MULTI-STEP METHODS FOR DETECTING NUCLEIC ACIDS

CROSS-REFERENCE

[0001] This application claims the benefit of U.S. Provisional Application No. 63/554,922, filed February 16, 2024, which application is incorporated herein by reference in its entirety.

BACKGROUND

[0002] Massively parallel sequencing (MPS), as the name implies, is a high-throughput technology that can generate an enormous amount of information about the genetic makeup of an organism. MPS is particularly useful for genomic studies that analyze sequences across a genome such as whole genome sequencing. MPS can be used to study cell-associated DNA, as well as cell-free DNA shed into a variety of samples including blood. MPS can also be used for metagenomic sequencing applications that detect microbial nucleic acids in a sample.

[0003] Detecting genetic markers is essential for diagnosing and treating diseases. In infectious diseases, identifying antimicrobial resistance (AMR) is particularly important for effective management. Early detection of AMR genes in microbes enables targeted antibiotic therapy, reducing the misuse of broad-spectrum antibiotics that may drive drug resistance. Rapid AMR identification can also aid in controlling outbreaks, especially in hospital settings where multidrug-resistant organisms pose serious risks. Additionally, AMR surveillance informs global health policies, guiding antibiotic stewardship programs and the development of new treatments. Without reliable AMR detection, infections become increasingly difficult to treat, leading to higher morbidity, mortality, and healthcare costs.

[0004] Current AMR detection methods often rely on traditional bacterial culture techniques, which are low throughput, time-consuming and labor-intensive. While sequencing-based approaches offer potential alternatives, they have limitations in sensitivity and accuracy. For example, molecular assays like PCR can identify known resistance genes but may not sufficiently provide information on phenotypic resistance of the organism. Therefore, there is a need for improved diagnostic tools that can rapidly and accurately detect a broad range of AMR determinants to inform effective treatment strategies. Provided herein are novel approaches for detecting AMR markers or carrier organisms harboring said AMR markers, determining antimicrobial resistance of the carrier organisms, and preparing samples useful for AMR detection.

SUMMARY

[0005] There remains a need to develop improved methods for detecting genetic markers in the context of disease diagnosis and treatment. The present disclosure provides novel approaches for detecting genetic markers using high-throughput sequencing.

[0006] This Summary introduces a selection of concepts that are described further below in the Detailed Description. This Summary is not intended to limit the scope of the claimed subject matter.

[0007] Disclosed herein are methods and compositions for performing a multi-step assay for detecting nucleic acids in a sample from a subject, the method comprising: providing a first aliquot of the sample from the subject, wherein the sample comprises nucleic acids comprising a first target nucleic acid and a second target nucleic acid that is different from the first target nucleic acid; performing a sequencing assay on the nucleic acids to produce sequence reads comprising sequence reads associated with the first target nucleic acid; analyzing the sequence reads, thereby obtaining an identification of the first target nucleic acid; after the identification of the first target nucleic acid is obtained, providing a second aliquot of the sample comprising nucleic acids comprising the first target nucleic acid and the second target nucleic acid; introducing primers into the second aliquot or to nucleic acids derived from the second aliquot wherein the primers specifically target the second target nucleic acid and do not target the first target nucleic acid; and conducting an amplification reaction on the second target nucleic acid with the primers, thereby amplifying the second target nucleic acid and producing amplicons associated with the second target nucleic acid. In some embodiments, the methods disclosed herein further comprise performing a high-throughput sequencing assay on the amplicons associated with the second target nucleic acid. In some embodiments, the second target nucleic acid is not detected by a sequencing assay. In some embodiments, the first nucleic acid is associated with a genome of an organism but not a phenotype of interest of the organism and the second nucleic acid is associated with the phenotype of interest of the organism. In some embodiments, the first target nucleic acid comprises at least 2, at least 3, at least 4, at least 5, at least 10, at least 15, at least 20, or at least 25 first target nucleic acids; or the second target nucleic acid comprises at least 2, at least 3, at least 4, at least 5, at least 10, at least 15, at least 20, or at least 25 second target nucleic acids. In some embodiments, the first target nucleic acid or the second target nucleic acid are not negative or positive controls for any step of the multi-step assay. In some embodiments, the primers comprise multiple primers targeting multiple target nucleic acids. In some embodiments, the first target nucleic acid comprises nucleic acids from an animal, a virus, a bacterium, a protozoa, a fungus, an archaea, or an algae. In some embodiments, the second target nucleic acid comprises nucleic acids from an animal, a virus, a

bacterium, a protozoa, a fungus, an archaea, or an algae. In some embodiments, the second target nucleic acid comprises a cancer marker. In some embodiments, the first target nucleic acid is associated with a carrier microbe harboring a target genetic marker and the second target nucleic acid comprises a sequence associated with the target genetic marker. In some embodiments, the carrier microbe comprises a virus, a bacterium, a protozoa, a fungus, an archaea, or an algae. In some embodiments, the target genetic marker comprises an antimicrobial resistance (AMR) genetic marker. In some embodiments, the methods disclosed herein further comprise comparing an abundance of the microbial sequence reads from the AMR genetic marker to a threshold value. In some embodiments, the methods disclosed herein further comprise comparing an abundance of the microbial sequence reads from the gene cassette to a threshold value. In some embodiments, the methods disclosed herein further comprise providing a second aliquot for the amplification reaction when the abundance of the microbial sequence reads from the AMR genetic marker is below the threshold. In some embodiments, the methods disclosed herein further comprise providing a second aliquot for the amplification reaction when the abundance of the microbial sequence reads from the gene cassette is below the threshold. In some embodiments, the methods disclosed herein further comprise quantifying mcfNA sequencing reads from the one or more microbes. In some embodiments, the methods disclosed herein further comprise comparing the abundance of the mcfNA sequencing reads from the one or more microbes to a threshold. In some embodiments, the methods disclosed herein further comprise providing a second aliquot for the amplification reaction when the abundance of the mcfNA sequencing reads from the one or more microbes is below the threshold. In some embodiments, the methods disclosed herein further comprise performing high-throughput sequencing on the amplicons associated with the AMR genetic marker. In some embodiments, the sequencing assay on the first nucleic acids comprises a high-throughput sequencing assay. In some embodiments, the methods disclosed herein further comprise conducting a polymerase chain reaction (PCR) to amplify the AMR genetic marker, thereby producing amplicons associated with the AMR genetic marker. In some embodiments, the PCR comprises multiplex PCR, random PCR (rPCR), non-biased PCR, Nested PCR, Hot Start PCR, or Assembly PCR. In some embodiments, the methods disclosed herein further comprise attaching an adapter sequence to the second nucleic acids. In some embodiments, the primers comprise an adapter sequence. In some embodiments, the methods disclosed herein further comprise physically manipulating the sample to produce a fraction of cfNA enriched for degraded cfNA, wherein the fraction of cfNA comprises the AMR genetic marker. In some embodiments, the degraded cfNA comprises ultra short cfNA, single stranded cfNA, or nicked double stranded cfNA. In some embodiments, the ultra short cfNA comprises cfNA fragments that are less than 100 nucleotides in length. In some embodiments, the ultra short cfNA comprises

cfNA fragments that are from 30 nucleotides to 70 nucleotides in length. In some embodiments, the methods disclosed herein further comprise performing size selection of the nucleic acids in the sample. In some embodiments, physically manipulating the sample comprises performing size selection of the nucleic acids in the sample. In some embodiments, performing size selection using a method selected from the group consisting of: chromatography, size-exclusion chromatography, electrophoresis, gel electrophoresis, automated electrophoresis, capillary electrophoresis, high-throughput size selection, centrifugation, density-gradient centrifugation, filtration, membrane ultrafiltration, magnetic beads, and affinity-based beads. In some embodiments, the second aliquot comprises at least 500 μl of plasma. In some embodiments, detecting the carrier microbe comprises determining an abundance of the carrier microbe. In some embodiments, the abundance is expressed as molecules of mcfNA per microliter of sample (MPM). In some embodiments, the methods disclosed herein further comprise calculating an AMR gene copy number from the amplification or sequencing of the AMR genetic marker. In some embodiments, the methods disclosed herein further comprise linking the AMR genetic marker to the carrier microbe using the abundance of the mcfNA from the carrier microbe and the AMR gene copy number. In some embodiments, the AMR gene copy number is an episomal gene copy number. In some embodiments, the methods disclosed herein comprise introducing at least 200 primers targeting the plurality of AMR genetic markers into the second aliquot of the sample. In some embodiments, the methods disclosed herein further comprise introducing primers targeting housekeeping genes into the second aliquot of the sample. In some embodiments, the methods disclosed herein further comprise determining antimicrobial resistance of the microbe infecting the subject. In some embodiments, the sample comprises plasma. In some embodiments, the first nucleic acids are DNA. In some embodiments, the second nucleic acids are DNA. In some embodiments, the sample is not subjected to a process that primarily causes cell lysis. In some embodiments, the methods disclosed herein further comprise preparing a library from the amplicons associated with the AMR genetic marker. In some embodiments, the primers are added directly to the second aliquot. In some embodiments, the sample comprises cell-free nucleic acids. In some embodiments, the sample comprises cell-free DNA. In some embodiments, the first nucleic acids comprise microbial cell-free nucleic acids. In some embodiments, the second nucleic acids comprise microbial cell-free nucleic acids. In some embodiments, intact microbes are not actively lysed prior to performing the sequencing assay. In some embodiments, the microbial nucleic acids comprise microbial cell-free nucleic acids (mcfNA) from the microbe. In some embodiments, the microbial nucleic acids comprise nucleic acids associated with a microbial cell. In some embodiments, the methods disclosed herein comprise enriching for at least at least 75%, at least 80%, at least 85%, or at least 90% of

the mcfNA in the sample. In some embodiments, the degraded cfNA comprise ultra short cfNA, single stranded cfNA, nicked double stranded cfNA, or any combination thereof. In some embodiments, the degraded cfNA comprise cfNA fragments that are less than 100 nucleotides in length. In some embodiments, the degraded cfNA comprise cfNA fragments that are from 30 nucleotides to 70 nucleotides in length. In some embodiments, physically manipulating the sample comprises performing size selection of the nucleic acids in the sample. In some embodiments, the methods disclosed herein further comprise generating sequence reads from the cfNA from the subject, wherein the sequence reads comprise microbial sequencing reads derived from the microbe infecting the subject. In some embodiments, the sequence reads further comprise microbial sequencing reads from one or more carrier microbe. In some embodiments, the methods disclosed herein further comprise calculating an abundance of the mcfNA from the microbe in the sample and an abundance of the mcfNA from the one or more carrier microbe. In some embodiments, the abundance is expressed as molecules of mcfNA per 100 nanoliters of sample. In some embodiments, the methods disclosed herein further comprise calculating an AMR gene copy number for each of the one or more carrier microbe. In some embodiments, the methods disclosed herein further comprise detecting one or more carrier microbe from the sequencing assay. In some embodiments, the methods disclosed herein further comprise identifying the one or more carrier microbe as the microbe infecting the subject. In some embodiments, the methods disclosed herein further comprise linking the AMR genetic marker to one of the carrier microbes or the microbe infecting the subject. In some embodiments, the methods disclosed herein further comprise obtaining an abundance of the one or more carrier microbes based on an abundance of carrier microbe sequences in the sample. In some embodiments, conducting a statistical analysis comprises using the abundance of the microbial nucleic acids of the one or more potential carrier microbes. In some embodiments, conducting a statistical analysis comprises using the AMR gene copy number. In some embodiments, the methods disclosed herein further comprise calculating a probability of each of the one or more potential carrier microbes being the microbe harboring the AMR genetic marker, thereby identifying the potential carrier microbe as the microbe infecting the subject. In some embodiments, the methods disclosed herein further comprise determining the antimicrobial resistance of the microbe infecting the subject based on the calculated probability. In some embodiments, the detecting comprises performing a sequencing assay. In some embodiments, the sequencing assay comprises a high-throughput sequencing assay. In some embodiments, the amplification reaction comprises introducing primers targeting a plurality of AMR genetic markers. In some embodiments, the amplification reaction produces amplicons associated with the AMR genetic marker. In some embodiments, the sample comprises plasma. In some

embodiments, the first nucleic acids are DNA. In some embodiments, the second nucleic acids are DNA. In some embodiments, the sample does not undergo cell lysis. In some embodiments, intact microbes are not actively lysed. In some embodiments, the amplicons associated with the AMR genetic marker undergoes direct library preparation. In some embodiments, the sample comprises cell-free nucleic acids. In some embodiments, the sample comprises cell-free DNA. In some embodiments, the first nucleic acids comprise cell-free nucleic acids. In some embodiments, the second nucleic acids comprise cell-free nucleic acids. In some embodiments, the methods disclosed herein comprise adding primers targeting a plurality of AMR genetic markers into the second aliquot of the sample. In some embodiments, the methods disclosed herein comprise adding at least 200 primers targeting a plurality of AMR genetic markers into the second aliquot of the sample. In some embodiments, the methods disclosed herein further comprise introducing primers targeting housekeeping genes into the second aliquot of the sample. In some embodiments, the second aliquot of the sample comprises at least 500 μl of plasma. In some embodiments, detecting the pathogen comprises detecting the abundance of mcfNA from the pathogen in the sample over a threshold. In some embodiments, the methods disclosed herein further comprise calculating a positive percent agreement (PPA), negative percent agreement (NPA), diagnostic yield (DY), or any combination thereof. In some embodiments, the methods disclosed herein further comprise detecting cfDNA from one or more housekeeping genes. In some embodiments, the methods disclosed herein further comprise spiking one or more control molecules into the sample at a known concentration. In some embodiments, the one or more control molecules are synthetic oligonucleotides. In some embodiments, the control molecules comprise whole assay internal control (WINC) molecules. In some embodiments, the methods disclosed herein further comprise spiking at least 25,000 unique WINC molecules at known concentrations. In some embodiments, the methods disclosed herein further comprise generating a report listing the carrier microbes or pathogen detected in the subject. In some embodiments, the report further comprises the abundance of microbial cell-free DNA (mcfDNA) from microbes detected in the sample or the antimicrobial resistance of the microbes infecting the subject. In some embodiments, the sample comprises a biological sample obtained from the subject. In some embodiments, the biological sample is a whole blood sample, a plasma sample, a serum sample, a cerebrospinal fluid sample, a synovial fluid sample, a urine sample, a stool sample, a bronchoalveolar lavage sample, or any combination thereof. In some embodiments, the sample is a plasma sample. In some embodiments, the AMR genetic marker comprises a gene, a genetic element, a genetic cassette, an allele, a mutation, or any combination thereof. In some embodiments, the AMR genetic marker is associated with one or more genes selected from the group consisting of: *SCCmec, mecA, mecC, vanA, vanB,* $bla_{CTX\text{-}M}$, bla_{KPC}, *OXA-48-like*, *OXA-23,*

NDM, VIM, IMP, or *mcr-1*. In some embodiments, the AMR genetic marker provides resistance to an anti-microbial agent selected from the group consisting of: methicillin, vancomycin, cephalosporin, carbapenem, and oxyimino-cephalosporin/aztreonam resistance. In some embodiments, the microbe infecting the subject or the carrier microbe comprises a virus, a bacterium, a protozoa, a fungus, an archaea, or an algae. In some embodiments, the microbe infecting the subject or the carrier microbe comprises a gram-positive bacterium or a fungus. In some embodiments, the microbe infecting the subject or the carrier microbe is a microbe listed in Table 1. In some embodiments, the microbe infecting the subject or the carrier microbe harbors at last two AMR genetic markers. In some embodiments, the antimicrobial resistance is phenotypic antimicrobial resistance. In some embodiments, the sequencing comprises next-generation sequencing or a sequencing method beyond next generation sequencing. In some embodiments, the sequencing comprises sequencing by synthesis. In some embodiments, the subject is an animal. In some embodiments, the subject is a human. In some embodiments, the subject has an infection by the microbe infecting the subject harboring the AMR genetic marker. In some embodiments, the methods disclosed herein further comprise administering an anti-infective agent to the subject. In some embodiments, the subject has been treated with an anti-infective agent for an infection. In some embodiments, the methods disclosed herein further comprise adjusting the anti-microbial agent received by the subject at least in part based on the antimicrobial resistance of the microbe infecting the subject. In some embodiments, the methods disclosed herein further comprise determining antimicrobial resistance of a microbe. In some embodiments, the microbe comprises a microbe infecting the subject. In some embodiments, the microbe comprises a carrier microbe. In some embodiments, the methods disclosed herein comprise determining whether an AMR gene is carried by a microbe infecting the subject or by a carrier microbe. In some embodiments, the determining comprises linking an AMR genetic marker to one of the carrier microbes or the microbe infecting the subject. In some embodiments, the determining comprises determining a copy number of the AMR genetic marker in the microbe. In some embodiments, the linking comprises performing a statistical analysis on the sequence reads. In some embodiments, the statistical analysis comprises a generalized linear model, a maximum-likelihood estimation, a probit model, a logistic regression, a linear probability, a linear regression, a complimentary log-log, a Poisson regression, a support vector machine, a decision tree, a random forest, a neural network, a gradient boosted model, a Bayesian model, a hidden Markov model, or any combination thereof. In some embodiments, the method further comprises comparing the sequence read data to the amplification data. In some embodiments, the method further comprises calculating estimated deduplicated templates (EDT) for an AMR gene, wherein calculating EDT comprises calculating a number of unique DNA sequencing reads from

a microorganism present in a sequenced library. In some embodiments, the method further comprises calculating an EDT for a housekeeping gene, wherein calculating EDT comprises calculating a number of unique DNA sequencing reads from a microorganism present in a sequenced library. In some embodiments, the methods disclosed herein further comprise determining a presence of an organism in a sample by calculating a ratio of an AMR estimated deduplicated templates (EDT) to a housekeeping gene EDT.

[0008] Disclosed herein are methods and compositions for performing a multi-step sample processing procedure for detecting nucleic acids in a sample from a subject: providing a first aliquot of the sample from the subject, wherein the sample comprises one or more first nucleic acids, wherein the one or more first nucleic acids comprise one or more microbial nucleic acids derived from one or more carrier microbes harboring an antimicrobial resistance (AMR) genetic marker; performing a sequencing assay on the one or more first nucleic acids to produce sequence reads comprising microbial nucleic acid sequence reads from the one or more carrier microbes; detecting the one or more carrier microbes by analyzing the microbial nucleic acid sequence reads; after at least one carrier microbe is detected, providing a second aliquot of the sample from the subject, wherein the second aliquot comprises second nucleic acids from at least one AMR genetic marker; annealing primers targeting at least one AMR genetic marker to the second nucleic acids; and amplifying the AMR genetic marker in an amplification reaction to produce amplicons associated with the at least one AMR genetic marker. In some embodiments, the methods disclosed herein further comprise performing a high-throughput sequencing assay on the amplicons associated with the second target nucleic acid. In some embodiments, the second target nucleic acid is not detected by a sequencing assay. In some embodiments, the first nucleic acid is associated with a genome of an organism but not a phenotype of interest of the organism and the second nucleic acid is associated with the phenotype of interest of the organism. In some embodiments, the first target nucleic acid comprises at least 2, at least 3, at least 4, at least 5, at least 10, at least 15, at least 20, or at least 25 first target nucleic acids; or the second target nucleic acid comprises at least 2, at least 3, at least 4, at least 5, at least 10, at least 15, at least 20, or at least 25 second target nucleic acids. In some embodiments, the first target nucleic acid or the second target nucleic acid are not negative or positive controls for any step of the multi-step assay. In some embodiments, the primers comprise multiple primers targeting multiple target nucleic acids. In some embodiments, the first target nucleic acid comprises nucleic acids from an animal, a virus, a bacterium, a protozoa, a fungus, an archaea, or an algae. In some embodiments, the second target nucleic acid comprises nucleic acids from an animal, a virus, a bacterium, a protozoa, a fungus, an archaea, or an algae. In some embodiments, the second target nucleic acid comprises a cancer marker. In some embodiments, the first target nucleic acid is

associated with a carrier microbe harboring a target genetic marker and the second target nucleic acid comprises a sequence associated with the target genetic marker. In some embodiments, the carrier microbe comprises a virus, a bacterium, a protozoa, a fungus, an archaea, or an algae. In some embodiments, the target genetic marker comprises an antimicrobial resistance (AMR) genetic marker. In some embodiments, the methods disclosed herein further comprise comparing an abundance of the microbial sequence reads from the AMR genetic marker to a threshold value. In some embodiments, the methods disclosed herein further comprise comparing an abundance of the microbial sequence reads from the gene cassette to a threshold value. In some embodiments, the methods disclosed herein further comprise providing a second aliquot for the amplification reaction when the abundance of the microbial sequence reads from the AMR genetic marker is below the threshold. In some embodiments, the methods disclosed herein further comprise providing a second aliquot for the amplification reaction when the abundance of the microbial sequence reads from the gene cassette is below the threshold. In some embodiments, the methods disclosed herein further comprise quantifying mcfNA sequencing reads from the one or more microbes. In some embodiments, the methods disclosed herein further comprise comparing the abundance of the mcfNA sequencing reads from the one or more microbes to a threshold. In some embodiments, the methods disclosed herein further comprise providing a second aliquot for the amplification reaction when the abundance of the mcfNA sequencing reads from the one or more microbes is below the threshold. In some embodiments, the methods disclosed herein further comprise performing high-throughput sequencing on the amplicons associated with the AMR genetic marker. In some embodiments, the sequencing assay on the first nucleic acids comprises a high-throughput sequencing assay. In some embodiments, the methods disclosed herein further comprise conducting a polymerase chain reaction (PCR) to amplify the AMR genetic marker, thereby producing amplicons associated with the AMR genetic marker. In some embodiments, the PCR comprises multiplex PCR, random PCR (rPCR), non-biased PCR, Nested PCR, Hot Start PCR, or Assembly PCR. In some embodiments, the methods disclosed herein further comprise attaching an adapter sequence to the second nucleic acids. In some embodiments, the primers comprise an adapter sequence. In some embodiments, the methods disclosed herein further comprise physically manipulating the sample to produce a fraction of cfNA enriched for degraded cfNA, wherein the fraction of cfNA comprises the AMR genetic marker. In some embodiments, the degraded cfNA comprises ultra short cfNA, single stranded cfNA, or nicked double stranded cfNA. In some embodiments, the ultra short cfNA comprises cfNA fragments that are less than 100 nucleotides in length. In some embodiments, the ultra short cfNA comprises cfNA fragments that are from 30 nucleotides to 70 nucleotides in length. In some embodiments, the methods disclosed herein further comprise performing size selection of the nucleic acids in

the sample. In some embodiments, physically manipulating the sample comprises performing size selection of the nucleic acids in the sample. In some embodiments, performing size selection using a method selected from the group consisting of: chromatography, size-exclusion chromatography, electrophoresis, gel electrophoresis, automated electrophoresis, capillary electrophoresis, high-throughput size selection, centrifugation, density-gradient centrifugation, filtration, membrane ultrafiltration, magnetic beads, and affinity-based beads. In some embodiments, the second aliquot comprises at least 500 μl of plasma. In some embodiments, detecting the carrier microbe comprises determining an abundance of the carrier microbe. In some embodiments, the abundance is expressed as molecules of mcfNA per microliter of sample (MPM). In some embodiments, the methods disclosed herein further comprise calculating an AMR gene copy number from the amplification or sequencing of the AMR genetic marker. In some embodiments, the methods disclosed herein further comprise linking the AMR genetic marker to the carrier microbe using the abundance of the mcfNA from the carrier microbe and the AMR gene copy number. In some embodiments, the AMR gene copy number is an episomal gene copy number. In some embodiments, the methods disclosed herein comprise introducing at least 200 primers targeting the plurality of AMR genetic markers into the second aliquot of the sample. In some embodiments, the methods disclosed herein further comprise introducing primers targeting housekeeping genes into the second aliquot of the sample. In some embodiments, the methods disclosed herein further comprise determining antimicrobial resistance of the microbe infecting the subject. In some embodiments, the sample comprises plasma. In some embodiments, the first nucleic acids are DNA. In some embodiments, the second nucleic acids are DNA. In some embodiments, the sample is not subjected to a process that primarily causes cell lysis. In some embodiments, the methods disclosed herein further comprise preparing a library from the amplicons associated with the AMR genetic marker. In some embodiments, the primers are added directly to the second aliquot. In some embodiments, the sample comprises cell-free nucleic acids. In some embodiments, the sample comprises cell-free DNA. In some embodiments, the first nucleic acids comprise microbial cell-free nucleic acids. In some embodiments, the second nucleic acids comprise microbial cell-free nucleic acids. In some embodiments, intact microbes are not actively lysed prior to performing the sequencing assay. In some embodiments, the microbial nucleic acids comprise microbial cell-free nucleic acids (mcfNA) from the microbe. In some embodiments, the microbial nucleic acids comprise nucleic acids associated with a microbial cell. In some embodiments, the methods disclosed herein comprise enriching for at least at least 75%, at least 80%, at least 85%, or at least 90% of the mcfNA in the sample. In some embodiments, the degraded cfNA comprise ultra short cfNA, single stranded cfNA, nicked double stranded cfNA, or any combination thereof. In some

embodiments, the degraded cfNA comprise cfNA fragments that are less than 100 nucleotides in length. In some embodiments, the degraded cfNA comprise cfNA fragments that are from 30 nucleotides to 70 nucleotides in length. In some embodiments, physically manipulating the sample comprises performing size selection of the nucleic acids in the sample. In some embodiments, the methods disclosed herein further comprise generating sequence reads from the cfNA from the subject, wherein the sequence reads comprise microbial sequencing reads derived from the microbe infecting the subject. In some embodiments, the sequence reads further comprise microbial sequencing reads from one or more carrier microbe. In some embodiments, the methods disclosed herein further comprise calculating an abundance of the mcfNA from the microbe in the sample and an abundance of the mcfNA from the one or more carrier microbe. In some embodiments, the abundance is expressed as molecules of mcfNA per 100 nanoliters of sample. In some embodiments, the methods disclosed herein further comprise calculating an AMR gene copy number for each of the one or more carrier microbe. In some embodiments, the methods disclosed herein further comprise detecting one or more carrier microbe from the sequencing assay. In some embodiments, the methods disclosed herein further comprise identifying the one or more carrier microbe as the microbe infecting the subject. In some embodiments, the methods disclosed herein further comprise linking the AMR genetic marker to one of the carrier microbes or the microbe infecting the subject. In some embodiments, the methods disclosed herein further comprise obtaining an abundance of the one or more carrier microbes based on an abundance of carrier microbe sequences in the sample. In some embodiments, conducting a statistical analysis comprises using the abundance of the microbial nucleic acids of the one or more potential carrier microbes. In some embodiments, conducting a statistical analysis comprises using the AMR gene copy number. In some embodiments, the methods disclosed herein further comprise calculating a probability of each of the one or more potential carrier microbes being the microbe harboring the AMR genetic marker, thereby identifying the potential carrier microbe as the microbe infecting the subject. In some embodiments, the methods disclosed herein further comprise determining the antimicrobial resistance of the microbe infecting the subject based on the calculated probability. In some embodiments, the detecting comprises performing a sequencing assay. In some embodiments, the sequencing assay comprises a high-throughput sequencing assay. In some embodiments, the amplification reaction comprises introducing primers targeting a plurality of AMR genetic markers. In some embodiments, the amplification reaction produces amplicons associated with the AMR genetic marker. In some embodiments, the sample comprises plasma. In some embodiments, the first nucleic acids are DNA. In some embodiments, the second nucleic acids are DNA. In some embodiments, the sample does not undergo cell lysis. In some embodiments,

