(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
22 February 2024 (22.02.2024)

WIPO | PCT

(10) International Publication Number
WO 2024/039600 A2

(51) International Patent Classification:
C12N 15/86 (2006.01) *A61K 35/15* (2015.01)

(21) International Application Number:
PCT/US2023/030141

(22) International Filing Date:
14 August 2023 (14.08.2023)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
63/397,981 15 August 2022 (15.08.2022) US

(71) Applicant: **THE BOARD OF TRUSTEES OF THE LELAND STANFORD JUNIOR UNIVERSITY** [US/US]; Office of the General Counsel, Building 170, 3rd Floor, Main Quad, P.O. Box 20386, STANFORD, California 94305-2038 (US).

(72) Inventors: **JAISWAL, Siddhartha**; 202 Virginia Avenue, San Francisco, California 94110 (US). **GOPAKUMAR, Jayakrishnan**; 18725 Metler Ct., Saratoga, California 95070 (US).

(74) Agent: **SHERWOOD, Pamela J.**; Bozicevic, Field & Francis LLP, 201 Redwood Shores Parkway, Suite 200, Redwood City, California 94065 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*

(54) **Title:** INDUCING TCL1A EXPRESSION TO INCREASE PROLIFERATION AND PROLONG STEMNESS OF HEMATOPOIETIC STEM CELLS

(57) **Abstract:** Compositions and methods are provided for increasing proliferation of hematopoietic stem cells (HSC), and increasing HSC capacity for self-renewal. It is shown herein that overexpression of the protein TCL1A in HSC allows the cells to expand in absolute number to a greater number than unmodified cells, and to maintain stem cell surface markers in culture in media supplemented with cytokines.

WO 2024/039600 A2

INDUCING TCL1A EXPRESSION TO INCREASE PROLIFERATION AND PROLONG STEMNESS OF HEMATOPOIETIC STEM CELLS

GOVERNMENT RIGHTS

[0001] This invention was made with Government support under contract HL157540 awarded by the National Institutes of Health. The Government has certain rights in the invention.

BACKGROUND

[0002] Effective ex vivo or in vivo expansion of hematopoietic stem cells (HSCs) would be a critical solution to overcome the material shortage of this tissue for transplantation purposes, genetic modification protocols, or for recovery after chemotherapy. One of the main shortcomings of the clinical use of HSCs is the limited number of cells that can be safely harvested from a patient. Thus, HSC expansion tools would unlock the potential of stem cell gene therapy for clinical applications.

[0003] Successful HSC expansion methods would provide for symmetric stem cell division and self-renewal without further differentiation, as well as a proliferative advantage that leads to increased resiliency and expedited engraftment. Such a tool would allow for accelerated expansion of true HSCs and improved engraftment of HSCs. Accomplishing these goals would greatly attenuate a patient's immunocompromised state during such a transplant.

[0004] Current clinical trials have only reported limited success in achieving this outcome, as current protocols lead to premature unwanted differentiation of the stem cells and progenitors. Combinations of different cytokines and growth factors, such as SCF, Flt3, TPO, IL-3, and IL6, are commonly used for supporting HSC and progenitor survival, proliferation and maintenance during in vitro culturing systems in clinics, but these have proven to be unsuccessful at improving HSC engraftment. Furthermore, new techniques such as co-culture systems, small molecule inhibitors, and chemical adjuvants, have been tested with limited success. It still requires 2-3 weeks for engraftment to occur following a depleting conditioning protocol.

[0005] Designing an HSC with improved engraftment kinetics and a proliferative advantage would allow a shortening of the engraftment period of the therapeutic cells and potentially eliminate the need for a destructive conditioning regime. The present disclosure addresses this need.

SUMMARY

[0006] Compositions and methods are provided for increasing proliferation of hematopoietic stem cells (HSC), and increasing HSC capacity for self-renewal. It is shown herein that overexpression of the protein TCL1A in HSC allows the cells to expand in absolute

number to a greater number than unmodified cells, and to maintain stem cell surface markers in culture in media supplemented with cytokines. In an in vivo model, transplantation of TCL1A overexpressing HSCs dramatically improved survival of an animal following a lethal total body radiation conditioning regime.

[0007] In some embodiments, a composition of modified HSC is provided, where the HSC have been genetically altered to express TCL1A at levels greater than non-modified HSC, e.g. expressing at a level 5-fold greater than unmodified cells, 10-fold great, 50-fold greater, 100-fold greater, or more. In some embodiments the HSC are human. In some embodiments the cells are genetically altered by infection with a lentiviral vector comprising human TCL1A operably linked to a promoter active in HSC. In some embodiments the cells are genetically modified by genome editing, e.g. with programmable gene editing tools (e.g., CRISPR/Cas RNA-guided proteins such as Cas9, CasX, CasY, and Cpf1, Zinc finger proteins such as Zinc finger nucleases, TALE proteins such as TALENs, CRISPR/Cas guide RNAs, PAMs, and the like) to increase expression.

[0008] In some embodiments, a composition of modified HSC is provided, where the HSC have been modified by introduction of mRNA encoding TCL1A, where TCL1A is then expressed at levels greater than non-modified HSC, e.g. expressing at a level 5-fold greater than unmodified cells, 10-fold great, 50-fold greater, 100-fold greater, or more. In some embodiments the HSC are human. In some embodiments the TCL1A sequence encodes human TCL1A. In some embodiments the mRNA is modified mRNA (mmRNA), e.g. comprising modified nucleosides, e.g. one or more of thiouridine, 5-methylcytidine (m5C), N1-methyl-pseudouridine (1mΨ), pseudouridine (Ψ), etc., particularly m5C and Ψ. The mmRNA may further comprise an m7GpppNm cap or an ARCA cap. The polyA tail may have a predetermined length other than that of the native mRNA. The sequence may be codon-optimized. The mmRNA may be packaged in lipid nanoparticles, *e.g.* comprising ionizable lipids, phospholipids, and cholesterol, where the outside of the particle is coated in pegylated lipids. Such modified HSC may be referred to as mmRNA modified HSC. In some embodiments an mmRNA composition encoding human TCL1A is provided, which is optionally packaged in a lipid nanoparticle.

[0009] In some embodiments a method is provided for the expansion of HSC *in vitro*, the method comprising genetically altering, or mmRNA altering, HSC to express TCL1A at levels greater than non-modified HSC, e.g. expressing at a level 5-fold greater than unmodified cells, 10-fold great, 50-fold greater, 100-fold greater, or more. The initial population of unmodified HSC may be selected from a suitable source, e.g. mobilized peripheral blood, bone marrow, and the like.

[0010] In some embodiments, a formulation comprising modified HSC that over-express TCL1A in a pharmaceutically acceptable excipient is provided. In some embodiments the HSC

are present in a unit dose. In some embodiments a unit dose is the dose sufficient for engraftment. For example, a dose of cells may be at least about 3 x 10^5 $CD34^+$ cells/kg, at least about 5 x 10^5 $CD34^+$ cells/kg, at least about 10^6 $CD34^+$ cells/kg. Higher doses can be administered with the proviso that not more than about 3 x 10^3 $CD3^+$ cells/kg are administered. The HSC are optionally expanded in vitro prior to formulation.

[0011] In some embodiments a method is provided for engraftment of HSC, the method comprising administering to an individual in need thereof an effective dose of a formulation comprising modified HSC that over-express TCL1A in a pharmaceutically acceptable excipient. In some embodiments the individual has been treated prior to administration with an agent that ablates endogenous HSC.

BRIEF DESCRIPTION OF THE DRAWINGS

[0012] The invention is best understood from the following detailed description when read in conjunction with the accompanying drawings. It is emphasized that, according to common practice, the various features of the drawings are not to-scale. On the contrary, the dimensions of the various features are arbitrarily expanded or reduced for clarity. Included in the drawings are the following figures.

[0013] FIGS. 1A-1C. A genome-wide association study (GWAS) of passenger counts identifies *TCL1A* as a genome-wide significant locus. B, The association between the genotypes of rs2887399 and PACER varied between *TET2* and *DNMT3A*. Alt-alleles were associated with decreased PACER score in *TET2* mutation carriers, in contrast to *DNMT3A* carriers, where no association was observed. C, The association between alt-alleles at rs2887399 and presence of specific CHIP mutations varies by CHIP mutations. Forest plot shows the effect estimates of a single T allele and two T-alleles respectively, estimating using Firth logistic regression. On the right of the forest plot, effect estimates and p-values are included from SAIGE, which uses an additive coding of the alt-alleles for hypothesis testing. In the additive tests, *SF3B1* and *SRSF2* were grouped together to aid convergence.

[0014] FIG. 2 The posterior inclusion probabilities (PIP) as estimated by $SuSIE_{25}$ are plotted on the y-axis, and the genomic position of a 0.8 Mb region including TCL1A is plotted on the x-axis. The linkage disequilibrium (LD) estimates are plotted on a color scale and are estimated on the genotypes used for association analyses.

[0015] FIGS. 3A-3F. A. Schematic of *TCL1A*-eGFP lentivirus construct (top). Comparison of control-eGFP lentivirus versus *TCL1A*-eGFP on *TCL1A* expression in CD34+ human HSPCs (bottom). B. Quantification of Lin-CD34+CD38-CD45RA- HSC/MPP counts after 14 days of in vitro expansion (left), and quantification of colony forming units in a methylcellulose assay after 14 days of liquid culture in vitro expansion; p-values were calculated using a two-sided t-test. C. Percent of cycling Lin-CD34+CD38-HSC/MPP cells

by DAPI staining after 10 days of in vitro expansion; p-values were calculated using a two-sided t-test. D. UMAP of cell clusters identified after in vitro expansion of sorted HSC/MPPs (left). UMAP split by the 4 independent samples analyzed (right). G/G or T/T refers to the rs2887399 genotype for the donor. E. Dot plot illustrating gene expression of representative marker genes across different cell clusters arranged by functional group. F. Forest plot of log2 fold-difference (Log2FD) in proportion of cells within each HSC/MPP cluster in TCL1A-eGFP sample versus control-eGFP using a permutation test. Results for each donor are shown separately and the FDR for each comparison is shown to the right.

[0016] FIG. 4. In vivo survival of TCL1A overexpressing hematopoietic stem cells. Kaplan Meier survival curve of mice irradiated with 9.5 Gy and transplanted with 1 million chimeric bone marrow cells harboring stem cells edited with control eGFP or TCL1A overexpression eGFP lentivirus.

[0017] FIG. 5. Long term in vivo engraftment. Shown are changes over time in the levels of engraftment for granulocytes.

DETAILED DESCRIPTION

[0018] Before the present methods and compositions are described, it is to be understood that this invention is not limited to particular method or composition described, as such may, of course, vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to be limiting, since the scope of the present invention will be limited only by the appended claims.

[0019] Where a range of values is provided, it is understood that each intervening value, to the tenth of the unit of the lower limit unless the context clearly dictates otherwise, between the upper and lower limits of that range is also specifically disclosed. Each smaller range between any stated value or intervening value in a stated range and any other stated or intervening value in that stated range is encompassed within the invention. The upper and lower limits of these smaller ranges may independently be included or excluded in the range, and each range where either, neither or both limits are included in the smaller ranges is also encompassed within the invention, subject to any specifically excluded limit in the stated range. Where the stated range includes one or both of the limits, ranges excluding either or both of those included limits are also included in the invention.

[0020] Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, some potential and preferred methods and materials are now described. All publications mentioned herein are incorporated herein by reference to disclose and describe the methods and/or materials in connection with which the publications are cited. It is understood that the

present disclosure supercedes any disclosure of an incorporated publication to the extent there is a contradiction.

[0021] It must be noted that as used herein and in the appended claims, the singular forms "a", "an", and "the" include plural referents unless the context clearly dictates otherwise. Thus, for example, reference to "a cell" includes a plurality of such cells and reference to "the peptide" includes reference to one or more peptides and equivalents thereof, e.g. polypeptides, known to those skilled in the art, and so forth.

[0022] The publications discussed herein are provided solely for their disclosure prior to the filing date of the present application. Nothing herein is to be construed as an admission that the present invention is not entitled to antedate such publication by virtue of prior invention. Further, the dates of publication provided may be different from the actual publication dates which may need to be independently confirmed.

[0023] Generally, conventional methods of protein synthesis, recombinant cell culture and protein isolation, cell engineering with synthetic messenger RNA, recombinant DNA techniques, and genome editing within the skill of the art are employed in the present invention, including, for example, programmable gene editing tools. For additional information related to programmable gene editing tools (e.g., CRISPR/Cas RNA-guided proteins such as Cas9, CasX, CasY, and Cpf1, Zinc finger proteins such as Zinc finger nucleases, TALE proteins such as TALENs, CRISPR/Cas guide RNAs, PAMs, and the like) refer to, for example, Dreier, et al., (2001) J Biol Chem 276:29466-78; Dreier, et al., (2000) J Mol Biol 303:489-502; Liu, et al., (2002) J Biol Chem 277:3850-6); Dreier, et al., (2005) J Biol Chem 280:35588-97; Jamieson, et al., (2003) Nature Rev Drug Discov 2:361-8; Durai, et al., (2005) Nucleic Acids Res 33:5978-90; Segal, (2002) Methods 26:76-83; Porteus and Carroll, (2005) Nat Biotechnol 23:967-73; Pabo, et al., (2001) Ann Rev Biochem 70:313-40; Wolfe, et al., (2000) Ann Rev Biophys Biomol Struct 29:183-212; Segal and Barbas, (2001) Curr Opin Biotechnol 12:632-7; Segal, et al., (2003) Biochemistry 42:2137-48; Beerli and Barbas, (2002) Nat Biotechnol 20:135-41; Carroll, et al., (2006) Nature Protocols 1:1329; Ordiz, et al., (2002) Proc Natl Acad Sci USA 99:13290-5; Guan, et al., (2002) Proc Natl Acad Sci USA 99:13296-301; Sanjana et al., Nature Protocols, 7:171-192 (2012); Zetsche et al, Cell. 2015 Oct 22;163(3):759-71; Makarova et al, Nat Rev Microbiol. 2015 Nov;13(11):722-36; Shmakov et al., Mol Cell. 2015 Nov 5;60(3):385-97; Jinek et al., Science. 2012 Aug 17;337(6096):816-21; Chylinski et al., RNA Biol. 2013 May;10(5):726-37; Ma et al., Biomed Res Int. 2013;2013:270805; Hou et al., Proc Natl Acad Sci U S A. 2013 Sep 24;110(39):15644-9; Jinek et al., Elife. 2013;2:e00471; Pattanayak et al., Nat Biotechnol. 2013 Sep;31(9):839-43; Qi et al, Cell. 2013 Feb 28;152(5):1173-83; Wang et al., Cell. 2013 May 9;153(4):910-8; Auer et. al., Genome Res. 2013 Oct 31; Chen et. al., Nucleic Acids Res. 2013 Nov 1;41(20):e19; Cheng et. al., Cell Res. 2013 Oct;23(10):1163-71; Cho et. al., Genetics. 2013

Nov;195(3):1177-80; DiCarlo et al., Nucleic Acids Res. 2013 Apr;41(7):4336-43; Dickinson et. al., Nat Methods. 2013 Oct;10(10):1028-34; Ebina et. al., Sci Rep. 2013;3:2510; Fujii et. al, Nucleic Acids Res. 2013 Nov 1;41(20):e187; Hu et. al., Cell Res. 2013 Nov;23(11):1322-5; Jiang et. al., Nucleic Acids Res. 2013 Nov 1;41(20):e188; Larson et. al., Nat Protoc. 2013 Nov;8(11):2180-96; Mali et. at., Nat Methods. 2013 Oct;10(10):957-63; Nakayama et. al., Genesis. 2013 Dec;51(12):835-43; Ran et. al., Nat Protoc. 2013 Nov;8(11):2281-308; Ran et. al., Cell. 2013 Sep 12;154(6):1380-9; Upadhyay et. al., G3 (Bethesda). 2013 Dec 9;3(12):2233-8; Walsh et. al., Proc Natl Acad Sci U S A. 2013 Sep 24;110(39):15514-5; Xie et. al., Mol Plant. 2013 Oct 9; Yang et. al., Cell. 2013 Sep 12;154(6):1370-9; Briner et al., Mol Cell. 2014 Oct 23;56(2):333-9; Burstein et al., Nature. 2016; Gao et al., Nat Biotechnol. 2016 Jul 34(7):768-73; Shmakov et al., Nat Rev Microbiol. 2017 Mar;15(3):169-182; as well as international patent application publication Nos. WO2002099084; WO00/42219; WO02/42459; WO2003062455; WO03/080809; WO05/014791; WO05/084190; WO08/021207; WO09/042186; WO09/054985; and WO10/065123; U.S. patent application publication Nos. 20030059767, 20030108880, 20140068797; 20140170753; 20140179006; 20140179770; 20140186843; 20140186919; 20140186958; 20140189896; 20140227787; 20140234972; 20140242664; 20140242699; 20140242700; 20140242702; 20140248702; 20140256046; 20140273037; 20140273226; 20140273230; 20140273231; 20140273232; 20140273233; 20140273234; 20140273235; 20140287938; 20140295556; 20140295557; 20140298547; 20140304853; 20140309487; 20140310828; 20140310830; 20140315985; 20140335063; 20140335620; 20140342456; 20140342457; 20140342458; 20140349400; 20140349405; 20140356867; 20140356956; 20140356958; 20140356959; 20140357523; 20140357530; 20140364333; 20140377868; 20150166983; and 20160208243; and U.S. Patent Nos. 6,140,466; 6,511,808; 6,453,242 8,685,737; 8,906,616; 8,895,308; 8,889,418; 8,889,356; 8,871,445; 8,865,406; 8,795,965; 8,771,945; and 8,697,359; all of which are hereby incorporated by reference in their entirety.

