as poor knock-in (KI) efficiency, insertional oncogenesis, unstable and/or anomalous expression of transgenes and/or adjacent genes, *etc.*.

SUMMARY

[0005] The present disclosure is directed, *inter alia*, to safe harbor loci that exhibit high knock-in efficiency and stable expression of their transgenes. These safe harbor loci can be used to alter T cells for immunotherapy. These safe harbor loci are useful for the treatment of various diseases, including cancer.

[0005a] In one aspect, the present disclosure refers to an engineered cell, comprising at least one sequence encoding a transgene, wherein the at least one sequence is inserted within a safe harbor locus; wherein the safe harbor locus is at any one or more of the sgRNA target loci selected from: chr11:128340000-128350000, chr10:33130000-33140000, chr10:72290000-72300000, chr11:65425000-65427000 (NEAT1), chr15:92830000-92840000, chr16:11220000-11230000, chr2:87460000-87470000, chr3:186510000-186520000, chr3:59450000-59460000, chr8:127980000-128000000, chr9:7970000-7980000, APRT, B2M, CAPNS1, CBLB, CD2, CD3E, CD3G, CD5, EDF1, FTL, PTEN, PTPN2, PTPN6, PTPRC, PTPRCAP, RPS23, RTRAF, SERF2, SLC38A1, SMAD2, SOCS1, SRP14, SRSF9, SUB1, TET2, TIGIT, TRAC, and TRIM28.

[0006] In one aspect, the present disclosure provides an engineered cell, comprising at least one sequence encoding a transgene, wherein the at least one sequence is inserted within a safe harbor locus, the safe harbor locus is at any one or more of the sgRNA target loci provided in Table 4; and wherein expression of the at least one sequence encoding the transgene is operatively linked to an endogenous promoter. In another aspect, the present disclosure provides an engineered cell, comprising at least one sequence encoding a transgene, wherein the at least one sequence is inserted within a safe harbor locus, the safe harbor locus is at any one or more of the sgRNA target loci provided in Table 4; and wherein expression of the at least one sequence encoding the transgene is operatively linked to an exogenous promoter.

[0007] In some embodiments, the target locus is selected from: chr10:33130000-33140000, chr10:72290000-72300000, chr11:128340000-128350000, chr11:65425000-65427000 (NEAT1), chr15:92830000-92840000, chr16:11220000-11230000, chr2:87460000-87470000, chr3:186510000-186520000, chr3:59450000-59460000, chr8:127980000-128000000, and chr9:7970000-7980000. In some embodiments, the target locus is selected from: chr10:72290000-72300000, chr11:128340000-128350000, chr15:92830000-92840000, and chr16:11220000-11230000. In some embodiments, the target locus is chr11:128340000-

128350000. In some embodiments, the target locus is chr15:92830000-92840000. In some embodiments, the target locus is a gene selected from: *APRT*, *B2M*, *CAPNS1*, *CBLB*, *CD2*, *CD3E*, *CD3G*, *CD5*, *EDF1*, *FTL*, *PTEN*, *PTPN2*, *PTPN6*, *PTPRC*, *PTPRCAP*, *RPS23*, *RTRAF*, *SERF2*, *SLC38A1*, *SMAD2*, *SOCS1*, *SRP14*, *SRSF9*, *SUB1*, *TET2*, *TIGIT*, *TRAC*, and *TRIM28*.

[0008] In some embodiments, the safe harbor locus is selected from any one of the integration sites in Table 4 designated: GS88, GS89, GS90, GS91, GS92, GS93, GS94, GS95, GS96, GS97, GS98, GS99, GS100, GS101, GS102, GS103, GS104, GS105, GS106, GS107, GS108, GS109, GS110, GS111, GS112, GS113, GS114, GS115, GS116, GS117, GS118, GS119, GS120. In some embodiments, the safe harbor locus is selected from any one

of the integration sites in Table 4 designated: GS91, GS92, GS93, GS94, GS95, GS96, GS100, GS101, GS102, GS103, GS104, and GS105. In some embodiments, the safe harbor locus is selected from any one of the integration sites in Table 4 designated: GS103, GS104, and GS105. In some embodiments, the safe harbor locus is selected from any one of the integration sites in Table 4 designated: GS94, GS95, and GS96. In some embodiments, the safe harbor locus is the GS94 integration site in Table 4. In some embodiments, the safe harbor locus is selected from any one of the integration sites in Table 4 designated: GS100, GS101, and GS102. In some embodiments, the safe harbor locus is the GS102 integration site in Table 4. In some embodiments, the safe harbor locus is selected from any one of the integration sites in Table 4 designated: GS91, GS92, and GS93.