intact microbes are not actively lysed. In some embodiments, the amplicons associated with the AMR genetic marker undergoes direct library preparation. In some embodiments, the sample comprises cell-free nucleic acids. In some embodiments, the sample comprises cell-free DNA. In some embodiments, the first nucleic acids comprise cell-free nucleic acids. In some embodiments, the second nucleic acids comprise cell-free nucleic acids. In some embodiments, the methods disclosed herein comprise adding primers targeting a plurality of AMR genetic markers into the second aliquot of the sample. In some embodiments, the methods disclosed herein comprise adding at least 200 primers targeting a plurality of AMR genetic markers into the second aliquot of the sample. In some embodiments, the methods disclosed herein further comprise introducing primers targeting housekeeping genes into the second aliquot of the sample. In some embodiments, the second aliquot of the sample comprises at least 500 μl of plasma. In some embodiments, detecting the pathogen comprises detecting the abundance of mcfNA from the pathogen in the sample over a threshold. In some embodiments, the methods disclosed herein further comprise calculating a positive percent agreement (PPA), negative percent agreement (NPA), diagnostic yield (DY), or any combination thereof. In some embodiments, the methods disclosed herein further comprise detecting cfDNA from one or more housekeeping genes. In some embodiments, the methods disclosed herein further comprise spiking one or more control molecules into the sample at a known concentration. In some embodiments, the one or more control molecules are synthetic oligonucleotides. In some embodiments, the control molecules comprise whole assay internal control (WINC) molecules. In some embodiments, the methods disclosed herein further comprise spiking at least 25,000 unique WINC molecules at known concentrations. In some embodiments, the methods disclosed herein further comprise generating a report listing the carrier microbes or pathogen detected in the subject. In some embodiments, the report further comprises the abundance of microbial cell-free DNA (mcfDNA) from microbes detected in the sample or the antimicrobial resistance of the microbes infecting the subject. In some embodiments, the sample comprises a biological sample obtained from the subject. In some embodiments, the biological sample is a whole blood sample, a plasma sample, a serum sample, a cerebrospinal fluid sample, a synovial fluid sample, a urine sample, a stool sample, a bronchoalveolar lavage sample, or any combination thereof. In some embodiments, the sample is a plasma sample. In some embodiments, the AMR genetic marker comprises a gene, a genetic element, a genetic cassette, an allele, a mutation, or any combination thereof. In some embodiments, the AMR genetic marker is associated with one or more genes selected from the group consisting of: *SCCmec, mecA, mecC, vanA, vanB,* $bla_{CTX\text{-}M}$, bla_{KPC}, *OXA-48-like*, *OXA-23, NDM, VIM, IMP*, or *mcr-1*. In some embodiments, the AMR genetic marker provides resistance to an anti-microbial agent selected from the group consisting of: methicillin, vancomycin,

cephalosporin, carbapenem, and oxyimino-cephalosporin/aztreonam resistance. In some embodiments, the microbe infecting the subject or the carrier microbe comprises a virus, a bacterium, a protozoa, a fungus, an archaea, or an algae. In some embodiments, the microbe infecting the subject or the carrier microbe comprises a gram-positive bacterium or a fungus. In some embodiments, the microbe infecting the subject or the carrier microbe is a microbe listed in Table 1. In some embodiments, the microbe infecting the subject or the carrier microbe harbors at last two AMR genetic markers. In some embodiments, the antimicrobial resistance is phenotypic antimicrobial resistance. In some embodiments, the sequencing comprises next-generation sequencing or a sequencing method beyond next generation sequencing. In some embodiments, the sequencing comprises sequencing by synthesis. In some embodiments, the subject is an animal. In some embodiments, the subject is a human. In some embodiments, the subject has an infection by the microbe infecting the subject harboring the AMR genetic marker. In some embodiments, the methods disclosed herein further comprise administering an anti-infective agent to the subject. In some embodiments, the subject has been treated with an anti-infective agent for an infection. In some embodiments, the methods disclosed herein further comprise adjusting the anti-microbial agent received by the subject at least in part based on the antimicrobial resistance of the microbe infecting the subject. In some embodiments, the methods disclosed herein further comprise determining antimicrobial resistance of a microbe. In some embodiments, the microbe comprises a microbe infecting the subject. In some embodiments, the microbe comprises a carrier microbe. In some embodiments, the methods disclosed herein comprise determining whether an AMR gene is carried by a microbe infecting the subject or by a carrier microbe. In some embodiments, the determining comprises linking an AMR genetic marker to one of the carrier microbes or the microbe infecting the subject. In some embodiments, the determining comprises determining a copy number of the AMR genetic marker in the microbe. In some embodiments, the linking comprises performing a statistical analysis on the sequence reads. In some embodiments, the statistical analysis comprises a generalized linear model, a maximum-likelihood estimation, a probit model, a logistic regression, a linear probability, a linear regression, a complimentary log-log, a Poisson regression, a support vector machine, a decision tree, a random forest, a neural network, a gradient boosted model, a Bayesian model, a hidden Markov model, or any combination thereof. In some embodiments, the method further comprises comparing the sequence read data to the amplification data. In some embodiments, the method further comprises calculating estimated deduplicated templates (EDT) for an AMR gene, wherein calculating EDT comprises calculating a number of unique DNA sequencing reads from a microorganism present in a sequenced library. In some embodiments, the method further comprises calculating an EDT for a housekeeping gene, wherein calculating EDT comprises

calculating a number of unique DNA sequencing reads from a microorganism present in a sequenced library. In some embodiments, the methods disclosed herein further comprise determining a presence of an organism in a sample by calculating a ratio of an AMR estimated deduplicated templates (EDT) to a housekeeping gene EDT.

[0009] Disclosed herein are methods and compositions for assaying nucleic acids in a sample from a subject comprising: providing a first aliquot of the sample from the subject, wherein the sample comprises one or more first nucleic acids, wherein the one or more first nucleic acids comprise one or more microbial nucleic acids derived from one or more carrier microbes harboring an antimicrobial resistance (AMR) genetic marker; performing a sequencing assay on the one or more first nucleic acids to produce sequence reads comprising microbial nucleic acid sequence reads from the one or more carrier microbes; detecting the one or more carrier microbes by analyzing the microbial sequence reads; providing a second aliquot of the sample from the subject, wherein the second aliquot comprises second nucleic acids, wherein the second aliquot comprises second nucleic acids from at least one AMR genetic marker; annealing primers to the second nucleic acids that target the at least one AMR genetic marker, wherein the primers that target the at least one AMR genetic marker are annealed prior to amplification of nucleic acids in the sample; and conducting an amplification reaction on the second nucleic acids using the primers targeting the at least one AMR genetic marker, thereby amplifying at least one AMR genetic marker and producing amplicons associated with the at least one AMR genetic marker. In some embodiments, the methods disclosed herein further comprise performing a high-throughput sequencing assay on the amplicons associated with the second target nucleic acid. In some embodiments, the second target nucleic acid is not detected by a sequencing assay. In some embodiments, the first nucleic acid is associated with a genome of an organism but not a phenotype of interest of the organism and the second nucleic acid is associated with the phenotype of interest of the organism. In some embodiments, the first target nucleic acid comprises at least 2, at least 3, at least 4, at least 5, at least 10, at least 15, at least 20, or at least 25 first target nucleic acids; or the second target nucleic acid comprises at least 2, at least 3, at least 4, at least 5, at least 10, at least 15, at least 20, or at least 25 second target nucleic acids. In some embodiments, the first target nucleic acid or the second target nucleic acid are not negative or positive controls for any step of the multi-step assay. In some embodiments, the primers comprise multiple primers targeting multiple target nucleic acids. In some embodiments, the first target nucleic acid comprises nucleic acids from an animal, a virus, a bacterium, a protozoa, a fungus, an archaea, or an algae. In some embodiments, the second target nucleic acid comprises nucleic acids from an animal, a virus, a bacterium, a protozoa, a fungus, an archaea, or an algae. In some embodiments, the second target nucleic acid comprises a cancer marker. In some

embodiments, the first target nucleic acid is associated with a carrier microbe harboring a target genetic marker and the second target nucleic acid comprises a sequence associated with the target genetic marker. In some embodiments, the carrier microbe comprises a virus, a bacterium, a protozoa, a fungus, an archaea, or an algae. In some embodiments, the target genetic marker comprises an antimicrobial resistance (AMR) genetic marker. In some embodiments, the methods disclosed herein further comprise comparing an abundance of the microbial sequence reads from the AMR genetic marker to a threshold value. In some embodiments, the methods disclosed herein further comprise comparing an abundance of the microbial sequence reads from the gene cassette to a threshold value. In some embodiments, the methods disclosed herein further comprise providing a second aliquot for the amplification reaction when the abundance of the microbial sequence reads from the AMR genetic marker is below the threshold. In some embodiments, the methods disclosed herein further comprise providing a second aliquot for the amplification reaction when the abundance of the microbial sequence reads from the gene cassette is below the threshold. In some embodiments, the methods disclosed herein further comprise quantifying mcfNA sequencing reads from the one or more microbes. In some embodiments, the methods disclosed herein further comprise comparing the abundance of the mcfNA sequencing reads from the one or more microbes to a threshold. In some embodiments, the methods disclosed herein further comprise providing a second aliquot for the amplification reaction when the abundance of the mcfNA sequencing reads from the one or more microbes is below the threshold. In some embodiments, the methods disclosed herein further comprise performing high-throughput sequencing on the amplicons associated with the AMR genetic marker. In some embodiments, the sequencing assay on the first nucleic acids comprises a high-throughput sequencing assay. In some embodiments, the methods disclosed herein further comprise conducting a polymerase chain reaction (PCR) to amplify the AMR genetic marker, thereby producing amplicons associated with the AMR genetic marker. In some embodiments, the PCR comprises multiplex PCR, random PCR (rPCR), non-biased PCR, Nested PCR, Hot Start PCR, or Assembly PCR. In some embodiments, the methods disclosed herein further comprise attaching an adapter sequence to the second nucleic acids. In some embodiments, the primers comprise an adapter sequence. In some embodiments, the methods disclosed herein further comprise physically manipulating the sample to produce a fraction of cfNA enriched for degraded cfNA, wherein the fraction of cfNA comprises the AMR genetic marker. In some embodiments, the degraded cfNA comprises ultra short cfNA, single stranded cfNA, or nicked double stranded cfNA. In some embodiments, the ultra short cfNA comprises cfNA fragments that are less than 100 nucleotides in length. In some embodiments, the ultra short cfNA comprises cfNA fragments that are from 30 nucleotides to 70 nucleotides in length. In some embodiments,

the methods disclosed herein further comprise performing size selection of the nucleic acids in the sample. In some embodiments, physically manipulating the sample comprises performing size selection of the nucleic acids in the sample. In some embodiments, performing size selection using a method selected from the group consisting of: chromatography, size-exclusion chromatography, electrophoresis, gel electrophoresis, automated electrophoresis, capillary electrophoresis, high-throughput size selection, centrifugation, density-gradient centrifugation, filtration, membrane ultrafiltration, magnetic beads, and affinity-based beads. In some embodiments, the second aliquot comprises at least 500 μl of plasma. In some embodiments, detecting the carrier microbe comprises determining an abundance of the carrier microbe. In some embodiments, the abundance is expressed as molecules of mcfNA per microliter of sample (MPM). In some embodiments, the methods disclosed herein further comprise calculating an AMR gene copy number from the amplification or sequencing of the AMR genetic marker. In some embodiments, the methods disclosed herein further comprise linking the AMR genetic marker to the carrier microbe using the abundance of the mcfNA from the carrier microbe and the AMR gene copy number. In some embodiments, the AMR gene copy number is an episomal gene copy number. In some embodiments, the methods disclosed herein comprise introducing at least 200 primers targeting the plurality of AMR genetic markers into the second aliquot of the sample. In some embodiments, the methods disclosed herein further comprise introducing primers targeting housekeeping genes into the second aliquot of the sample. In some embodiments, the methods disclosed herein further comprise determining antimicrobial resistance of the microbe infecting the subject. In some embodiments, the sample comprises plasma. In some embodiments, the first nucleic acids are DNA. In some embodiments, the second nucleic acids are DNA. In some embodiments, the sample is not subjected to a process that primarily causes cell lysis. In some embodiments, the methods disclosed herein further comprise preparing a library from the amplicons associated with the AMR genetic marker. In some embodiments, the primers are added directly to the second aliquot. In some embodiments, the sample comprises cell-free nucleic acids. In some embodiments, the sample comprises cell-free DNA. In some embodiments, the first nucleic acids comprise microbial cell-free nucleic acids. In some embodiments, the second nucleic acids comprise microbial cell-free nucleic acids. In some embodiments, intact microbes are not actively lysed prior to performing the sequencing assay. In some embodiments, the microbial nucleic acids comprise microbial cell-free nucleic acids (mcfNA) from the microbe. In some embodiments, the microbial nucleic acids comprise nucleic acids associated with a microbial cell. In some embodiments, the methods disclosed herein comprise enriching for at least at least 75%, at least 80%, at least 85%, or at least 90% of the mcfNA in the sample. In some embodiments, the degraded cfNA comprise ultra short cfNA,

single stranded cfNA, nicked double stranded cfNA, or any combination thereof. In some embodiments, the degraded cfNA comprise cfNA fragments that are less than 100 nucleotides in length. In some embodiments, the degraded cfNA comprise cfNA fragments that are from 30 nucleotides to 70 nucleotides in length. In some embodiments, physically manipulating the sample comprises performing size selection of the nucleic acids in the sample. In some embodiments, the methods disclosed herein further comprise generating sequence reads from the cfNA from the subject, wherein the sequence reads comprise microbial sequencing reads derived from the microbe infecting the subject. In some embodiments, the sequence reads further comprise microbial sequencing reads from one or more carrier microbe. In some embodiments, the methods disclosed herein further comprise calculating an abundance of the mcfNA from the microbe in the sample and an abundance of the mcfNA from the one or more carrier microbe. In some embodiments, the abundance is expressed as molecules of mcfNA per 100 nanoliters of sample. In some embodiments, the methods disclosed herein further comprise calculating an AMR gene copy number for each of the one or more carrier microbe. In some embodiments, the methods disclosed herein further comprise detecting one or more carrier microbe from the sequencing assay. In some embodiments, the methods disclosed herein further comprise identifying the one or more carrier microbe as the microbe infecting the subject. In some embodiments, the methods disclosed herein further comprise linking the AMR genetic marker to one of the carrier microbes or the microbe infecting the subject. In some embodiments, the methods disclosed herein further comprise obtaining an abundance of the one or more carrier microbes based on an abundance of carrier microbe sequences in the sample. In some embodiments, conducting a statistical analysis comprises using the abundance of the microbial nucleic acids of the one or more potential carrier microbes. In some embodiments, conducting a statistical analysis comprises using the AMR gene copy number. In some embodiments, the methods disclosed herein further comprise calculating a probability of each of the one or more potential carrier microbes being the microbe harboring the AMR genetic marker, thereby identifying the potential carrier microbe as the microbe infecting the subject. In some embodiments, the methods disclosed herein further comprise determining the antimicrobial resistance of the microbe infecting the subject based on the calculated probability. In some embodiments, the detecting comprises performing a sequencing assay. In some embodiments, the sequencing assay comprises a high-throughput sequencing assay. In some embodiments, the amplification reaction comprises introducing primers targeting a plurality of AMR genetic markers. In some embodiments, the amplification reaction produces amplicons associated with the AMR genetic marker. In some embodiments, the sample comprises plasma. In some embodiments, the first nucleic acids are DNA. In some embodiments, the second nucleic acids

are DNA. In some embodiments, the sample does not undergo cell lysis. In some embodiments, intact microbes are not actively lysed. In some embodiments, the amplicons associated with the AMR genetic marker undergoes direct library preparation. In some embodiments, the sample comprises cell-free nucleic acids. In some embodiments, the sample comprises cell-free DNA. In some embodiments, the first nucleic acids comprise cell-free nucleic acids. In some embodiments, the second nucleic acids comprise cell-free nucleic acids. In some embodiments, the methods disclosed herein comprise adding primers targeting a plurality of AMR genetic markers into the second aliquot of the sample. In some embodiments, the methods disclosed herein comprise adding at least 200 primers targeting a plurality of AMR genetic markers into the second aliquot of the sample. In some embodiments, the methods disclosed herein further comprise introducing primers targeting housekeeping genes into the second aliquot of the sample. In some embodiments, the second aliquot of the sample comprises at least 500 µl of plasma. In some embodiments, detecting the pathogen comprises detecting the abundance of mcfNA from the pathogen in the sample over a threshold. In some embodiments, the methods disclosed herein further comprise calculating a positive percent agreement (PPA), negative percent agreement (NPA), diagnostic yield (DY), or any combination thereof. In some embodiments, the methods disclosed herein further comprise detecting cfDNA from one or more housekeeping genes. In some embodiments, the methods disclosed herein further comprise spiking one or more control molecules into the sample at a known concentration. In some embodiments, the one or more control molecules are synthetic oligonucleotides. In some embodiments, the control molecules comprise whole assay internal control (WINC) molecules. In some embodiments, the methods disclosed herein further comprise spiking at least 25,000 unique WINC molecules at known concentrations. In some embodiments, the methods disclosed herein further comprise generating a report listing the carrier microbes or pathogen detected in the subject. In some embodiments, the report further comprises the abundance of microbial cell-free DNA (mcfDNA) from microbes detected in the sample or the antimicrobial resistance of the microbes infecting the subject. In some embodiments, the sample comprises a biological sample obtained from the subject. In some embodiments, the biological sample is a whole blood sample, a plasma sample, a serum sample, a cerebrospinal fluid sample, a synovial fluid sample, a urine sample, a stool sample, a bronchoalveolar lavage sample, or any combination thereof. In some embodiments, the sample is a plasma sample. In some embodiments, the AMR genetic marker comprises a gene, a genetic element, a genetic cassette, an allele, a mutation, or any combination thereof. In some embodiments, the AMR genetic marker is associated with one or more genes selected from the group consisting of: *SCCmec, mecA, mecC, vanA, vanB,* $bla_{CTX\text{-}M}$, bla_{KPC}, *OXA-48-like*, *OXA-23, NDM, VIM, IMP*, or *mcr-1*. In some embodiments, the AMR genetic marker provides resistance

to an anti-microbial agent selected from the group consisting of: methicillin, vancomycin, cephalosporin, carbapenem, and oxyimino-cephalosporin/aztreonam resistance. In some embodiments, the microbe infecting the subject or the carrier microbe comprises a virus, a bacterium, a protozoa, a fungus, an archaea, or an algae. In some embodiments, the microbe infecting the subject or the carrier microbe comprises a gram-positive bacterium or a fungus. In some embodiments, the microbe infecting the subject or the carrier microbe is a microbe listed in Table 1. In some embodiments, the microbe infecting the subject or the carrier microbe harbors at last two AMR genetic markers. In some embodiments, the antimicrobial resistance is phenotypic antimicrobial resistance. In some embodiments, the sequencing comprises next-generation sequencing or a sequencing method beyond next generation sequencing. In some embodiments, the sequencing comprises sequencing by synthesis. In some embodiments, the subject is an animal. In some embodiments, the subject is a human. In some embodiments, the subject has an infection by the microbe infecting the subject harboring the AMR genetic marker. In some embodiments, the methods disclosed herein further comprise administering an anti-infective agent to the subject. In some embodiments, the subject has been treated with an anti-infective agent for an infection. In some embodiments, the methods disclosed herein further comprise adjusting the anti-microbial agent received by the subject at least in part based on the antimicrobial resistance of the microbe infecting the subject. In some embodiments, the methods disclosed herein further comprise determining antimicrobial resistance of a microbe. In some embodiments, the microbe comprises a microbe infecting the subject. In some embodiments, the microbe comprises a carrier microbe. In some embodiments, the methods disclosed herein comprise determining whether an AMR gene is carried by a microbe infecting the subject or by a carrier microbe. In some embodiments, the determining comprises linking an AMR genetic marker to one of the carrier microbes or the microbe infecting the subject. In some embodiments, the determining comprises determining a copy number of the AMR genetic marker in the microbe. In some embodiments, the linking comprises performing a statistical analysis on the sequence reads. In some embodiments, the statistical analysis comprises a generalized linear model, a maximum-likelihood estimation, a probit model, a logistic regression, a linear probability, a linear regression, a complimentary log-log, a Poisson regression, a support vector machine, a decision tree, a random forest, a neural network, a gradient boosted model, a Bayesian model, a hidden Markov model, or any combination thereof. In some embodiments, the method further comprises comparing the sequence read data to the amplification data. In some embodiments, the method further comprises calculating estimated deduplicated templates (EDT) for an AMR gene, wherein calculating EDT comprises calculating a number of unique DNA sequencing reads from a microorganism present in a sequenced library. In some embodiments, the method further

comprises calculating an EDT for a housekeeping gene, wherein calculating EDT comprises calculating a number of unique DNA sequencing reads from a microorganism present in a sequenced library. In some embodiments, the methods disclosed herein further comprise determining a presence of an organism in a sample by calculating a ratio of an AMR estimated deduplicated templates (EDT) to a housekeeping gene EDT.

[0010] Disclosed herein are methods and compositions for performing a multi-step sample processing procedure for detecting nucleic acids in a sample from a subject: providing a first aliquot of the sample from the subject, wherein the sample comprises one or more first nucleic acids, wherein the one or more first nucleic acids comprise one or more microbial nucleic acids derived from one or more carrier microbes harboring an antimicrobial resistance (AMR) genetic marker; performing a sequencing assay on the one or more first nucleic acids to produce sequence reads comprising microbial nucleic acid sequence reads from the one or more carrier microbes; detecting the one or more carrier microbes by analyzing the microbial nucleic acid sequence reads; after at least one carrier microbe is detected, providing a second aliquot of the sample from the subject, wherein the second aliquot comprises second nucleic acids from at least one AMR genetic marker; annealing primers to the second nucleic acids that target at least one AMR genetic marker; and amplifying the AMR genetic marker and producing amplicons associated with the at least one AMR genetic marker. In some embodiments, the methods disclosed herein further comprise performing a high-throughput sequencing assay on the amplicons associated with the second target nucleic acid. In some embodiments, the second target nucleic acid is not detected by a sequencing assay. In some embodiments, the first nucleic acid is associated with a genome of an organism but not a phenotype of interest of the organism and the second nucleic acid is associated with the phenotype of interest of the organism. In some embodiments, the first target nucleic acid comprises at least 2, at least 3, at least 4, at least 5, at least 10, at least 15, at least 20, or at least 25 first target nucleic acids; or the second target nucleic acid comprises at least 2, at least 3, at least 4, at least 5, at least 10, at least 15, at least 20, or at least 25 second target nucleic acids. In some embodiments, the first target nucleic acid or the second target nucleic acid are not negative or positive controls for any step of the multi-step assay. In some embodiments, the primers comprise multiple primers targeting multiple target nucleic acids. In some embodiments, the first target nucleic acid comprises nucleic acids from an animal, a virus, a bacterium, a protozoa, a fungus, an archaea, or an algae. In some embodiments, the second target nucleic acid comprises nucleic acids from an animal, a virus, a bacterium, a protozoa, a fungus, an archaea, or an algae. In some embodiments, the second target nucleic acid comprises a cancer marker. In some embodiments, the first target nucleic acid is associated with a carrier microbe harboring a target genetic marker and the second target nucleic acid comprises a

sequence associated with the target genetic marker. In some embodiments, the carrier microbe comprises a virus, a bacterium, a protozoa, a fungus, an archaea, or an algae. In some embodiments, the target genetic marker comprises an antimicrobial resistance (AMR) genetic marker. In some embodiments, the methods disclosed herein further comprise comparing an abundance of the microbial sequence reads from the AMR genetic marker to a threshold value. In some embodiments, the methods disclosed herein further comprise comparing an abundance of the microbial sequence reads from the gene cassette to a threshold value. In some embodiments, the methods disclosed herein further comprise providing a second aliquot for the amplification reaction when the abundance of the microbial sequence reads from the AMR genetic marker is below the threshold. In some embodiments, the methods disclosed herein further comprise providing a second aliquot for the amplification reaction when the abundance of the microbial sequence reads from the gene cassette is below the threshold. In some embodiments, the methods disclosed herein further comprise quantifying mcfNA sequencing reads from the one or more microbes. In some embodiments, the methods disclosed herein further comprise comparing the abundance of the mcfNA sequencing reads from the one or more microbes to a threshold. In some embodiments, the methods disclosed herein further comprise providing a second aliquot for the amplification reaction when the abundance of the mcfNA sequencing reads from the one or more microbes is below the threshold. In some embodiments, the methods disclosed herein further comprise performing high-throughput sequencing on the amplicons associated with the AMR genetic marker. In some embodiments, the sequencing assay on the first nucleic acids comprises a high-throughput sequencing assay. In some embodiments, the methods disclosed herein further comprise conducting a polymerase chain reaction (PCR) to amplify the AMR genetic marker, thereby producing amplicons associated with the AMR genetic marker. In some embodiments, the PCR comprises multiplex PCR, random PCR (rPCR), non-biased PCR, Nested PCR, Hot Start PCR, or Assembly PCR. In some embodiments, the methods disclosed herein further comprise attaching an adapter sequence to the second nucleic acids. In some embodiments, the primers comprise an adapter sequence. In some embodiments, the methods disclosed herein further comprise physically manipulating the sample to produce a fraction of cfNA enriched for degraded cfNA, wherein the fraction of cfNA comprises the AMR genetic marker. In some embodiments, the degraded cfNA comprises ultra short cfNA, single stranded cfNA, or nicked double stranded cfNA. In some embodiments, the ultra short cfNA comprises cfNA fragments that are less than 100 nucleotides in length. In some embodiments, the ultra short cfNA comprises cfNA fragments that are from 30 nucleotides to 70 nucleotides in length. In some embodiments, the methods disclosed herein further comprise performing size selection of the nucleic acids in the sample. In some embodiments, physically manipulating the sample comprises performing

size selection of the nucleic acids in the sample. In some embodiments, performing size selection using a method selected from the group consisting of: chromatography, size-exclusion chromatography, electrophoresis, gel electrophoresis, automated electrophoresis, capillary electrophoresis, high-throughput size selection, centrifugation, density-gradient centrifugation, filtration, membrane ultrafiltration, magnetic beads, and affinity-based beads. In some embodiments, the second aliquot comprises at least 500 µl of plasma. In some embodiments, detecting the carrier microbe comprises determining an abundance of the carrier microbe. In some embodiments, the abundance is expressed as molecules of mcfNA per microliter of sample (MPM). In some embodiments, the methods disclosed herein further comprise calculating an AMR gene copy number from the amplification or sequencing of the AMR genetic marker. In some embodiments, the methods disclosed herein further comprise linking the AMR genetic marker to the carrier microbe using the abundance of the mcfNA from the carrier microbe and the AMR gene copy number. In some embodiments, the AMR gene copy number is an episomal gene copy number. In some embodiments, the methods disclosed herein comprise introducing at least 200 primers targeting the plurality of AMR genetic markers into the second aliquot of the sample. In some embodiments, the methods disclosed herein further comprise introducing primers targeting housekeeping genes into the second aliquot of the sample. In some embodiments, the methods disclosed herein further comprise determining antimicrobial resistance of the microbe infecting the subject. In some embodiments, the sample comprises plasma. In some embodiments, the first nucleic acids are DNA. In some embodiments, the second nucleic acids are DNA. In some embodiments, the sample is not subjected to a process that primarily causes cell lysis. In some embodiments, the methods disclosed herein further comprise preparing a library from the amplicons associated with the AMR genetic marker. In some embodiments, the primers are added directly to the second aliquot. In some embodiments, the sample comprises cell-free nucleic acids. In some embodiments, the sample comprises cell-free DNA. In some embodiments, the first nucleic acids comprise microbial cell-free nucleic acids. In some embodiments, the second nucleic acids comprise microbial cell-free nucleic acids. In some embodiments, intact microbes are not actively lysed prior to performing the sequencing assay. In some embodiments, the microbial nucleic acids comprise microbial cell-free nucleic acids (mcfNA) from the microbe. In some embodiments, the microbial nucleic acids comprise nucleic acids associated with a microbial cell. In some embodiments, the methods disclosed herein comprise enriching for at least at least 75%, at least 80%, at least 85%, or at least 90% of the mcfNA in the sample. In some embodiments, the degraded cfNA comprise ultra short cfNA, single stranded cfNA, nicked double stranded cfNA, or any combination thereof. In some embodiments, the degraded cfNA comprise cfNA fragments that are less than 100 nucleotides in

length. In some embodiments, the degraded cfNA comprise cfNA fragments that are from 30 nucleotides to 70 nucleotides in length. In some embodiments, physically manipulating the sample comprises performing size selection of the nucleic acids in the sample. In some embodiments, the methods disclosed herein further comprise generating sequence reads from the cfNA from the subject, wherein the sequence reads comprise microbial sequencing reads derived from the microbe infecting the subject. In some embodiments, the sequence reads further comprise microbial sequencing reads from one or more carrier microbe. In some embodiments, the methods disclosed herein further comprise calculating an abundance of the mcfNA from the microbe in the sample and an abundance of the mcfNA from the one or more carrier microbe. In some embodiments, the abundance is expressed as molecules of mcfNA per 100 nanoliters of sample. In some embodiments, the methods disclosed herein further comprise calculating an AMR gene copy number for each of the one or more carrier microbe. In some embodiments, the methods disclosed herein further comprise detecting one or more carrier microbe from the sequencing assay. In some embodiments, the methods disclosed herein further comprise identifying the one or more carrier microbe as the microbe infecting the subject. In some embodiments, the methods disclosed herein further comprise linking the AMR genetic marker to one of the carrier microbes or the microbe infecting the subject. In some embodiments, the methods disclosed herein further comprise obtaining an abundance of the one or more carrier microbes based on an abundance of carrier microbe sequences in the sample. In some embodiments, conducting a statistical analysis comprises using the abundance of the microbial nucleic acids of the one or more potential carrier microbes. In some embodiments, conducting a statistical analysis comprises using the AMR gene copy number. In some embodiments, the methods disclosed herein further comprise calculating a probability of each of the one or more potential carrier microbes being the microbe harboring the AMR genetic marker, thereby identifying the potential carrier microbe as the microbe infecting the subject. In some embodiments, the methods disclosed herein further comprise determining the antimicrobial resistance of the microbe infecting the subject based on the calculated probability. In some embodiments, the detecting comprises performing a sequencing assay. In some embodiments, the sequencing assay comprises a high-throughput sequencing assay. In some embodiments, the amplification reaction comprises introducing primers targeting a plurality of AMR genetic markers. In some embodiments, the amplification reaction produces amplicons associated with the AMR genetic marker. In some embodiments, the sample comprises plasma. In some embodiments, the first nucleic acids are DNA. In some embodiments, the second nucleic acids are DNA. In some embodiments, the sample does not undergo cell lysis. In some embodiments, intact microbes are not actively lysed. In some embodiments, the amplicons associated with the