TCL1A

[0024] TCL1A consists of 114 amino acids, and has a predicted molecular weight of 14 kDa. The protein has a unique symmetrical β-barrel structure. In the lymphoid compartment, TCL1A expression is limited to $CD4^-CD8^-CD3^-$ thymocytes as well as $CD34^+CD19^+$ pro-B cell through IgM-negative pre-B cells. TCL1 is an Akt kinase coactivator, which facilitates the oligomerization and activation of Akt in vivo. Consequently, it promotes Akt-dependent cell survival. Reference sequences for human TCL1A include Genbank mRNA NM_001098725 and NM_021966; protein NP_001092195 and NP_068801.

[0025] The TCL1 gene family, consisting of TCL1a (also called TCL1), TCL1b (also called TML1), MTCP1, TNG1 and TNG2 isoforms in human, are a group of proto-oncogenes whose proteins were initially identified in the translocation of human T-PLL. Under physiological conditions,TCL1 transcripts are preferentially expressed in cells of lymphoid lineages and mainly in immature $CD4^{-}CD8^{-}$ cells during development, but not in either $CD4^{+}$ or $CD8^{+}$ mature T cells in circulation. Studies have demonstrated the role of TCL1a as an Akt kinase co-activator that promotes kinase activity and transphosphorylation of Akt, thus promoting its nuclear transport. Activation of Akt leads to cell survival, which underlies the pathogenic mechanism of numerous neoplastic diseases such as lung, ovarian and prostate cancer. Therefore, over-expression of TCL1a could modulate and amplify Akt activation, allowing enhanced signal transduction, cell proliferation and survival, which forms the basis of malignancies.

[0026] The proteins encoded by genes in the TCL1 family are conserved between members, but none of them are matched with any known proteins, therefore they are characterized into a novel family of proteins. The structure of TCL1a protein is a β barrel with an internal hydrophobic core, which consists of two four-stranded β sheets connected by a long loop. Strands βA, βB, βE, and βF are 4 long boards forming one side of the barrel, while the other side of the barrel is composed of 4 short strands βC, βD, βG and βH. Approximately 40% homology has been found between the TCL1a and TCL1b protein, including most amino acids which forms the hydrophobic core. The A1 transcript is a small cysteine-rich coiled-coil protein composed of three α helices, among which two antiparallel helices form an α hairpin stabilized by two disulfide bridges and inter-helix hydrophobic contacts.

[0027] TCL1 proteins act as co-activators to influence the signaling transduction of Akt that might play a role in promoting cell survival, proliferation, growth and metabolism. In the Akt pathway, signal transduction is initiated by the activation of phosphatidylinositol 3-kinase (PI3K) via tyrosine kinase receptors. Activated PI3K forms phosphatidylinositol-3,4-biphosphate (PIP2) and phosphatidylinositol-3,4,5-triphosphate(PIP3) in the plasma membrane, which is tightly regulated by phosphatases. The combination of the pleckstrin homology (PH) domain of Akt with the inositol head group of PIP3 recruits Akt to the plasma membrane with conformational conversion. After being phosphorylated at the site of Thr-308 and Ser-473 by 3-phosphatidyinositol-dependentkinase 1 (PDK1) and another kinase, Akt is disassociated from the membrane into the cytosol to phosphorylate downstream proteins.

[0028] TCL1 proteins, including TCL1a, TCL1b and MTCP1, can bind to Akt and appear to have effects on promoting Akt kinase activation and nuclear translocation by interacting with Akt. For TCL1a, co-immunoprecipitation experiments have shown that the interaction of TCL1a with Akt facilitates Akt conformational exchange. TCL1a may induce Akt phosphorylation at the site of Ser-473 and Thr-308 and enhance Akt activity though

synergic effects instead of activating the Akt kinase directly. The structures of TCL1a and Akt suggest their interaction pattern. Akt kinase contains a polarized PH domain, which is critical for Akt activation by binding with PIP3. One terminal of the PH domain is capped by a C-terminal amphipathicα-helix with two antiparallel β sheets, while the other terminal is formed by three variable loops, VL1, VL2 and VL3, as the phospholipid-binding site. The β5 and β6 strand and the α-helix at the PH domain form a site where could be combined with the exposed 2AA hydrophobic patch at one terminal of the β barrel of TCL1a. Since a dimeric structure is required for TCL1a to have biological functions, two TCL1a-bound Akt kinases are then cross-linked with intactness of other PH-ligand interactions to form a TCL1a-Akt homodimer complex, which ultimately strengthens membrane association, promotes Akt phosphorylation and inhibits Akt inactivation. Therefore, by increasing the Akt-mediated phosphorylation of downstream substrates, such as BAD and GSK-3, TCL1a is able to promote cell proliferation, stabilize mitochondrial transmembrane potential and promote cell survival.

[0029] Furthermore, the interaction between TCL1a and Akt may also contributes to Akt nuclear translocation. Akt is mainly expressed in the cytoplasm, while TCL1a is distributed in both the cytoplasm and the nucleus. Immunofluorescence assays have indicated that Akt and TCL1a are co-localized in the cytoplasm and the nucleus in cells with co-expression ofTCL1a and Akt, meanwhile the TCL1a-Akt interaction in the cytoplasm contributes to the nuclear translocation of Akt.

[0030] SNP rs2887399 (at human genome position chr14:95714358 (GRCh38.p13)) is of interest for genotyping TCL1A. The reference allele of the SNP has forward strand G at the site of polymorphism, while the alt allele has T. Another SNP, 10 base pairs away from rs2887399, can also be used for genotyping (rs11846938). The REF allele for rs11846938 is a T, the alt allele is G. The two SNPs are strongly in linkage disequilibrium.

[0031] The terms "polypeptide," "peptide" and "protein" are used interchangeably herein to refer to a polymer of amino acid residues. The terms also apply to amino acid polymers in which one or more amino acid residue is an artificial chemical mimetic of a corresponding naturally occurring amino acid, as well as to naturally occurring amino acid polymers and non-naturally occurring amino acid polymer.

[0032] The term "sequence identity," as used herein in reference to polypeptide or DNA sequences, refers to the subunit sequence identity between two molecules. When a subunit position in both of the molecules is occupied by the same monomeric subunit (e.g., the same amino acid residue or nucleotide), then the molecules are identical at that position. The similarity between two amino acid or two nucleotide sequences is a direct function of the number of identical positions. In general, the sequences are aligned so that the highest order match is obtained. If necessary, identity can be calculated using

published techniques and widely available computer programs, such as the GCS program package (Devereux et al., Nucleic Acids Res. 12:387, 1984), BLASTP, BLASTN, FASTA (Atschul et al., J. Molecular Biol. 215:403, 1990).

[0033] By "protein variant" or "variant protein" or "variant polypeptide" herein is meant a protein that differs from a wild-type protein by virtue of at least one amino acid modification. The parent polypeptide may be a naturally occurring or wild-type (WT) polypeptide, or may be a modified version of a WT polypeptide. Variant polypeptide may refer to the polypeptide itself, a composition comprising the polypeptide, or the amino sequence that encodes it. Preferably, the variant polypeptide has at least one amino acid modification compared to the parent polypeptide, e.g. from about one to about ten amino acid modifications, and preferably from about one to about five amino acid modifications compared to the parent.

[0034] By "parent polypeptide", "parent protein", "precursor polypeptide", or "precursor protein" as used herein is meant an unmodified polypeptide that is subsequently modified to generate a variant. A parent polypeptide may be a wild-type (or native) polypeptide, or a variant or engineered version of a wild-type polypeptide. Parent polypeptide may refer to the polypeptide itself, compositions that comprise the parent polypeptide, or the amino acid sequence that encodes it.

[0035] The term "amino acid" refers to naturally occurring and synthetic amino acids, as well as amino acid analogs and amino acid mimetics that function in a manner similar to the naturally occurring amino acids. Naturally occurring amino acids are those encoded by the genetic code, as well as those amino acids that are later modified, e.g., hydroxyproline, gamma-carboxyglutamate, and O-phosphoserine. “Amino acid analogs” refers to compounds that have the same basic chemical structure as a naturally occurring amino acid, i.e., an α-carbon that is bound to a hydrogen, a carboxyl group, an amino group, and an R group, e.g., homoserine, norleucine, methionine sulfoxide, methionine methyl sulfonium. Such analogs have modified R groups (e.g., norleucine) or modified peptide backbones, but retain the same basic chemical structure as a naturally occurring amino acid. “Amino acid mimetics” refers to chemical compounds that have a structure that is different from the general chemical structure of an amino acid, but that functions in a manner similar to a naturally occurring amino acid.

[0036] Amino acid modifications disclosed herein may include amino acid substitutions, deletions and insertions, particularly amino acid substitutions. Variant proteins may also include conservative modifications and substitutions at other positions of the cytokine and/or receptor (e.g., positions other than those involved in the affinity engineering). Such conservative substitutions include those described by Dayhoff in The Atlas of Protein Sequence and Structure 5 (1978), and by Argos in EMBO J., 8:779-785 (1989). For example, amino acids belonging to one of the following groups represent conservative

changes: Group I: Ala, Pro, Gly, Gln, Asn, Ser, Thr; Group II: Cys, Ser, Tyr, Thr; Group III: Val, Ile, Leu, Met, Ala, Phe; Group IV: Lys, Arg, His; Group V: Phe, Tyr, Trp, His; and Group VI: Asp, Glu. Further, amino acid substitutions with a designated amino acid may be replaced with a conservative change.

[0037] The term "isolated" refers to a molecule that is substantially free of its natural environment. For instance, an isolated protein is substantially free of cellular material or other proteins from the cell or tissue source from which it is derived. The term refers to preparations where the isolated protein is sufficiently pure to be administered as a therapeutic composition, or at least 70% to 80% (w/w) pure, more preferably, at least 80%-90% (w/w) pure, even more preferably, 90-95% pure; and, most preferably, at least 95%, 96%, 97%, 98%, 99%, or 100% (w/w) pure. A "separated" compound refers to a compound that is removed from at least 90% of at least one component of a sample from which the compound was obtained. Any compound described herein can be provided as an isolated or separated compound.

[0038] The terms "RNA," or "polynucleotide" may refer to a polymer of ribonucleic acids, particularly encoding TCL1A, e.g. human TCL1A. The terms also apply to polymers in which one or more nucleotides are an artificial chemical mimetic of a corresponding naturally occurring nucleotide.

[0039] An RNA may be a naturally or non-naturally occurring RNA, e.g., mRNA. An mRNA may include one or more modified nucleobases, nucleosides, or nucleotides, as described below, in which case it may be referred to as a "chemically modified mRNA", also referred to herein as a "modified mRNA" or "mmRNA." As described herein "nucleoside" is defined as a compound containing a sugar molecule (e.g., a pentose or ribose) or derivative thereof in combination with an organic base (e.g., a purine or pyrimidine) or a derivative thereof (also referred to herein as "nucleobase"). As described herein, "nucleotide" is defined as a nucleoside including a phosphate group.

[0040] An mRNA may include a 5′ untranslated region (5′UTR), a 3′ untranslated region (3′UTR), and/or a coding region (e.g., an open reading frame). An mRNA may include any suitable number of base pairs, including hundreds (e.g., 200, 300, 400, 500, 600, 700, 800, or 900) or thousands (e.g., 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, 10,000) of base pairs. Any number (e.g., all, some, or none) of nucleobases, nucleosides, or nucleotides may be an analog of a canonical species, substituted, modified, or otherwise non-naturally occurring. In certain embodiments, all of a particular nucleotide or nucleobase type may be modified.

[0041] In one embodiment, the mRNA comprises a first flanking region located at the 5′ terminus of an open reading frame (coding region) and a second flanking region located at the 3′ terminus of the open reading frame (coding region), wherein the first flanking region

comprises a 5′ untranslated region (5′ UTR) and the second flanking region comprises a 3′ untranslated region (3′UTR).

[0042] In some embodiments, an mRNA as disclosed herein may comprise a 5′ cap structure, a chain terminating nucleotide, a Kozak sequence (also known as a Kozak consensus sequence), a stem loop, a polyA sequence, and/or a polyadenylation signal. In other embodiments, the mRNA lacks a poly A sequence and/or a polyadenylation signal but rather contains an alternative structure for stabilizing the mRNA.

[0043] A 5′ cap structure or cap species is a compound including two nucleoside moieties joined by a linker and may be selected from a naturally occurring cap, a non-naturally occurring cap or cap analog, or an anti-reverse cap analog (ARCA). A cap species may include one or more modified nucleosides and/or linker moieties. For example, a natural mRNA cap may include a guanine nucleotide and a guanine (G) nucleotide methylated at the 7 position joined by a triphosphate linkage at their 5′ positions, e.g., m^7G(5′)ppp(5′)G, commonly written as m^7GpppG. A cap species may also be an anti-reverse cap analog. A non-limiting list of possible cap species includes m^7GpppG, m^7Gpppm^{7}G, $m^7$3′dGpppG, m_2 $^{7,O3′}$GpppG, m_2 $^{7,O3′}$GppppG, m_2 $^{7,O2′}$GppppG, m^7Gpppm^{7}G, $m^7$3′dGpppG, m_2 $^{7,O3′}$GpppG, m_2 $^{7,O3′}$GppppG, and m_2 $^{7,O2′}$GppppG. In various embodiments, the mRNA can comprise a 5′ terminal cap selected from the group consisting of CapO, CapI, ARCA, inosine, N1-methyl-guanosine, 2′fluoro-guanosine, 7-deaza-guanosine, 8-oxo-guanosine, 2-amino-guanosine, LNA-guanosine, and 2-azido-guanosine. In one embodiment, the 5′ terminal cap is CapI.

[0044] An mRNA may instead or additionally include a chain terminating nucleoside. For example, a chain terminating nucleoside may include those nucleosides deoxygenated at the 2′ and/or 3′ positions of their sugar group. Such species may include 3′-deoxyadenosine (cordycepin), 3′-deoxyuridine, 3′-deoxycytosine, 3′-deoxyguanosine, 3′-deoxythymine, and 2′,3′-dideoxynucleosides, such as 2′,3′-dideoxyadenosine, 2′,3′-dideoxyuridine, 2′,3′-dideoxycytosine, 2′,3′-dideoxyguanosine, and 2′,3′-dideoxythymine. In some embodiments, incorporation of a chain terminating nucleotide into an mRNA, for example at the 3′-terminus may result in stabilization of the mRNA, as described, for example, in International Patent Publication No. WO 2013/103659.

[0045] An mRNA may instead or additionally include a polyA sequence and/or polyadenylation signal. A polyA sequence may be comprised entirely or mostly of adenine nucleotides or analogs or derivatives thereof. A polyA sequence may be a tail located adjacent to a 3′ untranslated region of an mRNA. In some embodiments, a polyA sequence may affect the nuclear export, translation, and/or stability of an mRNA.

[0046] In some embodiments, an mRNA of the disclosure comprises one or more modified nucleobases, nucleosides, or nucleotides (termed “chemically modified mRNAs”, also referred to herein as “modified mRNAs” or “mmRNAs”). The mmRNAs, of the disclosure

can include a combination of modifications to the sugar, the nucleobase, and/or the internucleoside linkage. In some embodiments, modified mRNAs may have useful properties, including enhanced stability, intracellular retention, enhanced translation, and/or the lack of a substantial induction of the innate immune response of a cell into which the mRNA is introduced, as compared to a reference unmodified mRNA. Therefore, use of modified mRNAs may enhance the efficiency of protein production, intracellular retention of nucleic acids, as well as possess reduced immunogenicity.

[0047] In some embodiments, an mRNA, includes one or more (e.g., 1, 2, 3 or 4) different modified nucleobases, nucleosides, or nucleotides. In some embodiments, an mRNA includes one or more (e.g., 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, or more) different modified nucleobases, nucleosides, or nucleotides. In some embodiments, the modified mRNA, may have reduced degradation in a cell into which the mmRN is introduced, relative to a corresponding unmodified mRNA.