[0009] In some embodiments, the exogenous promoter is an EF1a promoter. In some embodiments, the engineered cell is a stem cell, a human cell, a primary cell, an hematopoietic cell, an adaptive immune cell, an innate immune cell, a T cell or a T cell progenitor. In some embodiments, the transgene encodes a recombinant protein, optionally a therapeutic agent. In some embodiments, the transgene encodes a chimeric antigen receptor (CAR).

[0010] In another aspect, the present disclosure provides a composition comprising the engineered cell as described herein and a pharmaceutical excipient.

[0010a] In another aspect, the present disclosure refers to a guide ribonucleic acids (gRNA) for editing a cell at a safe harbor locus, wherein the gRNA comprises any one of SEQ ID NOS: 94, 102, 1-93, 95-101, or 103-120.

[0010b] In another aspect, the present disclosure refers to a method of editing a cell having chromosomal DNA, comprising inserting at least one sequence encoding a transgene within a safe harbor locus in the chromosomal DNA of the cell, wherein the safe harbor locus is at any one or more of the sgRNA target loci selected from: chr11:128340000-128350000, chr10:33130000-33140000, chr10:72290000-72300000, chr11:65425000-65427000 (NEAT1), chr15:92830000-92840000, chr16:11220000-11230000, chr2:87460000-87470000, chr3:186510000-186520000, chr3:59450000-59460000, chr8:127980000-128000000, chr9:7970000-7980000, *APRT*, *B2M*, *CAPNS1*, *CBLB*, *CD2*, *CD3E*, *CD3G*, *CD5*, *EDF1*, *FTL*, *PTEN*, *PTPN2*, *PTPN6*, *PTPRC*, *PTPRCAP*, *RPS23*, *RTRAF*, *SERF2*, *SLC38A1*, *SMAD2*, *SOCS1*, *SRP14*, *SRSF9*, *SUB1*, *TET2*, *TIGIT*, *TRAC*, and *TRIM28*.

[0011] In another aspect, the present disclosure provides a guide ribonucleic acids (gRNA) for editing a cell at a safe harbor locus, wherein gRNA comprises any one of the sgRNA sequences in Table 4.

[0012] In some embodiments, the gRNA comprises any one of SEQ ID NOS:1-120. In some embodiments, the gRNA comprises any one of SEQ ID NOS: 91-96 and 100-105. In some embodiments, the gRNA comprises SEQ ID NO:94 or SEQ ID NO:102. In some embodiments, the gRNA comprises SEQ ID NO:94. In some embodiments, the gRNA comprises SEQ ID NO:102. In some embodiments, the cell is a stem cell, a human cell, a primary cell, an hematopoietic cell, an adaptive immune cell, an innate immune cell, a T cell or a T cell progenitor.

[0013] In another aspect, the present disclosure provides a method of editing a cell having chromosomal DNA, comprising inserting at least one sequence encoding a transgene within a safe harbor locus in the chromosomal DNA of the cell, wherein the safe harbor locus is any one or more of the sgRNA target loci provided in Table 4.

RTRAF, *SERF2*, *SLC38A1*, *SMAD2*, *SOCS1*, *SRP14*, *SRSF9*, *SUB1*, *TET2*, *TIGIT*, *TRAC*, and *TRIM28*.

[0021] In some embodiments, the safe harbor locus is selected from any one of the integration sites in Table 4 designated: GS88, GS89, GS90, GS91, GS92, GS93, GS94, GS95, GS96, GS97, GS98, GS99, GS100, GS101, GS102, GS103, GS104, GS105, GS106, GS107, GS108, GS109, GS110, GS111, GS112, GS113, GS114, GS115, GS116, GS117, GS118, GS119, GS120. In some embodiments, the safe harbor locus is selected from any one of the integration sites in Table 4 designated: GS91, GS92, GS93, GS94, GS95, GS96, GS100, GS101, GS102, GS103, GS104, and GS105. In some embodiments, the safe harbor locus is selected from any one of the integration sites in Table 4 designated: GS103, GS104, and GS105. In some embodiments, the safe harbor locus is selected from any one of the integration sites in Table 4 designated: GS94, GS95, and GS96. In some embodiments, the safe harbor locus is the GS94 integration site in Table 4. In some embodiments, the safe harbor locus is selected from any one of the integration sites in Table 4 designated: GS100, GS101, and GS102. In some embodiments, the safe harbor locus is the GS102 integration site in Table 4. In some embodiments, the safe harbor locus is selected from any one of the integration sites in Table 4 designated: GS91, GS92, and GS93