AMR genetic marker undergoes direct library preparation. In some embodiments, the sample comprises cell-free nucleic acids. In some embodiments, the sample comprises cell-free DNA. In some embodiments, the first nucleic acids comprise cell-free nucleic acids. In some embodiments, the second nucleic acids comprise cell-free nucleic acids. In some embodiments, the methods disclosed herein comprise adding primers targeting a plurality of AMR genetic markers into the second aliquot of the sample. In some embodiments, the methods disclosed herein comprise adding at least 200 primers targeting a plurality of AMR genetic markers into the second aliquot of the sample. In some embodiments, the methods disclosed herein further comprise introducing primers targeting housekeeping genes into the second aliquot of the sample. In some embodiments, the second aliquot of the sample comprises at least 500 µl of plasma. In some embodiments, detecting the pathogen comprises detecting the abundance of mcfNA from the pathogen in the sample over a threshold. In some embodiments, the methods disclosed herein further comprise calculating a positive percent agreement (PPA), negative percent agreement (NPA), diagnostic yield (DY), or any combination thereof. In some embodiments, the methods disclosed herein further comprise detecting cfDNA from one or more housekeeping genes. In some embodiments, the methods disclosed herein further comprise spiking one or more control molecules into the sample at a known concentration. In some embodiments, the one or more control molecules are synthetic oligonucleotides. In some embodiments, the control molecules comprise whole assay internal control (WINC) molecules. In some embodiments, the methods disclosed herein further comprise spiking at least 25,000 unique WINC molecules at known concentrations. In some embodiments, the methods disclosed herein further comprise generating a report listing the carrier microbes or pathogen detected in the subject. In some embodiments, the report further comprises the abundance of microbial cell-free DNA (mcfDNA) from microbes detected in the sample or the antimicrobial resistance of the microbes infecting the subject. In some embodiments, the sample comprises a biological sample obtained from the subject. In some embodiments, the biological sample is a whole blood sample, a plasma sample, a serum sample, a cerebrospinal fluid sample, a synovial fluid sample, a urine sample, a stool sample, a bronchoalveolar lavage sample, or any combination thereof. In some embodiments, the sample is a plasma sample. In some embodiments, the AMR genetic marker comprises a gene, a genetic element, a genetic cassette, an allele, a mutation, or any combination thereof. In some embodiments, the AMR genetic marker is associated with one or more genes selected from the group consisting of: *SCCmec, mecA, mecC, vanA, vanB,* $bla_{CTX\text{-}M}$, bla_{KPC}, *OXA-48-like*, *OXA-23, NDM, VIM, IMP*, or *mcr-1*. In some embodiments, the AMR genetic marker provides resistance to an anti-microbial agent selected from the group consisting of: methicillin, vancomycin, cephalosporin, carbapenem, and oxyimino-cephalosporin/aztreonam resistance. In some

embodiments, the microbe infecting the subject or the carrier microbe comprises a virus, a bacterium, a protozoa, a fungus, an archaea, or an algae. In some embodiments, the microbe infecting the subject or the carrier microbe comprises a gram-positive bacterium or a fungus. In some embodiments, the microbe infecting the subject or the carrier microbe is a microbe listed in Table 1. In some embodiments, the microbe infecting the subject or the carrier microbe harbors at last two AMR genetic markers. In some embodiments, the antimicrobial resistance is phenotypic antimicrobial resistance. In some embodiments, the sequencing comprises next-generation sequencing or a sequencing method beyond next generation sequencing. In some embodiments, the sequencing comprises sequencing by synthesis. In some embodiments, the subject is an animal. In some embodiments, the subject is a human. In some embodiments, the subject has an infection by the microbe infecting the subject harboring the AMR genetic marker. In some embodiments, the methods disclosed herein further comprise administering an anti-infective agent to the subject. In some embodiments, the subject has been treated with an anti-infective agent for an infection. In some embodiments, the methods disclosed herein further comprise adjusting the anti-microbial agent received by the subject at least in part based on the antimicrobial resistance of the microbe infecting the subject. In some embodiments, the methods disclosed herein further comprise determining antimicrobial resistance of a microbe. In some embodiments, the microbe comprises a microbe infecting the subject. In some embodiments, the microbe comprises a carrier microbe. In some embodiments, the methods disclosed herein comprise determining whether an AMR gene is carried by a microbe infecting the subject or by a carrier microbe. In some embodiments, the determining comprises linking an AMR genetic marker to one of the carrier microbes or the microbe infecting the subject. In some embodiments, the determining comprises determining a copy number of the AMR genetic marker in the microbe. In some embodiments, the linking comprises performing a statistical analysis on the sequence reads. In some embodiments, the statistical analysis comprises a generalized linear model, a maximum-likelihood estimation, a probit model, a logistic regression, a linear probability, a linear regression, a complimentary log-log, a Poisson regression, a support vector machine, a decision tree, a random forest, a neural network, a gradient boosted model, a Bayesian model, a hidden Markov model, or any combination thereof. In some embodiments, the method further comprises comparing the sequence read data to the amplification data. In some embodiments, the method further comprises calculating estimated deduplicated templates (EDT) for an AMR gene, wherein calculating EDT comprises calculating a number of unique DNA sequencing reads from a microorganism present in a sequenced library. In some embodiments, the method further comprises calculating an EDT for a housekeeping gene, wherein calculating EDT comprises calculating a number of unique DNA sequencing reads from a microorganism present in a

sequenced library. In some embodiments, the methods disclosed herein further comprise determining a presence of an organism in a sample by calculating a ratio of an AMR estimated deduplicated templates (EDT) to a housekeeping gene EDT.

[0011] Disclosed herein are methods and compositions for assaying nucleic acids in a sample from a subject comprising: providing a first aliquot of the sample from the subject, wherein the first aliquot comprises first nucleic acids, and the first nucleic acids comprise an AMR genetic marker and a gene cassette comprising the AMR genetic marker derived from one or more carrier microbes harboring the AMR genetic marker; performing a sequencing assay on the first nucleic acids to produce sequence reads comprising microbial nucleic acid sequence reads from the one or more carrier microbes; detecting the AMR genetic marker or the gene cassette comprising the AMR genetic marker in the sample and quantifying microbial sequence reads from the AMR genetic marker and/or microbial sequence reads from the gene cassette; after and based on the quantification, providing a second aliquot of the sample from the subject, wherein the second aliquot comprises second nucleic acids, wherein the second nucleic acids comprise the AMR genetic marker; introducing primers targeting the AMR genetic marker to the second nucleic acids, wherein the primers targeting the AMR genetic marker are introduced prior to amplification of nucleic acids in the second aliquot; and conducting an amplification reaction on the second nucleic acids using the primers targeting the AMR genetic marker, thereby amplifying the AMR genetic marker and producing amplicons associated with the AMR genetic marker. In some embodiments, the methods disclosed herein further comprise performing a high-throughput sequencing assay on the amplicons associated with the second target nucleic acid. In some embodiments, the second target nucleic acid is not detected by a sequencing assay. In some embodiments, the first nucleic acid is associated with a genome of an organism but not a phenotype of interest of the organism and the second nucleic acid is associated with the phenotype of interest of the organism. In some embodiments, the first target nucleic acid comprises at least 2, at least 3, at least 4, at least 5, at least 10, at least 15, at least 20, or at least 25 first target nucleic acids; or the second target nucleic acid comprises at least 2, at least 3, at least 4, at least 5, at least 10, at least 15, at least 20, or at least 25 second target nucleic acids. In some embodiments, the first target nucleic acid or the second target nucleic acid are not negative or positive controls for any step of the multi-step assay. In some embodiments, the primers comprise multiple primers targeting multiple target nucleic acids. In some embodiments, the first target nucleic acid comprises nucleic acids from an animal, a virus, a bacterium, a protozoa, a fungus, an archaea, or an algae. In some embodiments, the second target nucleic acid comprises nucleic acids from an animal, a virus, a bacterium, a protozoa, a fungus, an archaea, or an algae. In some embodiments, the second target nucleic acid comprises a cancer marker. In some

embodiments, the first target nucleic acid is associated with a carrier microbe harboring a target genetic marker and the second target nucleic acid comprises a sequence associated with the target genetic marker. In some embodiments, the carrier microbe comprises a virus, a bacterium, a protozoa, a fungus, an archaea, or an algae. In some embodiments, the target genetic marker comprises an antimicrobial resistance (AMR) genetic marker. In some embodiments, the methods disclosed herein further comprise comparing an abundance of the microbial sequence reads from the AMR genetic marker to a threshold value. In some embodiments, the methods disclosed herein further comprise comparing an abundance of the microbial sequence reads from the gene cassette to a threshold value. In some embodiments, the methods disclosed herein further comprise providing a second aliquot for the amplification reaction when the abundance of the microbial sequence reads from the AMR genetic marker is below the threshold. In some embodiments, the methods disclosed herein further comprise providing a second aliquot for the amplification reaction when the abundance of the microbial sequence reads from the gene cassette is below the threshold. In some embodiments, the methods disclosed herein further comprise quantifying mcfNA sequencing reads from the one or more microbes. In some embodiments, the methods disclosed herein further comprise comparing the abundance of the mcfNA sequencing reads from the one or more microbes to a threshold. In some embodiments, the methods disclosed herein further comprise providing a second aliquot for the amplification reaction when the abundance of the mcfNA sequencing reads from the one or more microbes is below the threshold. In some embodiments, the methods disclosed herein further comprise performing high-throughput sequencing on the amplicons associated with the AMR genetic marker. In some embodiments, the sequencing assay on the first nucleic acids comprises a high-throughput sequencing assay. In some embodiments, the methods disclosed herein further comprise conducting a polymerase chain reaction (PCR) to amplify the AMR genetic marker, thereby producing amplicons associated with the AMR genetic marker. In some embodiments, the PCR comprises multiplex PCR, random PCR (rPCR), non-biased PCR, Nested PCR, Hot Start PCR, or Assembly PCR. In some embodiments, the methods disclosed herein further comprise attaching an adapter sequence to the second nucleic acids. In some embodiments, the primers comprise an adapter sequence. In some embodiments, the methods disclosed herein further comprise physically manipulating the sample to produce a fraction of cfNA enriched for degraded cfNA, wherein the fraction of cfNA comprises the AMR genetic marker. In some embodiments, the degraded cfNA comprises ultra short cfNA, single stranded cfNA, or nicked double stranded cfNA. In some embodiments, the ultra short cfNA comprises cfNA fragments that are less than 100 nucleotides in length. In some embodiments, the ultra short cfNA comprises cfNA fragments that are from 30 nucleotides to 70 nucleotides in length. In some embodiments,

the methods disclosed herein further comprise performing size selection of the nucleic acids in the sample. In some embodiments, physically manipulating the sample comprises performing size selection of the nucleic acids in the sample. In some embodiments, performing size selection using a method selected from the group consisting of: chromatography, size-exclusion chromatography, electrophoresis, gel electrophoresis, automated electrophoresis, capillary electrophoresis, high-throughput size selection, centrifugation, density-gradient centrifugation, filtration, membrane ultrafiltration, magnetic beads, and affinity-based beads. In some embodiments, the second aliquot comprises at least 500 µl of plasma. In some embodiments, detecting the carrier microbe comprises determining an abundance of the carrier microbe. In some embodiments, the abundance is expressed as molecules of mcfNA per microliter of sample (MPM). In some embodiments, the methods disclosed herein further comprise calculating an AMR gene copy number from the amplification or sequencing of the AMR genetic marker. In some embodiments, the methods disclosed herein further comprise linking the AMR genetic marker to the carrier microbe using the abundance of the mcfNA from the carrier microbe and the AMR gene copy number. In some embodiments, the AMR gene copy number is an episomal gene copy number. In some embodiments, the methods disclosed herein comprise introducing at least 200 primers targeting the plurality of AMR genetic markers into the second aliquot of the sample. In some embodiments, the methods disclosed herein further comprise introducing primers targeting housekeeping genes into the second aliquot of the sample. In some embodiments, the methods disclosed herein further comprise determining antimicrobial resistance of the microbe infecting the subject. In some embodiments, the sample comprises plasma. In some embodiments, the first nucleic acids are DNA. In some embodiments, the second nucleic acids are DNA. In some embodiments, the sample is not subjected to a process that primarily causes cell lysis. In some embodiments, the methods disclosed herein further comprise preparing a library from the amplicons associated with the AMR genetic marker. In some embodiments, the primers are added directly to the second aliquot. In some embodiments, the sample comprises cell-free nucleic acids. In some embodiments, the sample comprises cell-free DNA. In some embodiments, the first nucleic acids comprise microbial cell-free nucleic acids. In some embodiments, the second nucleic acids comprise microbial cell-free nucleic acids. In some embodiments, intact microbes are not actively lysed prior to performing the sequencing assay. In some embodiments, the microbial nucleic acids comprise microbial cell-free nucleic acids (mcfNA) from the microbe. In some embodiments, the microbial nucleic acids comprise nucleic acids associated with a microbial cell. In some embodiments, the methods disclosed herein comprise enriching for at least at least 75%, at least 80%, at least 85%, or at least 90% of the mcfNA in the sample. In some embodiments, the degraded cfNA comprise ultra short cfNA,

single stranded cfNA, nicked double stranded cfNA, or any combination thereof. In some embodiments, the degraded cfNA comprise cfNA fragments that are less than 100 nucleotides in length. In some embodiments, the degraded cfNA comprise cfNA fragments that are from 30 nucleotides to 70 nucleotides in length. In some embodiments, physically manipulating the sample comprises performing size selection of the nucleic acids in the sample. In some embodiments, the methods disclosed herein further comprise generating sequence reads from the cfNA from the subject, wherein the sequence reads comprise microbial sequencing reads derived from the microbe infecting the subject. In some embodiments, the sequence reads further comprise microbial sequencing reads from one or more carrier microbe. In some embodiments, the methods disclosed herein further comprise calculating an abundance of the mcfNA from the microbe in the sample and an abundance of the mcfNA from the one or more carrier microbe. In some embodiments, the abundance is expressed as molecules of mcfNA per 100 nanoliters of sample. In some embodiments, the methods disclosed herein further comprise calculating an AMR gene copy number for each of the one or more carrier microbe. In some embodiments, the methods disclosed herein further comprise detecting one or more carrier microbe from the sequencing assay. In some embodiments, the methods disclosed herein further comprise identifying the one or more carrier microbe as the microbe infecting the subject. In some embodiments, the methods disclosed herein further comprise linking the AMR genetic marker to one of the carrier microbes or the microbe infecting the subject. In some embodiments, the methods disclosed herein further comprise obtaining an abundance of the one or more carrier microbes based on an abundance of carrier microbe sequences in the sample. In some embodiments, conducting a statistical analysis comprises using the abundance of the microbial nucleic acids of the one or more potential carrier microbes. In some embodiments, conducting a statistical analysis comprises using the AMR gene copy number. In some embodiments, the methods disclosed herein further comprise calculating a probability of each of the one or more potential carrier microbes being the microbe harboring the AMR genetic marker, thereby identifying the potential carrier microbe as the microbe infecting the subject. In some embodiments, the methods disclosed herein further comprise determining the antimicrobial resistance of the microbe infecting the subject based on the calculated probability. In some embodiments, the detecting comprises performing a sequencing assay. In some embodiments, the sequencing assay comprises a high-throughput sequencing assay. In some embodiments, the amplification reaction comprises introducing primers targeting a plurality of AMR genetic markers. In some embodiments, the amplification reaction produces amplicons associated with the AMR genetic marker. In some embodiments, the sample comprises plasma. In some embodiments, the first nucleic acids are DNA. In some embodiments, the second nucleic acids

are DNA. In some embodiments, the sample does not undergo cell lysis. In some embodiments, intact microbes are not actively lysed. In some embodiments, the amplicons associated with the AMR genetic marker undergoes direct library preparation. In some embodiments, the sample comprises cell-free nucleic acids. In some embodiments, the sample comprises cell-free DNA. In some embodiments, the first nucleic acids comprise cell-free nucleic acids. In some embodiments, the second nucleic acids comprise cell-free nucleic acids. In some embodiments, the methods disclosed herein comprise adding primers targeting a plurality of AMR genetic markers into the second aliquot of the sample. In some embodiments, the methods disclosed herein comprise adding at least 200 primers targeting a plurality of AMR genetic markers into the second aliquot of the sample. In some embodiments, the methods disclosed herein further comprise introducing primers targeting housekeeping genes into the second aliquot of the sample. In some embodiments, the second aliquot of the sample comprises at least 500 µl of plasma. In some embodiments, detecting the pathogen comprises detecting the abundance of mcfNA from the pathogen in the sample over a threshold. In some embodiments, the methods disclosed herein further comprise calculating a positive percent agreement (PPA), negative percent agreement (NPA), diagnostic yield (DY), or any combination thereof. In some embodiments, the methods disclosed herein further comprise detecting cfDNA from one or more housekeeping genes. In some embodiments, the methods disclosed herein further comprise spiking one or more control molecules into the sample at a known concentration. In some embodiments, the one or more control molecules are synthetic oligonucleotides. In some embodiments, the control molecules comprise whole assay internal control (WINC) molecules. In some embodiments, the methods disclosed herein further comprise spiking at least 25,000 unique WINC molecules at known concentrations. In some embodiments, the methods disclosed herein further comprise generating a report listing the carrier microbes or pathogen detected in the subject. In some embodiments, the report further comprises the abundance of microbial cell-free DNA (mcfDNA) from microbes detected in the sample or the antimicrobial resistance of the microbes infecting the subject. In some embodiments, the sample comprises a biological sample obtained from the subject. In some embodiments, the biological sample is a whole blood sample, a plasma sample, a serum sample, a cerebrospinal fluid sample, a synovial fluid sample, a urine sample, a stool sample, a bronchoalveolar lavage sample, or any combination thereof. In some embodiments, the sample is a plasma sample. In some embodiments, the AMR genetic marker comprises a gene, a genetic element, a genetic cassette, an allele, a mutation, or any combination thereof. In some embodiments, the AMR genetic marker is associated with one or more genes selected from the group consisting of: *SCCmec, mecA, mecC, vanA, vanB,* $bla_{CTX\text{-}M}$, bla_{KPC}, *OXA-48-like*, *OXA-23, NDM, VIM, IMP*, or *mcr-1*. In some embodiments, the AMR genetic marker provides resistance

to an anti-microbial agent selected from the group consisting of: methicillin, vancomycin, cephalosporin, carbapenem, and oxyimino-cephalosporin/aztreonam resistance. In some embodiments, the microbe infecting the subject or the carrier microbe comprises a virus, a bacterium, a protozoa, a fungus, an archaea, or an algae. In some embodiments, the microbe infecting the subject or the carrier microbe comprises a gram-positive bacterium or a fungus. In some embodiments, the microbe infecting the subject or the carrier microbe is a microbe listed in Table 1. In some embodiments, the microbe infecting the subject or the carrier microbe harbors at last two AMR genetic markers. In some embodiments, the antimicrobial resistance is phenotypic antimicrobial resistance. In some embodiments, the sequencing comprises next-generation sequencing or a sequencing method beyond next generation sequencing. In some embodiments, the sequencing comprises sequencing by synthesis. In some embodiments, the subject is an animal. In some embodiments, the subject is a human. In some embodiments, the subject has an infection by the microbe infecting the subject harboring the AMR genetic marker. In some embodiments, the methods disclosed herein further comprise administering an anti-infective agent to the subject. In some embodiments, the subject has been treated with an anti-infective agent for an infection. In some embodiments, the methods disclosed herein further comprise adjusting the anti-microbial agent received by the subject at least in part based on the antimicrobial resistance of the microbe infecting the subject. In some embodiments, the methods disclosed herein further comprise determining antimicrobial resistance of a microbe. In some embodiments, the microbe comprises a microbe infecting the subject. In some embodiments, the microbe comprises a carrier microbe. In some embodiments, the methods disclosed herein comprise determining whether an AMR gene is carried by a microbe infecting the subject or by a carrier microbe. In some embodiments, the determining comprises linking an AMR genetic marker to one of the carrier microbes or the microbe infecting the subject. In some embodiments, the determining comprises determining a copy number of the AMR genetic marker in the microbe. In some embodiments, the linking comprises performing a statistical analysis on the sequence reads. In some embodiments, the statistical analysis comprises a generalized linear model, a maximum-likelihood estimation, a probit model, a logistic regression, a linear probability, a linear regression, a complimentary log-log, a Poisson regression, a support vector machine, a decision tree, a random forest, a neural network, a gradient boosted model, a Bayesian model, a hidden Markov model, or any combination thereof. In some embodiments, the method further comprises comparing the sequence read data to the amplification data. In some embodiments, the method further comprises calculating estimated deduplicated templates (EDT) for an AMR gene, wherein calculating EDT comprises calculating a number of unique DNA sequencing reads from a microorganism present in a sequenced library. In some embodiments, the method further

comprises calculating an EDT for a housekeeping gene, wherein calculating EDT comprises calculating a number of unique DNA sequencing reads from a microorganism present in a sequenced library. In some embodiments, the methods disclosed herein further comprise determining a presence of an organism in a sample by calculating a ratio of an AMR estimated deduplicated templates (EDT) to a housekeeping gene EDT.

[0012] Disclosed herein are methods and compositions for preparing a cell-free DNA fraction from a subject infected with a microbe useful for analyzing an antimicrobial resistance (AMR) genetic marker, comprising: providing a sample from the subject, wherein: the sample comprises cell-free nucleic acids (cfNA) from the subject and microbial nucleic acids from the microbe infecting the subject; the cfNA comprises degraded cfNA, single-stranded cfNA and intact cfNA; and the microbe harbors an AMR genetic marker; physically manipulating the sample in order to produce a fraction of cfNA enriched for degraded cfNA and single-stranded cfNA, wherein the fraction of cfNA comprises microbial nucleic acids comprising a genetic locus associated with the AMR genetic marker; performing a sequencing assay on the fraction of cfNA enriched for the degraded cfNA and single-stranded cfNA; and analyzing the genetic locus associated with the AMR genetic marker. In some embodiments, the methods disclosed herein further comprise performing a high-throughput sequencing assay on the amplicons associated with the second target nucleic acid. In some embodiments, the second target nucleic acid is not detected by a sequencing assay. In some embodiments, the first nucleic acid is associated with a genome of an organism but not a phenotype of interest of the organism and the second nucleic acid is associated with the phenotype of interest of the organism. In some embodiments, the first target nucleic acid comprises at least 2, at least 3, at least 4, at least 5, at least 10, at least 15, at least 20, or at least 25 first target nucleic acids; or the second target nucleic acid comprises at least 2, at least 3, at least 4, at least 5, at least 10, at least 15, at least 20, or at least 25 second target nucleic acids. In some embodiments, the first target nucleic acid or the second target nucleic acid are not negative or positive controls for any step of the multi-step assay. In some embodiments, the primers comprise multiple primers targeting multiple target nucleic acids. In some embodiments, the first target nucleic acid comprises nucleic acids from an animal, a virus, a bacterium, a protozoa, a fungus, an archaea, or an algae. In some embodiments, the second target nucleic acid comprises nucleic acids from an animal, a virus, a bacterium, a protozoa, a fungus, an archaea, or an algae. In some embodiments, the second target nucleic acid comprises a cancer marker. In some embodiments, the first target nucleic acid is associated with a carrier microbe harboring a target genetic marker and the second target nucleic acid comprises a sequence associated with the target genetic marker. In some embodiments, the carrier microbe comprises a virus, a bacterium, a protozoa, a fungus, an archaea, or an algae. In some embodiments, the target genetic marker

comprises an antimicrobial resistance (AMR) genetic marker. In some embodiments, the methods disclosed herein further comprise comparing an abundance of the microbial sequence reads from the AMR genetic marker to a threshold value. In some embodiments, the methods disclosed herein further comprise comparing an abundance of the microbial sequence reads from the gene cassette to a threshold value. In some embodiments, the methods disclosed herein further comprise providing a second aliquot for the amplification reaction when the abundance of the microbial sequence reads from the AMR genetic marker is below the threshold. In some embodiments, the methods disclosed herein further comprise providing a second aliquot for the amplification reaction when the abundance of the microbial sequence reads from the gene cassette is below the threshold. In some embodiments, the methods disclosed herein further comprise quantifying mcfNA sequencing reads from the one or more microbes. In some embodiments, the methods disclosed herein further comprise comparing the abundance of the mcfNA sequencing reads from the one or more microbes to a threshold. In some embodiments, the methods disclosed herein further comprise providing a second aliquot for the amplification reaction when the abundance of the mcfNA sequencing reads from the one or more microbes is below the threshold. In some embodiments, the methods disclosed herein further comprise performing high-throughput sequencing on the amplicons associated with the AMR genetic marker. In some embodiments, the sequencing assay on the first nucleic acids comprises a high-throughput sequencing assay. In some embodiments, the methods disclosed herein further comprise conducting a polymerase chain reaction (PCR) to amplify the AMR genetic marker, thereby producing amplicons associated with the AMR genetic marker. In some embodiments, the PCR comprises multiplex PCR, random PCR (rPCR), non-biased PCR, Nested PCR, Hot Start PCR, or Assembly PCR. In some embodiments, the methods disclosed herein further comprise attaching an adapter sequence to the second nucleic acids. In some embodiments, the primers comprise an adapter sequence. In some embodiments, the methods disclosed herein further comprise physically manipulating the sample to produce a fraction of cfNA enriched for degraded cfNA, wherein the fraction of cfNA comprises the AMR genetic marker. In some embodiments, the degraded cfNA comprises ultra short cfNA, single stranded cfNA, or nicked double stranded cfNA. In some embodiments, the ultra short cfNA comprises cfNA fragments that are less than 100 nucleotides in length. In some embodiments, the ultra short cfNA comprises cfNA fragments that are from 30 nucleotides to 70 nucleotides in length. In some embodiments, the methods disclosed herein further comprise performing size selection of the nucleic acids in the sample. In some embodiments, physically manipulating the sample comprises performing size selection of the nucleic acids in the sample. In some embodiments, performing size selection using a method selected from the group consisting of: chromatography, size-exclusion

chromatography, electrophoresis, gel electrophoresis, automated electrophoresis, capillary electrophoresis, high-throughput size selection, centrifugation, density-gradient centrifugation, filtration, membrane ultrafiltration, magnetic beads, and affinity-based beads. In some embodiments, the second aliquot comprises at least 500 μl of plasma. In some embodiments, detecting the carrier microbe comprises determining an abundance of the carrier microbe. In some embodiments, the abundance is expressed as molecules of mcfNA per microliter of sample (MPM). In some embodiments, the methods disclosed herein further comprise calculating an AMR gene copy number from the amplification or sequencing of the AMR genetic marker. In some embodiments, the methods disclosed herein further comprise linking the AMR genetic marker to the carrier microbe using the abundance of the mcfNA from the carrier microbe and the AMR gene copy number. In some embodiments, the AMR gene copy number is an episomal gene copy number. In some embodiments, the methods disclosed herein comprise introducing at least 200 primers targeting the plurality of AMR genetic markers into the second aliquot of the sample. In some embodiments, the methods disclosed herein further comprise introducing primers targeting housekeeping genes into the second aliquot of the sample. In some embodiments, the methods disclosed herein further comprise determining antimicrobial resistance of the microbe infecting the subject. In some embodiments, the sample comprises plasma. In some embodiments, the first nucleic acids are DNA. In some embodiments, the second nucleic acids are DNA. In some embodiments, the sample is not subjected to a process that primarily causes cell lysis. In some embodiments, the methods disclosed herein further comprise preparing a library from the amplicons associated with the AMR genetic marker. In some embodiments, the primers are added directly to the second aliquot. In some embodiments, the sample comprises cell-free nucleic acids. In some embodiments, the sample comprises cell-free DNA. In some embodiments, the first nucleic acids comprise microbial cell-free nucleic acids. In some embodiments, the second nucleic acids comprise microbial cell-free nucleic acids. In some embodiments, intact microbes are not actively lysed prior to performing the sequencing assay. In some embodiments, the microbial nucleic acids comprise microbial cell-free nucleic acids (mcfNA) from the microbe. In some embodiments, the microbial nucleic acids comprise nucleic acids associated with a microbial cell. In some embodiments, the methods disclosed herein comprise enriching for at least at least 75%, at least 80%, at least 85%, or at least 90% of the mcfNA in the sample. In some embodiments, the degraded cfNA comprise ultra short cfNA, single stranded cfNA, nicked double stranded cfNA, or any combination thereof. In some embodiments, the degraded cfNA comprise cfNA fragments that are less than 100 nucleotides in length. In some embodiments, the degraded cfNA comprise cfNA fragments that are from 30 nucleotides to 70 nucleotides in length. In some embodiments, physically manipulating the

sample comprises performing size selection of the nucleic acids in the sample. In some embodiments, the methods disclosed herein further comprise generating sequence reads from the cfNA from the subject, wherein the sequence reads comprise microbial sequencing reads derived from the microbe infecting the subject. In some embodiments, the sequence reads further comprise microbial sequencing reads from one or more carrier microbe. In some embodiments, the methods disclosed herein further comprise calculating an abundance of the mcfNA from the microbe in the sample and an abundance of the mcfNA from the one or more carrier microbe. In some embodiments, the abundance is expressed as molecules of mcfNA per 100 nanoliters of sample. In some embodiments, the methods disclosed herein further comprise calculating an AMR gene copy number for each of the one or more carrier microbe. In some embodiments, the methods disclosed herein further comprise detecting one or more carrier microbe from the sequencing assay. In some embodiments, the methods disclosed herein further comprise identifying the one or more carrier microbe as the microbe infecting the subject. In some embodiments, the methods disclosed herein further comprise linking the AMR genetic marker to one of the carrier microbes or the microbe infecting the subject. In some embodiments, the methods disclosed herein further comprise obtaining an abundance of the one or more carrier microbes based on an abundance of carrier microbe sequences in the sample. In some embodiments, conducting a statistical analysis comprises using the abundance of the microbial nucleic acids of the one or more potential carrier microbes. In some embodiments, conducting a statistical analysis comprises using the AMR gene copy number. In some embodiments, the methods disclosed herein further comprise calculating a probability of each of the one or more potential carrier microbes being the microbe harboring the AMR genetic marker, thereby identifying the potential carrier microbe as the microbe infecting the subject. In some embodiments, the methods disclosed herein further comprise determining the antimicrobial resistance of the microbe infecting the subject based on the calculated probability. In some embodiments, the detecting comprises performing a sequencing assay. In some embodiments, the sequencing assay comprises a high-throughput sequencing assay. In some embodiments, the amplification reaction comprises introducing primers targeting a plurality of AMR genetic markers. In some embodiments, the amplification reaction produces amplicons associated with the AMR genetic marker. In some embodiments, the sample comprises plasma. In some embodiments, the first nucleic acids are DNA. In some embodiments, the second nucleic acids are DNA. In some embodiments, the sample does not undergo cell lysis. In some embodiments, intact microbes are not actively lysed. In some embodiments, the amplicons associated with the AMR genetic marker undergoes direct library preparation. In some embodiments, the sample comprises cell-free nucleic acids. In some embodiments, the sample comprises cell-free DNA.