[0048] In some embodiments, the modified nucleobase is a modified uracil. Exemplary nucleobases and nucleosides having a modified uracil include pseudouridine (ψ), pyridin-4-one ribonucleoside, 5-aza-uridine, 6-aza-uridine, 2-thio-5-aza-uridine, 2-thio-uridine (s^2U), 4-thio-uridine (s^4U), 4-thio-pseudouridine, 2-thio-pseudouridine, 5-hydroxy-uridine (ho^5U), 5-aminoallyl-uridine, 5-halo-uridine (e.g., 5-iodo-uridineor 5-bromo-uridine), 3-methyl-uridine (m^3U), 5-methoxy-uridine (mo^5U), uridine 5-oxyacetic acid (cmo^5U), uridine 5-oxyacetic acid methyl ester ($mcmo^5U$), 5-carboxymethyl-uridine (cm^5U), 1-carboxymethyl-pseudouridine, 5-carboxyhydroxymethyl-uridine (chm^5U), 5-carboxyhydroxymethyl-uridine methyl ester ($mchm^5U$), 5-methoxycarbonylmethyl-uridine (mcm^5U), 5-methoxycarbonylmethyl-2-thio-uridine (mcm^5s^2U), 5-aminomethyl-2-thio-uridine (nm^5s^2U), 5-methylaminomethyl-uridine (mnm^5U), 5-methylaminomethyl-2-thio-uridine (mnm^5s^2U), 5-methylaminomethyl-2-seleno-uridine (mnm^5se^2U), 5-carbamoylmethyl-uridine (ncm^5U), 5-carboxymethylaminomethyl-uridine ($cmnm^5U$), 5-carboxymethylaminomethyl-2-thio-uridine ($cmnm^5s^2U$), 5-propynyl-uridine, 1-propynyl-pseudouridine, 5-taurinomethyl-uridine ($τm^5U$), 1-taurinomethyl-pseudouridine, 5-taurinomethyl-2-thio-uridine ($τm^5s^2U$), 1-taurinomethyl-4-thio-pseudouridine, 5-methyl-uridine (m^5U, i.e., having the nucleobase deoxythymine), 1-methyl-pseudouridine ($m^1ψ$, 5-methyl-2-thio-uridine ($m^5 s^2U$), 1-methyl-4-thio-pseudouridine ($m^1s^4ψ$), 4-thio-1-methyl-pseudouridine, 3-methyl-pseudouridine ($m^3ψ$), 2-thio-1-methyl-pseudouridine, 1-methyl-1-deaza-pseudouridine, 2-thio-1-methyl-1-deaza-pseudouridine, dihydrouridine (D), dihydropseudouridine, 5,6-dihydrouridine, 5-methyl-dihydrouridine (m^5D), 2-thio-dihydrouridine, 2-thio-dihydropseudouridine, 2-methoxy-uridine, 2-methoxy-4-thio-uridine, 4-methoxy-pseudouridine, 4-methoxy-2-thio-pseudouridine, N1-methyl-pseudouridine, 3-(3-amino-3-carboxypropyl)uridine (acp^3U), 1-methyl-3-(3-amino-3-carboxypropyl)pseudouridine (acp^3 w), 5-(isopentenylaminomethyl)uridine (inm^5U), 5-

(isopentenylaminomethyl)-2-thio-uridine (inm^{5}s^{2}U), α-thio-uridine, 2′-O-methyl-uridine (Um), 5,2′-O-dimethyl-uridine (m^{5}Um), 2′-O-methyl-pseudouridine (ψm), 2-thio-2′-O-methyl-uridine (s^{2}Um), 5-methoxycarbonylmethyl-2′-O-methyl-uridine (mcm^{5}Um), 5-carbamoylmethyl-2′-O-methyl-uridine (ncm^{5}Um), 5-carboxymethylaminomethyl-2′-O-methyl-uridine (cmnm^{5}Um), 3,2′-O-dimethyl-uridine (m^{3}Um), and 5-(isopentenylaminomethyl)-2′-O-methyl-uridine (inm^{5}Um), 1-thio-uridine, deoxythymidine, 2′-F-ara-uridine, 2′-F-uridine, 2′-OH-ara-uridine, 5-(2-carbomethoxyvinyl) uridine, and 5-[3-(1-E-propenylamino)]uridine.

[0049] In some embodiments, the modified nucleobase is a modified cytosine. Exemplary nucleobases and nucleosides having a modified cytosine include 5-aza-cytidine, 6-aza-cytidine, pseudoisocytidine, 3-methyl-cytidine (m^{3}C), N4-acetyl-cytidine (ac^{4}C), 5-formyl-cytidine (f^{5}C), N4-methyl-cytidine (m^{4}C), 5-methyl-cytidine (m^{5}C), 5-halo-cytidine (e.g., 5-iodo-cytidine), 5-hydroxymethyl-cytidine (hm^{5}C), 1-methyl-pseudoisocytidine, pyrrolo-cytidine, pyrrolo-pseudoisocytidine, 2-thio-cytidine (s^{2}C), 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, lysidine (k2C), α-thio-cytidine, 2′-O-methyl-cytidine (Cm), 5,2′-O-dimethyl-cytidine (m^{5}Cm), N4-acetyl-2′-O-methyl-cytidine (ac^{4}Cm), N4,2′-O-dimethyl-cytidine (m^{4}Cm), 5-formyl-2′-O-methyl-cytidine (f^{5}Cm), N4,N4,2′-O-trimethyl-cytidine (m$^4{}_2$Cm), 1-thio-cytidine, 2′-F-ara-cytidine, 2′-F-cytidine, and 2′-OH-ara-cytidine.

[0050] In some embodiments, the modified nucleobase is a modified adenine. Exemplary nucleobases and nucleosides having a modified adenine include α-thio-adenosine, 2-amino-purine, 2, 6-diaminopurine, 2-amino-6-halo-purine (e.g., 2-amino-6-chloro-purine), 6-halo-purine (e.g., 6-chloro-purine), 2-amino-6-methyl-purine, 8-azido-adenosine, 7-deaza-adenine, 7-deaza-8-aza-adenine, 7-deaza-2-amino-purine, 7-deaza-8-aza-2-amino-purine, 7-deaza-2,6-diaminopurine, 7-deaza-8-aza-2,6-diaminopurine, 1-methyl-adenosine (m′A), 2-methyl-adenine (m^{2}A), N6-methyl-adenosine (m^{6}A), 2-methylthio-N6-methyl-adenosine (ms^2 m^{6}A), N6-isopentenyl-adenosine (i^{6}A), 2-methylthio-N6-isopentenyl-adenosine (ms^{2}i^{6}A), N6-(cis-hydroxyisopentenyl)adenosine (io^{6}A), 2-methylthio-N6-(cis-hydroxyisopentenyl)adenosine (ms^{2}io^{6}A), N6-glycinylcarbamoyl-adenosine (g^{6}A), N6-threonylcarbamoyl-adenosine (t^{6}A), N6-methyl-N6-threonylcarbamoyl-adenosine (m^{6}t^{6}A), 2-methylthio-N6-threonylcarbamoyl-adenosine (ms^{2}g^{6}A), N6,N6-dimethyl-adenosine (m$^6{}_2$A), N6-hydroxynorvalylcarbamoyl-adenosine (hn^{6}A), 2-methylthio-N6-hydroxynorvalylcarbamoyl-adenosine (ms^{2}hn^{6}A), N6-acetyl-adenosine (ac^{6}A), 7-methyl-adenine, 2-methylthio-adenine, 2-methoxy-adenine, α-thio-adenosine, 2′-O-methyl-adenosine (Am), N6,2′-O-dimethyl-adenosine (m^{6}Am), N6,N6,2′-

O-trimethyl-adenosine ($m^6{}_2Am$), 1,2′-0-dimethyl-adenosine (m^1Am), 2′-O-ribosyladenosine (phosphate) (Ar(p)), 2-amino-N6-methyl-purine, 1-thio-adenosine, 8-azido-adenosine, 2′-F-ara-adenosine, 2′-F-adenosine, 2′-OH-ara-adenosine, and N6-(19-amino-pentaoxanonadecyl)-adenosine.

[0051] In some embodiments, the modified nucleobase is a modified guanine. Exemplary nucleobases and nucleosides having a modified guanine include α-thio-guanosine, inosine (I), 1-methyl-inosine (m^1I), wyosine (imG), methylwyosine (mimG), 4-demethyl-wyosine (imG-14), isowyosine (imG2), wybutosine (yW), peroxywybutosine (o_2yW), hydroxywybutosine (OhyW), undermodified hydroxywybutosine (OhyW*), 7-deaza-guanosine, queuosine (Q), epoxyqueuosine (oQ), galactosyl-queuosine (galQ), mannosyl-queuosine (manQ), 7-cyano-7-deaza-guanosine ($preQ_0$), 7-aminomethyl-7-deaza-guanosine ($preQ_1$), archaeosine (G^+), 7-deaza-8-aza-guanosine, 6-thio-guanosine, 6-thio-7-deaza-guanosine, 6-thio-7-deaza-8-aza-guanosine, 7-methyl-guanosine (m^7G), 6-thio-7-methyl-guanosine, 7-methyl-inosine, 6-methoxy-guanosine, 1-methyl-guanosine (m^1G), N2-methyl-guanosine (m^2G), N2,N2-dimethyl-guanosine ($m^2{}_2G$), N2,7-dimethyl-guanosine ($m^{2,7}G$), N2, N2,7-dimethyl-guanosine ($m^{2,2,7}G$), (8-oxo-guanosine, 7-methyl-8-oxo-guanosine, 1-methyl-6-thio-guanosine, N2-methyl-6-thio-guanosine, N2,N2-dimethyl-6-thio-guanosine, α-thio-guanosine, 2′-O-methyl-guanosine (Gm), N2-methyl-2′-O-methyl-guanosine (m^2Gm), N2,N2-dimethyl-2′-O-methyl-guanosine ($m^2{}_2Gm$), 1-methyl-2′-O-methyl-guanosine (m^1Gm), N2,7-dimethyl-2′-O-methyl-guanosine ($m^{2,7}Gm$), 2′-O-methyl-inosine (Im), 1,2′-O-dimethyl-inosine (m^1Im), 2′-O-ribosylguanosine (phosphate) (Gr(p)), 1-thio-guanosine, 06-methyl-guanosine, 2′-F-ara-guanosine, and 2′-F-guanosine.

[0052] In some embodiments, an mmRNA, of the disclosure includes a combination of one or more of the aforementioned modified nucleobases (e.g., a combination of 2, 3 or 4 of the aforementioned modified nucleobases).

[0053] In some embodiments, the mmRNA, comprises pseudouridine (ψ). In some embodiments, the mmRNA, comprises pseudouridine (ψ) and 5-methyl-cytidine (m^5C). In some embodiments, the mmRNA, comprises 1-methyl-pseudouridine ($m^1ψ$). In some embodiments, the mmRNA comprises 1-methyl-pseudouridine ($m^1ψ$) and 5-methyl-cytidine (m^5C). In some embodiments, the mmRNA, comprises 2-thiouridine (s^2U). In some embodiments, the mmRNA, comprises 2-thiouridine and 5-methyl-cytidine (m^5C). In some embodiments, the mmRNA, comprises 5-methoxy-uridine (mo^5U). In some embodiments, the RNA, e.g., comprises 5-methoxy-uridine (mo^5U) and 5-methyl-cytidine (m^5C). In some embodiments, the mmRNA, comprises 2′-O-methyl uridine. In some embodiments, the mmRNA, comprises 2′-O-methyl uridine and 5-methyl-cytidine (m^5C). In some embodiments, the mmRNA, comprises N6-methyl-adenosine (m^6A). In some

embodiments, the mmRNA, comprises N6-methyl-adenosine (m^6A) and 5-methyl-cytidine (m^5C).

[0054] In certain embodiments, an mmRNA, of the disclosure is uniformly modified (i.e., fully modified, modified through-out the entire sequence) for a particular modification. In some embodiments, an mRNA of the disclosure may be modified in a coding region (e.g., an open reading frame encoding a polypeptide). In other embodiments, an mRNA may be modified in regions besides a coding region. For example, in some embodiments, a 5′-UTR and/or a 3′-UTR are provided, wherein either or both may independently contain one or more different nucleoside modifications. In such embodiments, nucleoside modifications may also be present in the coding region.

[0055] The mRNAs of the present disclosure, or regions thereof, may be codon optimized. Codon optimization methods are known in the art and may be useful for a variety of purposes: matching codon frequencies in host organisms to ensure proper folding, bias GC content to increase mRNA stability or reduce secondary structures, minimize tandem repeat codons or base runs that may impair gene construction or expression, customize transcriptional and translational control regions, insert or remove proteins trafficking sequences, remove/add post translation modification sites in encoded proteins (e.g., glycosylation sites), add, remove or shuffle protein domains, insert or delete restriction sites, modify ribosome binding sites and mRNA degradation sites, adjust translation rates to allow the various domains of the protein to fold properly, or to reduce or eliminate problem secondary structures within the polynucleotide. Codon optimization tools, algorithms and services are known in the art.

[0056] The mRNAs of the present disclosure may be produced by means available in the art, including but not limited to in vitro transcription (IVT) and synthetic methods. Enzymatic (IVT), solid-phase, liquid-phase, combined synthetic methods, small region synthesis, and ligation methods may be utilized.

[0057] Non-natural modified nucleobases may be introduced into polynucleotides, e.g., mRNA, during synthesis or post-synthesis. In certain embodiments, modifications may be on internucleoside linkages, purine or pyrimidine bases, or sugar. In particular embodiments, the modification may be introduced at the terminal of a polynucleotide chain or anywhere else in the polynucleotide chain; with chemical synthesis or with a polymerase enzyme.

[0058] Either enzymatic or chemical ligation methods may be used to conjugate polynucleotides or their regions with different functional moieties, such as targeting or delivery agents, fluorescent labels, liquids, nanoparticles, etc. Conjugates of polynucleotides and modified polynucleotides are reviewed in Goodchild, Bioconjugate Chemistry, vol. 1(3), 165-187 (1990).

[0059] The mRNAs, of the disclosure may be formulated in nanoparticles or other delivery vehicles, e.g., to protect them from degradation when delivered to a subject. In certain

embodiments, an RNA, e.g., mRNA, of the disclosure is encapsulated within a nanoparticle. In particular embodiments, a nanoparticle is a particle having at least one dimension (e.g., a diameter) less than or equal to 1000 nM, less than or equal to 500 nM or less than or equal to 100 nM. In particular embodiments, a nanoparticle includes a lipid. Lipid nanoparticles include, but are not limited to, liposomes and micelles. Any of a number of lipids may be present, including cationic and/or ionizable lipids, anionic lipids, neutral lipids, amphipathic lipids, PEGylated lipids, and/or structural lipids. Such lipids can be used alone or in combination. In particular embodiments, a lipid nanoparticle comprises one or more RNAs, e.g., mRNAs, described herein, e.g., a mmRNA encoding TCL1A.

[0060] In some embodiments, the lipid nanoparticle formulations of the mRNAs, described herein may include one or more (e.g., 1, 2, 3, 4, 5, 6, 7, or 8) cationic and/or ionizable lipids. Such cationic lipids include, but are not limited to, 3-(didodecylamino)-N1,N1,4-tridodecyl-1-piperazineethanamine (KL10), N1-[2-(didodecylamino)ethyl]-N1,N4,N4-tridodecyl-1,4-piperazinediethanamine (KL22), 14,25-ditridecyl-15,18,21,24-tetraaza-octatriacontane (KL25), 1,2-dilinoleyloxy-N,N-dimethylaminopropane (DLin-DMA), 2,2-dilinoleyl-4-dimethylaminomethyl-[1,3]-dioxolane (DLin-K-DMA), heptatriaconta-6,9,28,31-tetraen-19-yl 4-(dimethylamino)butanoate (DLin-MC3-DMA), 2,2-dilinoleyl-4-(2-dimethylaminoethyl)[1,3]-dioxolane (DLin-KC2-DMA), 2-({8-[(3β)-cholest-5-en-3-yloxy]octyl}oxy)-N,N-dimethyl-3-[(9Z,12Z)-octadeca-9,12-dien-1-yloxy]propan-1-amine (Octyl-CLinDMA), (2R)-2-({8-[(3 (3)-cholest-5-en-3-yloxy]octyl}oxy)-N,N-dimethyl-3-[(9Z,12Z)-octadeca-9,12-dien-1-yloxy]propan-1-amine (Octyl-CLinDMA (2R)), (25)-2-({8-[(3 (3)-cholest-5-en-3-yloxy]octyl}oxy)-N,N-dimethyl-3-[(9Z,12Z)-octadeca-9,12-dien-1-yloxy]propan-1-amine (Octyl-CLinDMA (2S)).N,N-dioleyl-N,N-dimethylammonium chloride ("DODAC"); N-(2,3-dioleyloxy)propyl-N,N—N-triethylammonium chloride ("DOTMA"); N,N-distearyl-N,N-dimethylammonium bromide ("DDAB"); N-(2,3-dioleoyloxy)propyl)-N,N,N-trimethylammonium chloride ("DOTAP"); 1,2-Dioleyloxy-3-trimethylaminopropane chloride salt ("DOTAP.Cl"); 3-β-(N—(N',N'-dimethylaminoethane)-carbamoyl)cholesterol ("DC-Chol"), N-(1-(2,3-dioleyloxy)propyl)-N-2-(sperminecarboxamido)ethyl)-N,N-dimethyl-ammonium trifluoracetate ("DOSPA"), dioctadecylamidoglycyl carboxyspermine ("DOGS"), 1,2-dioleoyl-3-dimethylammonium propane ("DODAP"), N,N-dimethyl-2,3-dioleyloxy)propylamine ("DODMA"), and N-(1,2-dimyristyloxyprop-3-yl)-N,N-dimethyl-N-hydroxyethyl ammonium bromide ("DMRIE"). Additionally, a number of commercial preparations of cationic and/or ionizable lipids can be used, such as, e.g., LIPOFECTIN® (including DOTMA and DOPE, available from GIBCO/BRL), and LIPOFECTAMINE® (including DOSPA and DOPE, available from GIBCO/BRL). KL10, KL22, and KL25 are described, for example, in U.S. Pat. No.

