Gen3 clones identified in the LIMA assay (370P03, 370P07 and 378N03) (SEQ ID NOs: 102-104) was electroporated into K562s and assessed for on-target indel formation using next generation sequencing. Indel50s were calculated and  
25 reported in the table associated with **FIG. 9**.

Next, we examined the off-target profiles of the Gen3 clones. As in Example 2, we utilized both the oligo-capture assay and *in silico* predictions to identify putative off-targets. Verification of putative off-targets was performed as in Example 2 and the results are shown in

**FIG. 10.** Those sites that had an FDR corrected p-value  $<0.05$  were considered verified off targets (dark circles) all other sites were considered unverified (grey circles).

#### EXAMPLE 4

5

#### FURTHER OPTIMIZING MEGANUCLEASE SPECIFICITY

To improve specificity of the Gen3 enzymes, a fourth round of engineering was undertaken using the enzyme 370P03 (SEQ ID NO: 102) as the parent sequence. Using the off-target data generated from the Gen3 enzymes in Example 3, four regions with the megaTAL were selected to introduce amino acid variations. Using degenerate codons and assembly PCR, four libraries were constructed to cover these regions and cloned into the yeast-display vector system. Similar to Example 3, the libraries were screened by flow cytometry and underwent positive selection for on-target cleavage activity and negative selection against identified off-targets. At the conclusion of the yeast sorting, we identified 117 unique clones (Gen4 clones) and subjected them to the LIMA assay and assessed both on-target cleavage data and cleavage at 16 off-targets identified in Example 3, enabling selection of both the most active and specific Gen4 enzymes.

As in Example 3, a full mRNA titration for the top Gen4 clones identified in the LIMA assay (399A08, 399A10, 399B05, 399C12, 399D05, 399D09, 399D10 and 399H10) (SEQ ID NOs: 7, 8, 6, 9-13) was electroporated into K562s and assessed for on-target indel formation using next generation sequencing. Indel50s were calculated and reported in the table associated with **FIGs. 11A and 11B**.

Next, we examined the off-target profiles of the Gen4 clones. As in Example 3, we utilized both the oligo-capture assay and *in silico* predictions to identify putative off-targets. Verification of putative off-targets was performed as in Example 3 and the results are shown in **FIGs. 12A and 12B**. Those sites that had an FDR corrected p-value  $<0.05$  were considered verified off targets (black circles) all other sites were considered unverified (grey circles).

Of the 8 Gen4 clones tested in this Example, clone 399B05 (SEQ ID NO:6) had the lowest (best) indel50 and surprisingly had approximately an order of magnitude fewer verified putative off-target sites compared to the parent 370P03 (SEQ ID NO: 102) clone and fewer than all other tested Gen4 clones.

5

EXAMPLE 5

## CHARACTERIZING THERAPEUTIC GENE INSERTION AND ACTIVITY

Before the completion of the engineering cycles, the efficacy of an earlier generation enzyme (Gen2: 327H05.V26S) was tested in both a more translationally relevant *in vitro* cell model system as well as *in vivo* studies in non-human primates described in Example 6.

- 10 Primary human hepatocytes from an adult male donor were used as a clinically relevant *in vitro* model to test human albumin-targeted megaTAL-mediated insertion of an adeno-associated virus (AAV) encoding for the human factor VIII gene (FVIII) (therapeutic transgene of the donor repair template). Cells were cultured in a sandwich configuration on collagen I-coated wells with Matrigel overlay. Following a three-day recovery period to enhance
- 15 adherence and enable restoration of full hepatocyte cell function, cells were treated with *in vitro* transcribed Gen2 327H05V26S albumin megaTAL mRNA (SEQ ID NO: 106) encapsulated in a lipid nanoparticle at a static dose of 100µg/ml human. Concurrently, the cells were also treated with a serial dilution of AAV-FVIII at a maximum dose of 1e6 MOI down to 976.5625 MOI, as well as an untreated control. To assess genomic editing (% INDEL) and
- 20 AAV integration (% AAV integration), genomic DNA was isolated from the cultured cells and applied to amplicon sequencing (NGS for % INDEL measurement) and quantitative droplet digital polymerase chain reaction (ddPCR for % AAV integration). For editing activity, primers were designed to flank the albumin-targeting megaTAL binding site. For AAV integration, ddPCR primers and probes were designed to flank the junction of the AAV vector
- 25 sequence in the forward (therapeutic) direction and the albumin gene locus. At the static dose of 100µg/ml, the human album-targeted megaTAL yielded an average of 19.8% INDEL. This effect was coincident with a dose response to the serial dilution of AAV-FVIII applied, with 0.07% integration observed at the lowest dose of 9.765625 MOI increasing to a maximum of 4.19% integration at 1e6 MOI (**FIG. 13**). Accordingly, the experiment provides quantitative

