CLAIMS

1. A method for synthesizing a transcription activator-like effector (TALE) protein to target nucleic acid sequence comprising a 5-methyl-cytosine , said method comprising assembling a plurality of TALE-like repeat sequences, each of said sequences comprising a repeat variable-diresidue (RVD) specific to each nucleic acid base of said sequence, wherein the RVD(s) that specifically targets the 5-methyl-cytosine within said nucleic acid target sequence are selected from X*, where X represents one amino acid residue selected from the group of A,G, V, L, I, M, S, T, C, P, D, E, F, Y, W, Q, N, H, R and K and * represents a gap in one position of the RVD.

2. The method according to claim 1 wherein said method comprises:

(a) determining a nucleic acid target sequence comprising 5-methyl-cytosine in the genome of a cell;

(b) assembling TALE-like repeat polynucleotide sequences, each repeat being specific to each nucleic acid base of said nucleic acid target sequence by encoding a repeat variable-diresidue (RVD) comprising at least one RVD selected from the group consisting of:

- HD for recognizing C;
- NG for recognizing T;
- NI for recognizing A;
- NN for recognizing G or A;
- NS for recognizing A or C or G or T;
- HG for recognizing T;
- IG for recognizing T;
- NK for recognizing G;
- HA for recognizing C;
- ND for recognizing C;
- HI for recognizing C;
- HN for recognizing G;
- NA for recognizing G;
- SN for recognizing G or A; and
- YG for recognizing T;

wherein said RVD(s) specifically targeting the 5-methyl-cytosine(s) in the nucleic acid target sequence are selected from the RVDs X*, where

- X represents one amino acid residue selected from the group of A, G, V, L, I, M, S, T, C, P, D, E, F, Y, W, Q, N, H, R and K, and

- * represents a gap in one position of the RVD,

(c) expressing said polynucleotide sequence assembled in step b) in said cell.

3. The method according to claim 1 or 2 wherein said RVD specifically targeting the 5-methyl-cytosine(s) are preferentially selected from RVD N*, T*, Q* and H*.

4. A method to synthesize a chimeric protein to process nucleic acid at a locus defined by a nucleic acid target sequence that comprises a 5-methyl-cytosine , said method comprising:

(a) synthesizing a polynucleotide sequence comprising a fusion of:

(i) a first polynucleotide encoding a transcription activator-like effector (TALE) protein comprising a plurality of TALE-like repeat sequences, each repeat comprising a repeat variable-diresidue (RVD) specific to each nucleic acid base of said nucleic acid target sequence, wherein the RVD(s) that specifically targets the 5-methyl-cytosine within said nucleic acid target sequence are selected from X*, preferably N*, Q*, T* or H* where X represents one amino acid residue selected from the group of A,G, V, L, I, M, S, T, C, P, D, E, F, Y, W, Q, N, H, R and K and * represents a gap in one position of the RVD;

(ii) a second polynucleotide encoding an additional protein domain to process nucleic acid within or adjacent to said nucleic acid target sequence that comprises a 5-methyl-cytosine ;

(c) expressing said polynucleotide sequence of step a) into a host cell.

5. The method of claim 4 wherein said additional protein domain is an endonuclease.

6. A method to bind a nucleic acid target sequence comprising at least one 5-methyl-cytosine , said method comprising contacting:

(i) a nucleic acid target sequence comprising 5-methyl-cytosine and

(ii) a TALE protein comprising a repeat variable-diresidue (RVD) specific to each nucleic acid base of said nucleic acid target sequence, wherein the RVD(s) that specifically targets the 5-methyl-cytosine within said nucleic acid target sequence are selected from X*, preferably N*, Q*, T* or H* where X represents one amino acid residue selected from the group of A,G, V, L, I, M, S, T, C, P, D, E, F, Y, W, Q, N, H, R and K and * represents a gap in one position of the RVD.

7. A method to bind a nucleic acid target sequence comprising at least one 5-methyl-cytosine , said method comprising:

(a) providing a cell containing a nucleic acid target sequence that comprises a 5-methyl-cytosine ,

(b) synthesizing within said cell a TALE protein according to claims 1 to 4 directed to said nucleic acid target sequence and,

(c) testing the binding affinity of said TALE protein with said nucleic acid target sequence that comprises said 5-methyl-cytosine.

8. The method according to claim 6 or 7 wherein the said nucleic acid target sequence comprising at least one methylated dinucleotide selected from the group consisting of CpG, CpA, CpT, CpC.