8,691,750, which is incorporated herein by reference in its entirety. In particular embodiments, the lipid is DLin-MC3-DMA or DLin-KC2-DMA.

[0061] Anionic lipids suitable for use in lipid nanoparticles of the disclosure include, but are not limited to, phosphatidylglycerol, cardiolipin, diacylphosphatidylserine, diacylphosphatidic acid, N-dodecanoyl phosphatidylethanoloamine, N-succinyl phosphatidylethanolamine, N-glutaryl phosphatidylethanolamine, lysylphosphatidylglycerol, and other anionic modifying groups joined to neutral lipids.

[0062] Neutral lipids suitable for use in lipid nanoparticles of the disclosure include, but are not limited to, diacylphosphatidylcholine, diacylphosphatidylethanolamine, ceramide, sphingomyelin, dihydrosphingomyelin, cephalin, and cerebrosides. Lipids having a variety of acyl chain groups of varying chain length and degree of saturation are available or may be isolated or synthesized by well-known techniques. Additionally, lipids having mixtures of saturated and unsaturated fatty acid chains can be used. In some embodiments, the neutral lipids used in the disclosure are DOPE, DSPC, DPPC, POPC, or any related phosphatidylcholine. In some embodiments, the neutral lipid may be composed of sphingomyelin, dihydrosphingomyeline, or phospholipids with other head groups, such as serine and inositol.

[0063] In some embodiments, amphipathic lipids are included in nanoparticles of the disclosure. Exemplary amphipathic lipids suitable for use in nanoparticles of the disclosure include, but are not limited to, sphingolipids, phospholipids, and aminolipids. In some embodiments, a phospholipid is selected from the group consisting of 1,2-dilinoleoyl-sn-glycero-3-phosphocholine (DLPC), 1,2-dimyristoyl-sn-glycero-phosphocholine (DMPC), 1,2-dioleoyl-sn-glycero-3-phosphocholine (DOPC), 1,2-dipalmitoyl-sn-glycero-3-phosphocholine (DPPC), 1,2-distearoyl-sn-glycero-3-phosphocholine (DSPC), 1,2-diundecanoyl-sn-glycero-phosphocholine (DUPC), 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC), 1,2-di-O-octadecenyl-sn-glycero-3-phosphocholine (18:0 Diether PC), 1-oleoyl-2-cholesterylhemisuccinoyl-sn-glycero-3-phosphocholine (OChemsPC), 1-hexadecyl-sn-glycero-3-phosphocholine (C16 Lyso PC), 1,2-dilinolenoyl-sn-glycero-3-phosphocholine, 1,2-diarachidonoyl-sn-glycero-3-phosphocholine, 1,2-didocosahexaenoyl-sn-glycero-3-phosphocholine, 1,2-dioleoyl-sn-glycero-3-phosphoetha nolamine (DOPE), 1,2-diphytanoyl-sn-glycero-3-phosphoethanolamine (ME 16.0 PE), 1,2-distearoyl-sn-glycero-3-phosphoethanolamine, 1,2-dilinoleoyl-sn-glycero-3-phosphoethanolamine, 1,2-dilinolenoyl-sn-glycero-3-phosphoethanolamine, 1,2-diarachidonoyl-sn-glycero-3-phosphoethanolamine, 1,2-didocosahexaenoyl-sn-glycero-3-phosphoethanolamine, 1,2-dioleoyl-sn-glycero-3-phospho-rac-(1-glycerol) sodium salt (DOPG), and sphingomyelin. Other phosphorus-lacking compounds, such as sphingolipids, glycosphingolipid families, diacylglycerols,

and β-acyloxyacids, may also be used. Additionally, such amphipathic lipids can be readily mixed with other lipids, such as triglycerides and sterols.

[0064] In some embodiments, the lipid component of a nanoparticle of the disclosure may include one or more PEGylated lipids. A PEGylated lipid (also known as a PEG lipid or a PEG-modified lipid) is a lipid modified with polyethylene glycol. The lipid component may include one or more PEGylated lipids. A PEGylated lipid may be selected from the non-limiting group consisting of PEG-modified phosphatidylethanolamines, PEG-modified phosphatidic acids, PEG-modified ceramides, PEG-modified di alkyl amines, PEG-modified diacylglycerols, and PEG-modified dialkylglycerols. For example, a PEGylated lipid may be PEG-c-DOMG, PEG-DMG, PEG-DLPE, PEG-DMPE, PEG-DPPC, or a PEG-DSPE lipid.

[0065] A lipid nanoparticle of the disclosure may include one or more structural lipids. Exemplary, non-limiting structural lipids that may be present in the lipid nanoparticles of the disclosure include cholesterol, fecosterol, sitosterol, campesterol, stigmasterol, brassicasterol, ergosterol, tomatidine, tomatine, ursolic acid, or alpha-tocopherol).

[0066] In certain embodiments, it is desirable to target a nanoparticle, e.g., a lipid nanoparticle, of the disclosure using a targeting moiety that is specific to a cell type and/or tissue type. In some embodiments, a nanoparticle may be targeted to a particular cell, tissue, and/or organ using a targeting moiety. In particular embodiments, a nanoparticle comprises one or more mRNA described herein and a targeting moiety. Exemplary non-limiting targeting moieties include ligands, cell surface receptors, glycoproteins, vitamins (e.g., riboflavin) and antibodies (e.g., full-length antibodies, antibody fragments (e.g., Fv fragments, single chain Fv (scFv) fragments, Fab′ fragments, or F(ab′)2 fragments), single domain antibodies, camelid antibodies and fragments thereof, human antibodies and fragments thereof, monoclonal antibodies, and multispecific antibodies (e.g., bispecific antibodies)). In some embodiments, the targeting moiety may be a polypeptide. The targeting moiety may include the entire polypeptide (e.g., peptide or protein) or fragments thereof. A targeting moiety is typically positioned on the outer surface of the nanoparticle in such a manner that the targeting moiety is available for interaction with the target, for example, a cell surface receptor. A variety of different targeting moieties and methods are known and available in the art.

[0067] In some embodiments, a lipid nanoparticle (e.g., a liposome) may include a surface coating of hydrophilic polymer chains, such as polyethylene glycol (PEG) chains. In one approach, a targeting moiety for targeting the lipid nanoparticle is linked to the polar head group of lipids forming the nanoparticle. In another approach, the targeting moiety is attached to the distal ends of the PEG chains forming the hydrophilic polymer.

[0068] In some embodiments, a lipid nanoparticle of the disclosure includes a targeting moiety that targets the lipid nanoparticle to a hematopoietic stem cell, e.g. to CD117, CD34, etc..

Stem cells

[0069] The term stem cell is used herein to refer to a mammalian cell that has the ability both to self-renew, and to generate differentiated progeny (see Morrison et al. (1997) Cell 88:287-298). Generally, stem cells also have one or more of the following properties: an ability to undergo asynchronous, or symmetric replication, that is where the two daughter cells after division can have different phenotypes; extensive self-renewal capacity; capacity for existence in a mitotically quiescent form; and clonal regeneration of all the tissue in which they exist, for example the ability of hematopoietic stem cells to reconstitute all hematopoietic lineages.

[0070] "Hematopoietic stem cell," as used herein, refers to a pluripotent cell that is capable of both differentiating into all of the lineages of blood cells and self-renewing. As used herein, "self-renewal" refers to the process of producing at least one daughter during replication and division that has essentially the same pluripotency development potential as the parent cell. Self-renewal is one aspect of proliferation; the other aspect is the production of a cell committed to differentiation. Self-renewal of HSCs is critical to an organism's ability to have sufficient blood cells during its lifetime by sustaining a sufficient reservoir of pluripotent stem cell. As used herein, "enhancing self-renewal" encompasses producing more stem cells and/or producing them faster, compared to self-renewal under the same conditions and in the absence of genetic modification disclosed herein.

[0071] The skilled artisan will appreciate that the HSCs having enhanced self-renewal and proliferation as a result of the methods of the invention have numerous uses. They may be proliferated, ex vivo or in vivo, to provide an increased number of HSCs. They may be used therapeutically, e.g., either immediately after proliferation or after a period of storage. The HSCs may also be induced to differentiate, either ex vivo, using methods known to the skilled artisan, or in vivo, in response to endogenous differentiation signals. The HSCs may also be used ex vivo or in vivo in research applications.

[0072] CD34 is a marker of human HSC, and all colony-forming activity of human bone marrow (BM) cells is found in the $CD34^+$ fraction. Clinical transplantation studies that used enriched $CD34^+$ BM cells indicated the presence of HSC with long-term BM reconstitutional ability within this fraction. Other markers that can be used for positive or negative selection of HSC include, without limitation, CD117 and/or CD90 for positive selection, and lineage specific markers for negative selection, e.g. a lineage cocktail may comprise, without limitation, one or more of antibodies specific for CD45, CD3, CD4, CD8, MAC-1, TER-119 and Gr-1/Ly-6G.

[0073] In one embodiment of the invention, the stem cells are one or more of autologous hematopoietic stem cells, genetically modified hematopoietic stem cells, and allogeneic hematopoietic stem cells, for example and without limitation allogeneic or genetically

modified autologous cells. Such stem cells find use in the treatment of a variety of blood disorders, *e.g.* genetic disorders including aplastic anemia; sickle cell disease; thalassemias; severe immunodeficiency; bone marrow failure states, immune deficiencies, hemoglobinopathies, leukemias, lymphomas, immune-tolerance induction, genetic disorders treatable by bone marrow transplantation and other blood disorders, and the like. Allogeneic stem cells find use, for example and without limitation, in the treatment of hematologic malignancies, i.e. cancers and myelodysplastic syndromes, e.g. AML, MDS, CMML, multiple myeloma, CML, NHL, and the like, or non-malignant genetic disorders treatable by bone marrow transplantation such as cell disease; thalassemias; severe immunodeficiency, neurologic disorders and the like. The methods of the invention are also useful in the induction of tolerance in a patient, for example tolerance to donor tissue, *e.g.* in organ transplants; tolerance to autoantigens, *e.g.* in the context of treatment of autoimmune disease; and the like.

[0074] Hematopoietic stem cells can be obtained by harvesting from fetal liver, umbilical cord, bone marrow, peripheral blood, etc. Bone marrow is generally aspirated from the posterior iliac crests while the donor is under either regional or general anesthesia. Additional bone marrow can be obtained from the anterior iliac crest. A dose of 1 x 10^8 and 2 x 10^8 marrow mononuclear cells per kilogram is generally considered desirable to establish engraftment in autologous and allogeneic marrow transplants, respectively. Bone marrow can be primed with granulocyte colony-stimulating factor (G-CSF; filgrastim [Neupogen]) to increase the stem cell count. Reference to "whole bone marrow" for the purposes described herein generally refers to a composition of mononuclear cells derived from bone marrow that have not been selected for specific immune cell subsets. "Fractionated bone marrow" may be, for example, depleted of T cells, e.g. $CD8^+$ cells, $CD52^+$ cells, $CD3^+$ cells, etc.; enriched for CD34+ cells, etc.

[0075] Hematopoietic stem cells are also obtained from cord blood. Cord blood is an almost unlimited source of hematopoietic stem cells for allogeneic hematopoietic stem cell transplant. Cord blood banks (CBB) have been established for related or unrelated UCBT with more than 400,000 units available and more than 20,000 umbilical cord blood transplants performed in children and in adults. UCB hematopoietic progenitors are enriched in primitive stem/progenitor cells able to produce in vivo long-term repopulating stem cells. However, the number of cells available from any single donor can be relatively low in comparison with other sources.

[0076] Mobilization of stem cells from the bone marrow into peripheral blood by cytokines such as G-CSF or GM-CSF has led to the widespread adoption of peripheral blood progenitor cell collection by apheresis for hematopoietic stem cell transplantation. The dose of G-CSF used for mobilization is 10 μg/kg/day. In autologous donors who are heavily

pretreated, however, doses of up to 40 μg/kg/day can be given. Mozobil may be used in conjunction with G-CSF to mobilize hematopoietic stem cells to peripheral blood for collection.

[0077] Hematopoietic stem cells can also be generated in vitro, for example from pluripotent embryonic stem cells, induced pluripotent cells, and the like. For example, see Sugimura et al. (2017) Nature 545:432-438, herein specifically incorporated by reference, which details a protocol for generation of hematopoietic progenitors.

[0078] Cells may be collected from a subject or a donor may be separated from a mixture of cells by techniques that enrich for desired cells, or may be engineered and cultured without separation. An appropriate solution may be used for dispersion or suspension. Such solution will generally be a balanced salt solution, *e.g.* normal saline, PBS, Hank's balanced salt solution, *etc.*, conveniently supplemented with fetal calf serum or other naturally occurring factors, in conjunction with an acceptable buffer at low concentration, generally from 5-25 mM. Convenient buffers include HEPES, phosphate buffers, lactate buffers, *etc.*

[0079] Techniques for affinity separation may include magnetic separation, using antibody-coated magnetic beads, affinity chromatography, cytotoxic agents joined to a monoclonal antibody or used in conjunction with a monoclonal antibody, *e.g.,* complement and cytotoxic cells, and "panning" with antibody attached to a solid matrix, *e.g., a* plate, or other convenient technique. Techniques providing accurate separation include fluorescence activated cell sorters, which can have varying degrees of sophistication, such as multiple color channels, low angle and obtuse light scattering detecting channels, impedance channels, *etc.* The cells may be selected against dead cells by employing dyes associated with dead cells (*e.g.,* propidium iodide). Any technique may be employed which is not unduly detrimental to the viability of the selected cells. The affinity reagents may be specific receptors or ligands for the cell surface molecules indicated above. In addition to antibody reagents, peptide-MHC antigen and T cell receptor pairs may be used; peptide ligands and receptor; effector and receptor molecules, and the like.

[0080] The separated cells may be collected in any appropriate medium that maintain the viability of the cells, usually having a cushion of serum at the bottom of the collection tube. Various media are commercially available and may be used according to the nature of the cells, including dMEM, HBSS, dPBS, RPMI, Iscove's medium, *etc.*, frequently supplemented with fetal calf serum (FCS).

[0081] The HSC composition may be at least about 50% pure, as defined by the percentage of cells that are CD34+ in the population, may be at least about 75% pure, at least about 85% pure, at least about 95% pure, or more. The collected and optionally enriched cell population may be used immediately for genetic modification, or may be frozen at liquid

nitrogen temperatures and stored, being thawed and capable of being reused. The cells will usually be stored in 10% DMSO, 50% FCS, 40% RPMI 1640 medium.

Genetic Modification of Cells

[0082] HSC are altered in order to increase expression of TCL1A, e.g. by introducing mRNA or a coding sequence of TCL1A. The cells may additionally be modified to introduce genes useful in the differentiated cell, *e.g.* repair of a genetic defect in an individual, selectable marker, etc., or genes useful in selection against undifferentiated ES cells. For example, programmable gene editing tools such as CRISPR/cas9 and the like can be used to edit genomes. Cells may also be genetically modified to correct genetic defects, enhance survival, control proliferation, competitiveness, and the like. Cells may be genetically altering by transfection or transduction with a suitable vector, homologous recombination, or other appropriate technique, so that they express a gene of interest. In other embodiments, a selectable marker is introduced, to provide for greater purity of the desired differentiating cell. Cells may be genetically altered using vector containing supernatants over an 8-16 h period, and then exchanged into growth medium for 1-2 days. Genetically altered cells are selected using a drug selection agent such as puromycin, G418, or blasticidin, and then recultured.

[0083] The cells can also be genetically altered in order to enhance their ability to be involved in tissue regeneration, or to deliver a therapeutic gene to a site of administration. A vector is designed using the known encoding sequence for the desired gene, operatively linked to a promoter that is constitutive, pan-specific, specifically active in a differentiated cell type, *etc.* Suitable inducible promoters are activated in a desired target cell type, either the transfected cell, or progeny thereof. By transcriptional activation, it is intended that transcription will be increased above basal levels in the target cell by at least about 100-fold, more usually by at least about 1000 fold. Various promoters are known that are induced in different cell types.

[0084] In some cases, a cell is modified by a class 2 CRISPR/Cas effector protein (or a nucleic encoding the protein), e.g., as an endonuclease. In class 2 CRISPR systems, the functions of the effector complex (e.g., the cleavage of target DNA) are carried out by a single protein (which can be referred to as a CRISPR/Cas effector protein) – where the natural protein is an endonuclease (e.g., see Zetsche et al, Cell. 2015 Oct 22;163(3):759-71; Makarova et al, Nat Rev Microbiol. 2015 Nov;13(11):722-36; Shmakov et al., Mol Cell. 2015 Nov 5;60(3):385-97; and Shmakov et al., Nat Rev Microbiol. 2017 Mar;15(3):169-182: "Diversity and evolution of class 2 CRISPR-Cas systems"). As such, the term "class 2 CRISPR/Cas protein" or "CRISPR/Cas effector protein" is used herein to encompass the effector protein from class 2 CRISPR systems – for example, type II CRISPR/Cas proteins (e.g., Cas9), type V CRISPR/Cas proteins (e.g., Cpf1/Cas12a, C2c1/Cas12b,

C2C3/Cas12c), and type VI CRISPR/Cas proteins (e.g., C2c2/Cas13a, C2C7/Cas13c, C2c6/Cas13b). Class 2 CRISPR/Cas effector proteins include type II, type V, and type VI CRISPR/Cas proteins, but the term is also meant to encompass any class 2 CRISPR/Cas protein suitable for binding to a corresponding guide RNA and forming a ribonucleoprotein (RNP) complex.