evidence indicating functional therapeutic gene insertion (targeted integration) in a translationally relevant model via the editing action of the megaTAL.

#### EXAMPLE 6

##### 5 CHARACTERIZING IN VIVO EDITING AND THERAPEUTIC ACTIVITY

Non-human primates (NHP) of the species *Macaca fascicularis* were used to test the efficacy of Albumin targeted megaTAL mediated insertion of Adeno Associated Virus (AAV) encoding the human Factor VIII gene (FVIII) (therapeutic transgene of the donor repair template) as the genomic albumin-targeting megaTAL binding site is 100% conserved in this  
10 *in vivo* model. Four NHP animals were administered a bolus dose of the AAV-FVIII viral vector. The same four animals were then administered *in vitro* transcribed Gen2 327H05.V26S megaTAL mRNA (SEQ ID NO: 106) encapsulated in a lipid nanoparticle via vein infusion. Blood was collected 8 days after AAV and LNP administration to assess FVIII activity in blood plasma (**FIG. 14C**). The study was terminated, and animals necropsied at day 23, livers  
15 were collected to assess megaTAL editing and AAV integration in the target organ (**FIGs. 14A and 14B**). To assess genomic editing (% INDEL) and AAV integration (% AAV integration), genomic DNA was isolated from bulk liver samples from individual animals and quantitative digital drop polymerase chain reaction (ddPCR) was performed. ddPCR primer and probes for editing activity (%INDEL) were designed to flank the albumin-targeting megaTAL binding  
20 site and for AAV integration they were designed to flank the junction of the AAV vector sequence in the forward (therapeutic) direction and the albumin gene locus. The albumin-targeting megaTAL showed activity in NHPs livers with an average of 4.0 % INDELs in treated animals by ddPCR (**FIG. 14A**). The albumin-targeting megaTAL activity at the on-target binding site led to an ~3% average AAV integration rate in NHPs livers as measured by  
25 ddPCR (**FIG. 14B**). Albumin-targeting megaTAL activity and AAV-FVIII mediated integration led to an average of ~800 U/L of FVIII (80% normalized FVIII levels) in NHP blood serum 8 days after LNP-mRNA and AAV-FVIII administration by a chromogenic assay (**FIG. 14C**). For reference, 400 U/L of FVIII antigen or 40% of normalized FVIII levels is

considered therapeutic (Srivastava A, Santagostino E, Dougall A, et al. WFH Guidelines for the Management of Hemophilia, 3rd edition. Haemophilia. 2020; 26(Suppl 6): 1-158).

- 5           In general, in the following claims, the terms used should not be construed to limit the claims to the specific embodiments disclosed in the specification and the claims, but should be construed to include all possible embodiments along with the full scope of equivalents to which such claims are entitled. Accordingly, the claims are not limited by the disclosure.