In some embodiments, the first nucleic acids comprise cell-free nucleic acids. In some embodiments, the second nucleic acids comprise cell-free nucleic acids. In some embodiments, the methods disclosed herein comprise adding primers targeting a plurality of AMR genetic markers into the second aliquot of the sample. In some embodiments, the methods disclosed herein comprise adding at least 200 primers targeting a plurality of AMR genetic markers into the second aliquot of the sample. In some embodiments, the methods disclosed herein further comprise introducing primers targeting housekeeping genes into the second aliquot of the sample. In some embodiments, the second aliquot of the sample comprises at least 500 µl of plasma. In some embodiments, detecting the pathogen comprises detecting the abundance of mcfNA from the pathogen in the sample over a threshold. In some embodiments, the methods disclosed herein further comprise calculating a positive percent agreement (PPA), negative percent agreement (NPA), diagnostic yield (DY), or any combination thereof. In some embodiments, the methods disclosed herein further comprise detecting cfDNA from one or more housekeeping genes. In some embodiments, the methods disclosed herein further comprise spiking one or more control molecules into the sample at a known concentration. In some embodiments, the one or more control molecules are synthetic oligonucleotides. In some embodiments, the control molecules comprise whole assay internal control (WINC) molecules. In some embodiments, the methods disclosed herein further comprise spiking at least 25,000 unique WINC molecules at known concentrations. In some embodiments, the methods disclosed herein further comprise generating a report listing the carrier microbes or pathogen detected in the subject. In some embodiments, the report further comprises the abundance of microbial cell-free DNA (mcfDNA) from microbes detected in the sample or the antimicrobial resistance of the microbes infecting the subject. In some embodiments, the sample comprises a biological sample obtained from the subject. In some embodiments, the biological sample is a whole blood sample, a plasma sample, a serum sample, a cerebrospinal fluid sample, a synovial fluid sample, a urine sample, a stool sample, a bronchoalveolar lavage sample, or any combination thereof. In some embodiments, the sample is a plasma sample. In some embodiments, the AMR genetic marker comprises a gene, a genetic element, a genetic cassette, an allele, a mutation, or any combination thereof. In some embodiments, the AMR genetic marker is associated with one or more genes selected from the group consisting of: *SCCmec, mecA, mecC, vanA, vanB,* $bla_{CTX\text{-}M}$, bla_{KPC}, *OXA-48-like*, *OXA-23, NDM, VIM, IMP*, or *mcr-1*. In some embodiments, the AMR genetic marker provides resistance to an anti-microbial agent selected from the group consisting of: methicillin, vancomycin, cephalosporin, carbapenem, and oxyimino-cephalosporin/aztreonam resistance. In some embodiments, the microbe infecting the subject or the carrier microbe comprises a virus, a bacterium, a protozoa, a fungus, an archaea, or an algae. In some embodiments, the microbe

infecting the subject or the carrier microbe comprises a gram-positive bacterium or a fungus. In some embodiments, the microbe infecting the subject or the carrier microbe is a microbe listed in Table 1. In some embodiments, the microbe infecting the subject or the carrier microbe harbors at last two AMR genetic markers. In some embodiments, the antimicrobial resistance is phenotypic antimicrobial resistance. In some embodiments, the sequencing comprises next-generation sequencing or a sequencing method beyond next generation sequencing. In some embodiments, the sequencing comprises sequencing by synthesis. In some embodiments, the subject is an animal. In some embodiments, the subject is a human. In some embodiments, the subject has an infection by the microbe infecting the subject harboring the AMR genetic marker. In some embodiments, the methods disclosed herein further comprise administering an anti-infective agent to the subject. In some embodiments, the subject has been treated with an anti-infective agent for an infection. In some embodiments, the methods disclosed herein further comprise adjusting the anti-microbial agent received by the subject at least in part based on the antimicrobial resistance of the microbe infecting the subject. In some embodiments, the methods disclosed herein further comprise determining antimicrobial resistance of a microbe. In some embodiments, the microbe comprises a microbe infecting the subject. In some embodiments, the microbe comprises a carrier microbe. In some embodiments, the methods disclosed herein comprise determining whether an AMR gene is carried by a microbe infecting the subject or by a carrier microbe. In some embodiments, the determining comprises linking an AMR genetic marker to one of the carrier microbes or the microbe infecting the subject. In some embodiments, the determining comprises determining a copy number of the AMR genetic marker in the microbe. In some embodiments, the linking comprises performing a statistical analysis on the sequence reads. In some embodiments, the statistical analysis comprises a generalized linear model, a maximum-likelihood estimation, a probit model, a logistic regression, a linear probability, a linear regression, a complimentary log-log, a Poisson regression, a support vector machine, a decision tree, a random forest, a neural network, a gradient boosted model, a Bayesian model, a hidden Markov model, or any combination thereof. In some embodiments, the method further comprises comparing the sequence read data to the amplification data. In some embodiments, the method further comprises calculating estimated deduplicated templates (EDT) for an AMR gene, wherein calculating EDT comprises calculating a number of unique DNA sequencing reads from a microorganism present in a sequenced library. In some embodiments, the method further comprises calculating an EDT for a housekeeping gene, wherein calculating EDT comprises calculating a number of unique DNA sequencing reads from a microorganism present in a sequenced library. In some embodiments, the methods disclosed herein further comprise

determining a presence of an organism in a sample by calculating a ratio of an AMR estimated deduplicated templates (EDT) to a housekeeping gene EDT.

[0013] Disclosed herein are methods and compositions for determining antimicrobial resistance of a microbe infecting a subject, comprising: providing a sample comprising nucleic acids from the subject, wherein: the nucleic acids comprise microbial nucleic acids from one or more potential carrier microbes harboring an antimicrobial resistance (AMR) genetic marker; performing a sequencing assay on the nucleic acids from the subject to generate sequence reads and detecting the one or more potential carrier microbe based on the sequence reads; if a carrier microbe is detected in (c), performing an amplification reaction on an aliquot of the sample, wherein the amplification reaction targets one or more AMR genetic markers in order to provide a copy number for the one or more AMR genetic markers; conducting a statistical analysis to link one or more of the AMR genetic markers to one of the one or more potential carrier microbe in the subject, thereby determining that a potential carrier microbe comprises an AMR genetic marker. In some embodiments, the methods disclosed herein further comprise performing a high-throughput sequencing assay on the amplicons associated with the second target nucleic acid. In some embodiments, the second target nucleic acid is not detected by a sequencing assay. In some embodiments, the first nucleic acid is associated with a genome of an organism but not a phenotype of interest of the organism and the second nucleic acid is associated with the phenotype of interest of the organism. In some embodiments, the first target nucleic acid comprises at least 2, at least 3, at least 4, at least 5, at least 10, at least 15, at least 20, or at least 25 first target nucleic acids; or the second target nucleic acid comprises at least 2, at least 3, at least 4, at least 5, at least 10, at least 15, at least 20, or at least 25 second target nucleic acids. In some embodiments, the first target nucleic acid or the second target nucleic acid are not negative or positive controls for any step of the multi-step assay. In some embodiments, the primers comprise multiple primers targeting multiple target nucleic acids. In some embodiments, the first target nucleic acid comprises nucleic acids from an animal, a virus, a bacterium, a protozoa, a fungus, an archaea, or an algae. In some embodiments, the second target nucleic acid comprises nucleic acids from an animal, a virus, a bacterium, a protozoa, a fungus, an archaea, or an algae. In some embodiments, the second target nucleic acid comprises a cancer marker. In some embodiments, the first target nucleic acid is associated with a carrier microbe harboring a target genetic marker and the second target nucleic acid comprises a sequence associated with the target genetic marker. In some embodiments, the carrier microbe comprises a virus, a bacterium, a protozoa, a fungus, an archaea, or an algae. In some embodiments, the target genetic marker comprises an antimicrobial resistance (AMR) genetic marker. In some embodiments, the methods disclosed herein further comprise comparing an abundance of the microbial sequence

reads from the AMR genetic marker to a threshold value. In some embodiments, the methods disclosed herein further comprise comparing an abundance of the microbial sequence reads from the gene cassette to a threshold value. In some embodiments, the methods disclosed herein further comprise providing a second aliquot for the amplification reaction when the abundance of the microbial sequence reads from the AMR genetic marker is below the threshold. In some embodiments, the methods disclosed herein further comprise providing a second aliquot for the amplification reaction when the abundance of the microbial sequence reads from the gene cassette is below the threshold. In some embodiments, the methods disclosed herein further comprise quantifying mcfNA sequencing reads from the one or more microbes. In some embodiments, the methods disclosed herein further comprise comparing the abundance of the mcfNA sequencing reads from the one or more microbes to a threshold. In some embodiments, the methods disclosed herein further comprise providing a second aliquot for the amplification reaction when the abundance of the mcfNA sequencing reads from the one or more microbes is below the threshold. In some embodiments, the methods disclosed herein further comprise performing high-throughput sequencing on the amplicons associated with the AMR genetic marker. In some embodiments, the sequencing assay on the first nucleic acids comprises a high-throughput sequencing assay. In some embodiments, the methods disclosed herein further comprise conducting a polymerase chain reaction (PCR) to amplify the AMR genetic marker, thereby producing amplicons associated with the AMR genetic marker. In some embodiments, the PCR comprises multiplex PCR, random PCR (rPCR), non-biased PCR, Nested PCR, Hot Start PCR, or Assembly PCR. In some embodiments, the methods disclosed herein further comprise attaching an adapter sequence to the second nucleic acids. In some embodiments, the primers comprise an adapter sequence. In some embodiments, the methods disclosed herein further comprise physically manipulating the sample to produce a fraction of cfNA enriched for degraded cfNA, wherein the fraction of cfNA comprises the AMR genetic marker. In some embodiments, the degraded cfNA comprises ultra short cfNA, single stranded cfNA, or nicked double stranded cfNA. In some embodiments, the ultra short cfNA comprises cfNA fragments that are less than 100 nucleotides in length. In some embodiments, the ultra short cfNA comprises cfNA fragments that are from 30 nucleotides to 70 nucleotides in length. In some embodiments, the methods disclosed herein further comprise performing size selection of the nucleic acids in the sample. In some embodiments, physically manipulating the sample comprises performing size selection of the nucleic acids in the sample. In some embodiments, performing size selection using a method selected from the group consisting of: chromatography, size-exclusion chromatography, electrophoresis, gel electrophoresis, automated electrophoresis, capillary electrophoresis, high-throughput size selection, centrifugation, density-gradient centrifugation,

filtration, membrane ultrafiltration, magnetic beads, and affinity-based beads. In some embodiments, the second aliquot comprises at least 500 μl of plasma. In some embodiments, detecting the carrier microbe comprises determining an abundance of the carrier microbe. In some embodiments, the abundance is expressed as molecules of mcfNA per microliter of sample (MPM). In some embodiments, the methods disclosed herein further comprise calculating an AMR gene copy number from the amplification or sequencing of the AMR genetic marker. In some embodiments, the methods disclosed herein further comprise linking the AMR genetic marker to the carrier microbe using the abundance of the mcfNA from the carrier microbe and the AMR gene copy number. In some embodiments, the AMR gene copy number is an episomal gene copy number. In some embodiments, the methods disclosed herein comprise introducing at least 200 primers targeting the plurality of AMR genetic markers into the second aliquot of the sample. In some embodiments, the methods disclosed herein further comprise introducing primers targeting housekeeping genes into the second aliquot of the sample. In some embodiments, the methods disclosed herein further comprise determining antimicrobial resistance of the microbe infecting the subject. In some embodiments, the sample comprises plasma. In some embodiments, the first nucleic acids are DNA. In some embodiments, the second nucleic acids are DNA. In some embodiments, the sample is not subjected to a process that primarily causes cell lysis. In some embodiments, the methods disclosed herein further comprise preparing a library from the amplicons associated with the AMR genetic marker. In some embodiments, the primers are added directly to the second aliquot. In some embodiments, the sample comprises cell-free nucleic acids. In some embodiments, the sample comprises cell-free DNA. In some embodiments, the first nucleic acids comprise microbial cell-free nucleic acids. In some embodiments, the second nucleic acids comprise microbial cell-free nucleic acids. In some embodiments, intact microbes are not actively lysed prior to performing the sequencing assay. In some embodiments, the microbial nucleic acids comprise microbial cell-free nucleic acids (mcfNA) from the microbe. In some embodiments, the microbial nucleic acids comprise nucleic acids associated with a microbial cell. In some embodiments, the methods disclosed herein comprise enriching for at least at least 75%, at least 80%, at least 85%, or at least 90% of the mcfNA in the sample. In some embodiments, the degraded cfNA comprise ultra short cfNA, single stranded cfNA, nicked double stranded cfNA, or any combination thereof. In some embodiments, the degraded cfNA comprise cfNA fragments that are less than 100 nucleotides in length. In some embodiments, the degraded cfNA comprise cfNA fragments that are from 30 nucleotides to 70 nucleotides in length. In some embodiments, physically manipulating the sample comprises performing size selection of the nucleic acids in the sample. In some embodiments, the methods disclosed herein further comprise generating sequence reads from the

cfNA from the subject, wherein the sequence reads comprise microbial sequencing reads derived from the microbe infecting the subject. In some embodiments, the sequence reads further comprise microbial sequencing reads from one or more carrier microbe. In some embodiments, the methods disclosed herein further comprise calculating an abundance of the mcfNA from the microbe in the sample and an abundance of the mcfNA from the one or more carrier microbe. In some embodiments, the abundance is expressed as molecules of mcfNA per 100 nanoliters of sample. In some embodiments, the methods disclosed herein further comprise calculating an AMR gene copy number for each of the one or more carrier microbe. In some embodiments, the methods disclosed herein further comprise detecting one or more carrier microbe from the sequencing assay. In some embodiments, the methods disclosed herein further comprise identifying the one or more carrier microbe as the microbe infecting the subject. In some embodiments, the methods disclosed herein further comprise linking the AMR genetic marker to one of the carrier microbes or the microbe infecting the subject. In some embodiments, the methods disclosed herein further comprise obtaining an abundance of the one or more carrier microbes based on an abundance of carrier microbe sequences in the sample. In some embodiments, conducting a statistical analysis comprises using the abundance of the microbial nucleic acids of the one or more potential carrier microbes. In some embodiments, conducting a statistical analysis comprises using the AMR gene copy number. In some embodiments, the methods disclosed herein further comprise calculating a probability of each of the one or more potential carrier microbes being the microbe harboring the AMR genetic marker, thereby identifying the potential carrier microbe as the microbe infecting the subject. In some embodiments, the methods disclosed herein further comprise determining the antimicrobial resistance of the microbe infecting the subject based on the calculated probability. In some embodiments, the detecting comprises performing a sequencing assay. In some embodiments, the sequencing assay comprises a high-throughput sequencing assay. In some embodiments, the amplification reaction comprises introducing primers targeting a plurality of AMR genetic markers. In some embodiments, the amplification reaction produces amplicons associated with the AMR genetic marker. In some embodiments, the sample comprises plasma. In some embodiments, the first nucleic acids are DNA. In some embodiments, the second nucleic acids are DNA. In some embodiments, the sample does not undergo cell lysis. In some embodiments, intact microbes are not actively lysed. In some embodiments, the amplicons associated with the AMR genetic marker undergoes direct library preparation. In some embodiments, the sample comprises cell-free nucleic acids. In some embodiments, the sample comprises cell-free DNA. In some embodiments, the first nucleic acids comprise cell-free nucleic acids. In some embodiments, the second nucleic acids comprise cell-free nucleic acids. In some embodiments,

the methods disclosed herein comprise adding primers targeting a plurality of AMR genetic markers into the second aliquot of the sample. In some embodiments, the methods disclosed herein comprise adding at least 200 primers targeting a plurality of AMR genetic markers into the second aliquot of the sample. In some embodiments, the methods disclosed herein further comprise introducing primers targeting housekeeping genes into the second aliquot of the sample. In some embodiments, the second aliquot of the sample comprises at least 500 µl of plasma. In some embodiments, detecting the pathogen comprises detecting the abundance of mcfNA from the pathogen in the sample over a threshold. In some embodiments, the methods disclosed herein further comprise calculating a positive percent agreement (PPA), negative percent agreement (NPA), diagnostic yield (DY), or any combination thereof. In some embodiments, the methods disclosed herein further comprise detecting cfDNA from one or more housekeeping genes. In some embodiments, the methods disclosed herein further comprise spiking one or more control molecules into the sample at a known concentration. In some embodiments, the one or more control molecules are synthetic oligonucleotides. In some embodiments, the control molecules comprise whole assay internal control (WINC) molecules. In some embodiments, the methods disclosed herein further comprise spiking at least 25,000 unique WINC molecules at known concentrations. In some embodiments, the methods disclosed herein further comprise generating a report listing the carrier microbes or pathogen detected in the subject. In some embodiments, the report further comprises the abundance of microbial cell-free DNA (mcfDNA) from microbes detected in the sample or the antimicrobial resistance of the microbes infecting the subject. In some embodiments, the sample comprises a biological sample obtained from the subject. In some embodiments, the biological sample is a whole blood sample, a plasma sample, a serum sample, a cerebrospinal fluid sample, a synovial fluid sample, a urine sample, a stool sample, a bronchoalveolar lavage sample, or any combination thereof. In some embodiments, the sample is a plasma sample. In some embodiments, the AMR genetic marker comprises a gene, a genetic element, a genetic cassette, an allele, a mutation, or any combination thereof. In some embodiments, the AMR genetic marker is associated with one or more genes selected from the group consisting of: *SCCmec, mecA, mecC, vanA, vanB,* $bla_{CTX\text{-}M}$, bla_{KPC}, *OXA-48-like*, *OXA-23, NDM, VIM, IMP*, or *mcr-1*. In some embodiments, the AMR genetic marker provides resistance to an anti-microbial agent selected from the group consisting of: methicillin, vancomycin, cephalosporin, carbapenem, and oxyimino-cephalosporin/aztreonam resistance. In some embodiments, the microbe infecting the subject or the carrier microbe comprises a virus, a bacterium, a protozoa, a fungus, an archaea, or an algae. In some embodiments, the microbe infecting the subject or the carrier microbe comprises a gram-positive bacterium or a fungus. In some embodiments, the microbe infecting the subject or the carrier microbe is a microbe listed

in Table 1. In some embodiments, the microbe infecting the subject or the carrier microbe harbors at last two AMR genetic markers. In some embodiments, the antimicrobial resistance is phenotypic antimicrobial resistance. In some embodiments, the sequencing comprises next-generation sequencing or a sequencing method beyond next generation sequencing. In some embodiments, the sequencing comprises sequencing by synthesis. In some embodiments, the subject is an animal. In some embodiments, the subject is a human. In some embodiments, the subject has an infection by the microbe infecting the subject harboring the AMR genetic marker. In some embodiments, the methods disclosed herein further comprise administering an anti-infective agent to the subject. In some embodiments, the subject has been treated with an anti-infective agent for an infection. In some embodiments, the methods disclosed herein further comprise adjusting the anti-microbial agent received by the subject at least in part based on the antimicrobial resistance of the microbe infecting the subject. In some embodiments, the methods disclosed herein further comprise determining antimicrobial resistance of a microbe. In some embodiments, the microbe comprises a microbe infecting the subject. In some embodiments, the microbe comprises a carrier microbe. In some embodiments, the methods disclosed herein comprise determining whether an AMR gene is carried by a microbe infecting the subject or by a carrier microbe. In some embodiments, the determining comprises linking an AMR genetic marker to one of the carrier microbes or the microbe infecting the subject. In some embodiments, the determining comprises determining a copy number of the AMR genetic marker in the microbe. In some embodiments, the linking comprises performing a statistical analysis on the sequence reads. In some embodiments, the statistical analysis comprises a generalized linear model, a maximum-likelihood estimation, a probit model, a logistic regression, a linear probability, a linear regression, a complimentary log-log, a Poisson regression, a support vector machine, a decision tree, a random forest, a neural network, a gradient boosted model, a Bayesian model, a hidden Markov model, or any combination thereof. In some embodiments, the method further comprises comparing the sequence read data to the amplification data. In some embodiments, the method further comprises calculating estimated deduplicated templates (EDT) for an AMR gene, wherein calculating EDT comprises calculating a number of unique DNA sequencing reads from a microorganism present in a sequenced library. In some embodiments, the method further comprises calculating an EDT for a housekeeping gene, wherein calculating EDT comprises calculating a number of unique DNA sequencing reads from a microorganism present in a sequenced library. In some embodiments, the methods disclosed herein further comprise determining a presence of an organism in a sample by calculating a ratio of an AMR estimated deduplicated templates (EDT) to a housekeeping gene EDT.