[0085] In some cases, an RNA-guided endonuclease is a fusion protein that is fused to a heterologous polypeptide (also referred to as a "fusion partner"). In some cases, an RNA-guided endonuclease is fused to an amino acid sequence (a fusion partner) that provides for subcellular localization, i.e., the fusion partner is a subcellular localization sequence (e.g., one or more nuclear localization signals (NLSs) for targeting to the nucleus, two or more NLSs, three or more NLSs, etc.). In some embodiments, an RNA-guided endonuclease is fused to an amino acid sequence (a fusion partner) that provides a tag (i.e., the fusion partner is a detectable label) for ease of tracking and/or purification (e.g., a fluorescent protein, e.g., green fluorescent protein (GFP), YFP, RFP, CFP, mCherry, tdTomato, and the like; a histidine tag, e.g., a 6XHis tag; a hemagglutinin (HA) tag; a FLAG tag; a Myc tag; and the like). In some cases, the fusion partner can provide for increased or decreased stability (i.e., the fusion partner can be a stability control peptide, e.g., a degron, which in some cases is controllable (e.g., a temperature sensitive or drug controllable degron sequence).

[0086] An RNA-guided endonuclease (e.g., a Cas9 protein) can have multiple (1 or more, 2 or more, 3 or more, etc.) fusion partners in any combination of the above. As an illustrative example, an RNA-guided endonuclease (e.g., a Cas9 protein) can have a fusion partner that provides for tagging (e.g., GFP), and can also have a subcellular localization sequence (e.g., one or more NLSs). In some cases, such a fusion protein might also have a tag for ease of tracking and/or purification (e.g., a histidine tag, e.g., a 6XHis tag; a hemagglutinin (HA) tag; a FLAG tag; a Myc tag; and the like). As another illustrative example, an RNA-guided endonuclease (e.g., a Cas9 protein) can have one or more NLSs (e.g., two or more, three or more, four or more, five or more, 1, 2, 3, 4, or 5 NLSs). In some cases a fusion partner (or multiple fusion partners, e.g., 1, 2, 3, 4, or 5 NLSs) (e.g., an NLS, a tag, a fusion partner providing an activity, etc.) is located at or near the C-terminus of the RNA-guided endonuclease (e.g., Cas9 protein). In some cases a fusion partner (or multiple fusion partners, e.g., 1, 2, 3, 4, or 5 NLSs) (e.g., an NLS, a tag, a fusion partner providing an activity, etc.) is located at the N-terminus of the RNA-guided endonuclease (e.g., Cas9 protein). In some cases the genome editing nuclease (e.g., Cas9 protein) has a fusion partner (or multiple fusion partners, e.g., 1, 2, 3, 4, or 5 NLSs)(e.g., an NLS, a tag, a fusion partner providing an activity, etc.) at both the N-terminus and C-terminus.

[0087] Other vectors useful for transferring exogenous genes or mRNAs into target mammalian cells are available. The vectors may be episomal, *e.g.* plasmids, virus derived vectors such cytomegalovirus, adenovirus, *etc.*, or may be integrated into the target cell genome, through homologous recombination or random integration, *e.g.* retrovirus derived vectors such MMLV, HIV-1, ALV, *etc.* For modification of stem cells, lentiviral vectors are preferred. Lentiviral vectors such as those based on HIV or FIV gag sequences can be used to transfect non-dividing cells, such as the resting phase of human stem cells. Combinations of retroviruses and an appropriate packaging line may also find use, where the capsid proteins will be functional for infecting the target cells. Usually, the cells and virus will be incubated for at least about 24 hours in the culture medium. The cells are then allowed to grow in the culture medium for short intervals in some applications, *e.g.* 24-73 hours, or for at least two weeks, and may be allowed to grow for five weeks or more, before analysis. Commonly used retroviral vectors are “defective”, *i.e.* unable to produce viral proteins required for productive infection. Replication of the vector requires growth in the packaging cell line. The vectors may include genes that must later be removed, *e.g.* using a recombinase system such as Cre/Lox, or the cells that express them destroyed, *e.g.* by including genes that allow selective toxicity such as herpesvirus TK, bcl-xs, *etc.*

[0088] Nucleic acids are "operably linked" when placed into a functional relationship with another nucleic acid sequence. For example, DNA for a signal sequence is operably linked to DNA for a polypeptide if it is expressed as a preprotein that signals the secretion of the polypeptide; a promoter or enhancer is operably linked to a coding sequence if it affects the transcription of the sequence; and a ribosome binding site is operably linked to a coding sequence if it is positioned so as to facilitate translation. Generally, "operably linked" means that the DNA sequences being linked are contiguous, and, in the case of a secretory leader, contiguous and in reading phase. However, enhancers do not have to be contiguous.

[0089] Expression vectors will contain a promoter that is recognized by the host organism and is operably linked to the coding sequence. Promoters are untranslated sequences located upstream (5') to the start codon of a structural gene (generally within about 100 to 1000 bp) that control the transcription and translation of particular nucleic acid sequence to which they are operably linked. Such promoters typically fall into two classes, inducible and constitutive. Inducible promoters are promoters that initiate increased levels of transcription from DNA under their control in response to some change in culture conditions, e.g., the presence or absence of a nutrient or a change in temperature. A large number of promoters recognized by a variety of potential host cells are well known.

[0090] Transcription from vectors in mammalian host cells may be controlled, for example, by promoters obtained from the genomes of viruses such as polyoma virus, fowlpox virus,

adenovirus (such as Adenovirus 2), bovine papilloma virus, avian sarcoma virus, cytomegalovirus, a retrovirus LTR (such as murine stem cell virus), hepatitis-B virus and Simian Virus 40 (SV40), from heterologous mammalian promoters, e.g., the actin promoter, PGK (phosphoglycerate kinase), or an immunoglobulin promoter, or from heat-shock promoters, provided such promoters are compatible with the host cell systems. The early and late promoters of the SV40 virus are conveniently obtained as an SV40 restriction fragment that also contains the SV40 viral origin of replication.

[0091] Transcription by higher eukaryotes may be increased by inserting an enhancer sequence into the vector. Enhancers are cis-acting elements of DNA, usually about from 10 to 300 bp in length, which act on a promoter to increase its transcription. Enhancers are relatively orientation and position independent, having been found 5' and 3' to the transcription unit, within an intron, as well as within the coding sequence itself. Many enhancer sequences are now known from mammalian genes (globin, elastase, albumin, α-fetoprotein, and insulin). Typically, however, one will use an enhancer from a eukaryotic virus. Examples include the SV40 enhancer on the late side of the replication origin, the cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers. The enhancer may be spliced into the expression vector at a position 5' or 3' to the coding sequence, but is preferably located at a site 5' from the promoter.

[0092] Expression vectors for use in eukaryotic host cells will also contain sequences necessary for the termination of transcription and for stabilizing the mRNA. Such sequences are commonly available from the 5' and, occasionally 3', untranslated regions of eukaryotic or viral DNAs or cDNAs. Construction of suitable vectors containing one or more of the above-listed components employs standard techniques.

[0093] Cells may be cultured in conventional nutrient media modified as appropriate for inducing promoters, selecting transformants, or amplifying the genes encoding the desired sequences. Mammalian host cells may be cultured in a variety of media. Commercially available media such as Ham's F10 (Sigma), Minimal Essential Medium ((MEM), Sigma), RPMI 1640 (Sigma), and Dulbecco's Modified Eagle's Medium ((DMEM), Sigma) are suitable for culturing the host cells. Any of these media may be supplemented as necessary with hormones and/or other growth factors (such as insulin, transferrin, or epidermal growth factor), salts (such as sodium chloride, calcium, magnesium, and phosphate), buffers (such as HEPES), nucleosides (such as adenosine and thymidine), antibiotics, trace elements, and glucose or an equivalent energy source. Any other necessary supplements may also be included at appropriate concentrations that would be known to those skilled in the art. The culture conditions, such as temperature, pH and

the like, are those previously used with the host cell selected for expression, and will be apparent to the ordinarily skilled in the art.

Engraftment of HSC

[0094] The terms "subject," "individual," and "patient" are used interchangeably herein to refer to a mammal being assessed for treatment and/or being treated. In some embodiments, the mammal is a human. The terms "subject," "individual," and "patient" encompass, without limitation, individuals having a disease. Subjects may be human, but also include other mammals, particularly those mammals useful as laboratory models for human disease, e.g., mice, rats, etc.

[0095] The term "sample" with reference to a patient encompasses blood and other liquid samples of biological origin, solid tissue samples such as a biopsy specimen or tissue cultures or cells derived therefrom and the progeny thereof. The term also encompasses samples that have been manipulated in any way after their procurement, such as by treatment with reagents; washed; or enrichment for certain cell populations, such as diseased cells. The definition also includes samples that have been enriched for particular types of molecules, e.g., nucleic acids, polypeptides, etc. The term "biological sample" encompasses a clinical sample, and also includes tissue obtained by surgical resection, tissue obtained by biopsy, cells in culture, cell supernatants, cell lysates, tissue samples, organs, bone marrow, blood, plasma, serum, and the like. A "biological sample" includes a sample obtained from a patient's diseased cell, e.g., a sample comprising polynucleotides and/or polypeptides that is obtained from a patient's diseased cell (e.g., a cell lysate or other cell extract comprising polynucleotides and/or polypeptides); and a sample comprising diseased cells from a patient. A biological sample comprising a diseased cell from a patient can also include non-diseased cells.

[0096] The term "diagnosis" is used herein to refer to the identification of a molecular or pathological state, disease or condition in a subject, individual, or patient.

[0097] The term "prognosis" is used herein to refer to the prediction of the likelihood of death or disease progression, including recurrence, spread, and drug resistance, in a subject, individual, or patient. The term "prediction" is used herein to refer to the act of foretelling or estimating, based on observation, experience, or scientific reasoning, the likelihood of a subject, individual, or patient experiencing a particular event or clinical outcome. In one example, a physician may attempt to predict the likelihood that a patient will survive.

[0098] As used herein, the terms "treatment," "treating," and the like, refer to administering an agent, or carrying out a procedure, for the purposes of obtaining an effect on or in a subject, individual, or patient. The effect may be prophylactic in terms of completely or partially preventing a disease or symptom thereof and/or may be therapeutic in terms of effecting a partial or complete cure for a disease and/or symptoms of the disease.

"Treatment," as used herein, may include (a) inhibiting the disease, i.e., arresting its development; and (b) relieving the disease or its symptoms, i.e., causing regression of the disease or its symptoms.

[0099] Treating may refer to any indicia of success in the treatment or amelioration or prevention of a disease, including any objective or subjective parameter such as abatement; remission; diminishing of symptoms or making the disease condition more tolerable to the patient; slowing in the rate of degeneration or decline; or making the final point of degeneration less debilitating. The treatment or amelioration of symptoms can be based on objective or subjective parameters; including the results of an examination by a physician. Accordingly, the term "treating" includes the administration of engineered cells to prevent or delay, to alleviate, or to arrest or inhibit development of the symptoms or conditions associated with disease or other diseases. The term "therapeutic effect" refers to the reduction, elimination, or prevention of the disease, symptoms of the disease, or side effects of the disease in the subject.

[00100] As used herein, a "therapeutically effective amount" refers to that amount of the therapeutic agent or cells sufficient to treat or manage a disease or disorder. A therapeutically effective amount may also refer to the amount of the therapeutic agent or cells that provides a therapeutic benefit in the treatment or management of a disease. Further, a therapeutically effective amount with respect to a therapeutic agent or cells of the invention means the amount of therapeutic agent or cells alone, or in combination with other therapies, that provides a therapeutic benefit in the treatment or management of a disease.

[00101] As used herein, the term "dosing regimen" refers to a set of unit doses (typically more than one) that are administered individually to a subject, typically separated by periods of time. In some embodiments, a given therapeutic agent has a recommended dosing regimen, which may involve one or more doses. In some embodiments, a dosing regimen comprises a plurality of doses each of which are separated from one another by a time period of the same length; in some embodiments, a dosing regimen comprises a plurality of doses and at least two different time periods separating individual doses. In some embodiments, all doses within a dosing regimen are of the same unit dose amount. In some embodiments, different doses within a dosing regimen are of different amounts. In some embodiments, a dosing regimen comprises a first dose in a first dose amount, followed by one or more additional doses in a second dose amount different from the first dose amount. In some embodiments, a dosing regimen comprises a first dose in a first dose amount, followed by one or more additional doses in a second dose amount same as the first dose amount. In some embodiments, a dosing regimen is correlated with a desired or beneficial outcome when administered across a relevant population.

[00102] For engraftment of HSC, the host may be subjected to a conditioning regimen, as known in the art. The preparative or conditioning regimen can be a critical element in hematopoietic cell transplantation (HCT). In a successful transplantation, clearance of bone-marrow niches must be achieved for donor hematopoietic stem cell (HSC) to engraft. The preparative regimen may also provide immunosuppression sufficient to prevent rejection of transplanted genetically disparate grafts, and to eradicate the disease for which the transplantation is being performed. Current methods to clear niche space may utilize radiation and/or chemotherapy and or antibody-based ablation. Myeloablative regimens can be classified as radiation-containing or non – radiation-containing regimens:, therapies that were developed by escalating the dose of radiation or of a particular drug to the maximally tolerated dose. Total-body irradiation and cyclophosphamide or busulfan and cyclophosphamide are the commonly used myeloablative therapies. These regimens are especially used in aggressive malignancies, such as leukemias.

[00103] Reduced intensity conditioning (RIC) or non-myeloablative regimens are also used to obtain engraftment of HSC allowing HCT to be used for a broader array of patients. However, these regimens, while less intense still rely on radiation and/or chemotherapy to achieve engraftment.

[00104] In some embodiments, an individual is conditioned for HSC engraftment with an effective dose of an antibody specific for CD117, which may be combined with additional agents as known in the art, for example as disclosed in Srikanthan et al. Mol Ther Methods Clin Dev. 2020 Feb 8;17:455-464; Pang et al. Blood. 2019 May 9;133(19):2069-2078; Czechowicz et al. Nat Commun. 2019 Feb 6;10(1):617; Li et al. Nat Commun. 2019 Feb 6;10(1):616; Kwon et al. Blood. 2019 May 9;133(19):2104-2108; Devadasan et al. Biol Blood Marrow Transplant. 2018 Aug;24(8):1554-1562; Arai et al. Mol Ther. 2018 May 2;26(5):1181-1197; Yokoi et al. Mol Genet Metab. 2016 Nov;119(3):232-238; Czechowicz et al. Science. 2007 Nov 23;318(5854):1296-9.

[00105] Genetically or mRNA modified cells for transplantation may enriched for CD34+ hematopoietic stem cells. In some embodiments the donor cells are HLA-matched. In some embodiments the donor cells are haplotype matched. In some embodiments the donor cells are autologous, including without limitation genetically corrected autologous cells. In some embodiments the donor cells are mobilized peripheral blood cells; in other embodiments the donor cells are bone marrow cells. The dose of cells is at least about 3 x 10^5 $CD34^+$ cells/kg, at least about 5 x 10^5 $CD34^+$ cells/kg, at least about 10^6 $CD34^+$ cells/kg. Higher doses, if available, are generally not deleterious, with the proviso that not more than about 3 x 10^4 $CD3^+$ cells/kg are administered.

[00106] In some embodiments, success of the procedure is monitored by determining the presence of donor-derived myeloid cells, including without limitation, $CD15^+$ cells, in

circulation of the recipient. Blood myeloid chimerism is an indicator of true HSC engraftment due to the short-lived nature of myeloid cells. After about 8 weeks post-HCT, there can be measurable and sustained levels of blood myeloid chimerism, e.g. of at least about 1% donor type $CD15^+$ cells, at least about 2% donor type $CD15^+$ cells, at least about 4% donor type $CD15^+$ cells, at least about 8% donor type $CD15^+$ cells, or more. In some embodiments, long term HSC engraftment is evidenced by myeloid chimerism ≥5% at 24 weeks, reconstitution of T and B lymphoid compartments with reduced or eliminated dependence on immunoglobulin supplementation. Sustained chimerism may be achieved for greater than one year post-transplantation.

[00107] The modified cells which are employed may be fresh, frozen, or have been subject to prior culture. They may be fetal, neonate, adult, *etc.* Hematopoietic stem cells may be obtained from fetal liver, bone marrow, blood, particularly G-CSF or GM-CSF mobilized peripheral blood, or any other conventional source. Cells for engraftment are optionally isolated from other cells, where the manner in which the stem cells are separated from other cells of the hematopoietic or other lineage is not critical to this invention. If desired, a substantially homogeneous population of stem or progenitor cells may be obtained by selective isolation of cells free of markers associated with differentiated cells, while displaying epitopic characteristics associated with the stem cells.

[00108] Embodiments of the invention include transplantation into a patient suffering from a genetic blood disorder, where exogenous stem cells of a normal phenotype are transplanted into the patient. Such diseases include, without limitation, the treatment of anemias caused by defective hemoglobin synthesis (hemoglobinopathies). The stem cells may be allogeneic stem cells of a normal phenotype, or may be autologous cells that have been genetically engineered to delete undesirable genetic sequences, and/or to introduce genetic sequences that correct the genetic defect.

