CLAIMS

1. A system for modifying a polypyrimidine tract-binding protein 1 (BCL11A) gene target nucleic acid sequence, the system comprising a CasX variant protein and a guide ribonucleic acid (gRNA) variant wherein

a. the gRNA variant comprises:

(i) a scaffold stem loop sequence of SEQ ID NO: 25; and

(ii) a targeting sequence complementary to a target nucleic acid sequence comprising a region within a polypyrimidine tract-binding protein 1 (BCL11A) gene;

b. the Cas X variant protein is a chimeric CasX variant protein comprising:

(i) the NTSB domain of SEQ ID NO: 1, or a sequence with at least 90% sequence identity thereto

(ii) the helical Ib domain of SEQ ID NO: 1, or a sequence with at least 90% sequence identity thereto; and

(iii) the RuvC a and RuvC b domains of SEQ ID NO: 2, or sequences with at least 90% sequence identity thereto;

wherein the system is capable of modifying the BCL11A gene, wherein the modifying comprises introducing an insertion, deletion, substitution, duplication, or inversion of one or more nucleotides in the BCL11A gene.

2. The system of claim 1, wherein the region within the BCL11A gene is selected from the group consisting of:

a. a BCL11A intron;

b. a BCL11A exon;

c. a BCL11A intron-exon junction;

d. a BCL11A regulatory element; and

e. an intergenic region.

3. The system of claim 1 or claim 2, wherein the BCL11A gene comprises a wild-type sequence.

4. The system of any one of claims 1-3, wherein the gRNA variant is a single-molecule gRNA (sgRNA).

5. The system of any one of claims 1-4, wherein the targeting sequence of the gRNA variant comprises a sequence selected from the group consisting of SEQ ID NOS: 272-2100 and 2286-26789, or a sequence having at least about 95% sequence identity thereto.

6. The system of any one of claims 1-**Error! Reference source not found.**, wherein the targeting is linked to the 3’ end of the gRNA variant.

7. The system of any one of claims 1-6, wherein the targeting sequence has a single nucleotide removed from the 3’ end of the sequence, or has two, three, four, or five nucleotides removed from the 3’ end of the sequence.

8. The system of any one of claims 1-7, wherein the targeting sequence of the gRNA is complementary to a sequence of a BCL11A exon.

9. The system of claim 8, wherein the targeting sequence of the gRNA is complementary to a sequence selected from the group consisting of a BCL11A exon 1 sequence, BCL11A exon 2 sequence, BCL11A exon 3 sequence, BCL11A exon 4 sequence, BCL11A exon 5 sequence, BCL11A exon 6 sequence, BCL11A exon 7 sequence, BCL11A exon 8 sequence, and a BCL11A exon 9 sequence.

10. The system of claim 9, wherein the targeting sequence of the gRNA is complementary to a sequence selected from the group consisting of a BCL11A exon 1 sequence, BCL11A exon 2 sequence, and a BCL11A exon 3 sequence.

11. The system of any one of claims 1-7, wherein the targeting sequence of the gRNA is complementary to a sequence of a BCL11A regulatory element.

12. The system of claim 11, wherein the targeting sequence of the gRNA is complementary to a sequence of a promoter of the BCL11A gene.

13. The system of claim 11, wherein the targeting sequence of the gRNA is complementary to a sequence of an enhancer regulatory element.

14. The system of claim 13, wherein the targeting sequence of the gRNA is complementary to a sequence that comprises a GATA1 erythroid-specific enhancer binding site (GATA1) of the BCL11A gene or to a sequence that is 5’ or 3’ to the GATA1 binding site of the BCL11A gene.

15. The system of claim 14, wherein the targeting sequence of the gRNA variant comprises a sequence of UGGAGCCUGUGAUAAAAGCA (SEQ ID NO: 22), or a sequence having at least 90% sequence identity thereto.

16. The system of claim 14, wherein the targeting sequence of the gRNA comprises a sequence of UGCUUUUAUCACAGGCUCCA (SEQ ID NO: 23), or a sequence having at least 90% sequence identity thereto.