[0014] Disclosed herein are methods and compositions for determining antimicrobial resistance of a microbe infecting a subject, comprising: providing a first aliquot of the sample comprising first nucleic acids, wherein the first nucleic acids comprise microbial nucleic acids derived from a potential carrier microbe harboring an antimicrobial resistance (AMR) genetic marker; detecting the potential carrier microbe by analyzing the first nucleic acids; after the potential carrier microbe is detected, providing a second aliquot of the sample comprising second nucleic acids, wherein the second nucleic acids comprise microbial nucleic acids from the potential carrier microbe harboring an AMR genetic marker; conducting an amplification reaction on the second nucleic acids to detect the AMR genetic marker; and determining the antimicrobial resistance of the microbe infecting the subject based on the AMR genetic marker detection. In some embodiments, the methods disclosed herein further comprise performing a high-throughput sequencing assay on the amplicons associated with the second target nucleic acid. In some embodiments, the second target nucleic acid is not detected by a sequencing assay. In some embodiments, the first nucleic acid is associated with a genome of an organism but not a phenotype of interest of the organism and the second nucleic acid is associated with the phenotype of interest of the organism. In some embodiments, the first target nucleic acid comprises at least 2, at least 3, at least 4, at least 5, at least 10, at least 15, at least 20, or at least 25 first target nucleic acids; or the second target nucleic acid comprises at least 2, at least 3, at least 4, at least 5, at least 10, at least 15, at least 20, or at least 25 second target nucleic acids. In some embodiments, the first target nucleic acid or the second target nucleic acid are not negative or positive controls for any step of the multi-step assay. In some embodiments, the primers comprise multiple primers targeting multiple target nucleic acids. In some embodiments, the first target nucleic acid comprises nucleic acids from an animal, a virus, a bacterium, a protozoa, a fungus, an archaea, or an algae. In some embodiments, the second target nucleic acid comprises nucleic acids from an animal, a virus, a bacterium, a protozoa, a fungus, an archaea, or an algae. In some embodiments, the second target nucleic acid comprises a cancer marker. In some embodiments, the first target nucleic acid is associated with a carrier microbe harboring a target genetic marker and the second target nucleic acid comprises a sequence associated with the target genetic marker. In some embodiments, the carrier microbe comprises a virus, a bacterium, a protozoa, a fungus, an archaea, or an algae. In some embodiments, the target genetic marker comprises an antimicrobial resistance (AMR) genetic marker. In some embodiments, the methods disclosed herein further comprise comparing an abundance of the microbial sequence reads from the AMR genetic marker to a threshold value. In some embodiments, the methods disclosed herein further comprise comparing an abundance of the microbial sequence reads from the gene cassette to a threshold value. In some embodiments, the methods disclosed herein further

comprise providing a second aliquot for the amplification reaction when the abundance of the microbial sequence reads from the AMR genetic marker is below the threshold. In some embodiments, the methods disclosed herein further comprise providing a second aliquot for the amplification reaction when the abundance of the microbial sequence reads from the gene cassette is below the threshold. In some embodiments, the methods disclosed herein further comprise quantifying mcfNA sequencing reads from the one or more microbes. In some embodiments, the methods disclosed herein further comprise comparing the abundance of the mcfNA sequencing reads from the one or more microbes to a threshold. In some embodiments, the methods disclosed herein further comprise providing a second aliquot for the amplification reaction when the abundance of the mcfNA sequencing reads from the one or more microbes is below the threshold. In some embodiments, the methods disclosed herein further comprise performing high-throughput sequencing on the amplicons associated with the AMR genetic marker. In some embodiments, the sequencing assay on the first nucleic acids comprises a high-throughput sequencing assay. In some embodiments, the methods disclosed herein further comprise conducting a polymerase chain reaction (PCR) to amplify the AMR genetic marker, thereby producing amplicons associated with the AMR genetic marker. In some embodiments, the PCR comprises multiplex PCR, random PCR (rPCR), non-biased PCR, Nested PCR, Hot Start PCR, or Assembly PCR. In some embodiments, the methods disclosed herein further comprise attaching an adapter sequence to the second nucleic acids. In some embodiments, the primers comprise an adapter sequence. In some embodiments, the methods disclosed herein further comprise physically manipulating the sample to produce a fraction of cfNA enriched for degraded cfNA, wherein the fraction of cfNA comprises the AMR genetic marker. In some embodiments, the degraded cfNA comprises ultra short cfNA, single stranded cfNA, or nicked double stranded cfNA. In some embodiments, the ultra short cfNA comprises cfNA fragments that are less than 100 nucleotides in length. In some embodiments, the ultra short cfNA comprises cfNA fragments that are from 30 nucleotides to 70 nucleotides in length. In some embodiments, the methods disclosed herein further comprise performing size selection of the nucleic acids in the sample. In some embodiments, physically manipulating the sample comprises performing size selection of the nucleic acids in the sample. In some embodiments, performing size selection using a method selected from the group consisting of: chromatography, size-exclusion chromatography, electrophoresis, gel electrophoresis, automated electrophoresis, capillary electrophoresis, high-throughput size selection, centrifugation, density-gradient centrifugation, filtration, membrane ultrafiltration, magnetic beads, and affinity-based beads. In some embodiments, the second aliquot comprises at least 500 μl of plasma. In some embodiments, detecting the carrier microbe comprises determining an abundance of the carrier microbe. In

some embodiments, the abundance is expressed as molecules of mcfNA per microliter of sample (MPM). In some embodiments, the methods disclosed herein further comprise calculating an AMR gene copy number from the amplification or sequencing of the AMR genetic marker. In some embodiments, the methods disclosed herein further comprise linking the AMR genetic marker to the carrier microbe using the abundance of the mcfNA from the carrier microbe and the AMR gene copy number. In some embodiments, the AMR gene copy number is an episomal gene copy number. In some embodiments, the methods disclosed herein comprise introducing at least 200 primers targeting the plurality of AMR genetic markers into the second aliquot of the sample. In some embodiments, the methods disclosed herein further comprise introducing primers targeting housekeeping genes into the second aliquot of the sample. In some embodiments, the methods disclosed herein further comprise determining antimicrobial resistance of the microbe infecting the subject. In some embodiments, the sample comprises plasma. In some embodiments, the first nucleic acids are DNA. In some embodiments, the second nucleic acids are DNA. In some embodiments, the sample is not subjected to a process that primarily causes cell lysis. In some embodiments, the methods disclosed herein further comprise preparing a library from the amplicons associated with the AMR genetic marker. In some embodiments, the primers are added directly to the second aliquot. In some embodiments, the sample comprises cell-free nucleic acids. In some embodiments, the sample comprises cell-free DNA. In some embodiments, the first nucleic acids comprise microbial cell-free nucleic acids. In some embodiments, the second nucleic acids comprise microbial cell-free nucleic acids. In some embodiments, intact microbes are not actively lysed prior to performing the sequencing assay. In some embodiments, the microbial nucleic acids comprise microbial cell-free nucleic acids (mcfNA) from the microbe. In some embodiments, the microbial nucleic acids comprise nucleic acids associated with a microbial cell. In some embodiments, the methods disclosed herein comprise enriching for at least at least 75%, at least 80%, at least 85%, or at least 90% of the mcfNA in the sample. In some embodiments, the degraded cfNA comprise ultra short cfNA, single stranded cfNA, nicked double stranded cfNA, or any combination thereof. In some embodiments, the degraded cfNA comprise cfNA fragments that are less than 100 nucleotides in length. In some embodiments, the degraded cfNA comprise cfNA fragments that are from 30 nucleotides to 70 nucleotides in length. In some embodiments, physically manipulating the sample comprises performing size selection of the nucleic acids in the sample. In some embodiments, the methods disclosed herein further comprise generating sequence reads from the cfNA from the subject, wherein the sequence reads comprise microbial sequencing reads derived from the microbe infecting the subject. In some embodiments, the sequence reads further comprise microbial sequencing reads from one or more carrier microbe. In some embodiments,

the methods disclosed herein further comprise calculating an abundance of the mcfNA from the microbe in the sample and an abundance of the mcfNA from the one or more carrier microbe. In some embodiments, the abundance is expressed as molecules of mcfNA per 100 nanoliters of sample. In some embodiments, the methods disclosed herein further comprise calculating an AMR gene copy number for each of the one or more carrier microbe. In some embodiments, the methods disclosed herein further comprise detecting one or more carrier microbe from the sequencing assay. In some embodiments, the methods disclosed herein further comprise identifying the one or more carrier microbe as the microbe infecting the subject. In some embodiments, the methods disclosed herein further comprise linking the AMR genetic marker to one of the carrier microbes or the microbe infecting the subject. In some embodiments, the methods disclosed herein further comprise obtaining an abundance of the one or more carrier microbes based on an abundance of carrier microbe sequences in the sample. In some embodiments, conducting a statistical analysis comprises using the abundance of the microbial nucleic acids of the one or more potential carrier microbes. In some embodiments, conducting a statistical analysis comprises using the AMR gene copy number. In some embodiments, the methods disclosed herein further comprise calculating a probability of each of the one or more potential carrier microbes being the microbe harboring the AMR genetic marker, thereby identifying the potential carrier microbe as the microbe infecting the subject. In some embodiments, the methods disclosed herein further comprise determining the antimicrobial resistance of the microbe infecting the subject based on the calculated probability. In some embodiments, the detecting comprises performing a sequencing assay. In some embodiments, the sequencing assay comprises a high-throughput sequencing assay. In some embodiments, the amplification reaction comprises introducing primers targeting a plurality of AMR genetic markers. In some embodiments, the amplification reaction produces amplicons associated with the AMR genetic marker. In some embodiments, the sample comprises plasma. In some embodiments, the first nucleic acids are DNA. In some embodiments, the second nucleic acids are DNA. In some embodiments, the sample does not undergo cell lysis. In some embodiments, intact microbes are not actively lysed. In some embodiments, the amplicons associated with the AMR genetic marker undergoes direct library preparation. In some embodiments, the sample comprises cell-free nucleic acids. In some embodiments, the sample comprises cell-free DNA. In some embodiments, the first nucleic acids comprise cell-free nucleic acids. In some embodiments, the second nucleic acids comprise cell-free nucleic acids. In some embodiments, the methods disclosed herein comprise adding primers targeting a plurality of AMR genetic markers into the second aliquot of the sample. In some embodiments, the methods disclosed herein comprise adding at least 200 primers targeting a plurality of AMR genetic markers into

the second aliquot of the sample. In some embodiments, the methods disclosed herein further comprise introducing primers targeting housekeeping genes into the second aliquot of the sample. In some embodiments, the second aliquot of the sample comprises at least 500 μl of plasma. In some embodiments, detecting the pathogen comprises detecting the abundance of mcfNA from the pathogen in the sample over a threshold. In some embodiments, the methods disclosed herein further comprise calculating a positive percent agreement (PPA), negative percent agreement (NPA), diagnostic yield (DY), or any combination thereof. In some embodiments, the methods disclosed herein further comprise detecting cfDNA from one or more housekeeping genes. In some embodiments, the methods disclosed herein further comprise spiking one or more control molecules into the sample at a known concentration. In some embodiments, the one or more control molecules are synthetic oligonucleotides. In some embodiments, the control molecules comprise whole assay internal control (WINC) molecules. In some embodiments, the methods disclosed herein further comprise spiking at least 25,000 unique WINC molecules at known concentrations. In some embodiments, the methods disclosed herein further comprise generating a report listing the carrier microbes or pathogen detected in the subject. In some embodiments, the report further comprises the abundance of microbial cell-free DNA (mcfDNA) from microbes detected in the sample or the antimicrobial resistance of the microbes infecting the subject. In some embodiments, the sample comprises a biological sample obtained from the subject. In some embodiments, the biological sample is a whole blood sample, a plasma sample, a serum sample, a cerebrospinal fluid sample, a synovial fluid sample, a urine sample, a stool sample, a bronchoalveolar lavage sample, or any combination thereof. In some embodiments, the sample is a plasma sample. In some embodiments, the AMR genetic marker comprises a gene, a genetic element, a genetic cassette, an allele, a mutation, or any combination thereof. In some embodiments, the AMR genetic marker is associated with one or more genes selected from the group consisting of: *SCCmec, mecA, mecC, vanA, vanB,* $bla_{CTX\text{-}M}$, bla_{KPC}, *OXA-48-like*, *OXA-23, NDM, VIM, IMP*, or *mcr-1*. In some embodiments, the AMR genetic marker provides resistance to an anti-microbial agent selected from the group consisting of: methicillin, vancomycin, cephalosporin, carbapenem, and oxyimino-cephalosporin/aztreonam resistance. In some embodiments, the microbe infecting the subject or the carrier microbe comprises a virus, a bacterium, a protozoa, a fungus, an archaea, or an algae. In some embodiments, the microbe infecting the subject or the carrier microbe comprises a gram-positive bacterium or a fungus. In some embodiments, the microbe infecting the subject or the carrier microbe is a microbe listed in Table 1. In some embodiments, the microbe infecting the subject or the carrier microbe harbors at last two AMR genetic markers. In some embodiments, the antimicrobial resistance is phenotypic antimicrobial resistance. In some embodiments, the sequencing comprises next-

generation sequencing or a sequencing method beyond next generation sequencing. In some embodiments, the sequencing comprises sequencing by synthesis. In some embodiments, the subject is an animal. In some embodiments, the subject is a human. In some embodiments, the subject has an infection by the microbe infecting the subject harboring the AMR genetic marker. In some embodiments, the methods disclosed herein further comprise administering an anti-infective agent to the subject. In some embodiments, the subject has been treated with an anti-infective agent for an infection. In some embodiments, the methods disclosed herein further comprise adjusting the anti-microbial agent received by the subject at least in part based on the antimicrobial resistance of the microbe infecting the subject. In some embodiments, the methods disclosed herein further comprise determining antimicrobial resistance of a microbe. In some embodiments, the microbe comprises a microbe infecting the subject. In some embodiments, the microbe comprises a carrier microbe. In some embodiments, the methods disclosed herein comprise determining whether an AMR gene is carried by a microbe infecting the subject or by a carrier microbe. In some embodiments, the determining comprises linking an AMR genetic marker to one of the carrier microbes or the microbe infecting the subject. In some embodiments, the determining comprises determining a copy number of the AMR genetic marker in the microbe. In some embodiments, the linking comprises performing a statistical analysis on the sequence reads. In some embodiments, the statistical analysis comprises a generalized linear model, a maximum-likelihood estimation, a probit model, a logistic regression, a linear probability, a linear regression, a complimentary log-log, a Poisson regression, a support vector machine, a decision tree, a random forest, a neural network, a gradient boosted model, a Bayesian model, a hidden Markov model, or any combination thereof. In some embodiments, the method further comprises comparing the sequence read data to the amplification data. In some embodiments, the method further comprises calculating estimated deduplicated templates (EDT) for an AMR gene, wherein calculating EDT comprises calculating a number of unique DNA sequencing reads from a microorganism present in a sequenced library. In some embodiments, the method further comprises calculating an EDT for a housekeeping gene, wherein calculating EDT comprises calculating a number of unique DNA sequencing reads from a microorganism present in a sequenced library. In some embodiments, the methods disclosed herein further comprise determining a presence of an organism in a sample by calculating a ratio of an AMR estimated deduplicated templates (EDT) to a housekeeping gene EDT.

[0015] In some embodiments, the method further comprises administering an alternative therapy to the subject if an AMR marker is detected. In some embodiments, the method further comprises detecting methicillin-resistant Staph aureus by detecting SCCmec, mecA, or mecC. In some embodiments, the method further comprises detecting methicillin-resistant

Staph aureus and administering ceftaroline, ceftobiprole, vancomycin, linezolid, daptomycin, ceftobipole or combination thereof to the subject in order to treat the methicillin-resistant staph aureus. In some embodiments, the method further comprises detecting vancomycin resistant enterococcus by detecting VanA or VanB. In some embodiments, the method further comprises detecting vancomycin resistant enterococcus and administering linezolid or daptomycin to the subject in order to treat the vancomycin resistant enterococcus.

[0016] Described herein in some embodiments is a method to detect phenotypic antimicrobial susceptibility of a human host comprising: a) providing a sample comprising cell-free nucleic acids from a pathogen harboring at least one antimicrobial resistance (AMR) marker, wherein the cell-free nucleic acids from the pathogen are obtained from a biological sample of the human host; b) performing high-throughput sequencing of said cell-free nucleic acids from the pathogen to provide cell-free sequence reads; c) using said cell-free sequence reads from the pathogen in the biological sample to identify said at least one AMR marker. In one embodiment, the disclosed method further comprises adding adapters to said cell-free nucleic acids and producing a sequencing library comprising said cell-free nucleic acids. In one embodiment, the disclosed method further comprises enriching for ultra short, ultra rare target fragments in the biological sample prior to sequencing. In one embodiment, the disclosed method further comprises spiking whole assay internal control (WINC) molecules into each sample at a known concentration as a control for sequencing depth or human background levels, wherein the control molecules are synthetic oligonucleotides. In one embodiment, the control molecules comprise a minimum of 25,000 unique WINC molecules. In one embodiment of the disclosed method, the AMR marker provides methicillin resistance. In one embodiment of the disclosed method, the AMR marker provides vancomycin resistance. In one embodiment of the disclosed method, the AMR marker provides cephalosporin resistance. In one embodiment of the disclosed method, the AMR marker provides oxyimino-cephalosporin and aztreonam resistance. In one embodiment of the disclosed method, the AMR marker provides carbapenem resistance. In one embodiment of the disclosed method, the AMR markers include SCCmec, mecA and mecC for methicillin resistance, vanA and vanB for vancomycin resistance, blaCTX-M for oxyimino-cephalosporin and aztreonam resistance, and blaKPC for carbapenem resistance. In one embodiment of the disclosed method, the human host is a patient having an infection by the pathogen harboring at least one AMR marker. In one embodiment, the disclosed method further comprises adjusting the antibiotic treatment of the patient to compensate of the presence of the AMR in the pathogen. In one embodiment of the disclosed method, the biological sample is selected from the group consisting of blood, serum, cerebrospinal fluid, synovial fluid, urine, and stool. In one embodiment of the

disclosed method, the sequence analysis comprises whole-genome sequencing (WGS). In one embodiment of the disclosed method, the pathogen comprises a target bacteria listed in Table 1.

INCORPORATION BY REFERENCE

[0017] All publications, patents, and patent applications mentioned in this specification are herein incorporated by reference in their entireties to the same extent as if each individual publication, patent, or patent application was specifically and individually indicated to be incorporated by reference.

BRIEF DESCRIPTION OF THE DRAWINGS

[0018] The novel features of the disclosure are set forth with particularity in the appended claims. A better understanding of the features and advantages of the present disclosure will be obtained by reference to the following detailed description that sets forth illustrative embodiments, in which the principles of the disclosure are utilized, and the accompanying drawings of which:

[0019] **FIG. 1** depicts an overview of the Karius Test AMR workflow.

[0020] **FIG. 2** shows blaCTX-M variant prevalence (%) taken from CARD Resistomes database. Dark grey represents CTX-M variants with >95% sequence identity to the reference CTX-M sequence used in primer design and light gray represents CTX-M variants with <75% sequence identity.

[0021] **FIG. 3**. Distribution of index correct read counts across 108 clinical samples.

[0022] **FIG. 4**. Distribution of EDTs in negative control samples for each AMR target gene.

[0023] **FIGs. 5A-5B.** Probit fits for estimating CTX-M presence LoD using the (FIG. 5A) original dilution series and (FIG. 5B) simulated dilution series assuming one CTX-M gene copy. Light grey lines reflect the probit model fit and dark grey lines reflect the observed data of fraction positive calls per concentration.

[0024] **FIG. 6** shows a computer control system that is programmed or otherwise configured to implement the methods provided herein.

[0025] **FIG. 7** shows a table of clinical validation performance summary.

DETAILED DESCRIPTION

[0026] The following passages describe different aspects of the disclosure in greater detail. Each aspect, embodiment, or feature of the disclosure can be combined with any other aspect, embodiment, or feature of the disclosure unless clearly indicated to the contrary. There remains a need for developing improved methods for identifying genetic markers in the context of disease diagnosis and treatment.

Overview

[0027] The present disclosure provides, in some embodiments, methods of performing a plurality of assays in order to detect whether a subject is infected with a microbe that contains an antimicrobial resistance (AMR) genetic marker. The methods provided herein are particularly useful for determining whether a detected microbe carries an AMR genetic marker. In some embodiments, the assays are performed sequentially. For example, in some embodiments the methods comprise performing unbiased high-throughput sequencing (HTS) to detect a potential carrier microbe, followed by performing targeted sequencing in order to detect an AMR genetic marker in a sample. In some cases, the method comprises performing a targeted sequencing assay to detect both a carrier microbe and its associated AMR genetic marker, if present. The methods provided herein, in some embodiments, can comprise performing a statistical analysis to link a carrier microbe to a particular AMR genetic marker. The statistical analysis can be particularly helpful when nucleic acids from multiple different carrier microbes are detected, which can make it difficult to determine which carrier microbe comprises a detected AMR genetic marker.

[0028] The methods provided herein, in some embodiments, comprise performing a multi-step assay for detecting an AMR genetic marker by analyzing at least two aliquots of the same sample in at least two sequential assays. In some embodiments, the first assay is a high throughput sequencing assay, and the second assay is a targeted PCR amplification assay, that can, in some cases, be followed by a second high throughput sequencing assay. In some embodiments, carrier microbes and AMR genetic markers are detected in the same assay. In some embodiments, the carrier microbes are detected using primers specific for housekeeping genes and the AMR markers are detected using primers specific for sets of AMR genes. In some embodiments, targeted amplification for multiple targets is performed, followed by, in some embodiments, performing a statistical analysis to link detected carrier microbes with AMR genetic markers.

[0029] The present disclosure provides, in some embodiments, methods of detecting an antimicrobial resistance (AMR) genetic marker by analyzing cell-free nucleic acids in a biological sample from a patient using high-throughput sequencing and/or targeted sequencing in order to quantify or detect an abundance of a carrier microbe and/or an AMR genetic marker. In some embodiments, the methods provided herein enable AMR marker detection by identifying one or more potential carrier microbes and one or more AMR genetic markers, and conducting a statistical analysis to link the AMR genetic marker to the carrier microbe harboring the AMR genetic marker. In some embodiments, the methods comprise calculating an abundance of mcfNA (e.g., expressed as molecules per microliter (MPM)) from the one or more potential

carrier microbes and/or calculating an abundance or copy number of mcfNA associated with the AMR genetic marker. In some embodiments, the methods comprise calculating a copy number of the AMR genetic marker and/or calculating abundances of one or more microbial housekeeping genes. In some embodiments, the methods comprise linking the AMR genetic marker to its carrier microbe based on the abundance of the mcfNA associated with the carrier and the copy number of the AMR genetic marker. In some embodiments, the methods comprise linking the AMR genetic marker to its carrier microbe based on the copy number of both the AMR genetic marker and the one or more microbe-specific housekeeping genes. In some embodiments, the method comprises analyzing certain features or criteria to determine that a carrier microbe harbors the AMR genetic marker. In some embodiments, such criteria can comprise a copy number of the AMR marker, the size of the AMR marker, the size of a gene cassette comprising the AMR marker, an abundance of mcfNA associated with the AMR marker, an abundance of mcfNA associated with the carrier microbe, or an abundance of mcfNA associated with the AMR marker in comparison to an abundance of mcfNA associated with one or more microbial-specific housekeeping genes.

[0030] In some embodiments, the methods comprise performing an HTS assay such as a sequencing-by-synthesis assay in order to detect and/or quantify nucleic acids associated with the one or more potential carrier microbes. In some embodiments, a targeted assay is then performed to detect the AMR genetic marker, particularly if a carrier microbe was identified in the first high throughput sequencing assay. For example, in some embodiments, detection of a carrier microbe in the first high throughput sequencing assay triggers obtaining a second aliquot of the sample and performing a targeted assay to detect nucleic acids in the second aliquot associated with the AMR genetic marker. The targeted assay can, for example, comprise performing a PCR reaction with primers specific for one or more AMR genetic markers, optionally followed by high throughput sequencing to detect the AMR genetic markers. In some embodiments, the methods comprise calculating abundances of microbial nucleic acids detected in the targeted assay, such as an estimated deduplicated transcripts or templates (EDT) or a copy number. In some embodiments, the methods comprise calculating an EDT of one or more AMR genetic markers, a copy number of the one or more AMR genetic markers, or an EDT of one or more microbial-specific housekeeping genes. In some embodiments, the methods comprise linking an AMR genetic marker to its carrier microbe based on the results from the targeted assay. In some embodiments, the methods comprise linking an AMR genetic marker to its carrier microbe using the copy number of the AMR genetic marker and the abundances of the one or more microbial-specific housekeeping genes.

[0031] In some embodiments, criteria associated with an HTS assay that detected a carrier microbe or an AMR gene cassette are analyzed in order to determine whether a subsequent assay such as a targeted sequencing assay is performed. In some cases, the methods comprise comparing the amount or abundance of mcfNA (e.g., MPM) associated with a carrier microbe, an AMR gene cassette, and/or an AMR gene with a threshold value. In some embodiments, the methods comprise proceeding to the targeted assay for the AMR genetic marker based on the comparison with the threshold value. In some instances, the methods comprise the use of primers targeting the AMR genetic marker for the targeted amplification of the AMR genetic marker, which is optionally followed by high throughput sequencing to detect the AMR genetic marker.

[0032] In some embodiments, a nucleic acid fraction enriched for ultra-short, degraded, nicked double-stranded, and/or single-stranded cfNA is used in the assays provided herein. In some cases, the methods comprise an extraction step. In some cases, the cfNA are amplified within a sample prior to extraction. In some cases, adapters or primers for the cfNA are added directly to the sample prior to extraction.

[0033] The high-throughput nature of the methods provided herein generally enables comprehensive profiling of the subject's microbiome, infectome, and/or resistome. For example, the methods provided herein can further detect co-infections in a patient with complex infections. The methods can also comprise treating patients based on the identification of the AMR genetic markers, which can reduce or inform the use of broad-spectrum antibiotics and allow for a more targeted approach to treatment of infectious diseases. For example, the methods can comprise treating the patients who have received one class of antimicrobial treatment with a different class of antimicrobial treatments to overcome AMR.

Genetic markers

[0034] Disclosed herein in some embodiments are methods for detecting one or more genetic markers associated with a disease or a disorder using high-throughput sequencing. In some embodiments, the one or more genetic markers are associated with an infectious disease. In some embodiments, the one or more genetic markers are associated with a non-communicable disease. In some embodiments, the one or more genetic markers comprise a genetic marker from a prokaryotic organism. In some embodiments, the one or more genetic markers comprise a genetic marker from a eukaryotic organism.

[0035] In some embodiments, the one or more genetic markers comprises a genetic locus, a gene, an allele, a genetic variant, a gene element (e.g., a regulatory element), a gene cassette, an operon, an integron, a resistance island, a resistance supercluster, a chromosome, a plasmid, or any combination thereof. In some embodiments, the one or more genetic markers comprise a

mutation associated with a disease or disorder. In some embodiments, the mutation comprises a point mutation (e.g., missense, nonsense, silent, frameshift), a structural mutation (e.g., insertion, deletion, duplication, inversion, translocation), repeat expansions (e.g., trinucleotide repeats, MSI), a gain-of-function mutation, a loss-of-function mutation, a dominant negative mutation, or any combination thereof.

[0036] In some embodiments, the genetic markers comprise a microbial genetic marker. In some embodiments, the microbial genetic marker is from a virus, a bacterium, a protozoa, a fungus, an archaea, or an algae. In some embodiments, the microbial genetic marker comprises an antimicrobial resistance (AMR) marker. In some embodiments, the genetic markers comprise a mammalian genetic marker. In some embodiments, the mammalian genetic marker is from a human. In some embodiments, the genetic marker is associated with a human disease. In some embodiments, the genetic marker comprises a cancer marker.

Antimicrobial Resistance (AMR) Markers

[0037] Disclosed herein in some embodiments are methods for detecting one or more antimicrobial resistance (AMR) genetic markers. An AMR genetic marker can be associated with one or more AMR genes that enable microorganisms to survive exposure to anti-microbial agents, leading to treatment failure and persistent infections.

[0038] The AMR genetic marker can be located anywhere in a microbial genome. In some embodiments, the AMR genetic marker is integrated in a microbial genome (e.g., chromosome), an episomal genome (e.g., plasmid), a transposon, or any combination thereof. In some embodiments, the AMR genetic marker is located on an episomal genome.

[0039] The AMR genetic marker can be of any size within a microbial genome. In some embodiments, the AMR genetic marker comprises a genetic locus, a gene, an allele, a genetic variant, a gene element (e.g., a regulatory element), a gene cassette, an operon, an integron, a resistance island, a resistance supercluster, a chromosome, a plasmid, or any combination thereof.

[0040] In some embodiments, the AMR genetic marker is associated with one or more AMR genes. In some embodiments, the AMR genetic marker is associated with at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 20, at least 30, at least 40, at least 50, at least 60, at least 70, at least 80, at least 90, or at least 100 AMR genes.

[0041] In some embodiments, the AMR genetic marker is originated from one or more carrier microbes harboring the AMR genetic marker. In some embodiments, the AMR genetic marker is originated from at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 20, at least 30, at least 40, at least 50, at least 60, at least 70, at least 80, at least 90, or at least 100 AMR carrier microbes harboring the AMR genetic marker. For example,

the *mecA* gene can be carried by multiple Staphylococcus species, such as S. aureus and S. epidermidis.

[0042] In some embodiments, the AMR genetic marker is linked to the one or more carrier microbes harboring the AMR genetic marker. In some embodiments, the AMR genetic marker is linked to the one or more carrier microbes detected by first sequencing assay in the method. In some embodiments, the methods detect a carrier microbe and then detect a phenotypic characteristic of the carrier microbe. In some cases, the phenotypic characteristic is antimicrobial resistance.

[0043] In some embodiments, the AMR genetic marker is associated with resistance to one or more anti-microbial or anti-infective agents. The AMR genetic marker can be associated with resistance to any class of anti-microbial agents. In some embodiments, the AMR genetic marker is associated with resistance to a β-Lactam drug (e.g., carbapenems, methicillin), oxyimino-cephalosporins/aztreonam, an aminoglycoside drug (e.g., gentamicin), a Macrolide-Lincosamide-Streptogramin (MLS) drug (e.g., clindamycin), a tetracycline drug (e.g., doxycycline), a fluoroquinolone drug (e.g., ciprofloxacin), a glycopeptide drug (e.g., vancomycin), a polymyxin drug (e.g., colistin), rifamycin (e.g., rifampin), or a sulfonamide drug (e.g., trimethoprim-sulfamethoxazole). In some embodiments, the AMR genetic marker is associated with resistance to a broad-spectrum anti-microbial drug. In some embodiments, the AMR genetic marker is associated with resistance to a narrow-spectrum anti-microbial drug.

[0044] In some embodiments, the AMR genetic marker is associated with or comprises one or more AMR genes selected from the group consisting of: mecA, mecC, bla_TEM, bla_SHV, bla_CTX-M, bla_KPC, bla_NDM, bla_OXA-48, bla_OXA-23, VIM, IMP aac(6')-Ib, aph(3')-IIIa, ant(2'')-Ia, erm(A), erm(B), erm(C), mef(A), mef(E), tet(A), tet(B), tet(M), gyrA/parC, qnrA, qnrB, qnrS, aac(6')-Ib-cr, vanA, vanB, mcr-1, pmrA/pmrB, rpoB, sul1, sul2, sul3, dfrA genes, acrAB-TolC, mexAB-oprM., katG, inhA, rpoB, pncA, embB, rrs, gyrA, and gyrB.

[0045] In some embodiments, the AMR genetic marker is associated with a gene cassette. A gene cassette is a mobile genetic element that can harbor one or more AMR genes. In some embodiments, the gene cassette can be "empty" and does not harbor any AMR gene. In some embodiments, the gene cassette can integrate into a genome. In some embodiments, the gene cassette comprises SCCmec. SCCmec are mobile genetic elements that can integrate into coagulase-negative staphylococci species (e.g., Staphylococcus aureus). In some embodiments, the SCCmec can harbor one or more AMR genes. For example, SCCmec can harbor mecA/mecC in Staphylococcus aureus and additional AMR genes such as blaZ or aminoglycoside resistance genes. In some embodiments, SCCmec can be "empty" and does not harbor one or more AMR

genes. In some embodiments, SCCmec comprises Type I, Type II, Type III, Type IV, Type V or Type VIII SCCmec.