[00109] Sickle cell diseases include HbS Disease; drepanocytic anemia; meniscocytosis. Chronic hemolytic anemia occurring almost exclusively in blacks and characterized by sickle-shaped RBCs caused by homozygous inheritance of Hb S. Homozygotes have sickle cell anemia; heterozygotes are not anemic, but the sickling trait (sicklemia) can be demonstrated *in vitro*. In Hb S, valine is substituted for glutamic acid in the sixth amino acid of the beta chain. Deoxy-Hb S is much less soluble than deoxy-Hb A; it forms a semisolid gel of rodlike tactoids that cause RBCs to sickle at sites of low PO_2. Distorted, inflexible RBCs adhere to vascular endothelium and plug small arterioles and capillaries, which leads to occlusion and infarction. Because sickled RBCs are too fragile to withstand the mechanical trauma of circulation, hemolysis occurs after they enter the circulation. In homozygotes, clinical manifestations are caused by anemia and vaso-occlusive events resulting in tissue ischemia and infarction. Growth and development are impaired, and susceptibility to infection increases. Anemia is usually severe but varies highly among

patients. Anemia may be exacerbated in children by acute sequestration of sickled cells in the spleen.

[00110] Thalassemias are a group of chronic, inherited, microcytic anemias characterized by defective Hb synthesis and ineffective erythropoiesis, particularly common in persons of Mediterranean, African, and Southeast Asian ancestry. Thalassemia is among the most common inherited hemolytic disorders. It results from unbalanced Hb synthesis caused by decreased production of at least one globin polypeptide chain (β, α, γ, δ).

[00111] Aplastic anemia results from a loss of RBC precursors, either from a defect in stem cell pool or an injury to the microenvironment that supports the marrow, and often with borderline high MCV values. The term aplastic anemia commonly implies a panhypoplasia of the marrow with associated leukopenia and thrombocytopenia.

[00112] Combined immunodeficiency is a group of disorders characterized by congenital and usually hereditary deficiency of both B- and T-cell systems, lymphoid aplasia, and thymic dysplasia. The combined immunodeficiencies include severe combined immunodeficiency, Swiss agammaglobulinemia, combined immunodeficiency with adenosine deaminase or nucleoside phosphorylase deficiency, and combined immunodeficiency with immunoglobulins (Nezelof syndrome). Most patients have an early onset of infection with thrush, pneumonia, and diarrhea. If left untreated, most die before age 2. Most patients have profound deficiency of B cells and immunoglobulin. The following are characteristic: lymphopenia, low or absent T-cell levels, poor proliferative response to mitogens, cutaneous anergy, an absent thymic shadow, and diminished lymphoid tissue. Pneumocystis pneumonia and other opportunistic infections are common.

[00113] "In combination with", "combination therapy" and "combination products" refer, in certain embodiments, to the concurrent administration to a patient of the engineered proteins and cells described herein in combination with additional therapies, e.g. surgery, radiation, chemotherapy, and the like. When administered in combination, each component can be administered at the same time or sequentially in any order at different points in time. Thus, each component can be administered separately but sufficiently closely in time so as to provide the desired therapeutic effect.

[00114] "Concomitant administration" means administration of one or more components, such as engineered proteins and cells, known therapeutic agents, etc. at such time that the combination will have a therapeutic effect. Such concomitant administration may involve concurrent (*i.e.* at the same time), prior, or subsequent administration of components. A person of ordinary skill in the art would have no difficulty determining the appropriate timing, sequence and dosages of administration.

[00115] The use of the term "in combination" does not restrict the order in which prophylactic and/or therapeutic agents are administered to a subject with a disorder. A first prophylactic or therapeutic agent can be administered prior to (e.g., 5 minutes, 15 minutes, 30 minutes, 45 minutes, 1 hour, 2 hours, 4 hours, 6 hours, 12 hours, 24 hours, 48 hours, 72 hours, 96 hours, 1 week, 2 weeks, 3 weeks, 4 weeks, 5 weeks 6 weeks, 8 weeks, or 12 weeks before), concomitantly with, or subsequent to (e.g., 5 minutes, 15 minutes, 30 minutes, 45 minutes, 1 hour, 2 hours, 4 hours, 6 hours, 12 hours, 24 hours, 48 hours, 72 hours, 96 hours, 1 week, 2 weeks, 3 weeks, 4 weeks, 5 weeks, 6 weeks, 8 weeks, or 12 weeks after) the administration of a second prophylactic or therapeutic agent to a subject with a disorder.

Cell Compositions

[00116] In some embodiments a cell composition is provided. The cell can be provided in a unit dose for therapy, and can be allogeneic, autologous, etc. with respect to an intended recipient. Methods may include a step of obtaining desired cells, e.g., T cells, hematopoietic stem cells, etc., which may be isolated from a biological sample, or may be derived *in vitro* from a source of progenitor cells. The cells are transduced or transfected with a vector comprising a sequence encoding TCL1A or TCL1A mRNA, which step may be performed in any suitable culture medium. For example, cells may be collected from a patient, modified *ex vivo*, and reintroduced into the subject. The cells collected from the subject may be collected from any convenient and appropriate source, including e.g., peripheral blood (e.g., the subject's peripheral blood), a biopsy (e.g., a biopsy from the subject), and the like.

[00117] Engineered cells can be provided in pharmaceutical compositions suitable for therapeutic use, e.g. for human treatment. Therapeutic formulations comprising such cells can be frozen, or prepared for administration with physiologically acceptable carriers, excipients or stabilizers (Remington's Pharmaceutical Sciences 16th edition, Osol, A. Ed. (1980)), in the form of aqueous solutions. The cells will be formulated, dosed, and administered in a fashion consistent with good medical practice. Factors for consideration in this context include the particular disorder being treated, the particular mammal being treated, the clinical condition of the individual patient, the cause of the disorder, the site of delivery of the agent, the method of administration, the scheduling of administration, and other factors known to medical practitioners.

[00118] The cells can be administered by any suitable means, usually parenteral. Parenteral infusions include intramuscular, intravenous (bolus or slow drip), intraarterial, intraperitoneal, intrathecal or subcutaneous administration.

[00119] Kits may be provided. Kits may further include cells or reagents suitable for isolating and culturing cells in preparation for conversion; reagents suitable for culturing HSC; and reagents useful for determining the expression of TCL1A genes in the contacted cells. Kits may also include tubes, buffers, etc., and instructions for use.

EXPERIMENTAL

[00120] The following examples are put forth so as to provide those of ordinary skill in the art with a complete disclosure and description of how to make and use the present invention, and are not intended to limit the scope of what the inventors regard as their invention nor are they intended to represent that the experiments below are all or the only experiments performed. Efforts have been made to ensure accuracy with respect to numbers used (e.g. amounts, temperature, etc.) but some experimental errors and deviations should be accounted for. Unless indicated otherwise, parts are parts by weight, molecular weight is weight average molecular weight, temperature is in degrees Centigrade, and pressure is at or near atmospheric.

EXAMPLE 1

TCL1A and Stem Cell Proliferation

[00121] *Genome wide association study identifies inherited determinants of clonal expansion.* We performed a genome-wide association study (GWAS) in CHIP carriers to identify inherited genetic variation that causally associates with clonal expansion. Association analyses were performed using the SAIGE statistical package. The GWAS identified a single locus at genome-wide significance overlapping *TCL1A* (FIG. 1a). We used SuSIE to perform genetic fine-mapping to identify the most likely causal set of variants, which further narrowed down the associated region to a credible set containing a single variant, rs2887399 (Fig. 3). The alt-allele is common, occurring in 26% of haplotypes sequenced in TOPMed. rs2887399 lies in the core promoter of *TCL1A* as defined by the Ensembl regulatory build, 162 base-pairs from the canonical transcription start site (TSS) and in a CpG island. Analysis of the variant by the Open Targets variant-to-gene prediction algorithm also nominated *TCL1A* as the causal gene.

[00122] *TCL1A* has been implicated in lymphoid malignancies as a translocation partner in T-prolymphocytic leukemia, but it has not been studied in the context of HSC biology. *TCL1A* is also the only gene in the duplicated region of chromosome 14q32 associated with an inherited predisposition to develop myeloid malignancies shared by all kindreds. Of note, the region in the *TCL1A* promoter where rs2887399 resides is only partially conserved between humans and other primates, and poorly conserved with non-primate species.

[00123] *TCL1A expression in hematopoietic cells.* Next, we sought to establish how rs2887399 might shape the hematologic phenotypes observed. We first asked if the variant was associated with *TCL1A* expression in any cell type. As identified in the GTEx v8 eQTL release, the alt-allele reduces expression of *TCL1A* in whole blood (normalized effect size = -0.13, pvalue = 1.4×10^{5}). The GWAS colocalized with cis-expression quantitative trait loci (eQTLs) for *TCL1A* in whole blood (posterior probability of a single shared causal variant = 97.1%). The association in whole blood is likely driven by B-cells, as *TCL1A* is highly expressed in B-cells but appears to have absent or low expression in all other cell types in blood except for rare plasmacytoid dendritic cells.

[00124] Little is known about *TCL1A* expression in HSCs. We examined whether CHIP-associated mutations altered the regulation of the *TCL1A* locus in human hematopoietic stem and progenitor cells (HSPCs) using publicly available single-cell RNA sequencing (scRNA-seq) and ATAC-sequencing (ATAC-seq) datasets of normal and malignant hematopoiesis. *TCL1A* was expressed in fewer than 1 in 1000 cells identified as HSC/MPPs in scRNA-seq data from 6 normal human marrow samples (range 0-0.17%). In contrast, *TCL1A* was expressed in a much higher fraction of HSC/MPPs in 3 out of 5 samples from persons with *TET2* or *ASXL1*-mutated myeloid malignancies (range 2.7-7%). Next, using a dataset of ATAC-seq in normal and pre-leukemic HSCs (pHSCs), we evaluated chromatin accessibility at the *TCL1A* promotor. Consistent with the lack of *TCL1A* transcripts in normal HSCs, we observed that the promoter was not accessible in either normal human donor HSCs or in HSCs from patients with AML that were not part of the mutant clone.

[00125] We next asked if the neighboring genes *TCL6* or *TCL1B* either became expressed or had accessible chromatin in HSCs carrying CHIP mutations in these same datasets. In contrast to the result for *TCL1A*, no RNA expression or promoter accessible chromatin could be found at these genes in HSCs, indicating that rs2887399 was marking *TCL1A* as the likely causal gene for clonal expansion in this locus.

[00126] *Functional effect of rs2887399 on normal and CHIP-mutated HSCs.* Our GWAS data strongly implicated *TCL1A* as a causal factor for clonal expansion in CHIP, likely through effects on HSCs. We next wished to determine how rs2887399 impacted the expression of *TCL1A*, and consequently clonal expansion, in HSCs. We proposed the following mechanistic model: Normally, the *TCL1A* promoter is inaccessible and gene expression is absent or very low in HSCs. In the presence of driver mutations in *TET2*, *ASXL1*, *SF3B1*, *SRSF2*, or LOY, *TCL1A* is aberrantly expressed and drives clonal expansion of the mutated HSCs. The presence of the alt-allele of rs2887399 inhibits accessibility of chromatin at the *TCL1A* promoter, leading to reduced expression of *TCL1A* RNA and protein and abrogation of the clonal advantage due to the mutations.

[00127] To test our model experimentally, we first obtained human CD34+ mobilized peripheral blood cells from donors who were G/G (homozygous reference), T/T (homozygous alternate), or G/T (heterozygous) genotype at rs2887399. The three donors were healthy and between 29-32 years old at the time of donation. To mimic CHIP-associated mutations, we used CRISPR to introduce insertion-deletion mutations in *DNMT3A*, *TET2*, or *ASXL1* in HSPCs for each rs2887399 genotype. Editing at the adeno-associated virus integration site 1 (AAVS1) was done as a control for each rs2887399 genotype. High efficiency of editing was confirmed by Sanger sequencing.

[00128] First, we examined whether the accessibility of the *TCL1A* promoter seen in the setting of *TET2* mutations was altered by rs2887399 genotype. We edited bulk CD34 cells from each genotype for *TET2*, sorted cells with a marker profile of HSCs and multipotent progenitors (MPPs) (Lineage- CD34+ CD38- CD45RA-), cultured them for 5 days in cytokine-supported media, and then performed ATAC-seq. Consistent with the pre-leukemic HSC data, we detected accessibility at the *TCL1A* promoter in *TET2*-edited cells from the rs2887399 G/G donor. However, accessibility at the *TCL1A* promoter was decreased in the *TET2*-edited cells in samples from carriers of the T allele in a dose-dependent manner, indicating that the protective effect of the alt-allele of rs2887399 is mediated by blocking promoter accessibility. We did not observe any change in accessible chromatin at the *TCL1A* promoter in *DNMT3A*- or AAVS1-edited samples.

[00129] Next, we asked if the differential chromatin accessibility due to rs2887399 altered TCL1A protein expression in HSC/MPPs. We edited CD34+ cells from donors with the three rs2887399 genotypes at AAVS1, *DNMT3A*, *TET2*, and *ASXL1*. After 11 days in culture, we performed a flow cytometry-based assay for TCL1A protein expression. We found that ~1% of HSCs/MPPs from AAVS1 or *DNMT3A* edited samples were positive for TCL1A, which did not vary by rs2887399 genotype. In contrast, 4.6-9.3% of HSC/MPPs from the G/G donor that had been edited for *ASXL1* or *TET2* expressed TCL1A, and the proportion of TCL1A positive HSC/MPPs decreased in donor samples with each additional T allele (4 biological replicates per condition). There was minimal expression of TCL1A in any non-HSC/MPP CD34+ population in any of the samples. Notably, the proportion of TCL1A expressing HSC/MPPs was less than 10% in all samples even though the proportion of mutant cells was >90%. This suggests that even in the presence of driver mutations in *TET2* or *ASXL1*, only a fraction of HSC/MPPs are capable of expressing TCL1A at any given time and is consistent with the single-cell RNA sequencing data from hematological malignancy samples.

[00130] We then asked if rs2887399 had any effect on expansion of HSPCs in vitro. For this experiment, we edited the CD34+ cells from GG and TT donors, sorted HSCs (Lin- CD34+ CD38- CD45RA- CD90+), and allowed the cultures to grow for 14 days, at which time cells were counted and analyzed for HSPC markers by flow cytometry. There was a

notable expansion of cells bearing markers of HSC/MPPs in the *ASXL1* and *TET2* edited samples from the rs2887399 G/G donor compared to the AAVS1 edited sample, but this effect was abrogated in edited samples from the rs2887399 T/T donor (4 biological replicates per donor and edit). A population of cells that was Lin-/lo CD34+ CD38- CD45RA dim ($CD45RA_{dim}$ HSPCs), presumably progenitors descended from the HSC/MPP population, was also markedly expanded in the *ASXL1* and *TET2* edited samples from the G/G donor, but the degree of expansion was partially reversed in the edited samples from the T/T donor. The ratio of CD34+ CD45RA-/lo progenitors to CD34- cells was also increased in the *ASXL1* and *TET2*-edited samples from the G/G donor compared to the T/T donor, indicating either less retention of stem/progenitor cell activity or faster differentiation in the absence of *TCL1A* expression. There were no differences in any populations in the AAVS1 or *DNMT3A* edited samples based on rs2887399 genotype.

[00131] To orthogonally validate the necessity of *TCL1A* for *TET2* mutation-mediated clonal expansion, we edited CD34+ cells from a rs2887399 G/G donor with AAVS1 or *TET2* guides, followed by lentiviral delivery of shRNA targeting *TCL1A* or scramble control. The *TCL1A* shRNA construct we used was validated to knockdown TCL1A protein by ~90%. Following these manipulations, we sorted GFP+ HSCs and performed the same 14-day *in vitro* expansion assay. The increase in *TET2* mutated HSC/MPP counts seen after 14 days was nearly completely attenuated by *TCL1A* knockdown, indicating that *TCL1A* expression is necessary for expansion of *TET2*-mutant HSCs in this assay.

[00132] *TCL1A expression is sufficient for HSC expansion and resistance to cell cycle arrest.* Forced expression of *TCL1A* in unmutated HSCs should be sufficient to recapitulate clonal expansion phenotypes. To test this hypothesis, we transduced human MPB CD34+ cells with lentivirus containing the *TCL1A* open reading frame (*TCL1A*-eGFP) or empty vector control (control-eGFP) (FIG. 3a) and performed in vitro clonal expansion assays on purified HSCs. After 14 days, cultures from HSCs that received *TCL1A*-eGFP virus had ~4-fold higher counts of phenotypic HSC/MPPs and colony forming cells compared to cultures from HSCs that received control-eGFP (FIG. 3b), indicating that *TCL1A* expression is sufficient for HSC clonal expansion. To further characterize the effect of *TCL1A*, we assessed cell cycle status by DAPI staining for DNA content in control- or *TCL1A*-eGFP transduced HSC/MPPs after 10 days of culture. *TCL1A* expressing HSC/MPPs were ~2-fold more likely to be cycling compared to control cells (FIG. 3c).