[0021a] In another aspect, the present disclosure refers to an ex vivo method of obtaining an engineered cell or population thereof, comprising: a.obtaining a cell; and b. genetically modifying the cell by inserting at least one sequence encoding a transgene within a safe harbor locus, wherein the safe harbor locus is at any one or more of the sgRNA target loci selected from: chr11:128340000-128350000, chr10:33130000-33140000, chr10:72290000-72300000, chr11:65425000-65427000 (NEAT1), chr15:92830000-92840000, chr16:11220000-11230000, chr2:87460000-87470000, chr3:186510000-186520000, chr3:59450000-59460000, chr8:127980000-128000000, chr9:7970000-7980000, APRT, B2M, CAPNS1, CBLB, CD2, CD3E, CD3G, CD5, EDF1, FTL, PTEN, PTPN2, PTPN6, PTPRC, PTPRCAP, RPS23, RTRAF, SERF2, SLC38A1, SMAD2, SOCS1, SRP14, SRSF9, SUB1, TET2, TIGIT, TRAC, and TRIM28.

[0022] In another aspect, the present disclosure provides an *ex vivo* method of obtaining an engineered cell or population thereof, comprising (a) obtaining a cell; and (b) genetically modifying the cell by inserting at least one sequence encoding a transgene within a safe harbor locus, wherein the safe harbor locus is selected from any one of the sgRNA target loci in Table 4.

[0023] In some embodiments, obtaining the cell comprises: (i) collecting a tissue sample from a subject, (ii) isolating the cells from the tissue samples, and (iii) culturing the cells in vitro. In some embodiments, the tissue sample is a blood sample. In some embodiments, the cell is a stem cell, a human cell, a primary cell, an hematopoietic cell, an adaptive immune cell, an innate immune cell, a T cell or T cell precursor (T cell progenitor). In some embodiments, the at least one sequence is inserted using a homology-directed repair. In some embodiments, the at least one sequence is inserted using a homology independent targeted insertion.

[0024] In some embodiments, the genetically modifying in step (b) comprises contacting the cell with one or more guide ribonucleic acids (gRNAs), the at least one sequence, and one or more Cas9 endonucleases, wherein the one or more gRNAs and Cas9 endonucleases facilitate

the insertion of the at least one sequence into chromosomal DNA within the safe harbor locus. In some embodiments, the one or more gRNAs comprises a sequence selected from any one of the sgRNA sequences in Table 4.

[0025] In some embodiments, the transgene encodes a recombinant protein, optionally a therapeutic agent. In some embodiments, the transgene encodes a chimeric antigen receptor (CAR). In some embodiments, the at least one sequence comprises an exogenous promoter and the exogenous promoter is operably linked to the transgene. In some embodiments, the exogenous promoter is an EF1a promoter.

[0025a] In another aspect, the present disclosure refers to the use of the engineered cell as described herein in the manufacture of medicament for treating a disease, wherein the medicament is to be administered in an effective amount to a subject having or at risk of having the disease.

[0025b] In another aspect, the present disclosure refers to the use of a composition comprising the cell or population thereof as described herein in the manufacture of medicament for treating a disease, wherein the medicament is obtained using method as described herein, wherein the medicament is to be administered in an effective amount to a subject having or at risk of having the disease.

[0025c] In another aspect, the present disclosure refers to an engineered cell as described herein for use in therapy.

[0026] In yet another aspect, the present disclosure provides a method of treating a subject having or at risk of having a disease, comprising administering to the subject an effective amount of an engineered cell as described herein, a population thereof, or a composition as described herein. In some embodiments, the cell, the population thereof, or the composition is administered to the subject by infusion.

[0027] In yet another aspect, the present disclosure provides a method of treating a subject having or at risk of having a disease, comprising (a) conducting any one of the methods described *supra*; and (b) administering to the subject an effective amount of a composition comprising the cell or a population thereof. In some embodiments, the composition is administered to the subject by infusion. In some embodiments, the disease is cancer. In some embodiments, the disease is blood cancer.

[0028] In another aspect, the present disclosure provides a method of identifying a safe harbor locus, comprising: (a) identifying genes or non-coding regions in a chromosome that are above a threshold level for expression across developmental cell states and/or a threshold

level for accessibility of chromatin; (b) generating a linear model that correlates the gene or non-coding region from step (a) with knock-in (KI) efficiency and estimates the KI efficiency of any gene or coding region on the chromosome; and (c) selecting the safe harbor locus based on threshold parameters; wherein the safe harbor locus is selected for insertion of at least one sequence encoding a transgene within a cell.

[0029] In some embodiments, the threshold parameters include one or more of: stable expression of a transgene, knockout of the gene confers benefit to the function of the cell, no known function within the cell, stable transgene expression *in vitro* with or without CD3/CD28 stimulation, negligible off-target cleavage as detected by iGuide-Seq or CRISPR-Seq, less off-target cleavage relative to other loci as detected by iGuide-Seq or CRISPR-Seq, negligible transgene-independent cytotoxicity, negligible transgene-independent cytokine