[0046] In some embodiments, a gene cassette comprises an integron. In some embodiments, an integron comprises a class 1 integron, a class 2 integron, or a class 2 integron.

[0047] In some embodiments, the AMR genetic markers can be present in a carrier microbe at any copy number. As used herein, "copy number" can refer to the number of times a microbial gene or genomic region (e.g., the AMR genetic marker) is present in a microbe. In some embodiments, the copy number of the AMR genetic marker comprises a chromosomal copy number, an integron copy number, or an episomal (or plasmid) copy number. In some embodiments, the copy number of the AMR genetic marker is 1. In some embodiments, the copy number of the AMR genetic marker is at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10. In some embodiments, the copy number of the AMR genetic marker is calculated by PCR.

Methods and Assay Workflow

[0048] The methods described herein can comprise performing a multi-step assay on a sample from a subject. In some embodiments, the multi-step assay comprises performing a first high-throughput sequencing (HTS) assay on nucleic acids from the sample followed by performing a targeted sequencing assay. In some instances, the targeted sequencing assay can comprise performing a PCR reaction on the sample, optionally followed by a second high-throughput sequencing assay. In some cases, performing a targeted sequencing assay is predicated on the results of a first HTS assay. In some embodiments, results from a first HTS assay may meet certain criteria before a targeted sequencing assay is performed. In some embodiments, after the targeted sequencing assay is performed, an analysis is performed to link the nucleic acids detected in the targeted sequencing assay with nucleic acids detected in the first high-throughput sequencing assay. In some cases, linking results from the two assays can provide phenotypic characteristics related to a genome identified in the first high throughput sequencing assay.

[0049] In some embodiments, the methods provided herein comprise performing a multi-step assay on a sample in order to detect a microbe of interest. In some instances, a first HTS assay is performed in order to detect a carrier microbe; in some embodiments, the first HTS is an unbiased sequencing assay. In some embodiments, a targeted sequencing assay is performed if the results of the first HTS assay meet certain criteria. In some cases, the targeted sequencing assay is performed if one or more microbes known to carry an AMR marker is detected in the first HTS. In some embodiments, if a carrier microbe or AMR cassette is identified in the first HTS, a targeted sequencing assay is performed in order to detect the AMR marker. If an AMR marker

is detected, then, in some embodiments, a linkage analysis is performed in order to link the AMR marker with the carrier microbe detected in the first HTS. In some embodiments, both a carrier microbe and an AMR marker are detected in a targeted assay. In such cases, the methods may comprise performing an analysis in order to link the AMR marker with the carrier microbe detected in the targeted assay.

[0050] In some embodiments, the at least two sequencing assays can be the same sequencing assay. In some embodiments, the at least two sequencing assays can be different sequencing assays. In some embodiments, the at least two sequencing assays are performed in parallel. In some embodiments, the at least two sequencing assays are performed sequentially. In some embodiments, methods comprise performing a first high throughput sequencing assay before performing a targeted amplification or a second high throughput sequencing assay. In some embodiments, the second high throughput assay is a targeted sequencing assay. In some embodiments, only a targeted amplification or sequencing assay is performed. In some cases, an unbiased sequencing method is performed prior to the targeted assay.

Sample Aliquots

[0051] In some embodiments, the methods comprise performing a high throughput sequencing assay or amplification assay on an aliquot of a sample. Generally, as used herein, an aliquot is a portion of a sample; but in some cases, the term aliquot as used herein refers to the entire sample. In some embodiments, the methods comprise obtaining multiple aliquots from a single sample. In some cases, the methods comprise performing an assay on a first sample and then performing an assay on a second sample that does not derive from the same sample as the first sample. In some embodiments, an aliquot is a remaining portion of the sample (or a fraction of the remaining portion of the sample), after an initial aliquot has been taken out of the sample. For example, a first aliquot can be taken out of the subject's sample and analyzed by a first sequencing assay in the methods described herein. After that, in some embodiments, a second aliquot can be taken out of the subject's sample and analyzed by a second sequencing assay (e.g., a targeted sequencing assay). In some cases, after the first aliquot is removed from the sample, the entire remaining sample is analyzed as the second aliquot in an amplification or a second sequencing assay. In some cases, after the first aliquot is removed from the sample, a portion or fraction of the remaining sample is analyzed as the second aliquot in the amplification or the second sequencing assay.

[0052] In some embodiments, assays within workflow analyze a different aliquot of the sample. In some embodiments, an aliquot of the sample can be any measured portion of the sample. In some cases, an aliquot is a portion of an original sample, such as at least 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90% percent of the sample. In some cases, an aliquot is a portion of an

original sample, such as less than 10%, 20%, 30%, 40%, 50%, 60%, 70%, or 80% percent of the sample. In some embodiments, the aliquot is at most 50 μl, at most 100 μl, at most 150 μl, at most 200 μl, at most 250 μl, at most 300 μl, at most 350 μl, at most 400 μl, at most 450 μl, at most 500 μl, at most 550 μl, at most 600 μl, at most 650 μl, at most 700 μl, at most 750 μl, at most 800 μl, at most 850 μl, at most 900 μl, at most 950 μl, or at most 1 ml of the sample. In some embodiments, the aliquot is 500 μl of a plasma sample obtained from a human. In some embodiments, the aliquot is 200 μl of the plasma sample.

[0053] In some embodiments, the methods comprise performing a first sequencing assay to analyze a first aliquot of the sample from the subject in order to detect first target nucleic acids, wherein the first aliquot comprises the first target nucleic acids. The methods described herein further comprise, in some embodiments, after detecting the first target nucleic acids, performing an amplification or a second sequencing assay to analyze a second aliquot of the sample in order to detect second target nucleic acids, wherein the second aliquot comprises the second target nucleic acids. In some embodiments, the first target nucleic acids and the second target nucleic acids comprise any one of the genetic markers provided herein. In some cases, the first target nucleic acids are associated with a carrier microbe or an AMR cassette; in some cases, the second target nucleic acids are AMR markers. In some embodiments, the first target nucleic acids and the second target nucleic acids comprise a genetic marker associated with an infection or a non-communicable disease or disorder. In some embodiments, the first target nucleic acids and the second target nucleic acids comprise the same genetic marker. In some embodiments, the first target nucleic acids and the second target nucleic acids comprise different genetic markers.

HTS to Detect Carrier Microbes

[0054] In some embodiments, the methods comprise detecting a microbe using an HTS assay followed by performing a targeted sequencing assay. In some embodiments, an unbiased HTS is not performed prior to performing the targeted sequencing assay. In some embodiments, the methods comprise using unbiased HTS to detect a carrier microbe potentially harboring an AMR genetic marker. In some embodiments, the carrier microbe, but not the AMR marker, is detected by the unbiased HTS. In some embodiments, the methods comprise detecting two or more carrier microbes harboring an AMR genetic marker. In some embodiments, the methods comprise detecting at least three carrier microbes. In some embodiments, the methods comprise detecting at least 5, least 6, least 7, least 8, least 9, at least 10, least 15, at least 20, at least 25, at least 30, at least 35, at least 40, at least 45, at least 50, at least 55, at least 60, at least 65, at least 70, at least 75, at least 80, at least 85, at least 90, at least 95, or at least 100 carrier microbes.

[0055] In some embodiments, the methods comprise detecting nucleic acids in the sequencing assay. In some embodiments, the methods comprise detecting microbial nucleic acids. In some

embodiments, the methods comprise detecting microbial cell-free nucleic acids (mcfNA). In some embodiments, the methods comprise detecting mcfNA from one or more carrier microbes harboring an AMR genetic marker. In some embodiments, the methods comprise detecting mcfNA associated with or comprising an AMR genetic marker. In some embodiments, the methods comprise performing unbiased HTS to detect an AMR cassette (e.g., SCCmec cassette).

[0056] In some embodiments, the methods comprise generating nucleic acids sequence reads using the sequencing assay. In some embodiments, the methods comprise generating microbial nucleic acids sequence reads. In some embodiments, the methods comprise generating mcfNA sequence reads. In some embodiments, the methods comprise generating mcfNA sequence reads associated with one or more carrier microbes harboring an AMR genetic marker. In some embodiments, the methods comprise generating mcfNA sequence reads associated with or comprising an AMR genetic marker. In some embodiments, the methods comprise generating sequence reads associated with one or more housekeeping genes. In some embodiments, the methods comprise generating microbial sequence reads associated with one or more microbial-specific housekeeping genes. In some embodiments, the methods further comprise mapping nucleic acid sequence reads from the sequencing assay to a reference genome. In some embodiments, the methods further comprise mapping microbial nucleic acid sequence reads to one or more microbial genomes. In some embodiments, the methods further comprise mapping microbial nucleic acid sequence reads to an AMR reference sequence or AMR cassette reference sequence.

Abundance

[0057] In some embodiments, the methods further comprise quantifying the nucleic acids detected in a sequencing assay, e.g., in an unbiased HTS or in a targeted sequencing assay. In some embodiments, quantifying the nucleic acids comprises calculating an abundance of the mcfNA detected in the sequencing assay. In some embodiments, the abundance of the mcfNA is an absolute quantity expressed in molecules per microliter (MPM). In some embodiments, the abundance of the mcfNA is an absolute quantity expressed in molecules per 100 nanoliters. In some embodiments, the methods further comprise calculating a copy number. In some embodiments, quantifying the nucleic acids comprises calculating an abundance of unique sequencing reads. In some embodiments, calculating an abundance of unique sequencing reads comprises estimating a copy number. In some embodiments, the abundance of the unique sequencing reads is a relative quantity expressed in estimated deduplicated reads (EDR). As used herein “estimated deduplicated reads” or “EDR” generally refers to the number of sequence reads after removing duplicate reads and then multiplying the total deduplicated reads by the estimated fraction of reads generated from each microbe.

[0058] In some embodiments, EDR is calculated per microbe. In some embodiments, calculating an EDR comprises determining an estimated fraction of reads. An estimated fraction of reads may be the calculated proportion of total sequencing reads that fall within a region of interest. The region of interest may be a genomic region. In some embodiments, the EDR is calculated for one or more microbes. In some embodiments, an estimated fraction of reads may be calculated by dividing the number of reads aligned to the region of interest by the total number of reads in the sequence data. In some embodiments, the EDR comprises identifying duplicates. In some embodiments, the EDR per microbe is calculated by removing duplicate reads and then multiplying the total deduplicated reads by the estimated fraction of reads generated from a microbe.

[0059] In some embodiments, quantifying the nucleic acids comprises calculating an abundance of one or more spike-in nucleic acids. In some embodiments, the spike-in nucleic acids are synthetic oligonucleotides directly added to the initial sample. In some embodiments, the synthetic spike-in nucleic acids comprise whole assay internal control (WINC). In some embodiments, the read depth of the region of interest may be normalized. Examples of normalization includes GC content normalization, mappability normalization, control genome normalization, Poisson latent factor models, and tangent normalization. In some embodiments, the copy number calculation comprises comparing the read depth of a genomic region of interest to a standard. A genomic region may be a window of the genome within a given range of continuous nucleotides between a start point and an end point. The range may be 100 bp, 200 bp, 300, bp, 400 bp, 500 bp, 600 bp, 700 bp, 800 bp, 900 bp, 1 kbp,10 kbp, 20 kbp, 30 kbp, 40 kbp, 50 kbp, 60 kbp, 70 kbp, 80 kbp, 90 kbp, 100 kbp, 1 mbp, 2 mbp, 3 mbp, 4 mbp, 5 mbp, or any combination thereof. In some embodiments, the standard has a known copy number. In some embodiments, the standard is the synthetic spike-in nucleic acids. In some embodiments, the standard is a selected region of the genome. In some embodiments, the genomic region of interest may be compared to the standard. Examples of comparisons include read depth (coverage) analysis, ration of normalized read counts, b-allele frequency analysis, control-FREEC (free copy number estimation from coverage), read depth per probe, hidden Markov model analysis, or sequencing depth windowing.

Detection of AMR Cassettes and Markers

[0060] In some embodiments, the methods comprise detecting a gene cassette (such as a gene cassette described herein) using a sequencing assay. In some cases, a gene cassette (e.g., AMR cassette) is detected by an unbiased HTS; in some cases, the gene cassette is detected by a targeted sequencing assay. In some embodiments, the methods comprise detecting at least one gene cassette. In some embodiments, the methods detect two or more gene cassettes. In some

embodiments, the methods detect at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 15, at least 20, at least 25, at least 30, at least 35, at least 40, at least 45, at least 50, at least 60, at least 70, at least 90, at least 100, at least 200, at least 300, at least 400, at least 500, at least 600, at least 700, at least 800, at least 900, or at least 1000 gene cassettes. In some embodiments, the methods comprise detecting a gene cassette harboring at least one AMR genetic marker. In some embodiments, the methods comprise detecting an empty gene cassette that does not harbor any AMR genetic marker. In some embodiments, the methods comprise detecting gene cassettes from one or more carrier microbes harboring at least one AMR genetic marker. In some embodiments, the methods comprise detecting gene cassettes from the same carrier microbe. In some embodiments, the methods comprise detecting an integrated gene cassette. In some embodiments, the methods comprise detecting an episomal gene cassette.

[0061] In some embodiments, the methods described herein comprise detecting an AMR genetic marker using a sequencing assay. In some embodiments, the methods comprise detecting two or more AMR genetic markers. In some embodiments, the methods comprise detecting at least three AMR genetic markers. In some embodiments, the methods comprise detecting at least 5, at least 10, at least 20, at least 25, at least 30, at least 35, at least 40, at least 45, at least 50, at least 55, at least 60, at least 65, at least 70, at least 75, at least 80, at least 85, at least 90, at least 95, at least 100, at least 200, at least 300, at least 400, at least 500, at least 600, at least 700, at least 800, at least 900, at least 1000, at least 1100, at least 1200, at least 1300, at least 1400, at least 1500, at least 1600, at least 1700, at least 1800, at least 1900, at least 2000, at least 2100, at least 2200, at least 2300, at least 2400, at least 2500, at least 2600, at least 2700, at least 2800, at least 2900, or at least 3000 AMR genetic markers. In some embodiments, the methods comprise detecting AMR genetic markers from one or more carrier microbes harboring at least one AMR genetic marker. In some embodiments, the methods comprise detecting AMR genetic markers detected are from the same carrier microbe. In some embodiments, the methods comprise detecting an integrated AMR genetic marker. In some embodiments, the methods comprise detecting an episomal AMR genetic marker.

[0062] In some embodiments, the methods provided herein comprise detecting a carrier microbe harboring a first genetic marker in a sample using a first sequencing assay and then detecting a second genetic marker in the same sample using a second sequencing assay. In some embodiments, the first genetic marker and the second genetic marker are associated with the same phenotype.

[0063] In some embodiments, the methods comprise 1) detecting a carrier microbe harboring an AMR genetic marker in a first sequencing assay, wherein the first sequencing assay is performed on a first aliquot of a sample; and then 2) detecting the AMR genetic marker in a

second sequencing assay, wherein the second sequencing assay is performed on a second aliquot of a sample. In some embodiments, the methods comprise 1) detecting an AMR genetic cassette in the first sequencing assay, wherein the first sequencing assay is performed on a first aliquot of a sample; and then 2) detecting the AMR genetic marker in a second sequencing assay; wherein the second sequencing assay is performed on a second aliquot of a sample. In some embodiments, the methods comprise 1) detecting a carrier microbe and not detecting an AMR genetic marker in the first sequencing assay, wherein the first sequencing assay is performed on a first aliquot of a sample; and then 2) detecting the AMR genetic marker in a second sequencing assay; wherein the second sequencing assay is performed on a second aliquot of a sample. In some embodiments, the methods comprise 1) detecting a carrier microbe and not detecting an AMR gene cassette in the first sequencing assay, wherein the first sequencing assay is performed on a first aliquot of a sample; and then 2) detecting the AMR genetic marker in a second sequencing assay; wherein the second sequencing assay is performed on a second aliquot of a sample. In some embodiments, the methods further comprise providing sufficient amount of the second aliquot for the second sequencing assay. In some embodiments, the second aliquot is the entire remaining sample after the first aliquot has been removed from the sample.

[0064] In some embodiments, the methods comprise detecting an AMR genetic marker harbored by a microbe selected from the group consisting of: *Staphylococcus aureus, S. epidermidis, S. lugdunensis, Enterococcus faecalis, E. faecium, Enterobacter cloacae* complex, *Escherichia coli, Klebsiella aerogenes, K. pneumoniae, K. oxytoca, Proteus mirabilis, P. vulgaris, Salmonella bongori, S. enterica, Serratia marcescens, Pseudomonas aeruginosa, Acinetobacter baumannii,* and *A. calcoaceticus* in the second sequencing assay. In some embodiments, the methods comprise performing a second sequencing assay after the first sequencing assay detects a microbe selected from the group consisting of: *Staphylococcus aureus, S. epidermidis, S. lugdunensis, Enterococcus faecalis, E. faecium, Enterobacter cloacae* complex, *Escherichia coli, Klebsiella aerogenes, K. pneumoniae, K. oxytoca, Proteus mirabilis, P. vulgaris, Salmonella bongori, S. enterica, Serratia marcescens, Pseudomonas aeruginosa, Acinetobacter baumannii,* and *A. calcoaceticus*.

Determining whether to Perform a Second Sequencing Assay

[0065] In some embodiments, performing a second sequencing assay (e.g., a targeted sequencing assay) is predicated on the first sequencing assay meeting certain criteria or filters. In some embodiments, satisfaction of the criteria can mean that there will be a greater likelihood of detecting an AMR target in the second sequencing assay or of assigning linkage of the AMR marker to the carrier microbe based on the second assay. In some embodiments, the criteria comprise taxa of the potential carrier microbe detected in the first sequencing assay, abundance

of the potential carrier microbe, presence or amount of confounding microbes, presence or absence of an AMR cassette (e.g., SCCmec), indeterminant results, abundance statistically significant above background, microbial sequencing reads spread uniformly across a genome, abundance above a level that could be explained by sequence misattribution from another high abundance microbe, abundance above background expected in healthy samples, or any combination thereof. In some embodiments, criteria that are met before performing a second sequencing assay comprise (1) presence of a microbe known to harbor an AMR marker higher than a threshold value (e.g., a threshold value provided herein) and AMR cassette presence is "indeterminate"; (2) presence of a microbe known to harbor an AMR marker lower than a threshold value (e.g., a threshold value provided herein) and with confounding microbes present below a threshold value (e.g., a threshold value provided herein), and AMR cassette presence indeterminate; or (3) gram negative microbes greater than a threshold value (e.g., a threshold value provided herein).

[0066] In some embodiments, the methods comprise performing a first sequencing assay on a first aliquot of a sample and based on the results of the first sequencing assay, and determining whether to perform the second sequencing assay on a second aliquot of the sample. In some embodiments, the methods comprise 1) detecting nucleic acids from a potential carrier microbe harboring an AMR genetic marker in a first sequencing assay, wherein the first sequencing assay is performed on a first aliquot of a sample; 2) quantifying the nucleic acids from the potential carrier microbe in the first aliquot; 3) determining the likelihood that the potential carrier microbe comprises the AMR genetic marker; and 4) determining to perform an amplification or a second sequencing assay on a second aliquot of the sample to detect the AMR genetic marker. In some embodiments, determining the likelihood in 3) comprises quantifying microbial sequence reads associated with the potential carrier microbe. In some embodiments, determining the likelihood in 3) comprises quantifying sequence reads associated with the AMR genetic marker. In some embodiments, determining the likelihood in 3) comprises comparing the sequence reads associated with the AMR genetic marker to the microbial sequence reads associated with the potential carrier microbe. In some embodiments, determining the likelihood in 3) comprises using a statistical model.

[0067] In some embodiments, detecting an abundance of a microbe in a first sequencing assay that exceeds an abundance threshold value determines whether a second sequencing assay is performed. In some embodiments, the threshold abundance value is at least about 50 MPM, at least about 100 MPM, at least about 150 MPM, at least about 200 MPM, at least about 300 MPM, at least about 400 MPM, at least about 500 MPM, at least about 600 MPM, at least about 700 MPM, at least about 800 MPM, at least about 900 MPM, at least about 1000 MPM, at least about

1100 MPM, at least about 1200 MPM, at least about 1300 MPM, at least about 1400 MPM, at least about 1500 MPM, at least about 1600 MPM, at least about 1700 MPM, at least about 1800 MPM, at least about 1900 MPM, at least about 2000 MPM, at least about 2500 MPM, at least about 3000 MPM, at least about 3500 MPM, at least about 4000 MPM, at least about 4500 MPM, at least about 5000 MPM, at least about 5500 MPM, at least about 6000 MPM, at least about 6500 MPM, at least about 7000 MPM, at least about 7500 MPM, at least about 8000 MPM, at least about 8500 MPM, at least about 9000 MPM, at least about 9500 MPM, or at least about 10,000 MPM.

[0068] In some embodiments, the threshold abundance value for gram-positive bacteria is at least about 50 MPM, at least about 100 MPM, at least about 150 MPM, at least about 200 MPM, at least about 300 MPM, at least about 400 MPM, at least about 500 MPM, at least about 600 MPM, at least about 700 MPM, at least about 800 MPM, at least about 900 MPM, at least about 1000 MPM, at least about 1100 MPM, at least about 1200 MPM, at least about 1300 MPM, at least about 1400 MPM, at least about 1500 MPM, at least about 1600 MPM, at least about 1700 MPM, at least about 1800 MPM, at least about 1900 MPM, at least about 2000 MPM, at least about 2500 MPM, at least about 3000 MPM, at least about 3500 MPM, at least about 4000 MPM, at least about 4500 MPM, at least about 5000 MPM, at least about 5500 MPM, at least about 6000 MPM, at least about 6500 MPM, at least about 7000 MPM, at least about 7500 MPM, at least about 8000 MPM, at least about 8500 MPM, at least about 9000 MPM, at least about 9500 MPM, or at least about 10,000 MPM.

[0069] In some embodiments, the abundance threshold value for gram-negative bacteria is at least about 4500 MPM, at least about 5000 MPM, at least about 5500 MPM, at least about 6000 MPM, at least about 6500 MPM, at least about 7000 MPM, at least about 7500 MPM, at least about 8000 MPM, at least about 8500 MPM, at least about 9000 MPM, at least about 9500 MPM, at least about 10,000 MPM, at least about 15,000 MPM, at least about 20,000 MPM, at least about 25,000 MPM, at least about 30,000 MPM, at least about 35,000 MPM, at least about 40,000 MPM, at least about 45,000 MPM, or at least about 50,000 MPM. In some embodiments, the threshold for a carrier microbe harboring $bla_{\text{CTX-M}}$ or bla_{KPC} is at least about 5000 MPM, at least about 5500 MPM, at least about 6000 MPM, at least about 6500 MPM, at least about 7000 MPM, at least about 7500 MPM, at least about 8000 MPM, at least about 8500 MPM, at least about 9000 MPM, at least about 9500 MPM, at least about 10,000 MPM, at least about 15,000 MPM, or at least about 20,000 MPM.

[0070] In some embodiments, the abundance threshold value is from 50 MPM to 20,000 MPM, from 50 MPM to 10,000 MPM, from 50 MPM to 9,500 MPM, from 50 MPM to 5,000 MPM, from 100 MPM to 10,000 MPM, from 200 MPM to 8,000 MPM, from 300 MPM to 7,000 MPM,

from 400 MPM to 6,000 MPM, from 500 MPM to 5,000 MPM, from 1,000 MPM to 4,000 MPM, from 1,500 MPM to 3,000 MPM, from 2,000 MPM to 2,500 MPM, or from 2,500 MPM to 2,000 MPM.

[0071] In some embodiments, the methods provided herein comprise detecting one or more carrier microbes harboring an AMR genetic marker using a first sequencing assay on a first aliquot of a sample form a subject, and then based on the results from the first sequencing assay, determining to perform an amplification assay or a second sequencing assay on a second aliquot of the same sample. In some embodiments, the methods comprise, after the first sequencing assay on the first aliquot of sample, providing a second aliquot of sample with sufficient amount of cfNA for the amplification assay or the second sequencing assay. In some embodiments, the second aliquot of sample comprises at least 0.05 ng, at least 0.10 ng, at least 0.15 ng, at least 0.20 ng, at least 0.25 ng, at least 0.30 ng, at least 0.35 ng, at least 0.40 ng, at least 0.45 ng, at least 0.50 ng, at least 0.55 ng, at least 0.60 ng, at least 0.65 ng, at least 0.70 ng, at least 0.75 ng, at least 0.80 ng, at least 0.85 ng, at least 0.90 ng, at least 0.95 ng, at least 1.00 ng, at least 1.25 ng, at least 1.50 ng, at least 1.75 ng, at least 2.00 ng, at least 2.25 ng, at least 2.50 ng, at least 2.75 ng, at least 3.00 ng, at least 3.25 ng, at least 3.50 ng, at least 3.75 ng, at least 4.00 ng, at least 4.25 ng,at least 4.50 ng, at least 4.75 ng, or at least 5.00 ng of cfNA.

[0072] In some embodiments, the methods comprise, after the first sequencing assay on the first aliquot of sample, providing a second aliquot of sample with sufficient volume for the amplification assay or the second sequencing assay. In some embodiments, the second aliquot of sample comprises at least 50 µl, at least 100 µl, at least 150 µl, at least 200 µl, at least 250 µl, at least 300 µl, at least 350 µl, at least 400 µl, at least 450 µl, at least 500 µl, at least 550 µl, at least 600 µl, at least 650 µl, at least 700 µl, at least 750 µl, at least 800 µl, at least 850 µl, at least 900 µl, at least 950 µl, or at least 1 ml of the sample. In some embodiments, the second aliquot is at least 500 µl of plasma obtained from a human. In some embodiments, the second aliquot is at least 250 µl of plasma obtained from a human. In some embodiments, the second aliquot is at least 100 µl of plasma obtained from a human.

[0073] In some embodiments, the one or more carrier microbes harboring an AMR genetic marker comprise one or more Gram-positive bacteria. In some embodiments, the one or more carrier microbes harboring an AMR genetic marker comprise one or more Gram-negative bacteria. In some embodiments, the one or more Gram-positive bacteria comprise a Staphylococcus spp. In some embodiments, the one or more Gram-positive bacteria comprise one or more Methicillin-resistance carriers. In some embodiments, the one or more Methicillin-resistance carriers comprise In some embodiments, the methicillin-resistance carrier comprises Staphylococcus pseudintermedius, Staphylococcus fleurettii, Staphylococcus epidermidis,

Staphylococcus schleiferi, Staphylococcus lugdunensis, Staphylococcus cohnii, Staphylococcus caprae, Staphylococcus warneri, Staphylococcus saprophyticus, Staphylococcus aureus, Staphylococcus haemolyticus, Staphylococcus capitis, Staphylococcus hominis, Staphylococcus pettenkoferi, or Staphylococcus simulans. For example, the methods provided herein can comprise detecting *Staphylococcus aureus* (*S. aureus*) in a high throughput sequencing assay performed on an aliquot of a sample from a patient, wherein the aliquot of the sample comprises mcfNA from the *S. aureus*.

[0074] I*n* some embodiments, the methods comprise performing an amplification assay or a second sequencing assay on a second aliquot of the same sample based on the results from the first sequencing assay. For example, if the methods detect a quantity of mcfNA from *S. aureus* below a first threshold (e.g., below 100 MPM, 200 MPM, 300 MPM, 400 MPM, 500 MPM); and 2) detect a quantity of mcfNA from other Staphylococci spp. (a confounding spp.) below a second threshold (e.g., below 1, 5, 10, 15, 20 MPM); the methods can then determine to perform an amplification assay or a second sequencing assay targeting *mecA* and *mecC*. In some embodiments, the methods comprise performing the amplification assay or the second sequencing assay when the results from the first sequencing assay is indeterminate. In some embodiments, the methods comprise determining the result from the first sequencing assay is indeterminate and performing a second sequencing assay or an amplification assay if the SCC*mec* cassette in the sample is likely to be a non-mec SCC, as determined from alignments to alleles known to lack mecA and mecC. In some embodiments, the result from the first sequencing assay is indeterminate if the SCC fragments in the clinical sample have originated from a microbe other than S. aureus, as determined by comparing the abundances of known cross-reacting or interfering species. In some embodiments, the interfering species comprise coagulase-negative *staphylococci* that harbor sequences highly homologous to SCCmec.