[00133] To uncover the mechanism by which *TCL1A* promotes clonal expansion and proliferation of HSCs, we transduced *TCL1A*-eGFP or control-eGFP into MPB CD34+ cells from two normal donors that were G/G or T/T at rs2887399, sorted GFP+ HSCs, cultured them for 7 days, and then performed CITE-seq on the four samples. After integration, dimensionality reduction, and clustering, we annotated four clusters of HSC/MPPs as well

as two populations of myeloid progenitors using the cell surface markers CD34, CD38, CD45RA, CD49f, and CD11a. We focused on the 4 HSC/MPP clusters for further analysis. Pseudotime analysis supported a trajectory of progression from HSC/MPP 1 (initial state) to 4 (most 'differentiated' state). HSC/MPP 1 expressed stem cell identity genes such as *MECOM*, *FAM30A*, and *HEMGN*, as well as high levels of proliferative markers such as *MKI67*, *TOP2A*, *PCNA*, and *CENPA* (FIG. 3E). In contrast, HSC/MPP 2-4 expressed lower levels of stem cell identity genes and proliferative markers, and cell cycle analysis confirmed these clusters contained cells that were predominantly in G0 or G1 phase. HSC/MPP 2-4 also displayed a progressive increase in genes associated with the integrated stress response such as *PPP1R15A* (GADD34), *DDIT3* (CHOP), and *ATF4*, as well as FOXO target genes such as *CDKN1A* (p21), *CDKN1B* (p27), *SOD2*, *CCNG2*, and *TXNIP*. FOXO transcription factors can drive downstream target gene expression in an adaptive response to stressors to preserve cell viability, but prolonged activation of this response can lead to a terminal state of cell cycle arrest or apoptosis. Indeed, cells in HSC/MPP 4 expressed the highest levels of the apoptosis effector genes *BAD*, *BCL2L11* (BIM), and *BBC3* (PUMA). We then asked how *TCL1A* expression affected the relative proportion of cells in these four HSC/MPP clusters. Strikingly, we found that *TCL1A* expression led to a significant increase in the proportion of cells in the HSC/MPP 1 cluster, and a significant decrease in the proportion of cells in the HSC/MPP 3 and 4 clusters, an effect that was consistent in both donors (FIG. 3F). This indicates that *TCL1A* functions to preserve HSCs in a proliferative state by avoiding prolonged, deleterious stress responses.

[00134] *In Vivo Long Term Chimerism.* Shown in, FIG. 5, cKit enriched bone marrow cells from B6 mice were cultured in mouse HSPC Expansion media (StemSpanII + 10 ng/mL SCF, 100 ng/mL TPO + 1% Penicillin/Streptomycin) for 24 hours before lentivirus editing. 2.8 million cells were collected, spun down, and resuspended in a final volume of 800 uL HSPC Lentivirus Media (StemSpanII + 10 ng/mL SCF, 100 ng/mL TPO + 1% Penicillin/Streptomycin + 10 uM prostaglandin E2 + 100 ng/uL poloxamer 407) with virus added at an MOI of 20 for both. Cells were plated in a 96 well u-bottom plate for 16 hours. Following 16-hour incubation, lentivirus edited cells were washed in PBS, and then mixed with B6 wildtype whole bone marrow cells for subsequent bone marrow transplant.

[00135] Eighteen B6 CD45.1 mice were irradiated with 6.45 Gy, and then immediately transplanted with donor B6 CD45.2 bone marrow. Nine control B6 CD45.1 mice were transplanted with 1 million cells consisting of 25% cKIT+ eGFP control edited cells and 75% wildtype whole bone marrow cells. Nine experimental B6 CD45.1 mice were transplanted with 1 million cells consisting of 25% cKIT+ TCL1A overexpression edited cells and 75% wildtype whole bone marrow cells. Mice were bled 4 weeks and 7 weeks

post transplant to assess the chimerism of granulocytes, monocytes, B-cells, and T-cells derived from donor bone marrow. Granulocyte engraftment is shown in the figure.

[00136] We observed that expression of *TCL1A*, even in the absence of CHIP driver mutations, was sufficient to promote clonal expansion of HSCs in vitro. Several studies have shown that HSCs are especially sensitive to various stressors, such as the integrated stress response, unfolded protein response, and loss of autophagic capacity. *TCL1A* has a known role in promoting AKT oligomerization and activation, which could promote downstream growth pathways such as mTORC1 while also inhibiting FOXO transcriptional activity. We found that stress response and FOXO target genes, including genes involved in cell cycle arrest and apoptosis, were suppressed in HSCs expressing *TCL1A*, consistent with AKT mediated inhibition of FOXO in these cells. Studies in mice have shown that constitutive activation of PI3K/AKT or deletion of FOXO can push HSCs into cycle, but also lead to loss of functional HSCs. In accordance, directly activating mutations in genes of the PI3K/AKT pathway are not seen in CHIP, possibly due to the same deleterious effect in human HSCs. Our results indicate that expression of *TCL1A* does not lead to these deleterious effects. It is possible that *TCL1A* expression activates AKT in moderation, which allows for clonal expansion without long-term negative effects on HSCs. Alternatively, *TCL1A* may alter other aspects of AKT activity that preserve stemness, or even have AKT-independent effects.

METHODS

[00137] *Lentivirus Plasmid Construction.* An insert containing the TCL1A coding region followed in frame with GFP (TCLA1 -T2A Linker-GFP) under the control of mammalian EF1a promoter, as well as a control sequence composed of GFP under the EF1a promoter, was synthetized by Gene Universal. The insert was cloned into a second-generation lentivirus backbone, adapted from the addgene vector pMH0001, using enzymatic cloning. Briefly both the insert and backbone were digested with MluI and SbfI enzymes (NEB) and ligated using the T4 ligase (NEB). NEB DH5a competent bacteria were transformed with the ligation product. The transformed bacteria were screened by Ampilicin resistance and grown in liquid culture in LB media to amplify the plasmid. Maxiprep plasmid purification (Macherey-Nagel NucleoBond Xtra Maxi) was performed to obtain the final purified plasmid used for lentivirus production.

[00138] *Lentivirus Production.* The plasmid was transfected into 293T HEK cells at roughly 80% confluency in 10 cm tissue culture plates coated with poly-d-lysine using Lipofectamine 3000. The lipofectamine media was exchanged 16 hours later, and the viral supernatant was collected at 72h post-transfection. The collected viral supernatant was filtered via a 0.45 µm filtration unit, and concentrated using the LentiX concentrator (Takara) for 2

hours at 4 C and then spun down at 1500 x g for 45 minutes at 4 C. The concentrated supernatant was subsequently aliquoted, flash frozen, and stored in –80°C until use.

[00139] *Combined CRISPR and shRNA Assay.* CD34+ cells were thawed and cultured in HSPC Expansion media (StemSpanII + 10% CD34+ Expansion Supplement + 0.1% Penicillin/Streptomycin) for 48 hours before CRISPR editing. Editing of AAVS, TET2, DNMT3A, and ASXL1 was performed by electroporation of Cas9 ribonucleoprotein complex (RNP). For each combination of rs2887399 genotype and gRNA, 100,000 cells were incubated with 3.26 ug of Synthego synthetic sgRNA guide and 8.332 ug of IDT Alt-R S.p. Cas9 Nuclease V3 for 15 minutes at room temperature before electroporation. CD34+ cells were resuspended in 18 uL of Lonza P3 solution and mixed with the ribonucleoprotein complex, and then transferred to Nucleocuvette strips for electroporation with program DZ-100 (Lonza 4D Nucleofector). Immediately following electroporation, each condition of 500,000 cells was transferred to 2 mLs of HSPC Expansion media and allowed to recover for 8 hours. Later that same day, 250,000 CRISPR edited cells were collected, spun down, and resuspended in a final volume of HSPC Lentivirus Media (StemSpanII + 10% CD34+ Expansion Supplement + 0.1% Penicillin/Streptomycin + 10 uM prostaglandin E2 + 100 ng/uL poloxamer 407) with virus added at an MOI of 20. Cells were plated in a 96 well u-bottom plate for 16 hours. shRNA-A and the scramble-shRNA from Origene CAT#: TL301172V were used for this experiment. Following a 16-hour incubation, cells were washed in PBS, and then plated in 2 mL of HSPC Expansion media. After 72 hours, previously described liquid culture expansion assay was done on sorted Lineage- CD34+ CD38- CD90+ CD45RA- GFP+ cells.

[00140] *Lentivirus TCL1A Expression Assay.* CD34+ cells were thawed and cultured in HSPC Expansion media (StemSpanII + 10% CD34+ Expansion Supplement + 0.1% Penicillin/Streptomycin) for 48 hours before overexpression lentivirus editing. 750,000 edited cells were collected, spun down, and resuspended in a final volume of HSPC Lentivirus Media (StemSpanII + 10% CD34+ Expansion Supplement + 0.1% Penicillin/Streptomycin + 10 uM prostaglandin E2 + 100 ng/uL poloxamer 407) with virus added at an MOI of 100. Cells were plated in a 96 well u-bottom plate for 16 hours. eGFP control was Origene CAT#: PS100093V or the eGFP we produced, and the TCL1A-eGFP was Origene CAT#: RC204243L4V or the TCL1A-eGFP we produced. Following 16-hour incubation, cells were washed in PBS, and then plated in 2 mL of HSPC Expansion media. After 72 hours, previously described liquid culture expansion assay was done on sorted Lineage- CD34+ CD38- CD90+ CD45RA- GFP+ cells. After 14 days, cells were harvested and assessed for HSC/MPP frequency using flow cytometry as previously described. The total HSC/MPP count was determined by multiplying the percentage of live cells that were in the HSC/MPP gate by the total live cell count for each replicate.

[00141] *In Vivo Survival of the Cells/Transplanted Mice Methods:* cKit enriched bone marrow cells from B6 mice were cultured in mouse HSPC Expansion media (StemSpanII + 10 ng/mL SCF, 100 ng/mL TPO + 1% Penicillin/Streptomycin) for 24 hours before lentivirus editing. 2.8 million cells were collected, spun down, and resuspended in a final volume of 800 uL HSPC Lentivirus Media (StemSpanII + 10 ng/mL SCF, 100 ng/mL TPO + 1% Penicillin/Streptomycin + 10 uM prostaglandin E2 + 100 ng/uL poloxamer 407) with virus added at an MOI of 11 for TCL1A overexpression virus and MOI of 50 for control eGFP virus. Cells were plated in a 96 well u-bottom plate for 16 hours. Following 16-hour incubation, lentivirus edited cells were washed in PBS, and then mixed with B6 wildtype whole bone marrow cells for subsequent bone marrow transplant.

[00142] Twenty B6 CD45.1 mice were irradiated with 9.5 Gy, and then immediately transplanted with donor B6 CD45.2 bone marrow. Ten control B6 CD45.1 mice were transplanted with 1 million cells consisting of 25% cKIT+ eGFP control edited cells and 75% wildtype whole bone marrow cells. Ten experimental B6 CD45.1 mice were transplanted with 1 million cells consisting of 10% cKIT+ TCL1A overexpression edited cells and 90% wildtype whole bone marrow cells. Survival of mice following irradiation and transplant was tracked for 9 weeks.

[00143] *Study Samples.* Whole genome sequencing (WGS) was performed on 127,946 samples as part of 51 studies contributing to Freeze 8 NHLBI TOPMed program as previously described. None of the TOPMed studies included selected individuals for sequencing because of hematologic malignancy. Each of the included studies provided informed consent. Age was obtained for 82,807 of the samples, and the median age was 55, the mean age 52.5, and the maximum age 98. The samples have diverse reported ethnicity (40% European, 32% African, 16% Hispanic/Latino, 10% Asian).

[00144] *WGS Processing, Variant Calling and CHIP annotation.* BAM files were remapped and harmonized through the functionally equivalent pipeline. SNPs and indels were discovered across TOPMed and were jointly genotyped across samples using the GotCloud pipeline[59]. An SVM filter was trained to discriminate between high- and low-quality variants. Variants were annotated with snpEff 4.3. Sample quality was assessed through mendelian discordance, contamination estimates, sequencing converge, and among other quality control metrics.

[00145] Putative somatic SNPs were called with GATK Mutect2, which searches for sites where there is evidence for alt-reads that support evidence for variation, and then performs local haplotype assembly. We used a panel of normals to filter sequencing artifacts and used an external reference of germline variants to exclude germline calls. We deployed this pipeline on Google Cloud using Cromwell.

[00146] Samples were annotated as having CHIP if the Mutect2 output contained at least one variant in a curated list of leukemogenic driver mutations with at least three alt-reads

supporting the call. We expanded the list of driver mutations to include those in recently identified CHIP genes, increasing the number of CHIP cases from our previous report.

[00147] *Fine-mapping of the TCL1A region*. We applied the SuSIE algorithm to the genotypes included in a 200kb region surrounding *TCL1A*. We used the same covariates as the single variant association analysis. We used the posterior inclusion probabilities (PIP) and credible sets identified by SuSIE to identify the putative causal variant. We used LD directly calculated on the genotypes as opposed to an external reference.

[00148] *CRISPR–Cas9 editing of CD34⁺ human HSPCs*. $CD34_+$ HSPCs from adult donors were purchased from the Cooperative Center of Excellence in Hematology (CCEH) at the Fred Hutch Cancer Research Center, Seattle, USA. TCL1A rs2887399 genotyping was performed using ThermoFisher SNP assay. CD34+ cells were thawed and cultured in HSPC Expansion media (StemSpanII + 10% CD34+ Expansion Supplement + 0.1% Penicillin/Streptomycin) for 48 hours before CRISPR editing. Editing of *AAVS*, *TET2*, *DNMT3A*, and *ASXL1* was performed by electroporation of Cas9 ribonucleoprotein complex (RNP). For each combination of rs2887399 genotype and gRNA, 100,000 cells were incubated with 3.2 ug of Synthego synthetic sgRNA guide and 8.18 ug of IDT Alt-R S.p. Cas9 Nuclease V3 for 15 minutes at room temperature before electroporation. CD34+ cells were resuspended in 18 uL of Lonza P3 solution and mixed with the ribonucleoprotein complex, and then transferred to Nucleocuvette strips for electroporation with program DZ-100 (Lonza 4D Nucleofector). Immediately following electroporation, each condition of 100,000 cells was transferred to 2 mL of HSPC Expansion media and allowed to recover for 24 hours. CRISPR editing efficiency was measured using Sanger Sequencing and ICE Analysis.

	Guide Sequence	Sanger Forward Primer	Sanger Reverse Primer
AAVS	GCCAGTAGCCAGCCCCG TCC	GGGTCCAGGCCAAGTAGGT	TGGCTCTTCACCTTTCT AGTCCC
TET2	TCATGGAGCATGTACTAC AA	GGTTATGCCACAGCTTAATA CAGA	TGACACCCCTTTAAAA CTTTGG
DNMT3A	GCCCGTGGGGTCCGATG CTG	GGAGCTCCATCTGAATGAG G	GGCTGGAATTGTGTGA CTTG
ASXL1	GTATCCGTGGACTCACCG TG	TACCCATCCCATCGAATGAT	GCAGCAACTGCATCAC AAGT

[00149] The preceding merely illustrates the principles of the invention. It will be appreciated that those skilled in the art will be able to devise various arrangements which, although not explicitly described or shown herein, embody the principles of the invention and are included within its spirit and scope. Furthermore, all examples and conditional language

recited herein are principally intended to aid the reader in understanding the principles of the invention and the concepts contributed by the inventors to furthering the art, and are to be construed as being without limitation to such specifically recited examples and conditions. Moreover, all statements herein reciting principles, aspects, and embodiments of the invention as well as specific examples thereof, are intended to encompass both structural and functional equivalents thereof. Additionally, it is intended that such equivalents include both currently known equivalents and equivalents developed in the future, i.e., any elements developed that perform the same function, regardless of structure. The scope of the present invention, therefore, is not intended to be limited to the exemplary embodiments shown and described herein. Rather, the scope and spirit of the present invention is embodied by the appended claims.

THAT WHICH IS CLAIMED IS:

1. A method for enhancing one or both of self-renewal and proliferation of a hematopoietic stem cell (HSC), the method comprising altering the HSC by introducing genetic sequences to induce over-expression of TCL1A.

2. The method of claim 1, wherein the genetic sequences comprise a vector encoding TL1A operably linked to a promoter active in HSC.

3. The method of claim 2, wherein the vector is a lentiviral vector.

4. The method of claim 1, wherein the genetic sequences comprise modified mRNA (mmRNA) encoding TCL1A.

5. The method of claim 4, wherein the mmRNA comprises modified nucleosides.

6. The method of claim 1, wherein genetic sequences that increase TCL1A expression are integrated into the genome.

7. The method of any of claims 1-6, wherein the cells are human cells and TCL1A is a human sequence.

8. A composition comprising one or more HSC modified by the method of any of claims 1-7.

9. A formulation comprising a population of HSC of claim 8, and a pharmaceutically acceptable excipient.

10. The formulation of claim 9, in a therapeutically effective dose.

11. A method of treating a subject in need of hematopoietic stem cells (HSCs), the method comprising:

obtaining a population of HSCs from a biological sample;

genetically altering the HSC by introducing sequences to induce over-expression of TCL1A compared to an unmodified population of HSCs, thereby enhancing self-renewal and/or proliferation of the modified HSCs; and

administering a therapeutically effective amount of the modified HSCs to the subject in need thereof.

12. The method of claim 11, wherein the modified HSC are expanded in vitro prior to administering.

13. A modified mRNA (mmRNA) in an effective dose sufficient to induce over-expression of TCL1A in a population of HSC, the mmRNA comprising modified nucleosides and encoding human TCL1A.

