[0007] Zhou *et al.* (WO 2020/198664; PCT US2020/025425) teach a method of preparing sequencing libraries from cell-free DNA to facilitate 'genomic and epigenomic profiling of microbiome' but, here again, the aim is to separate mammalian nucleic acid molecules from non-mammalian such that most downstream sequencing reads are of microbial origin. Furthermore, while the method of Zhou *et al.* provides a means of preparing sequencing libraries that may be amenable to microbial epigenomic analyses, the manner of epigenomic analysis or the epigenetic features to be analyzed are not taught.

[0008] In contrast to the foregoing art wherein epigenetic features of exclusively mammalian or non-mammalian origin—but not of both—are subject to analysis, the methods of the present invention harness and combine the epigenetic data derived from taxonomically diverse life forms manifest within a nucleic acid sample. As microbes are increasingly implicated in mammalian disease processes and disease-specific mammalian epigenetic features have proven a robust source of diagnostic biomarkers, we reasoned that combining the epigenetic content from both mammalian and microbial taxonomic sources within a nucleic acid sample would enable the creation of highly sensitive and specific 'metaepigenomic' diagnostic signatures. In this manner we diverge sharply from all existing art and produce a novel method of identifying disease-diagnostic biomarkers.

[0008a] In one aspect, provided herein is a method of determining a disease of a subject, comprising: (a) providing a biological sample of a subject; (b) enriching one or more nucleic acid molecules of the biological sample by affinity targeting of an epigenetic feature common to the one or more nucleic acid molecules; (c) sequencing the enriched one or more nucleic acid molecules to generate one or more nucleic acid molecule sequencing reads; and (d) determining the disease of the subject as an output of a predictive model when the predictive model is provided the enriched one or more nucleic acid molecules as an input.

[0008b] In another aspect, provided herein is a method of training a predictive model, comprising: (a) providing a biological sample of one or more subjects with a disease; (b) enriching the biological sample of the one or more subjects by affinity targeting of an epigenetic feature common to one or more nucleic acid molecules of the biological sample; (c) sequencing the enriched one or more nucleic acid molecules to generate one or more nucleic acid molecule sequencing reads; and (d) training the predictive model with one or more features of the one or more nucleic acid molecule sequencing reads and the disease of the one or more subjects.

[0008c] In another aspect, provided herein is a computer system to determine a disease of a subject, comprising: (a) one or more processors; and (b) a non-transient computer readable

storage medium including software, wherein the software comprises executable instruction that, as a result of execution, cause the one or more processors of the computer system to: (i) receive a subject's one or more nucleic acid molecule sequencing reads of one or more nucleic acid molecules of a biological sample, wherein the one or more nucleic acid molecules are enriched by affinity targeting of an epigenetic feature common to the one or more nucleic acid molecules; and (ii) determine a disease of the subject as an output of a predictive model when the predictive model is provided the one or more nucleic acid molecule sequencing.

[0009] Aspects disclosed herein provide a method of creating a diagnostic model for diagnosing disease in a subject based on the combination of mammalian and non-mammalian epigenetic information contained in a nucleic acid sample, comprising: (a) enriching one or more mammalian and non-mammalian nucleic acid molecules by affinity targeting of an epigenetic feature shared by both the one or more mammalian and non-mammalian nucleic acid molecules; (b) sequencing the enriched nucleic acid compositions to generate sequencing reads; (c) filtering the sequencing reads with a build of a genome database to isolate non-mammalian sequencing reads and produce a mammalian alignment file; (d) analyzing the mammalian alignment file to generate mammalian feature abundance tables; (e) analyzing the non-mammalian sequencing reads to generate non-mammalian feature abundance tables; (f) combining the mammalian and non-mammalian feature abundance tables to generate combined meta-epigenomic machine learning feature sets; (g) training and testing predictive models on the meta-epigenomic feature sets to produce a trained predictive model; and (h) using an output of the trained predictive model to provide a diagnosis of a presence or absence of the disease in the subject. In some embodiments, the nucleic acid sample may be derived from a tissue, liquid biopsy sample or any combination thereof. In some embodiments, the subject may be a human or a non-human mammal. In some embodiments, the nucleic acids may comprise a total