

CLAIMS

What is claimed is:

1. A fusion protein comprising, in N-terminal to C-terminal order: a DNA targeting domain and a first transposase domain comprising the sequence set forth in SEQ ID NO: 544, wherein the first transposase domain comprises a deletion of amino acids 1-83, 1-84, 1-85, 186, 1-87, 1-88, 1-89, 1-90, 1-91, 1-92, 1-93, 1-94, 1-95, 1-96, 1-97, 1-98, 1-99, 1-100, 1-101, 1-102 or 1-103 of SEQ ID NO: 544.
2. The fusion protein of claim 1, wherein the DNA targeting domain comprises (a) three Zinc Finger Motifs; or (b) one or more TAL domains.
3. The fusion protein of claim 2, wherein the TAL domain comprises the sequence set forth in any one of SEQ ID NOs: 107-110.
4. The fusion protein of any one of claims 1-3, wherein the DNA targeting domain binds to a nucleic acid sequence encoding GFP, zinc finger 268 (ZFM268), phenylalanine hydroxylase (PAH), beta-2-microglobulin (B2M) or a LINE1 repeat element.
5. The fusion protein of any one of claims 1-4, wherein the first transposase domain and the DNA targeting domain are connected by a linker.
6. The fusion protein of claim 5, wherein the linker comprises the sequence GGGGS (SEQ ID NO: 78).
7. The fusion protein of any one of claims 1-6, wherein the first transposase domain comprises an N-terminal deletion of amino acids 1-85 of SEQ ID NO: 544.
8. The fusion protein of any one of claims 1-7, wherein the transposase domain comprises the sequence set forth in any one of SEQ ID NOs: 86-106.
9. The fusion protein of any one of claims 1-8, wherein the first transposase domain comprises (a) at least one mutation selected from the group consisting of M185R, M185K, D197K, D197R, D198K, D198R, D201K, and D201R; or (b) at least one mutation selected from the group consisting of L204D, L204E, K500D, K500E, R504E, and R504D.

10. The fusion protein of any one of claim 1-9, further comprising a second transposase domain C-terminal to the first transposase domain, wherein the second transposase domain comprises the sequence set forth in SEQ ID NO: 544.
11. The fusion protein of claim 10, wherein the second transposase domain comprises a deletion of N-terminal amino acids 1-83, 1-84, 1-85, 186, 1-87, 1-88, 1-89, 1-90, 1-91, 1-92, 1-93, 1-94, 1-95, 1-96, 1-97, 1-98, 1-99, 1-100, 1-101, 1-102 or 1-103 of SEQ ID NO: 544.
12. The fusion protein of claim 10 or 11, wherein the second transposase domain comprises (a) at least one mutation selected from the group consisting of M185R, M185K, D197K, D197R, D198K, D198R, D201K, and D201R; or (b) at least one mutation selected from the group consisting of L204D, L204E, K500D, K500E, R504E, and R504D.
13. A polynucleotide comprising a nucleic acid sequence encoding the fusion protein of any one of claims 1-12.
14. A vector comprising the polynucleotide of claim 13.
15. A method of integrating a transgene into a genomic target site of a cell, the method comprising introducing into the cell the fusion protein of any one of claims 1-12 and a transposon, wherein the transposon comprises, in 5' to 3' order: a 5' ITR, the transgene, and a 3' ITR.
16. The method of claim 15, wherein the transposon further comprises an exogenous promoter between the 5' ITR and the transgene.
17. The method of claim 15 or 16, wherein the transgene encodes a detectable marker.
18. The method of claim 17, wherein the detectable marker is GFP.
19. The method of any one of claims 15-18, wherein
 - (a) the genomic target site is located on chromosome 17 or 21;
 - (b) the genomic target site is located in the B2M gene;
 - (c) the genomic target site is located in a repetitive element; or
 - (d) the genomic target site is located in an intron of a gene.

20. The method of claim 19, wherein the repetitive element is a LINE element or the intron of the gene is the intron of the PAH gene.
21. The method of any one of claims 15-20, wherein the cell is in vivo.
22. A method of modifying the genome of a cell, the method comprising: providing the cell with the fusion protein of any one of claims 1-12, wherein the cell comprises a modified binding site comprising, in 5' to 3' order, the reverse of the sequence of a target site for the DNA targeting domain, a first spacer, a TTAA target integration site for SPB, a second spacer, and the complement of the sequence of the target site for the DNA targeting domain.