[0075] In some cases, if the methods do not comprise detecting an abundance of sequencing reads from *S. aureus* over a threshold provided herein, the amplification assay or the second sequencing assay targeting *mecA* and *mecC* will be performed. In some cases, if the methods do not comprise detecting a carrier microbe harboring $bla_{\text{CTX-M}}$ or bla_{KPC} above the threshold, the amplification assay or the second sequencing assay targeting $bla_{\text{CTX-M}}$ or bla_{KPC} will be performed. In some cases, if the methods do not comprise detecting a carrier microbe harboring *vanA or van B* above the threshold, the amplification assay or the second sequencing assay targeting *vanA or van B* will be performed.

[0076] In some cases, the amplification assay or the second sequencing assay will only be performed if sufficient amount of a second aliquot of the sample is provided. In some cases, the amplification assay or the second sequencing assay will only be performed if detect a low

quantity of mcfNA from a confounding species is detected. In some embodiments, the amplification assay or the second sequencing assay will only be performed if the first sequencing assay detects a quantity of mcfNA from a confounding species below a threshold. In some embodiments, the threshold of mcfNA from a confounding species comprises at most 200 MPM, at most 180 MPM, at most 160 MPM, at most 140 MPM, at most 120 MPM, at most 100 MPM, at most 80 MPM, at most 60 MPM, at most 40 MPM, at most 20 MPM, at most 10 MPM, at most 8 MPM, at most 6 MPM, at most 4 MPM, or at most 2 MPM, or at most 1 MPM.

[0077] In some embodiments, the methods comprise detecting one or more Gram-negative bacteria using the first sequencing assay on the first aliquot of a sample form a subject, wherein the one or more Gram-negative bacteria are potential carrier microbes of an AMR genetic marker, and then based on the results from the first sequencing assay, determining to perform an amplification assay or a second sequencing assay on a second aliquot of the same sample. In some embodiments, the one or more Gram-negative bacteria comprise Enterobacter cloacae complex, Escherichia coli, Klebsiella pneumoniae, Klebsiella oxytoca, Proteus mirabilis, Proteus vulgaris, Salmonella enterica, Salmonella bongori, Serratia marcescens, Enterobacter aerogenes, Pseudomonas aeruginosa, Acinetobacter baumannii, Acinetobacter calcoaceticus, or a combination thereof. In some embodiments, the amplification assay or the second sequencing assay is performed only when the methods detect a quantity of mcfNA from the one or more Gram-negative bacteria passes a threshold provided herein.

[0078] In some embodiments, the methods comprise discounting a result from a step of the multi-step process if there are multiple contaminants detected by the step. In some embodiments, the methods comprise discounting a result from a sequencing assay if there are multiple contaminants detected by the sequencing assay. In some embodiments, contaminants comprise contaminating nucleic acids or microbes. In some embodiments, contaminating nucleic acids arise from sample collection, sample site (e.g., skin), or reagents (e.g., buffers, buffer components, water, beads, etc.).

Detection of MRSA based on detection of SCC*mec* cassette and abundance of *S. aureus reads*

[0079] In some embodiments, the methods provided herein comprise using abundance of *S. aureus* combined with detection of a SCCmec cassette in order to identify whether the *S. aureus* harbors an AMR genetic marker. In some cases, a number of reads mapping to SCCmec is compared to the number of reads mapping to the *S. aureus* genome and a statistical model is used to determine the likelihood of either the presence or absence of methicillin resistance. In some embodiments, the methods further comprise quantifying sequencing reads from the high throughput sequencing assay, wherein the sequencing reads comprise *S. aureus* mcfNA

sequencing reads. The abundance of the *S. aureus* mcfNA in the sample can be calculated based on the abundance of the *S. aureus* mcfNA sequencing reads.

[0080] In some embodiments, the methods comprise detecting SCC*mec* cassette from the sample and quantifying microbial sequencing reads from the SCC*mec* cassette. In some embodiments the SCC*mec* cassette is detected by unbiased HTS; in some embodiments, it is detected by targeted sequencing. The SCC*mec* cassette is an AMR gene cassette of a known size found in *S. aureus*. The SCC*mec* cassette in certain strains of *S. aureus* can harbor *mecA* or *mecC*, which are AMR genes associated with resistance to methicillin. In some embodiments, the *SCCmec* cassette can be empty and does not comprise the AMR genetic marker of interest. In some embodiments, the methods comprise detecting a non-empty *SCCmec* cassette.

[0081] The methods provided herein can further comprise determining the likelihood that *mecA* or *mecC* is present in the *S. aureus* in the patient based on the amount of microbial sequencing reads mapped to *S. aureus* and the amount of microbial sequencing reads from the SCC*mec* cassette. In some embodiments, the SCC*mec* cassette in *S. aureus* is an integrated region of the genome, the sequencing read count from *S. aureus* and the sequencing read count from the SCC*mec* cassette can be reliably correlated. For example, when there are more reads (e.g., above the limit of detection) an estimated deduplicated reads (EDR) of >1000, 2000, 3000, 4000 or 5000 can allow the methods to determine whether the SCC*mec* cassette likely harbors *mecA* or *mecC* by analyzing the *S. aureus* sequencing reads. The methods can compare the SCC*mec* cassette sequencing reads with the *S. aureus* genome sequencing reads to determine this likelihood. In some embodiments, if there is a high likelihood that the SCC*mec* cassette harbors *mecA* or *mecC*, the method will determine that the *S. aureus* is methicillin-resistant *S. aureus* (MRSA). In some embodiments, if there is a high likelihood that the SCC*mec* cassette is empty and does not harbor *mecA* or *mecC*, the method will determine that the *S. aureus* is methicillin-susceptible *S. aureus* (MSSA).

[0082] In some embodiments, a beta-binomial distribution can be used to determine the probability of getting X SCC*mec* reads given Y *S. aureus* reads. In some cases, a distribution comprises a distribution of ratio of SCC*mec* length to *S. aureus* genome length for MRSA. In some cases, a similar model can be built for MSSA. In some cases, Bayesian interference can be used to calculate probability of MRSA given the observed *S. aureus* reads and the observed number of SCC*mec* reads.

Detecting AMR Markers using a Targeted Sequencing Assay

[0083] In some embodiments, the methods comprise detecting a nucleic acid (e.g., AMR genetic markers) using a targeted sequencing assay. In some embodiments, the targeted sequencing comprises performing an amplification reaction on a sample using primers specific for target

sequences such as target AMR sequences. In some embodiments, amplicons generated from the amplification reaction can be analyzed in order to determine the abundance of a target sequence (e.g., AMR genetic marker). In some cases, an HTS assay is performed on the amplicons in order to determine the abundance of the target nucleic acid.

[0084] In some embodiments, the methods comprise calculating an estimated deduplicated templates (EDT) from a sequencing assay, e.g., a targeted sequencing assay. As used herein "estimated deduplicated templates" or "EDT" refers to an estimate of the number of original template molecules present prior to amplification. EDT can be calculated by leveraging the primer plus read length or primer plus unique molecular identifier tags to deduplicate PCR amplicons generated from non-amplified nucleic acids.

[0085] In some embodiments, linkage of an AMR marker with a carrier microbe is based on analyzing an EDT value of an AMR marker and the abundance of a carrier microbe detected by HTS. In some embodiments, linkage of an AMR marker with a carrier microbe is based on analyzing an EDT value of an AMR marker and an EDT value of a microbe-specific gene, such as a microbe-specific housekeeping gene.

[0086] An EDT may comprise a calculation of the number of unique template molecules from the amplicon alignments for each targeted marker and from the unique molecular identifiers for the quality control oligonucleotides. An EDT may use a method of deduplicating PCR amplicons. Deduplication of a PCR amplicon may comprise incorporation of unique molecular identifiers. Deduplication of a PCR amplicon may comprise the use of primer+read length or primer+unique molecular identifier tags to deduplicate PCR amplicons. Deduplication of a PCR amplicon may comprise the use of software such as picard or SAMtools to aid in identification of duplicates and removal of duplicates. EDT may be used to calculate the number of original template molecules vs. deduplicating randomly sheared HTS reads.

Detecting Microbe-specific Genes using a Targeted Sequencing Assay

[0087] In some embodiments, the methods comprise introducing primers targeting housekeeping genes, amplifying the housekeeping genes to produce amplicons and quantifying the housekeeping genes (e.g., by conducting HTS on the amplicons). In some embodiments, the housekeeping genes are microbe-specific in that identification of the housekeeping gene also enables identification of the corresponding microbe. Housekeeping genes may be chosen from species-specific core protein coding genes and filtered to those where the nucleotide sequence would not cross-react with genomes of other pathogens. In some embodiments, the housekeeping genes target *S. aureus, S. epidermidis*, and E. *faecium*. The targeted housekeeping genes may be used to determine the presence or absence of the organism they are specific to one or more carrier

microbes. For example, a housekeeping gene targeted to *S. aureus* may be used in the determination of whether nucleic acids derived from *S. aureus* are present in a sample.

[0088] In some embodiments, the methods comprise calculating an estimated deduplicated templates (EDT) corresponding to housekeeping genes detected in a sequencing assay, e.g., a targeted sequencing assay. In some embodiments, an EDT value of a housekeeping gene is determined. In some embodiments, linkage of an AMR marker with a carrier microbe is based on analyzing an EDT value of an AMR marker and an EDT value of a microbe-specific gene, such as a microbe-specific housekeeping gene.

[0089] An EDT may comprise a calculation of the number of unique template molecules from the amplicon alignments for each targeted marker and from the unique molecular identifiers for the quality control oligonucleotides. An EDT may use a method of deduplicating PCR amplicons. Deduplication of a PCR amplicon may comprise incorporation of unique molecular identifiers. Deduplication of a PCR amplicon may comprise the use of primer+read length or primer+unique molecular identifier tags to deduplicate PCR amplicons. Deduplication of a PCR amplicon may comprise the use of software such as picard or SAMtools to aid in identification of duplicates and removal of duplicates. EDT may be used to calculate the number of original template molecules vs. deduplicating randomly sheared HTS reads.

Linking an AMR marker with a carrier microbe via unbiased HTS and a targeted sequencing assay

[0090] The methods provided herein can be especially useful for linking an AMR marker associated with a carrier microbe. In some embodiments, targeted sequencing of an AMR marker is based on the sequence of the AMR marker itself (e.g., primers directed to the AMR marker), without giving information about its associated carrier microbe. In some embodiments, linking a carrier microbe with an AMR genetic marker is especially useful for samples in which nucleic acids from multiple carrier microbes are detected, nucleic acids from multiple AMR markers, or a combination thereof. In such cases, it can be challenging to determine which carrier microbe is linked to which AMR genetic marker. In some cases, when multiple species of carrier microbes are detected, the methods provided herein identify which species of the carrier microbe is associated with a particular AMR marker.

[0091] In some embodiments, the methods provided herein comprise linking an AMR genetic marker with its carrier microbe via an unbiased high throughput sequencing (HTS) assay and a targeted sequencing assay. In some embodiments, the methods comprise 1) detecting one or more potential carrier microbes harboring an AMR genetic marker via an unbiased HTS assay; 2) detecting the AMR genetic marker via a targeted sequencing assay; and 3) performing statistical analysis to link the AMR genetic marker with its carrier microbe. In some embodiments, the

methods comprise detecting an episomal AMR genetic marker in the targeted sequencing assay. In some embodiments, performing the statistical analysis comprises calculating an abundance of mcfNA from the one or more potential carrier microbes detected in the unbiased HTS assay. In some embodiments, performing the statistical analysis comprises calculating an abundance of mcfNA associated with the AMR genetic marker.

[0092] In some embodiments, the methods comprise linking the AMR genetic marker with its carrier microbe based on the abundance of mcfNA from the one or more potential carrier microbes, the abundance of mcfNA associated with the AMR genetic marker, the copy number of the AMR genetic marker, the abundances of microbial-specific housekeeping genes detected in the targeted sequencing assay, or any combination thereof. In some embodiments, the methods comprise linking the AMR genetic marker with its carrier microbe based on the abundance of mcfNA from the one or more potential carrier microbes and the copy number of the AMR genetic marker. In some embodiments, the methods comprise linking the AMR genetic marker with its carrier microbe based on results from the unbiased HTS assay and results from the targeted sequencing assay.

[0093] In some embodiments, the methods provided herein comprise performing a statistical analysis to link one or more AMR genetic markers to one or more carrier microbes harboring the AMR genetic markers. In some embodiments, the methods comprise linking one AMR genetic marker to one carrier microbe. In some embodiments, the methods comprise linking multiple AMR genetic markers to multiple carrier microbes; for example, the methods can comprise linking one or more carrier microbes to its corresponding AMR marker. In some cases, the statistical analysis comprises a generalized linear model. In some cases, the statistical analysis comprises a maximum-likelihood estimation. In some cases, the statistical analysis comprises a probit model. In some cases, the statistical analysis comprises a maximum-likelihood probit model. Examples of other models that may be used in the statistical analysis include logistic regression, linear probability, linear regression, complimentary log-log, Poisson regression, support vector machine, decision tree, random forest, neural network, gradient boosted model, Bayesian models, hidden Markov models, or any combination thereof.

[0094] In some embodiments, to link the HTS data to the AMR genetic marker detected in the targeted sequencing assay, the methods comprise using a model that uses the observed abundance of the carrier microbes (e.g., MPM), copy number of AMR gene, ratio of AMR gene length to microbe genome length, a measure of PCR reaction efficiency from the abundance of an internal control (PCR WINC), or any combination thereof. With those known variables, in some embodiments, a model is prepared that gives an expected AMR abundance, such as estimated

deduplicated templates or transcripts (EDT)). In some cases, copy number is calculated and used in the statistical analysis. In some cases, copy number is not used in the statistical analysis.

[0095] In some embodiments, the methods comprise using a model (e.g., a model described herein) and/or probabilities for AMR prevalence. In some embodiments, the model or probabilities for AMR prevalence are derived from empirical observations for combinations of observed carrier microbe and AMR gene copy number. In some embodiments, the model or probabilities for AMR prevalence are derived from known observations reported in scientific literature for combinations of observed carrier microbe and AMR gene copy number. In some embodiments, a combination of observed carrier microbe and AMR gene copy number is used to determine the likelihood of this combination given the observed AMR EDT. In some embodiments, thresholds to this final likelihood for microbe-AMR combination are applied in order to determine presence, indeterminate, or absence.

[0096] In some embodiments, the methods comprise performing a statistical analysis to link an AMR genetic marker with a carrier microbe based on detecting or quantifying an AMR marker that is integrated into a chromosome or based on quantifying an AMR marker that is an episomal marker. In some embodiments, such integrated or episomal AMR markers are detected by an unbiased HTS assay or a targeted sequencing assay. In some embodiments, the methods comprise performing a statistical analysis to link an AMR genetic marker with a carrier microbe based on detecting an abundance of an AMR cassette. In some embodiments, the methods comprise detecting an integrated AMR marker (e.g., SCC*mec*) in a first assay (e.g., unbiased HTS assay) and then detecting a set of integrated or episomal AMR marker in a second assay (e.g., targeted sequencing assay). In some cases, at least 1, 2, 3, 4, 5, 10, 15, or 20 integrated or episomal AMR markers are detected in the targeted sequencing assay. In some embodiments, the methods comprise linking a full set, or a portion of a set, of such AMR markers with carrier microbes detected in a different assay (e.g., unbiased HTS assay).

[0097] In some embodiments, a first assay (e.g., unbiased HTS assay) detects larger genomic targets, such as SCCmec. If the larger genomic targets are known to be integrated in the microbial genome, the methods comprise identifying the linkage to a carrier microbe using a statistical model or without a statistical model. In some embodiments, the methods comprise detecting genes or portions of plasmids by an assay (e.g., unbiased HTS assay) and then using a statistical model to assign linkage without necessarily performing a targeted sequencing assay.

[0098] In some embodiments, the methods comprise detecting an integrated AMR genetic marker that has a copy number that tracks the copy number of a microbial genome; in some cases, the AMR genetic marker is integrated and has one copy per microbial genome. In some cases, multiple AMR genetic markers are integrated into the same microbial genome. In some

embodiments, the methods comprise detecting an episomal AMR genetic marker. In some cases, the copy number of the episomal AMR genetic marker dose not present in a one-to-one ratio with the microbial genome (e.g., microbial DNA, RNA), microbial gene, or microbial chromosome. In some embodiments, the methods detect an integrated AMR cassette that may comprise the AMR genetic marker. In some embodiments, the methods detect an episomal AMR cassette. In some embodiments, the methods detect a non-empty AMR cassette. In some embodiments, the methods detect the integrated AMR genetic marker or the integrated AMR cassette in the unbiased HTS assay. In some embodiments, the methods detect the episomal AMR genetic marker or the episomal AMR cassette in the unbiased HTS assay. In some embodiments, the methods detect the episomal AMR genetic marker or the episomal AMR cassette in the targeted HTS assay. In some embodiments, the methods detect the integrated AMR genetic marker or the integrated AMR cassette in the targeted HTS assay. In some embodiments, the methods can detect the AMR genetic marker or the carrier microbes by high throughput sequencing, next generation sequencing (NGS), massively parallel sequencing, or sequencing by synthesis. In some embodiments, the methods detect the AMR genetic marker by an unbiased sequencing assay. In some embodiments, the methods detect the AMR genetic marker by a targeted sequencing assay.

[0099] The general framework of these methods has potential applications for general biomarkers of disease. In some embodiments, a first assay can detect whether disease state is present or absent and a second assay can define the specific disease state (e.g., inflammation) or site of disease (e.g., a site of infection).

Linking an AMR marker with a carrier microbe using a targeted sequencing assay

[00100] In some embodiments, the methods provided herein comprise linking an AMR genetic marker with its carrier microbe via a targeted sequencing assay that detects both the carrier microbe as well as the AMR genetic marker. In some embodiments, the methods comprise 1) detecting the AMR genetic marker via a targeted sequencing assay; 2) detecting one or more microbial-specific genes such as microbe-specific housekeeping genes via the targeted assay; and 3) performing statistical analysis to link the AMR genetic marker with its carrier microbe. In some embodiments, the AMR marker detected by the targeted sequencing assay is an episomal AMR genetic marker. In some embodiments, the AMR marker detected by the targeted sequencing assay is an integrated AMR genetic marker. In some embodiments, performing the statistical analysis comprises calculating a copy number of the AMR genetic marker or calculating abundances of the microbial-specific housekeeping genes detected in the targeted sequencing assay. In some embodiments, a statistical model for linking a carrier microbe with an AMR genetic marker using a targeted assay is similar to a statistical model used for linking a

carrier microbe detected by HTS with an AMR marker detected by a targeted assay. Some examples of such models are provided herein. In some embodiments, housekeeping gene abundance (EDT) and AMR-to-housekeeping-gene ratio are used in a statistical model. And in some embodiments, the MPM value in a statistical model described herein can be substituted with the EDT value when using the targeted sequencing assay to detect a carrier microbe and its associated AMR marker.

[00101] In some embodiments, the method comprises detecting microbe-specific housekeeping genes by a targeted sequencing assay. In some cases, the microbe-specific housekeeping genes are detected by HTS or unbiased HTS. In some embodiments, the microbe-specific housekeeping genes are specific for different species of a microbe. In some embodiments, the microbe-specific housekeeping genes are specific for different species of Staphylococcus, e.g., *S. aureus* and/or *S. epidermidis*. In some embodiments, the microbe-specific genes comprise A-SA, B-SA, C-SA, D-SA or any combination thereof. In some embodiments, the microbe-specific genes comprise a housekeeping (HK) gene listed in Supplemental Table 6 in the Examples.

[00102] In some embodiments, the methods comprise detecting one or more microbial-specific housekeeping genes in the targeted sequencing assay in order to detect potential carrier microbes. In some embodiments, the methods comprise detecting a carrier microbe based on the detection of the one or more microbial-specific housekeeping genes in the targeted assay. In some embodiments, the methods comprise performing a statistical analysis in order to link a carrier microbe (detected by a housekeeping gene) with an AMR genetic marker.

[00103] In some embodiments, performing the statistical analysis comprises calculating an abundance of the AMR genetic marker, a copy number of the AMR genetic marker, and/or abundances of the microbial-specific housekeeping genes detected in the targeted sequencing assay. In some embodiments, the abundance of the AMR genetic marker or the microbial-specific housekeeping genes comprises a relative abundance, such as an estimated deduplicated template or transcript (EDT). In some embodiments, performing the statistical analysis comprises calculating a ratio of the EDT of AMR genetic marker to the EDT of the microbial-specific housekeeping genes. In some embodiments, performing the statistical analysis comprises detecting a carrier microbe based on the abundances of the one or more microbial-specific housekeeping genes. In some embodiments, the methods comprise linking the AMR genetic marker with its carrier microbe based on the abundance of mcfNA associated with the AMR genetic marker, the copy number of the AMR genetic marker, the abundances of the microbial-specific housekeeping genes detected in the targeted sequencing assay, the ratio of the AMR EDT to housekeeping gene EDT, or any combination thereof. In some embodiments, the methods comprise linking the AMR genetic marker with its carrier microbe based on the copy number of

the AMR genetic marker and the abundances of the microbial-specific housekeeping genes detected in the targeted assay. In some embodiments, the methods comprise linking the AMR genetic marker with its carrier microbe based only on results from the targeted sequencing assay.

[00104] In some embodiments, a determination of "indeterminate" may be arrived at if the ratio of AMR EDT to housekeeping gene EDT for a carrier microbe departs from the expected ratio. For example, a lack of adequate housekeeping gene EDT would lower confidence in an AMR absence call, making the call indeterminate.

Sample Preparation

[00105] In some embodiments, the methods further comprise preparing the sample using any of the sample preparation methods. In some embodiments, the methods further comprise preparing the nucleic acids in the sample for a sequencing assay. In some embodiments, the methods further comprise preparing the nucleic acids in the sample before a sequencing assay. In some embodiments, the methods further comprise preparing the nucleic acids in the sample after a first sequencing assay but before a second sequencing assay. In some embodiments, the methods further comprise performing a sequencing assay without preparing the nucleic acids. In some embodiments, preparing the nucleic acids in the sample comprises extracting, physically enriching, concentrating, amplifying the nucleic acids, or attaching one or more adapter to the nucleic acids. In some embodiments, amplifying the nucleic acids comprises performing a polymerase chain reaction (PCR). In some embodiments, amplifying the nucleic acids comprises attaching one or more adapter to the nucleic acids. In some embodiments, attaching one or more adapter to the nucleic acids comprises performing a ligase reaction or a PCR.

[00106] In some embodiments, the methods provided herein comprise performing a sequencing assay on nucleic acids in a sample that have not been extracted from the sample before the sequencing assay. In some embodiments, the methods provided herein comprise performing a sequencing assay on nucleic acids in a sample that have been extracted from the sample before the sequencing assay. In some embodiments, the methods comprise 1) performing a firsthigh throughput sequencing assay on nucleic acids in a first aliquot of a sample, wherein the nucleic acids have not been extracted from the first aliquot; and 2) performing a second high throughput sequencing assay on nucleic acids in a second aliquot of the sample, wherein the nucleic acids have been extracted from the second aliquot. In some embodiments, the methods comprise 1) performing a first high throughput sequencing assay on nucleic acids in a first aliquot of a sample, wherein the nucleic acids have been extracted from the first aliquot; and 2) performing a second high throughput sequencing assay on nucleic acids in a second aliquot of the sample, wherein the nucleic acids have been extracted from the second aliquot. In some embodiments, the methods comprise 1) performing a first high throughput sequencing assay on nucleic acids in a first aliquot

of a sample, wherein the nucleic acids have not been extracted from the first aliquot; and 2) performing a second high throughput sequencing assay on nucleic acids in a second aliquot of the sample, wherein the nucleic acids have not been extracted from the second aliquot. In some embodiments, the methods comprise 1) performing a first high throughput sequencing assay on nucleic acids in a first aliquot of a sample, wherein the nucleic acids have been extracted from the first aliquot; and 2) performing a second high throughput sequencing assay on nucleic acids in a second aliquot of the sample, wherein the nucleic acids have not been extracted from the second aliquot. In some embodiments, the first high throughput sequencing assay is not a targeted sequencing assay. In some embodiments, the second high throughput sequencing assay is a targeted sequencing assay.

[00107] In some embodiments, the methods further comprise extracting nucleic acids from a nucleic acid library before the sequencing assay.

[00108] In some embodiments, the methods further comprise extracting nucleic acids from a sample before the sequencing assay. In some embodiments, the methods provided herein comprise 1) performing a first sequencing assay on nucleic acids in a first aliquot of a sample, wherein the nucleic acids in the first aliquot have not been extracted; 2) extracting nucleic acids from a second aliquot of the sample; and after 2), 3) performing a second sequencing assay on the nucleic acids extracted from the second aliquot. In some embodiments, the methods provided herein comprise 1) performing a first sequencing assay on nucleic acids in a first aliquot of a sample, wherein the nucleic acids in the first aliquot have been extracted; 2) extracting nucleic acids from a second aliquot of the sample; and after 2), 3) performing a second sequencing assay on the nucleic acids extracted from the second aliquot. In some embodiments, the methods provided herein comprise 1) extracting nucleic acids from a first aliquot of a sample; 2) performing a first sequencing assay on the nucleic acids extracted from the first aliquot; and after 2), 3) performing a second sequencing assay on nucleic acids from a second aliquot, wherein the nucleic acids from the second aliquot have not been extracted. In some embodiments, the methods provided herein comprise 1) extracting nucleic acids from a first aliquot of a sample; 2) performing a first sequencing assay on the nucleic acids extracted from the first aliquot; and after 2), 3) performing a second sequencing assay on nucleic acids from a second aliquot, wherein the nucleic acids from the second aliquot have been extracted.

[00109] In some embodiments, the methods provided herein comprise 1) performing a first high throughput sequencing assay on nucleic acids from a first aliquot of a sample in order to detect one or more carrier microbes harboring an AMR genetic marker, wherein the nucleic acids form the first aliquot have not been extracted and the nucleic acids comprise microbial nucleic acids from the one or more carrier microbes; 2) after detecting at least one carrier microbe, extracting

nucleic acids from a second aliquot of the sample, thereby producing a concentrated sample comprising the microbial nucleic acids from the carrier microbe; and 3) performing a second high throughput sequencing assay on the nucleic acids from the concentrated sample to detect the AMR marker. In some embodiments, the first high throughput sequencing assay is a non-biased sequencing and not a targeted sequencing assay. In some embodiments, the second high throughput sequencing assay is a targeted sequencing assay.

[00110] In some embodiments, the methods provided herein further comprise attaching one or more adapters to the nucleic acids in an aliquot of the sample before performing a sequencing assay. In some embodiments, the methods comprise attaching the one or more adapters to the nucleic acids that have been extracted from the sample. In some embodiments, the methods comprise attaching the one or more adapters to the nucleic acids that have not been extracted from the sample. In some embodiments, the methods comprise adding the one or more adapters directly to an aliquot of the initial sample. In some embodiments, the one or more adapters are attached to the nucleic acids by a ligase. In some embodiments, the ligase comprises a DNA ligase, a RNA ligase, a DNA/RNA ligase, a splint ligase, a high-fidelity ligase, a thermostable ligase, a truncated ligase, a natural ligase, an engineered ligase, or any combination thereof. In some embodiments, the ligase comprises CircLigase II, CircLigase ssDNA ligase, CircLigase RNA ligase, Thermostable App-DNA/RNA ligase, T4 DNA ligase, T4 RNA ligase 1, T4 RNA Ligase 2, T4 RNA Ligase 2 truncated, Splint ligase, Splint-R ligase, Pfu ligase, any engineered variants thereof, any natural variants thereof, or any combination thereof. In some embodiments, the one or more adapters are attached to the nucleic acids by PCR. In some embodiments, attaching the one or more adapters can comprise using one or more primers comprising an adapter sequence. In some embodiments, the primers comprising an adapter sequence further comprise a sequence targeting a genetic marker provided herein for producing adapted nucleic acids comprising the genetic marker. In some embodiments, the primers comprise an adapter sequence and a sequence targeting an AMR genetic marker for producing adapted nucleic acids comprising the AMR genetic marker.

[00111] In some embodiments, the methods further comprise amplifying nucleic acids and producing amplicons associated with the nucleic acids. In some embodiments, amplifying the nucleic acids comprise amplifying with PCR or a sequencing by synthesis assay. In some embodiments, amplifying nucleic acids comprises generating a nucleic acid library for a sequencing assay provided herein. In some embodiments, the amplified nucleic acids or amplicons are extracted. In some embodiments, amplifying comprises amplifying the nucleic acids comprising the one or more genetic markers provided herein. In some embodiments, amplifying comprises amplifying adapted nucleic acids comprising one or more adapters. In

some embodiments, the methods comprise amplifying nucleic acids comprising the AMR genetic marker to generate amplicons comprising or associated with the AMR genetic marker. In some embodiments, the methods further comprise performing a high throughput sequencing assay on the amplicons comprising the AMR genetic marker to detect the AMR genetic marker. In some embodiments, the methods further comprise extracting the amplicons comprising the AMR genetic marker before the high throughput sequencing assay.