17. The system of claim 14, wherein the targeting sequence of the gRNA comprises a sequence of CAGGCUCCAGGAAGGGUUUG (SEQ ID NO: 2949), or a sequence having at least 90% sequence identity thereto.

18. The system of claim 14, wherein the targeting sequence of the gRNA comprises a sequence of GAGGCCAAACCCUUCCUGGA (SEQ ID NO: 2948), or a sequence having at least 90% sequence identity thereto.

19. The system of claim 14, wherein the targeting sequence of the gRNA comprises a sequence of AGUGCAAGCUAACAGUUGCU (SEQ ID NO: 15747), or a sequence having at least 90% sequence identity thereto.

20. The system of claim 14, wherein the targeting sequence of the gRNA comprises a sequence of AUACAACUUUGAAGCUAGUC (SEQ ID NO: 15748), or a sequence having at least 90% sequence identity thereto.

21. The system of any one of claims 1-20, wherein the gRNA variant has a sequence comprising the sequence of SEQ ID NO: 2238, or a sequence with at least 90% sequence identity thereto.

22. The system of any one of claims 1-20, wherein the gRNA variant has a sequence comprising the sequence of SEQ ID NO: 26800, or a sequence having at least 90% sequence identity thereto.

23. The system of any one of claims 1-22, wherein the gRNA variant is chemically modified.

24. The system of any one of claims 1-23, wherein the chimeric CasX variant protein comprises a sequence with at least 95% sequence identity to SEQ ID NO: 126.

25. The system of claim 24, wherein the chimeric CasX variant protein comprises a sequence with at least 99% sequence identity to SEQ ID NO: 126.

26. The system of any one of claims 1-23, wherein the CasX variant protein comprises a sequence of SEQ ID NO: 123, 133, or 143, or sequences having at least about 95% sequence identity thereto.

27. The system of any one of claims 1-26, wherein the CasX variant protein further comprises one or more nuclear localization signals (NLS).

28. The system of claim 27, wherein the one or more NLS are selected from the group of sequences consisting of PKKKRKV (SEQ ID NO: 168), KRPAATKKAGQAKKKK (SEQ ID NO: 169), PAAKRVKLD (SEQ ID NO: 170), RQRRNELKRSP (SEQ ID NO: 171), NQSSNFGPMKGGNFGGRSSGPYGGGGQYFAKPRNQGGY (SEQ ID NO: 172), RMRIZFKNKGKDTAELRRRRVEVSVELRKAKKDEQILKRRNV (SEQ ID NO: 173), VSRKRPRP (SEQ ID NO: 174), PPKKARED (SEQ ID NO: 175), PQPKKKPL (SEQ ID NO: 176), SALIKKKKKMAP (SEQ ID NO: 177), DRLRR (SEQ ID NO: 178), PKQKKRK (SEQ ID NO: 179), RKLKKKIKKL (SEQ ID NO: 180), REKKKFLKRR (SEQ ID NO: 181), KRKGDEVDGVDEVAKKKSKK (SEQ ID NO: 182), RKCLQAGMNLEARKTKK (SEQ ID NO: 183), PRPRKIPR (SEQ ID NO: 184), PPRKKRTVV (SEQ ID NO: 185), NLSKKKKRKREK (SEQ ID NO: 186), RRPSRPFRKP (SEQ ID NO: 187), KRPRSPSS (SEQ ID NO: 188), KRGINDRNFWRGENERKTR (SEQ ID NO: 189), PRPPKMARYDN (SEQ ID NO: 190), KRSFSKAF (SEQ ID NO: 191), KLKIKRPVK (SEQ ID NO: 192), PKKKRKVPPPPAAKRVKLD (SEQ ID NO: 193), PKTRRRPRRSQRKRPPT (SEQ ID NO:26792), SRRRKANPTKLSENAKKLAKEVEN (SEQ ID NO: 194), KTRRRPRRSQRKRPPT (SEQ ID NO: 195), RRKKRRPRRKKRR (SEQ ID NO: 196), PKKKSRKPKKKSRK (SEQ ID NO: 197), HKKKHPDASVNFSEFSK (SEQ ID NO: 198), QRPGPYDRPQRPGPYDRP (SEQ ID NO: 199), LSPSLSPLLSPSLSPL (SEQ ID NO: 200), RGKGGKGLGKGGAKRHRK (SEQ ID NO: 201), PKRGRGRPKRGRGR (SEQ ID NO: 202), PKKKRKVPPPPAAKRVKLD (SEQ ID NO: 203), PKKKRKVPPPPKKKRKV (SEQ ID NO: 204), PAKRARRGYKC (SEQ ID NO: 27199), KLGPRKATGRW (SEQ ID NO: 27200), PRRKREE (SEQ ID NO: 27201), PYRGRKE (SEQ ID NO: 27202), PLRKRPRR (SEQ ID NO: 27203), PLRKRPRRGSPLRKRPRR (SEQ ID NO: 27204), PAAKRVKLDGGKRTADGSEFESPKKKRKV (SEQ ID NO: 27205), PAAKRVKLDGGKRTADGSEFESPKKKRKVGIHGVPAA (SEQ ID NO: 27206), PAAKRVKLDGGKRTADGSEFESPKKKRKVAEAAAKEAAAKEAAAKA (SEQ ID NO: 207), PAAKRVKLDGGKRTADGSEFESPKKKRKVPG (SEQ ID NO: 27208), KRKGSPERGERKRHW (SEQ ID NO: 27209), KRTADSQHSTPPKTKRKVEFEPKKKRKV (SEQ ID NO: 27210), and PKKKRKVGGSKRTADSQHSTPPKTKRKVEFEPKKKRKV (SEQ ID NO: 27211), wherein the one or more NLS are linked to the CasX variant protein or to adjacent NLS with a linker peptide wherein the linker peptide is selected from the group consisting of RS, (G)n (SEQ ID NO: 27212), (GS)n (SEQ ID NO: 27213), (GSGGS)n (SEQ ID

