WHAT IS CLAIMED IS:

1. A method of modifying a target polynucleotide, the target polynucleotide comprising cytosine (C) and methylcytosine (mC), the method comprising:

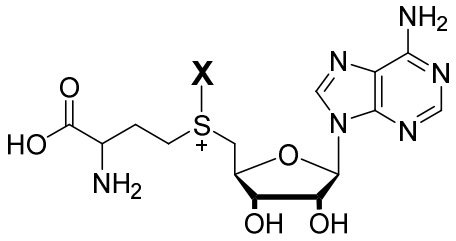
(a) protecting the C in the target polynucleotide from deamination;

(b) after step (a), deaminating the mC in the target polynucleotide to form thymine (T).

2. The method of claim 1, wherein protecting the C from deamination comprises adding a first protective group to the 5 position of the C; optionally wherein a first methyltransferase enzyme adds the first protective group to the 5 position of the C; and

further optionally

(i) wherein the first methyltransferase enzyme adds the first protective group from an S-adenosyl-L-methionine analog (xSAM) having the structure:



xSAM

where X includes the first protective group and a methylene group via which the first protective group is coupled to the sulfonium ion (S+); and/or

(ii) wherein the first methyltransferase enzyme is selected from the group consisting of: DNMT1, DNMT3A, DNMT3B, *dam*, and CpG (M.SssI).

3. The method of claim 1 or claim 2, wherein the first protective group comprises an alkyne group, a carboxyl group, an amino group, a hydroxymethyl group, an isopropyl group, or a dye; and/or

wherein the methyl group of mC inhibits addition of X to the 5 position of the mC.

4. The method of any one of claims 1 to 3, wherein a cytidine deaminase enzyme deaminates the mC;

optionally

wherein X fits within the first methyltransferase enzyme and inhibits activity of the cytidine deaminase enzyme; and/or

wherein the cytidine deaminase enzyme comprises APOBEC, further optionally wherein the APOBEC is selected from the group consisting of: APOBEC 1, APOBEC2, APOBEC3A, APOBEC3B, APOBEC3C, APOBEC3E, APOBEC3F, APOBEC3G, APOBEC3H, and APOBEC4.

5. The method of any one of claims 1 to 4, wherein the target polynucleotide further comprises hydroxymethylcytosine (hmC), and step (b) comprises deaminating the hmC in the target polynucleotide to form hydroxythymine (hT).

6. The method of any one of claims 1 to 4, wherein the target polynucleotide further comprises hydroxymethylcytosine (hmC), the method further comprising:

(c) before step (b), protecting the hmC in the target polynucleotide from deamination.

7. The method of claim 6,

(i) wherein step (c) is performed after step (a); and/or

(ii) comprising performing steps (a) and (b) on a first sample including the target polynucleotide, and performing steps (a), (b), and (c) on a second sample including the target polynucleotide; and/or

(iii) wherein protecting the hmC from deamination comprises adding a second protective group to the hydroxymethyl group of the hmC;

optionally

wherein an enzyme adds the second protective group to the hydroxymethyl group of the hmC, and further optionally wherein the enzyme is selected from the group consisting of: β-glucosyltransferase (βGT) and β-arabinosyltransferase (βAT); and/or

wherein the second protective group comprises a sugar.

8. The method of any one of claims 1 to 7, wherein the target polynucleotide further comprises formylcytosine (fC), wherein the formyl group of the fC inhibits deamination of the fC during step (b).

9. The method of any one of claims 1 to 7, wherein the target polynucleotide further comprises formylcytosine (fC), the method further comprising:

(d) before step (b), converting the fC to an unprotected C that is deaminated during step (b) to form uracil (U).

10. The method of claim 9, wherein a thymine deglycosylase enzyme replaces the base of fC with C; and/or

comprising performing steps (a) and (b) on a first sample including the target polynucleotide, and performing steps (a), (b), and (d) on a third sample including the target polynucleotide.

11. The method of any one of claims 1 to 10, wherein the target polynucleotide further comprises carboxylcytosine (caC), wherein the carboxyl group of the caC inhibits deamination of the fC during step (b).

12. The method of any one of claims 1 to 10, wherein the target polynucleotide further comprises carboxylcytosine (caC), the method further comprising:

(e) before step (b), converting the caC to unprotected C that is deaminated during step (b) to form uracil (U).

13. The method of claim 12,

(i) wherein a second methyltransferase enzyme removes the carboxyl group from caC, or a thymine deglycosylase enzyme replaces the base of caC with C; and/or

(ii) comprising performing steps (a) and (b) on a first sample including the target polynucleotide, and performing steps (a), (b), and (e) on a fourth sample including the target polynucleotide.

14. The method of any one of claims 1 to 13, wherein the target polynucleotide comprises DNA;

optionally wherein the target polynucleotide comprises first and second adapters; and

further optionally wherein the first and second adapters are added to the target polynucleotide before step (a) or after step (b).

15. A method of sequencing a target polynucleotide, the method comprising:

modifying the target polynucleotide in accordance with any one of claims 1 to 14;

generating a first amplicon of the modified target nucleotide, the first amplicon including a first guanine (G) at a location complementary to the protected C, and a first adenine (A) at a location complementary to the T;

generating a second amplicon of the first amplicon, the second amplicon including a first unprotected C at a location complementary to the first G, and a first thymine (T) at a location complementary to the first A;

sequencing the first amplicon, the second amplicon, or both the first amplicon and the second amplicon; and

identifying the mC based on the first A in the first amplicon, the first T in the second amplicon, or both the first A in the first amplicon and the first T in the second amplicon.

16. The method of claim 15 wherein:

(i) as dependent on claim 5, the first amplicon includes a second A at a location complementary to the hT and the second amplicon includes a second T at a location complementary to the second A, the method further comprising:

identifying the hmC based on the second A in the first amplicon, the second T in the second amplicon, or both the second A in the first amplicon and the second T in the second amplicon; or

(ii) as dependent on claim 6, the first amplicon includes a second G at a location complementary to the hmC and the second amplicon includes a second unprotected C at a location complementary to the second G, the method further comprising:

identifying the hmC based on the second G in the first amplicon, the second unprotected C in the second amplicon, or both the second G in the first amplicon and the second unprotected C in the second amplicon; or

(iii) as dependent on claim 8, the first amplicon includes a third G at a location complementary to the fC and the second amplicon includes a third unprotected C at a location complementary to the third G, the method further comprising:

identifying the fC based on the third G in the first amplicon, the third unprotected C in the second amplicon, or both the third G in the first amplicon and the third unprotected C in the second amplicon; or

(iv) as dependent on claim 9, the first amplicon includes a third A at a location complementary to the U and the second amplicon includes a third T at a location complementary to the third A, the method further comprising:

identifying the fC based on the third A in the first amplicon, the third T in the second amplicon, or both the third A in the first amplicon and the third T in the second amplicon; or

(v) as dependent on claim 11, the first amplicon includes a fourth G at a location complementary to the caC and the second amplicon includes a fourth unprotected C at a location complementary to the fourth G, the method further comprising:

identifying the caC based on the fourth G in the first amplicon, the fourth unprotected C in the second amplicon, or both the fourth G in the first amplicon and the fourth unprotected C in the second amplicon; or

(vi) as dependent on claim 12, the first amplicon includes a fourth A at a location complementary to the U and the second amplicon includes a fourth T at a location complementary to the fourth A, the method further comprising:

identifying the caC based on the fourth A in the first amplicon, the fourth T in the second amplicon, or both the fourth A in the first amplicon and the fourth T in the second amplicon.