**[0005]** 5mC and 5hmC may be referred to as epigenetic markers, and it can be desirable to detect them in a genomic sequence. The current golden standard method for detecting 5mC and 5hmC is bisulfite sequencing, which converts any unmethylated C in the sequence to uracil (U), but does not convert 5mC or 5hmC to the corresponding uracil derivatives. When the sequence is amplified using polymerase chain reaction (PCR), the uracil is amplified as thymidine (T), and as such the unmethylated C is sequenced as T. In comparison, the 5mC and 5hmC are amplified as C, and as such are sequenced as C. Thus, any Cs in the sequence may be identified as corresponding to 5mC or 5hmC because they had not been converted to U. Such a scheme may be referred to as a "three-base" sequencing scheme because any unmethylated C is converted to T. However, this type of scheme reduces sequence complexity and may lead to reduced sequencing quality, lower mapping rates, and relatively uneven coverage of the sequence.

## SUMMARY

**[0006]** Examples provided herein are related to detecting methylcytosine and its derivatives using S-adenosyl-L-methionine analogs (xSAMs). Compositions and methods for performing such detection are disclosed.

**[0007]** Some examples herein provide a method of modifying a target polynucleotide. The target polynucleotide may include cytosine (C) and methylcytosine (mC). The method may include (a) protecting the C in the target polynucleotide from deamination. The method may include (b) after step (a), deaminating the mC in the target polynucleotide to form thymine (T). In one aspect, the present disclosure refers to 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).

**[0008]** In some examples, protecting the C from deamination includes adding a first protective group to the 5 position of the C. In some examples, a first methyltransferase enzyme adds the first protective group to the 5 position of the C. In some examples, the first methyltransferase enzyme adds the first protective group from an S-adenosyl-L-methionine analog (xSAM) having the structure:

**[0023]** In some examples, the target polynucleotide includes first and second adapters. In some examples, the first and second adapters are added to the target polynucleotide before step (a). In some examples, the first and second adapters are added to the target polynucleotide after step (b).

**[0024]** Some examples herein provide a method of sequencing a target polynucleotide. The method may include modifying the target polynucleotide in accordance with any of the foregoing methods. The method may include generating a first amplicon of the modified target nucleotide. The first amplicon may include a first guanine (G) at a location complementary to the protected C, and a first adenine (A) at a location complementary to the T. The method may include 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. The method may include sequencing the first amplicon, the second amplicon, or both the first amplicon and the second amplicon. The method may include 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. In one aspect, the present disclosure refers to a method of sequencing a target polynucleotide, the method comprising: modifying the target polynucleotide in accordance with any one of the foregoing methods as disclosed herein; 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.

**[0025]** In some examples, 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 may include 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.

**[0026]** In some examples, 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 may include 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.

**[0027]** In some examples, 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 may include identifying the fC based on the third G in the