## CLAIMS

What is claimed is:

1. A polypeptide comprising an engineered I-OnuI homing endonuclease (HE) that binds and cleaves DNA at a target site within a double-stranded DNA (dsDNA) molecule,  
wherein the target site is within intron 1 of the human albumin (ALB) gene,  
wherein the I-OnuI HE cleaves both strands of the dsDNA molecule,  
wherein the engineered I-OnuI HE comprises an amino acid sequence that is at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to the amino acid sequence set forth in any one of SEQ ID NOs: 6-13.
2. The polypeptide of claim 1, wherein the target site comprises a nucleic acid sequence as set forth in any one of SEQ ID NOs: 32-34, preferably a nucleic acid sequence as set forth in SEQ ID NO: 34.
3. The polypeptide of claim 1 or claim 2, wherein the engineered I-OnuI HE comprises the following amino acid substitutions in relation to the numbering of SEQ ID NO: 4:
  - (a) S20C, L22S, R24V, R26Q, N27K, N28A, K30G, S31W, S32T, V33T, G34R, S36N, E38K, G40V, Q42S, T44G, V64T, A66S, N67K, N71T, A72Q, S74R, K76T, C176S, F178M, N180I, I182A, K185R, S186T, K187G, L188A, G189R, Q193R, V195C, S197E, T199G, Y219H, K221S, K225L, F228S, W230F, D232I, and V234R;
  - (b) S20C, L22S, R24T, N27K, N28F, K30G, S31R, S32T, V33T, G34R, S36T, E38K, G40M, Q42S, T44G, V64T, A66S, N67K, N71T, A72Q, S74R, K76T, C176S, F178M, N180I, I182A, K185R, S186T, K187G, L188A, G189R, Q193R, V195C, S197E, T199G, Y219H, K221S, K225L, F228S, W230F, D232I, and V234R;
  - (c) S20C, L22S, R24A, R26G, N27K, N28L, K30G, S31G, S32T, V33T, G34R, E38K, G40V, Q42S, T44G, V64T, A66S, N67K, N71T, A72Q, S74R, K76T, C176S, F178M, N180I,

I182A, K185R, S186T, K187G, L188A, G189R, Q193R, V195C, S197E, T199G, Y219H, K221S, K225L, F228S, W230F, D232I, and V234R;

(d) S20C, L22S, R24A, R26G, N27K, N28L, K30G, S31G, S32T, V33T, G34R, E38K, G40V, Q42S, T44G, V64T, A66S, N67K, N71T, A72Q, S74R, K76T, C176S, F178M, N180I, I182A, K185R, S186T, K187G, L188A, G189R, Q193R, V195C, S197E, T199G, Y219H, K221S, K225L, F228S, W230F, D232I, and V234R;

(e) S20C, L22S, R24V, R26A, N27K, N28G, K30G, S31G, S32T, V33T, G34R, S36T, E38K, G40V, Q42S, T44G, V64T, A66S, N67K, N71T, A72Q, S74R, K76T, C176S, F178M, N180I, I182A, K185R, S186T, K187G, L188A, G189R, Q193R, V195C, S197E, T199G, Y219H, K221S, K225L, F228S, W230F, D232I, and V234R;

(f) S20C, L22S, R24A, R26G, N27K, N28C, K30G, S31C, S32T, V33T, G34R, E38K, G40V, Q42S, T44G, V64T, A66S, N67K, N71T, A72Q, S74R, K76T, C176S, F178M, N180I, I182A, K185R, S186T, K187G, L188A, G189R, Q193R, V195C, S197E, T199G, Y219H, K221S, K225L, F228S, W230F, D232I, and V234R;

(g) S20C, L22S, R24A, R26G, N27K, N28L, K30G, S31A, S32T, V33T, G34R, S36V, E38K, G40V, Q42S, T44G, V64T, A66S, N67K, N71T, A72Q, S74R, K76T, C176S, F178M, N180I, I182A, K185R, S186T, K187G, L188A, G189R, Q193R, V195C, S197E, T199G, Y219H, K221S, K225L, F228S, W230F, D232I, and V234R; or

(h) S20C, L22S, R24H, R26G, N27K, N28M, K30G, S31R, S32T, V33T, G34R, S36L, E38K, G40V, Q42S, T44G, V64T, A66S, N67K, N71T, A72Q, S74R, K76T, C176S, F178M, N180I, I182A, K185R, S186T, K187G, L188A, G189R, Q193R, V195C, S197E, T199G, Y219H, K224S, K225L, F228S, W230F, D232I, and V234R.

4. The polypeptide of any one of the preceding claims, wherein the engineered HE comprises the amino acid sequence set forth in SEQ ID NO: 6.