NO: 214), (GGSGGS)n (SEQ ID NO: 215), (GGGS)n (SEQ ID NO: 216), GGSG (SEQ ID NO: 217), GGSGG (SEQ ID NO: 218), GSGSG (SEQ ID NO: 219), GSGGG (SEQ ID NO: 220), GGGSG (SEQ ID NO: 221), GSSSG (SEQ ID NO: 222), GPGP (SEQ ID NO: 223), GGP, PPP, PPAPPA (SEQ ID NO: 224), PPPG (SEQ ID NO: 27214), PPPGPPP (SEQ ID NO: 225), PPP(GGGS)n (SEQ ID NO: 27215), (GGGS)nPPP (SEQ ID NO: 27216), AEAAAKEAAAKEAAAKA (SEQ ID NO: 27217), and TPPKTKRKVEFE (SEQ ID NO: 27218), wherein n is 1 to 5.

29. The system of claim 27 or claim 28, wherein the one or more NLS are located at or near the C-terminus of the CasX variant protein, at or near the N-terminus of the CasX variant protein, or at or near the N-terminus and at or near the C-terminus of the CasX variant protein.

30. The system of any one of claims 1-29, wherein the CasX variant protein forms a ribonuclear protein complex (RNP) with the gRNA variant.

31. The system of claim 30, wherein the RNP exhibits greater editing efficiency and/or binding of a target nucleic acid sequence when any one of the PAM sequences TTC, ATC, GTC, or CTC is located 1 nucleotide 5’ to the non-target strand of a protospacer having identity with the targeting sequence of the gRNA in a cellular assay system compared to the editing efficiency and/or binding of an RNP comprising a reference CasX protein and a reference gRNA in a comparable assay system.

32. The system of claim 31, wherein the PAM sequence is TTC.

33. The system of claim 32, wherein the targeting sequence of the gRNA comprises a sequence selected from the group consisting of SEQ ID NOS: 17904-26789.

34. The system of claim 33, wherein the PAM sequence is ATC.

35. The system of claim 34, wherein the targeting sequence of the gRNA comprises a sequence selected from the group consisting of SEQ ID NOS: 272-2100 and 2286-5625.