14. The mmRNA of claim 13, wherein the mmRNA is packaged in lipid nanoparticles.

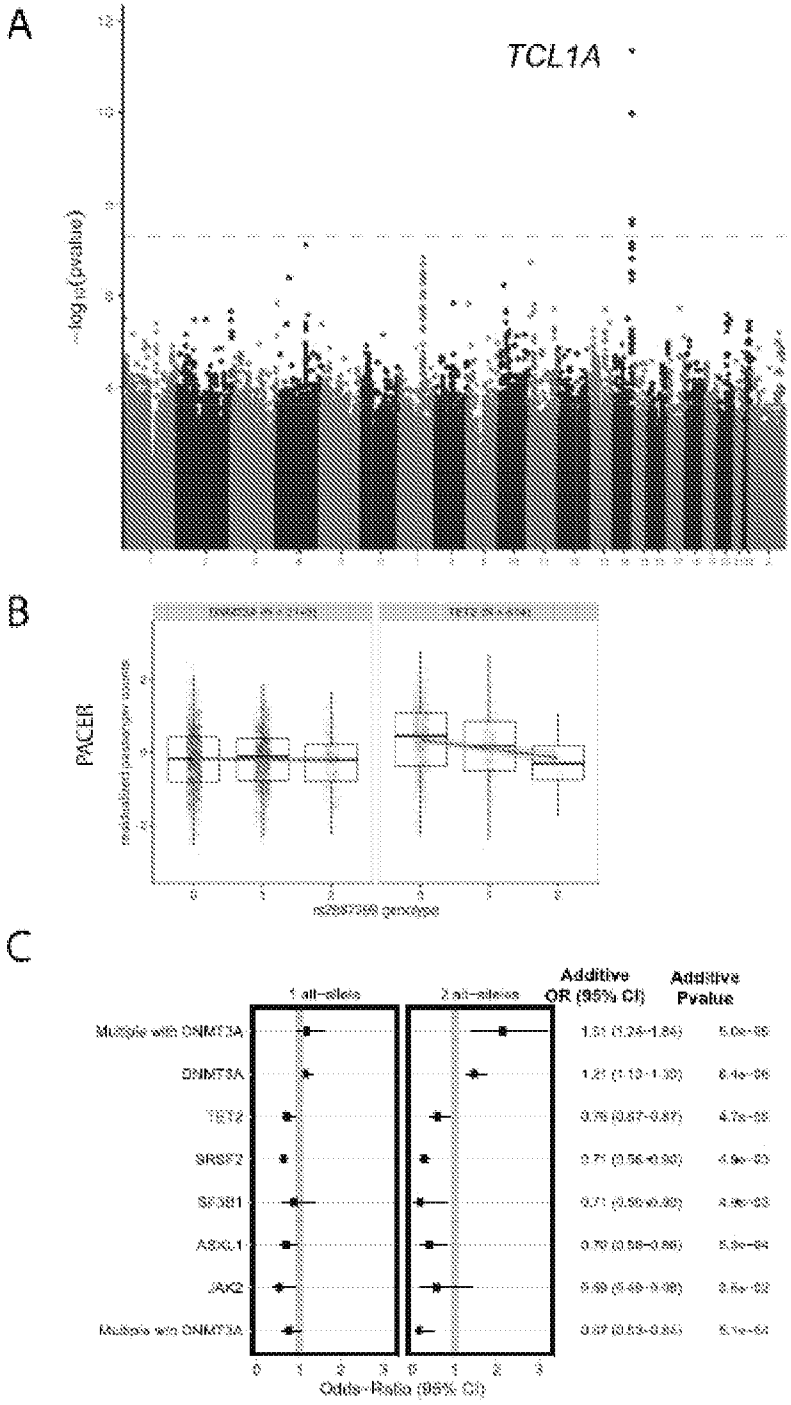

A
TCL1A
B
PACER
C
Additive OR (95% CI)
Additive Pvalue
Multiple with DNMT3A
DNMT3A
TET2
SRSF2
SF3B1
ASXL1
JAK2
Multiple w/o DNMT3A
Odds-Ratio (95% CI)

FIG. 1

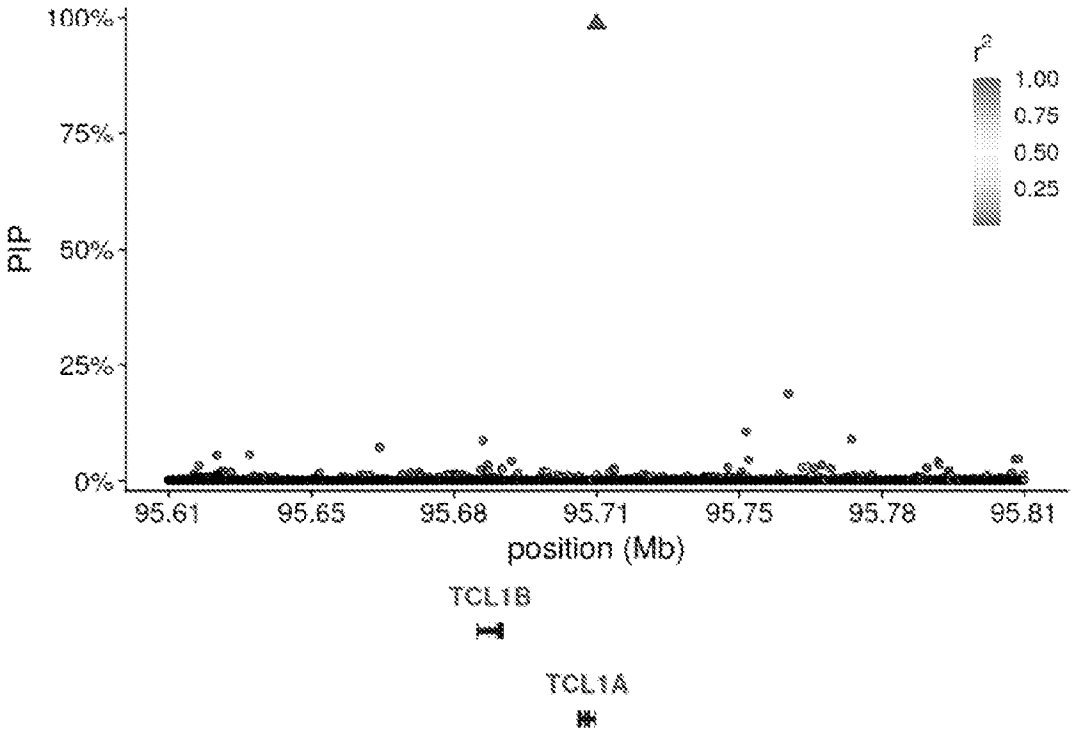

100%
75%
50%
25%
0%
PIP
r^2
1.00
0.75
0.50
0.25
95.61
95.65
95.68
95.71
95.75
95.78
95.81
position (Mb)
TCL1B
TCL1A

FIG. 2

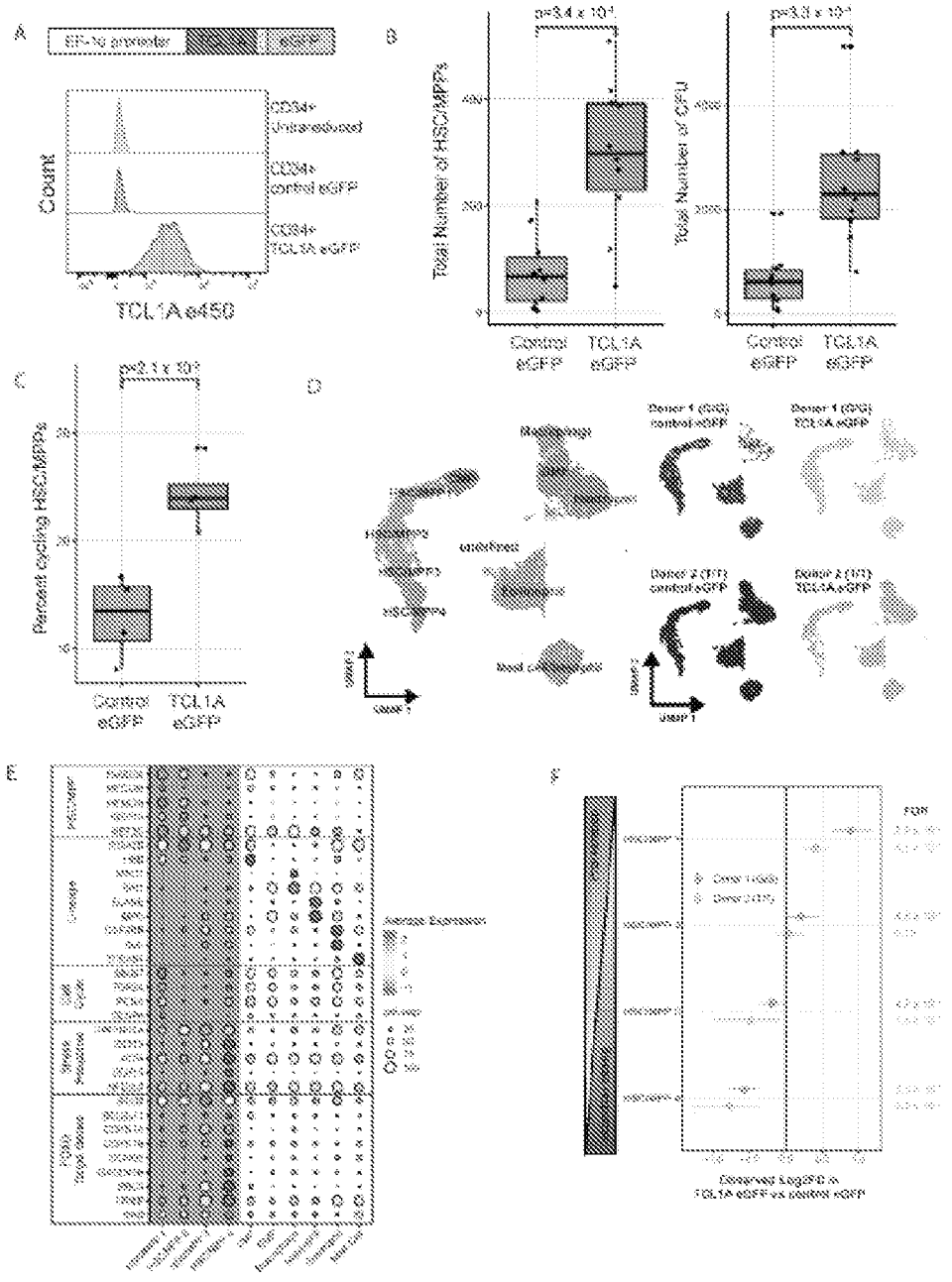


FIG. 3

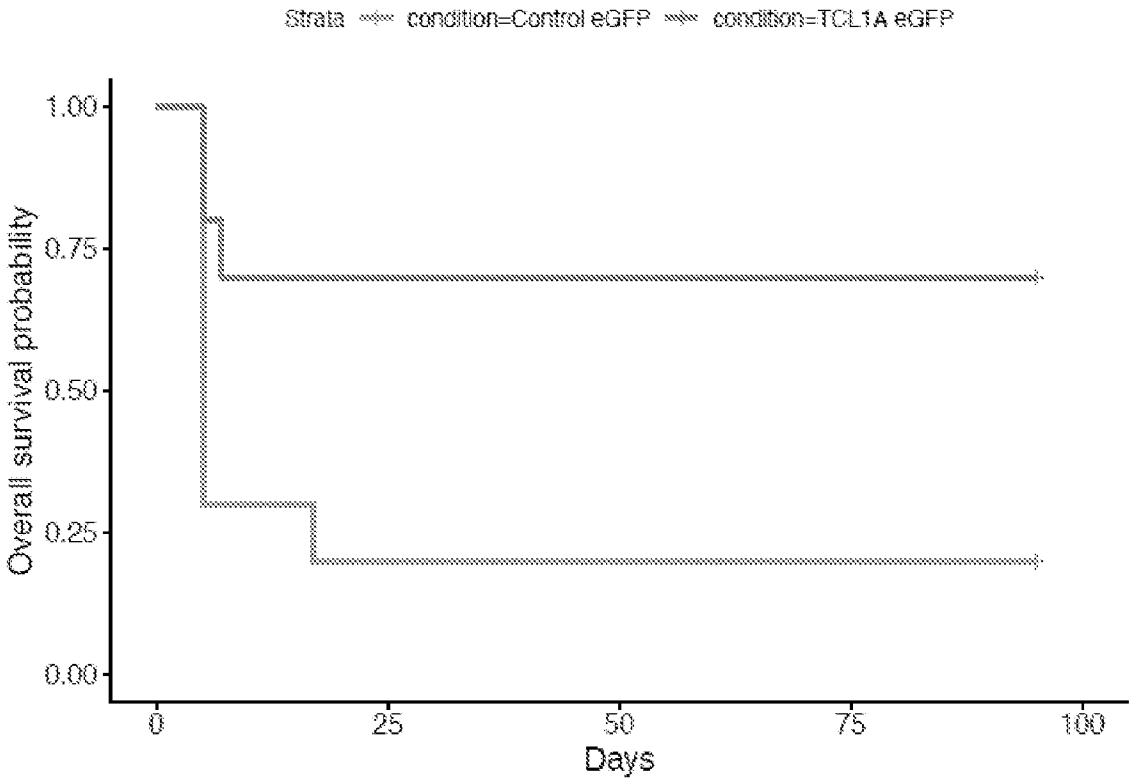

Strata
condition=Control eGFP
condition=TCL1A eGFP
Overall survival probability
1.00
0.75
0.50
0.25
0.00
0
25
50
75
100
Days

FIG. 4

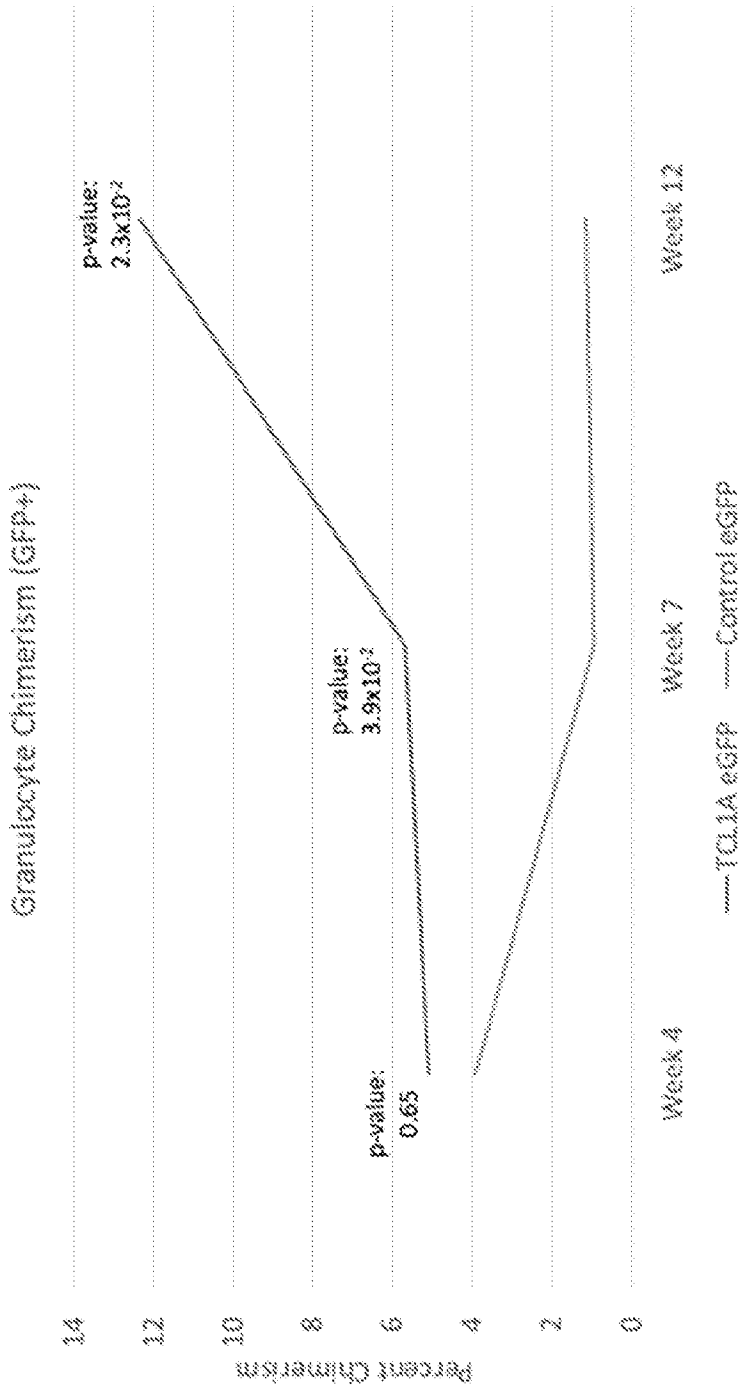

Granulocyte Chimerism (GFP+)
Percent Chimerism
14
12
10
8
6
4
2
0
p-value:
2.3x10-2
p-value:
3.9x10-2
p-value:
0.65
Week 4
Week 7
Week 12
TCL1A eGFP
Control eGFP

FIG. 5