5. The polypeptide of any one of the preceding claims, further comprising one or more nuclear localization signal (NLS) and/or a TALE DNA binding domain, preferably the TALE

DNA binding domain comprises about 8.5 TALE repeat units to about 15.5 TALE repeat units and/or preferably the NLS comprises a nucleic acid sequence set forth in SEQ ID NO: 86.

6. The polypeptide of claim 5, wherein the TALE DNA binding domain binds the polynucleotide sequence set forth in any one of SEQ ID NOs: 35-42, preferably the polypeptide binds and cleaves the polynucleotide sequence set forth in SEQ ID NO: 43.

7. The polypeptide of any one of the preceding claims, wherein the polypeptide comprises an amino acid sequence that is at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to the amino acid sequence set forth in any one of SEQ ID NOs: 14-21, preferably the polypeptide comprises the amino acid sequence set forth in SEQ ID NO: 14.

8. The polypeptide of any one of the preceding claims, further comprising a peptide linker and an end-processing enzyme or biologically active fragment thereof, preferably wherein the end-processing enzyme comprises Trex2, more preferably, the end-processing enzyme comprises an amino acid sequence as set forth in SEQ ID NO: 44.

9. A polynucleotide encoding the polypeptide of any one of the preceding claims, preferably wherein the polynucleotide is an mRNA or a cDNA.

10. The polynucleotide of claim 9, wherein the polynucleotide is an mRNA, wherein the mRNA comprises a nucleic acid sequence that is at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to the nucleic acid sequence set forth in any one of SEQ ID NOs: 22-29 and 108, preferably wherein the mRNA comprises a nucleic acid sequence set forth in SEQ ID NO: 22.

11. A pharmaceutical composition comprising a polynucleotide encoding the polypeptide of any one of claims 1-8 or the polynucleotide of claim 9 or claim 10, and a physiologically acceptable carrier, preferably the polynucleotide is an mRNA comprising a nucleic acid



sequence set forth in SEQ ID NO: 22, and preferably wherein the pharmaceutically acceptable carrier comprises a lipid nanoparticle and the polynucleotide is encapsulated in said lipid nanoparticle.

12. The pharmaceutical composition of claim 11 further comprising a donor repair template, preferably an AAV vector particle comprises the donor repair template, and more preferably the donor repair template comprises a FVIII transgene or a FIX transgene.

13. A method of editing a human ALB gene in a cell comprising: introducing a polynucleotide encoding the polypeptide of any one of claims 1 to 8 or the polynucleotide of claim 9 or claim 10 into the cell, wherein the polynucleotide is translated in the cell to produce a polypeptide, wherein the polypeptide creates a double-strand break (DSB) at the target site; optionally further comprising introducing a donor repair template into the cell and wherein the donor repair template is incorporated into the human ALB gene by homology directed repair (HDR) at the site of DSB.

14. A method of treating a blood disorder, lysosomal storage disorder, or metabolic disorder, or condition associated therewith, comprising: introducing into the cell (a) a polynucleotide encoding the polypeptide of any one of claims 1 to 8 or the polynucleotide of claim 9 or claim 10, and (b) a donor repair template into a cell;

wherein the polynucleotide is translated in the cell to produce a polypeptide, wherein the polypeptide creates a double-strand break (DSB) at the target site and the donor repair template is incorporated into the human ALB gene by homology directed repair (HDR) at the site of the DSB,

preferably wherein the cell is a hepatocyte and the donor repair template comprises a FVIII transgene or FIX transgene.

15. The pharmaceutical composition of claim 12 for use in editing a human ALB gene in a cell or for use in treating a blood disorder, lysosomal storage disorder, or metabolic disorder, or condition associated therewith,

wherein the composition delivers the polynucleotide and the donor repair template into the cell,

wherein the polynucleotide is translated in the cell to produce a polypeptide, wherein the polypeptide creates a double-strand break (DSB) at the target site, and

wherein the donor repair template is incorporated into the human ALB gene by homology directed repair (HDR) at the site of the DSB,

preferably wherein the cell is a hepatocyte and the donor repair template is FVIII or FIX.