36. The system of claim 35, wherein the PAM sequence is CTC.

37. The system of claim 36, wherein the targeting sequence of the gRNA variant comprises a sequence selected from the group consisting of SEQ ID NOS: 5626-13616.

38. The system of claim 37, wherein the PAM sequence is GTC.

39. The system of claim 38, wherein the targeting sequence of the gRNA variant comprises a sequence selected from the group consisting of SEQ ID NOS: 13617-17903.

40. The system of any one of claims 31-39, wherein the increased editing and/or binding affinity for the one or more PAM sequences is at least 1.5-fold greater compared to the binding affinity of any one of the reference CasX proteins of SEQ ID NOS: 1-3 for the PAM sequences.

41. The system of any one of claims 31-39, wherein the RNP has at least a 5%, at least a 10%, at least a 15%, or at least a 20% higher percentage of cleavage-competent RNP compared to an RNP of the reference CasX protein and the gRNA of SEQ ID NOs: 4-16.

42. One or more nucleic acids comprising or encoding the CasX variant protein and the gRNA variant of the system of any one of claims 1-41.

43. The one or more nucleic acids of claim 42, wherein the sequence that encodes the CasX variant protein is codon optimized for expression in a eukaryotic cell.

44. A vector comprising the one or more nucleic acids of claim 42 or claim 43, wherein the vector is selected from the group consisting of a retroviral vector, a lentiviral vector, an adenoviral vector, an adeno-associated viral (AAV) vector, a herpes simplex virus (HSV) vector, a plasmid, a minicircle, a nanoplasmid, a DNA vector, an RNA vector, a lipid nanoparticle and a liposome.

45. The vector of claim 44, wherein the AAV vector is selected from AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV-Rh74, or AAVRh10.

46. A host cell comprising the vector of claim 44 or claim 45.

47. A method of modifying a BCL11A target nucleic acid sequence in cell in vitro or ex vivo, the method comprising introducing into cells:

a. the system of any one of claims 1-41;
b. the one or more nucleic acids of claim 42 or claim 43;
c. the vector of claim 44 or claim 45;
d. the XDP of any one of claims **Error! Reference source not found.**-**Error! Reference source not found.**; or
e. combinations thereof,

wherein the BCL11A target nucleic acid sequence of a cell targeted by the gRNA variant is modified by the CasX variant protein.

48. The method of claim 47, wherein the modifying comprises introducing an insertion, deletion, substitution, duplication, or inversion of one or more nucleotides in the BCL11A target nucleic acid sequence of the cell.

49. The method of claim 48, wherein a GATA1 binding site sequence of the target nucleic acid is modified.

50. The method of any one of claims 47-49, wherein the cells are eukaryotic.

51. The method of claim 50, wherein the eukaryotic cells are selected from the group consisting of rodent cells, mouse cells, rat cells, and non-human primate cells.

52. The method of claim 51, wherein the eukaryotic cells are human cells selected from the group consisting of a hematopoietic stem cell (HSC), a hematopoietic progenitor cell (HPC), a CD34+ cell, a mesenchymal stem cell (MSC), induced pluripotent stem cell (iPSC), a common myeloid progenitor cell, a proerythroblast cell, and an erythroblast cell.

53. The system of any one of claims 1-41, the nucleic acid of claim 42 or 43, or the vector of claim 44 or claim 45, for use in the treatment of a hemoglobinopathy.

54. A system for modifying a polypyrimidine tract-binding protein 1 (BCL11A) gene target nucleic acid sequence, the system comprising a CasX variant protein and a guide ribonucleic acid (gRNA) variant, wherein the gRNA variant comprises:

(i) a scaffold stem loop sequence of SEQ ID NO: 25; and

(ii) a targeting sequence complementary to a target nucleic acid sequence comprising a region within a polypyrimidine tract-binding protein 1 (BCL11A) gene, wherein the targeting sequence of the gRNA variant comprises a sequence selected from the group consisting of SEQ ID NOS: 22, 23, 2949, 2948, 15747, and 15748,

wherein the system is capable of modifying the BCL11A gene, wherein the modifying comprises introducing an insertion, deletion, substitution, duplication, or inversion of one or more nucleotides in the BCL11A gene.