LIBRARIES FOR METHYLATION ANALYSIS

CROSS-REFERENCE

[0001] This application claims the benefit of U.S. provisional patent application number 63/326,647 filed on April 1, 2022 which is incorporated by reference in its entirety.

BACKGROUND

[0002] Nucleic acid analysis with high fidelity and low cost has a central role in biotechnology and medicine, and in basic biomedical research. While various methods are known for analyzing complex nucleic acid samples via hybridization-based processes, these techniques often suffer from scalability, automation, speed, accuracy, and cost.

INCORPORATION BY REFERENCE

[0003] All publications, patents, and patent applications mentioned in this specification are herein incorporated by reference to the same extent as if each individual publication, patent, or patent application was specifically and individually indicated to be incorporated by reference.

BRIEF SUMMARY

[0003a] In one aspect, provided herein is a sequencing method, comprising: contacting a methylation control with a plurality of genomic fragments, thereby generating a sample, wherein the methylation control comprises a target methylation level; converting unmethylated cytosines in the sample to uracil, thereby generating converted fragments; contacting the converted fragments with a probe library comprising a plurality of polynucleotide probes configured to hybridize to a genomic region comprising a CpG site; enriching a converted fragment that binds to a polynucleotide probe of the plurality of polynucleotide probes to generate an enriched target polynucleotide; and sequencing the enriched target polynucleotide.

[0003b] In another aspect, provided herein is a methylation detection system comprising: a probe library comprising a plurality of polynucleotide probes configured to hybridize to genomic fragments comprising CpG sites; a blocking library; and a methylation control comprising: a first polynucleotide comprising one or more CpG sites, wherein all of the one or more CpG sites of the first polynucleotide are methylated, a second polynucleotide comprising one or more CpG sites and a nucleotide sequence that is identical to a nucleotide sequence of the first polynucleotide, wherein the one or more CpG sites of the second polynucleotide are unmethylated, and wherein the first

polynucleotide is mixed with the second polynucleotide at a defined ratio to produce a target methylation level.

[0004] Provided herein are methods for detection of methylated nucleic acids, comprising: (a) contacting a probe library comprising a plurality of polynucleotide probes with a plurality of genomic fragments comprising methylated nucleic acids; (b) enriching at least one genomic fragment that binds to at least one polynucleotide probe in the probe library to generate at least one enriched target polynucleotide; and (c) sequencing the at least one enriched target polynucleotide. Further provided herein are methods wherein the method further comprises deamination of cytosine in the plurality of genomic fragments prior to step (a). Further provided herein are methods wherein deamination comprises treatment with bisulfite or one or more enzymes. Further provided herein are methods wherein the enzyme is APOBEC ("apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like"). Further provided herein are methods wherein the one or more enzymes are APOBEC and TET2. Further provided herein are methods wherein the probe library is configured to hybridize to at least one genomic fragment comprising a CpG island. Further provided herein are methods wherein the probe library is configured to hybridize to at least one genomic fragment comprising 5-methylcytosine or 5-hydroxymethlycytosine. Further provided herein are methods wherein the probe library comprises at least 5000 polynucleotide probes. Further provided herein are methods wherein the probe library comprises at least 1 million polynucleotide probes. Further