9. A method to process a nucleic acid target sequence comprising at least one 5-methyl-cytosine comprising:

(a) providing a cell containing a nucleic acid target sequence that comprises a 5-methyl-cytosine;

(b) synthesizing within said cell a chimeric protein according to claims 4 to 5 directed to said nucleic acid sequence, so that said chimeric protein processes the nucleic acid within or adjacent to said nucleic acid target sequence independently of methylation and,

(c) testing the nucleic acid processing at the locus of said nucleic acid target sequence.

10. The method according to claim 9 wherein said protein domain has a catalytical activity selected from the group consisting of nuclease activity, polymerase activity, kinase activity, phosphatase activity, methylase activity, topoisomerase activity, integrase activity, transposase activity, ligase activity, helicase activity, recombinase activity.

11. The method according to claim 9 wherein said additional protein domain is a transcription activator.

12. The method according to claim 9 wherein said nucleic acid target sequence is a methylated promoter sequence.

13. The method according to claim 9 wherein said additional protein domain is an endonuclease.

14. The method according to claim 13 further comprising providing to the cell an exogenous nucleic acid comprising a sequence homologous to at least a portion of the nucleic acid target sequence, such that homologous recombination occurs between the nucleic acid target sequence and the exogenous nucleic acid.

15. A method to detect at least one 5-methyl-cytosine in a nucleic acid target sequence comprising:

(a) binding said nucleic acid target sequence with a transcription activator-like effector (TALE) protein comprising a plurality of TALE-like repeat sequences, each of said sequences comprising a repeat variable-diresidue (RVD) specific to each nucleic acid base of said nucleic acid target sequence wherein at least one RVD selected from the group consisting of:
    - HD for recognizing C;
    - NG for recognizing T;
    - NI for recognizing A;
    - NN for recognizing G or A;
    - NS for recognizing A or C or G or T;
    - HG for recognizing T;
    - IG for recognizing T;
    - NK for recognizing G;

- HA for recognizing C;
- ND for recognizing C;
- HI for recognizing C;
- HN for recognizing G;
- NA for recognizing G;
- SN for recognizing G or A; and
- YG for recognizing T;

(b) binding the same nucleic acid target sequence with another transcription activator-like effector (TALE) protein comprising a plurality of TALE-like repeat sequences, similar to that used in step a), wherein at least one RVD has been replaced by a RVD consisting of X*, preferably H*, T*, Q* or N*, wherein

- X represents one amino acid residue selected from the group of A,G, V, L, I, M, S, T, C, P, D, E, F, Y, W, Q, N, H, R and K,
- and * represents a gap in one position of the RVD,

(c) determining the binding affinity with said nucleic acid sequence under a) and b),

(d) calculating the ratio of binding activities determined under c), wherein said ratio, when close to 0, indicates the presence of methylated nucleic acid in said nucleic acid target sequence and, when close to 1, the absence of methylated nucleic acid in said nucleic acid target sequence.

16. The method according to claim 15 wherein binding affinity is measured by a nuclease activity.

17. The method according to claim 15 wherein binding affinity is measured by a transcriptional activity.

18. The method according to claim 15 wherein binding affinity is measured by detecting signal of fluorescent proteins fused to said TALE proteins of (a) and (b).

19. A composition for processing a nucleic acid target sequence comprising at least one 5-methyl-cytosine, wherein said composition comprises a TALE-nuclease comprising a plurality of TALE-like repeat sequences, each of said sequences comprising a repeat variable-

diresidue (RVD) specific to each nucleic acid base of said target sequence, wherein the RVD(s) that specifically target(s) the 5-methyl-cytosine within said nucleic acid target sequence are selected from X*, where X represents one amino acid residue selected from the group of A,G, V, L, I, M, S, T, C, P, D, E, F, Y, W, Q, N, H, R and K and * represents a gap in one position of the RVD.

20. The composition of claim 19, wherein said TALE-nuclease comprises or consist of a sequence selected from the group consisting of SEQ ID Nos: 39, 41, 44, 47, 49-51, 55 and 56.

21. A TALE-nuclease comprising at least one TALE-like repeat sequences with a repeat variable-diresidue (RVD) selected from the group consisting of T* and Q*, where * represents a gap in one position of the RVD.

22. The TALE-nuclease of claim 21, wherein said TALE-nuclease comprises or consist of a sequence selected from the group consisting of SEQ ID Nos: 55 and 56.