[00112] In some embodiments, the methods comprise performing a first high throughput sequencing assay on nucleic acids that have not been extracted from the sample comprising the nucleic acids. For example, without extracting the nucleic acids from the initial sample, the methods can directly perform a high throughput sequencing assay on the nucleic acids in the initial sample to generate sequence reads, wherein the nucleic acids can comprise microbial nucleic acids from one or more carrier microbe harboring an AMR genetic marker. The first high throughput sequencing assay can detect the microbial nucleic acids from at least one carrier microbe and identify the carrier microbe. After detecting the carrier microbe, the methods further comprise preparing nucleic acid from another portion or aliquot of the same sample. For example, nucleic acids from a second aliquot of the sample can be extracted to provide a concentrated sample comprising the microbial nucleic acids from the carrier microbe. Additionally, the method can further comprise attaching one or more adapters to the nucleic acids in the concentrated sample. The one or more adapters can be attached to the nucleic acids by a ligase, thereby producing adapted nucleic acids. The one or more adapters can also be attached to the nucleic acids by PCR, thereby producing adapted nucleic acids. In some instances, the attaching the one or more adapters can comprise using one or more primers targeting the AMR genetic marker, wherein the primers also contain the adapter sequences, thereby generating adapted nucleic acids comprising the AMR genetic marker. Additionally, the method can further comprise amplifying the adapted nucleic acids by PCR, thereby generating amplicons comprising the AMR genetic marker. In some instances, one or more additional adapters can be attached to the adapted nucleic acids during amplification. The methods can therefore generate adapted amplicons comprising the AMR genetic marker useful for a second sequencing assays to analyze the AMR genetic marker. After this, the methods can further comprise performing a high throughput sequencing assay on the adapted amplicons comprising the AMR genetic marker.

Samples

[00113] Disclosed herein in some embodiments are samples derived from subjects. In some embodiments, a sample provided herein can comprise a nucleic acid molecule to be sequenced

by a method described herein. As used herein, a "sample" generally refers to any material comprising nucleic acids that has been derived from a subject. A sample can comprise a raw biological sample, such as whole blood. As used herein, the phrase "raw biological sample" refers to an unmanipulated or unprocessed sample obtained from a subject, *e.g.,* a host, containing or presumed to contain target nucleic acids. In some embodiments, a raw biological sample has not been subjected to any extraction methods after being obtained from a subject. In some embodiments, a raw biological sample can be processed or manipulated to produce an initial sample. For example, a raw biological sample can comprise whole blood which is centrifuged to produce an initial sample of plasma for a sequencing assay. As used herein, the term "initial sample" refers to a sample comprising nucleic acids derived from a raw biological sample. In some embodiments, an initial sample can comprise a sample that has been processed or manipulated, such as plasma or serum. In some embodiments, an initial sample can comprise target or desired nucleic acids obtained or extracted from a raw biological sample. In some embodiments, an initial sample can be subjected to a sequencing assay as. In some embodiments, a raw biological sample or an initial sample can be used directly in a sequencing assay without extraction of a nucleic acid. In some embodiments, a nucleic acid as described herein can be extracted from a raw biological sample or an initial sample for use in a sequencing assay. In some embodiments, an extraction method can comprise an alcohol-based extraction, a column purification, a filtration, a size separation, or any combination thereof. As used herein, "removal" or "extraction," and their cognates, of nucleic acids refers to steps prior to the start of generating or preparing a nucleic acid library that separate nucleic acids from at least one component with which they are normally associated. In some embodiments, removal or extraction of nucleic acids can refer to the process of creating an initial sample from a raw biological sample. For example, without limitation, the fractionation of whole blood into its component parts, such as plasma, can be considered to involve removal or extraction. Similarly, purification or isolation of DNA from a sample (e.g., plasma sample) can be considered extraction. In some embodiments, a nucleic acid extracted from a sample can be subjected to a sequencing assay. In some embodiments, a raw biological sample or an initial sample can comprise a biological sample.

[00114] In some embodiments, a sample can comprise a biological sample obtained or collected from a subject. In some embodiments, a biological sample can comprise cells. In some embodiments, a biological sample can be substantially cell-free. In some embodiments, a biological sample can comprise a biological fluid. In some embodiments, a biological fluid can comprise a bodily fluid of a subject (e.g., blood), a fluid obtained from the subject via a medical procedure (e.g., lavage), or any fluid obtained from processing a biopsy of the subject (e.g., serous fluid). In some embodiments, a biological fluid can comprise a bodily fluid. In some

embodiments, a bodily fluid can comprise a whole blood, a plasma, a serum, a lymph, a synovial fluid, a cerebrospinal fluid (CSF), a saliva, a gastric juice, a bile, a pancreatic juice, an intestinal fluid, a respiratory tract mucosal secretion, a semen, a cervical mucus, a vaginal secretion, a urine, a sebum, a breast milk, an amniotic fluid, a pericardial fluid, a pleural fluid, a peritoneal fluid, or any combination thereof. In some embodiments, a biological fluid can be processed from a bodily fluid. For example, blood from a subject can be processed to generate a plasma sample, a serum sample, or a platelet sample. In some embodiments, a biological fluid can comprise a plasma sample. In some embodiments, a biological fluid can comprise a lavage from diagnosing, treating, or cleaning an area of a body of a subject. In some embodiments, a lavage can comprise a bronchoalveolar lavage (BAL), a gastric lavage, a peritoneal lavage, a nasal lavage, a bladder lavage, a rectal lavage, a wound lavage, a joint lavage (arthrocentesis), an eye lavage, a sinus lavage, or any combination thereof. In some embodiments, a biological fluid can comprise an amniotic fluid. In some embodiments, a biological fluid can comprise a BAL. In some embodiments, a biological fluid can comprise a joint lavage. In some embodiments, a biological fluid can comprise a fluid obtained from processing a biopsy of a subject. In some embodiments, a biological fluid can comprise a needle aspiration fluid, a serous fluid, a microdialysis fluid, an exudate fluid, or any combination thereof. As used herein, “plasma” or “blood plasma” refers to the liquid component or fraction of blood. Plasma is generally obtained by spinning a whole blood sample and removing the liquid component.

Microbes

[00115] In some embodiments, the methods described herein comprise detecting a microbe or microbial nucleic acids from a microbe. In some embodiments, the microbe comprises a carrier microbe harboring the one or more AMR genetic markers. In some embodiments, the one or more AMR genetic markers are associated with one or more carrier microbes that harbor the one or more AMR genetic markers. In some embodiments, the one or more carrier microbes comprise a virus, a bacterium, a protozoa, a fungus, an archaea, an algae, or any combination thereof. In some embodiments, the one or more carrier microbes comprise a pathogen. In some embodiments, the one or more carrier microbes comprise a commensal microbe.

[00116] In some embodiments, the one or more carrier microbes comprise one or more bacteria. In some embodiments, the one or more bacteria comprise a gram-positive or a gram-negative bacterium. In some embodiments, the one or more bacteria comprise *Staphylococcus aureus, S. epidermidis, S. lugdunensis, Enterococcus faecalis, E. faecium, Enterobacter cloacae* complex, *Escherichia coli, Klebsiella aerogenes, K. pneumoniae, K. oxytoca, Proteus mirabilis, P. vulgaris, Salmonella bongori, S. enterica, Serratia marcescens, Pseudomonas aeruginosa,*

Acinetobacter baumannii, or *A. calcoaceticus*. In some embodiments, the one or more carrier microbes comprise a methicillin-resistance carrier. In some embodiments, the methicillin-resistance carrier comprises Staphylococcus pseudintermedius, Staphylococcus fleurettii, Staphylococcus epidermidis, Staphylococcus schleiferi, Staphylococcus lugdunensis, Staphylococcus cohnii, Staphylococcus caprae, Staphylococcus warneri, Staphylococcus saprophyticus, Staphylococcus aureus, Staphylococcus haemolyticus, Staphylococcus capitis, Staphylococcus hominis, Staphylococcus pettenkoferi, or Staphylococcus simulans.

[00117] In some embodiments, the one or more carrier microbes comprise a vancomycin-resistance carrier. In some embodiments, the vancomycin-resistance carrier comprises Enterococcus faecium, Enterococcus raffinosus, Enterococcus casseliflavus, Enterococcus gallinarum, Enterococcus avium, Enterococcus durans, Enterococcus faecalis, or Streptococcus gallolyticus.

[00118] In some embodiments, the one or more carrier microbes comprise a Gram-negative carbapenem or extended-spectrum beta-lactamase (ESBL) resistance carrier. In some embodiments, the Gram-negative carbapenem or extended-spectrum beta-lactamase (ESBL) resistance carrier comprises Enterobacter cloacae complex, Escherichia coli, Klebsiella pneumoniae, Klebsiella oxytoca, Proteus mirabilis, Proteus vulgaris, Salmonella enterica, Salmonella bongori, Serratia marcescens, Enterobacter aerogenes, Pseudomonas aeruginosa, Acinetobacter baumannii, Acinetobacter calcoaceticus, Aeromonas caviae, Aeromonas hydrophila, Citrobacter amalonaticus, Citrobacter freundii, Citrobacter koseri, Pluralibacter gergoviae, Enterobacter hormaechei, Morganella morganii, Pantoea agglomerans, Providencia rettgeri, Providencia stuartii, Shigella flexneri, Shigella sonnei, Serratia liquefaciens, Stenotrophomonas maltophilia, Vibrio cholerae, Kluyvera georgiana, Kluyvera cryocrescens, Kluyvera ascorbata, Klebsiella variicola, Raoultella planticola, Raoultella ornithinolytica, Citrobacter braakii, Kluyvera intermedia, Pantoea sp. PSNIH2, Pantoea sp. PSNIH1, or Clavibacter cf. michiganensis LMG 26808.

[00119] In some embodiments, a confounding microbe comprises a carrier microbe. A confounding microbe can refer to any microbe that can harbor nucleic acids that are homologous to the target nucleic acids (e.g., AMR gene or housekeeping gene). In some embodiments, the confounding microbe is a carrier microbe harboring an AMR cassette or an AMR genetic marker.

[00120] As used herein, a “microbe” can refer to a living microorganism or a non-living microscopic entity. In some embodiments, a living microorganism can comprise a bacterium, a protozoa, a fungus, an archaea, an algae, a parasite, or any other living microorganism. In some embodiments, a non-living microscopic entity can comprise a virus, a live virus, a replicating

virus, or an attenuated virus. In some embodiments, non-host nucleic acids can be derived from a microbe. In some embodiments, target nucleic acids can be derived from a plurality of microbes.

[00121] In some embodiments, a microbe can be pathogenic to a subject (e.g., a pathogen), a commensal microbe of a subject, or a microbe present in a general environment. In some embodiments, a pathogen can comprise any pathogenic or virulent microbe. In some embodiments, a commensal microbe of a subject can comprise a microbe that inhabits any location in or on a subject without causing any symptom of a disease or disorder. In some embodiments, a microbe of a general environment can comprise a microbe at or near a sample collection site or a microbe at or near a location of a subject. In some embodiments, a microbe of a general environment comprises a commensal microbe. In some embodiments, a commensal microbe of a subject may become a pathogen to the subject. In some embodiments, a commensal microbe of a first subject may be a pathogen to a second subject. In some embodiments, a microbe of a general environment of a subject may become a pathogen to a subject. In some embodiments, a microbe of a general environment of a first subject may be a pathogen to a second subject.

[00122] In some embodiments, a pathogen can cause an infection or disease comprising gastrointestinal infections (e.g., *Escherichia coli*, *Salmonella* spp., *Clostridioides difficile*), urinary infections (e.g., *Escherichia coli*), skin infections (e.g., *Staphylococcus aureus*, including MRSA), strep throat (or scarlet fever, rheumatic fever) (e.g., *Streptococcus pyogenes*), tuberculosis (e.g., *Mycobacterium tuberculosis*), gonorrhea (e.g., *Neisseria gonorrhoeae*), cholera (e.g., *Vibrio cholerae*), Lyme disease (e.g., *Borrelia burgdorferi*), ulcers or stomach cancer (e.g., *Helicobacter pylori*), syphilis (e.g., *Treponema pallidum*), anthrax (e.g., *Bacillus anthracis*), seasonal flu (e.g., *Influenza* viruses), acquired immunodeficiency syndrome (AIDS) (e.g., *Human Immunodeficiency Virus* (HIV)), liver infections (e.g., *Hepatitis B* and *Hepatitis C* viruses), cervical cancer (e.g., *Human Papillomavirus* (HPV)), respiratory infections (e.g., *SARS-CoV-2*), oral or genital herpes, (e.g., *Herpes Simplex Virus* (HSV-1 and HSV-2)), chickenpox or shingles (e.g., *Varicella Zoster Virus*), measles (e.g., *Measles virus*), neurological disorders (e.g., *Rabies virus*, *Trypanosoma brucei*), dengue fever (e.g., *Dengue virus*), Ebola disease (e.g., *Ebola virus*), candidiasis (e.g., *Candida albicans*), deep lung infections (e.g., *Aspergillus* spp., *Pneumocystis jirovecii*), histoplasmosis (e.g., *Histoplasma capsulatum*), meningitis (e.g., *Cryptococcus neoformans*), ringworm (e.g., *Trichophyton* spp.), malaria (e.g., *Plasmodium* spp.), intestinal infection (e.g., *Ascaris lumbricoides*), giardiasis (e.g., *Giardia lamblia*), amebiasis (e.g., *Entamoeba histolytica*), toxoplasmosis (e.g., *Toxoplasma gondii*), leishmaniasis (e.g., *Leishmania* spp.), schistosomiasis (e.g., *Schistosoma* spp.), strongyloidiasis (e.g., *Strongyloides stercoralis*), tapeworm, taeniasis, or cysticercosis (e.g., *Taenia solium*).

[00123] In some embodiments, commensal microbes can inhabit a gastrointestinal tract, a skin, a respiratory tract, a urogenital tract, or an oral cavity of a subject. In some embodiments, commensal microbes can comprise an endogenous virus (e.g., endogenous retroviruses (ERVs)) of a subject. In some embodiments, a commensal microbe can comprise a *Bacteroides fragilis,* a *Lactobacillus acidophilus,* a *Bifidobacterium bifidum,* an *Escherichia coli (non-pathogenic strains),* a *Staphylococcus epidermidis,* a *Streptococcus salivarius,* a *Propionibacterium acnes,* a *Candida albicans (under normal conditions),* a *Enterococcus faecalis,* a *Clostridium difficile (non-toxigenic strains),* a *Rothia mucilaginosa,* a *Fusobacterium nucleatum,* a *Peptostreptococcus anaerobius,* a *Prevotella melaninogenica*, or any combination thereof. In some embodiments, a commensal microbe can comprise a *Lactobacillus spp.*, a *Bacteroides spp.*, a *Faecalibacterium prausnitzii,* an *Escherichia coli,* a *Clostridium spp.*, an *Enterococcus spp.*, a *Staphylococcus spp.*, a *Candida spp.*, an *Aspergillus spp.*, a *Porcine Endogenous Retroviruses (PERVs),* an *Eimeria spp.*, a *Bifidobacterium spp.*, a *usobacterium spp.*, a *Simian Immunodeficiency Virus (SIV),* a *Simian Retrovirus (SRV),* a *Entamoeba spp*, or any combination thereof.

[00124] In some embodiments, a microbe disclosed herein can comprise one or more microbes comprising: Coniosporium, Hantavirus, Talaromyces, Machlomovirus, Betatetravirus, Raoultella, Aeromonas, Ephemerovirus, Empedobacter, Loa, Macluravirus, Stenotrophomonas, Alfamovirus, Rosavirus, Emmonsia, Aggregatibacter, Orthopneumovirus, Weeksella, Nairovirus, Salivirus, Weissella, Mosavirus, Gammapartitivirus, Strongyloides, Passerivirus, Erysipelatoclostridium, Bacillarnavirus, Iotatorquevirus, Taenia, Trypanosoma, Olsenella, Cladosporium, Rhizobium, Prevotella, Leclercia, Paracoccus, Ilarvirus, Lagovirus, Rasamsonia, Plasmodium, Acremonium, Chlamydia, Clonorchis, Vibrio, Bartonella, Nakazawaea, Franconibacter, Anisakis, Norovirus, Nocardia, Solobacterium, Parechovirus, Avenavirus, Orthohepevirus, Aphthovirus, Hepandensovirus, Microbacterium, Lichtheimia, Lomentospora, Achromobacter, Ipomovirus, Tsukamurella, Elizabethkingia, Hepevirus, Seadornavirus, Alternaria, Trueperella, Gammatorquevirus, Bifidobacterium, Chrysosporium, Thogotovirus, Curtovirus, Deltatorquevirus, Balamuthia, Mastrevirus, Bdellomicrovirus, Mupapillomavirus, Pseudozyma, Wickerhamiella, Aquamavirus, Alloscardovia, Thielavia, Idaeovirus, Henipavirus, Coxiella, Haemophilus, Gammacoronavirus, Negevirus Brevibacterium, Peptoniphilus, Alphacarmotetravirus, Nosema, Trichovirus, Arenavirus, Thermomyces, Necator, Waikavirus, Blosnavirus, Jonesia, Tetraparvovirus, Emaravirus, Plectrovirus, Sclerodamavirus, Toxocara, Umbravirus, Burkholderia, Chromobacterium, Paracoccidioides, Brugia, Eragrovirus, Macrococcus, Absidia, Colletotrichum, Inovirus, Phycomyces, Wickerhamomyces, Acidaminococcus, Moraxella, Rothia, Phlebovirus, Slackia, Purpureocillium,

Betapapillomavirus, Tupavirus, Cryspovirus, Saksenaea, Erysipelothrix, Kobuvirus, Mimoreovirus, Echinococcus, Mannheimia, Bergeyella, Cyclospora, Xylanimonas, Leptospira, Finegoldia, Curvularia, Cryptosporidium, Babuvirus, Pecluvirus, Lambdatorquevirus, Pythium, Carlavirus, Entomobimavirus, Kocuria, Anaplasma, Ampelovirus, Avihepatovirus, Nepovirus, Rhodococcus, Bordetella, Mischivirus, Scedosporium, Gardnerella, Maculavirus, Trichoderma, Aveparvovirus, Salmonella, Avastrovirus, Copiparvovirus, Trachipleistophora, Clostridioides, Nanovirus, Siccibacter, Leptotrichia, Citrivirus, Odoribacter, Sanguibacter, Novirhabdovirus, Acremonium, Hafnia, Chaetomium, Tenuivirus, Yokenella, Rubulavirus, Varicellovirus, Alphamesonivirus, Sicinivirus, Leuconostoc, Microvirus, Gallantivirus, Morbillivirus, Lolavirus, Pantoea, Hepatovirus, Nupapillomavirus, Metschnikowia, Bamavirus, Kytococcus, Tritimovirus, Tannerella, Respirovirus, Pneumocystis, Dirofilaria, Pediococcus, Lactococcus, Blastomyces, Dianthovirus, Actinobacillus, Teschovirus, Oscivirus, Begomovirus, Potyvirus, Byssochlamys, Alphacoronavirus, Molluscipoxvirus, Lymphocryptovirus, Sapelovirus, Parabacteroides, Pyrenochaeta, Listeria, Senecavirus, Brevidensovirus, Potexvirus, Parvimonas, Flavivirus, Recovirus, Toxoplasma, Yatapoxvirus, Opisthorchis, Trichuris, Cyphellophora, Morganella, Perhabdovirus, Micrococcus, Pequenovirus, Mastadenovirus, Anaeroglobus, Tropheryma, Dolosigranulum, Wolbachia, Lelliottia, Mycoplasma Tobravirus, Shewanella, Paeniclostridium, Erythroparvovirus, Sutterella, Sporopachydermia, Namavirus, Nyavirus, Francisella, Arthroderma, Epsilontorquevirus, Sigmavirus, Amdoparvovirus, Actinomyces, Alphapermutotetravirus, Cardiobacterium, Influenzavirus C, Orthopoxvirus, Poacevirus, Phialophora, Lactobacillus, Polyomavirus, Debaryomyces, Foveavirus, Bymovirus, Mycoflexivirus, Grimontia, Mucor, Rhytidhysteron, Quadrivirus, Thermoascus, Aureusvirus, Trichosporon, Myceliophthora, Dermacoccus, Dysgonomonas, Pseudoramibacter, Becurtovirus, Gordonia, Sapovirus, Orthobunyavirus, Spiromicrovirus, Pomovirus, Exophiala, Sneathia, Helicobacter, Photorhabdus, Mogibacterium, Betapartitivirus, Avibirnavirus, Ambidensovirus, Oleavirus, Orientia, Deltacoronavirus, Anulavirus, Trichomonasvirus, Budvicia, Geotrichum, Enamovirus, Lachnoclostridium, Schistosoma, Paecilomyces, Panicovirus, Rhizoctonia, Brevibacillus, Beauveria, Pestivirus, Tombusvirus, Cilevirus, Cokeromyces, Peptostreptococcus, Phanerochaete, Proteus, Idnoreovirus, Aspergillus, Pasteurella, Malassezia, Hanseniaspora, Endornavirus, Azospirillum, Velarivirus, Cystovirus, Avisivirus, Bacteroides, Picobirnavirus, Myroides, Circovirus, Arterivirus, Aquaparamyxovirus, Onchocerca, Cosavirus, Kluyveromyces, Fijivirus, Candida, Hepacivirus, Dermabacter, Ourmiavirus, Allexivirus, Enterobacter, Acidovorax, Bracorhabdovirus, Carmovirus, Pluralibacter, Coltivirus, Fonsecaea, Streptobacillus, Corynebacterium, Macrophomina, Marburgvirus, Comovirus, Fabavirus, Alphanodavirus, Cellulomonas, Enterobius, Catabacter, Moellerella, Nakaseomyces,

Cucumovirus, Valsa, Deltapartitivirus, Plesiomonas, Pseudomonas, Torovirus, Cuevavirus, Hypovirus, Trichomonas, Influenzavirus D, Giardiavirus, Crinivirus, Tepovirus, Sakobuvirus, Cyberlindnera, Paenalcaligenes, Bafinivirus, Rymovirus, Pegivirus, Yarrowia, Treponema, Borreliella, Rubivirus, Aureobasidium, Angiostrongylus, Filobasidium, Photobacterium, Rhizopus, Orthoreovirus, Ustilago, Simplexvirus, Aquareovirus, Protoparvovirus, Propionibacterium, Sprivivirus, Hunnivirus, Apophysomyces, Meyerozyma, Alphapapillomavirus, Candida, Brucella, Gallivirus, Dinovernavirus, Anaerobiospirillum, Eubacterium, Tatlockia, Terrisporobacter, Quaranjavirus, Sobemovirus, Dicipivirus, Arcanobacterium, Macanavirus, Atopobium, Vesivirus, Lodderomyces, Dinornavirus, Betatorquevirus, Kerstersia, Aparavirus, Neisseria, Agrobacterium, Edwardsiella, Labyrnavirus, Totivirus, Actinomadura, Tobamovirus, Influenzavirus B, Mandarivirus, Anaerococcus, Kunsagivirus, Naegleria, Campylobacter, Veillonella, Yamadazyma, Filobasidiella, Oerskovia, Penicillium, Anncaliia, Leptosphaeria, Pneumovirus, Psychrobacter, Isavirus, Granulicatella, Torradovirus, Cladophialophora, Influenzavirus A, Ophiostoma, Aerococcus, Ureaplasma, Etatorquevirus, Bocaparvovirus, Megasphaera, Reptarenavirus, Comamonas, Capnocytophaga, Alphatorquevirus, Syncephalastrum, Wallemia, Betacoronavirus, Hyphopichia, Nocardiopsis, Legionella, Trichinella, Paraburkholderia, Mammarenavirus, Echinostoma, Sphingobacterium, Enterovirus, Methanobrevibacter, Ochroconis, Cheravirus, Pasivirus, Enterococcus, Mycoreovirus, Tospovirus, Betanodavirus, Phytoreovirus, Enterocytozoon, Ferlavirus, StemphyliumFilifactor, Leishmaniavirus, Gemella, Bromovirus, Alloiococcus, Cunninghamella, Cronobacter, Oribacterium, Orbivirus, Chrysovirus, Cripavirus, Tatumella, Pandoraea, Ogataea, Dracunculus, Volvariella, flavirus, Benyvirus, Rhadinovirus, Histoplasma, Rahnella, Morococcus, Verticillium, Janibacter, Gyrovirus, Alphapartitivirus, Mycobacterium, Roseomonas, Varicosavirus, Chryseobacterium, Parapoxvirus, Rhizomucor, Aureimonas, Levivirus, Leishmania, Luteovirus, Cypovirus, Ochrobactrum, Microsporum, Piscihepevirus, Ceratocystis, Sporothrix, Vesiculovirus, Cupriavidus, Cryptococcus, Metapneumovirus, Alphanecrovirus, Eikenella, Brevundimonas, Escherichia, Leifsonia, Schizophyllum, Granulibacter, Gordonibacter, Lachancea, Madurella, Ophiovirus, Phellinus, Nebovirus, Acanthamoeba, Fusobacterium, Pichia, Verruconis, Ehrlichia, Tibrovirus, Higrevirus, Wohlfahrtiimonas, Rhinocladiella, Neorickettsia, Sadwavirus, Roseobacter, Sequivirus, Pannonibacter, Rotavirus, Turicella, Cardiovirus, Propionimicrobium, Furovirus, Naumovozyma, Closterovirus, Fluoribacter, Zeavirus, Clavispora, Megrivirus, Gammapapillomavirus, Rickettsia, Polemovirus, Corynespora, Encephalitozoon, Shimwellia, Fusarium, Yersinia, Capronia, Delftia, Victorivirus, Marafivirus, Kluyvera, Iteradensovirus, Isoptericola, Vitivirus, Roseolovirus, Conidiobolus, Abiotrophia, Babesia, Phoma,

Sanguibacteroides, Staphylococcus, Rhodotorula, Zetatorquevirus, Hymenolepis, Fasciola, Cytorhabdovirus, Cardoreovirus, Memnoniella, Trichophyton, Mitovirus, Phaeoacremonium, Providencia, Lysinibacillus, Giardia, Oligella, Streptomyces, Paraclostridium, Ralstonia, Coccidioides, Brambyvirus, Biatriospora, Allolevivirus, Acinetobacter, Starmerella, Omegatetravirus, Porphyromonas, Avulavirus, Streptococcus, Arcobacter, Topocuvirus, Mamastrovirus, Ancylostoma, Bomavirus, Capillovirus, Alphavirus, Tymovirus, Nucleorhabdovirus, Diaporthe, Chlamydiamicrovirus, Tumcurtovirus, Saccharomyces, Riemerella, Betanecrovirus, Clostridium, Mobiluncus, Cercospora, Mamavirus, Mortierella, Aquabimavirus, Xanthomonas, Dependoparvovirus, Ebolavirus, Neofusicoccum, Borrelia, Leminorella, Klebsiella, Blastocystis, Alcaligenes, Citrobacter, Eggerthella, Cedecea, Serratia, Penstyldensovirus, Bacillus, Laribacter, Wuchereria, Hordeivirus, Cytomegalovirus, Actinomucor, Ascaris, Shigella, Vittaforma, Torulaspora, Kingella, Oryzavirus, Polerovirus, Tremovirus, Erbovirus, Entamoeba, Lyssavirus, Paenibacillus, Facklamia, Kappatorquevirus, Metarhizium, Stachybotrys, Okavirus, Botrexvirus, Thetatorquevirus, or Basidiobolus. In some embodiments, a microbe disclosed herein can comprise one or more microbes disclosed in the Drawings or the Examples.

Cell-Free Nucleic acids

[00125] In some embodiments, the nucleic acids described herein comprise cell-free nucleic acids (cfNAs). Cell-free nucleic acids are generally fragments of nucleic acids that float freely outside of cells in any body fluid of a subject. In some embodiments, cfNA can comprise plasma cfNA, serum cfNA, whole-blood cfNA, cerebrospinal fluid (CSF) cfNA, saliva cfNA, bronchoalveolar lavage (BAL) cfNA, urine cfNA, amniotic cfNA, fetal cfNA, synovial fluid cfNA, lymphatic cfNA, or any combination thereof. In some cases, cfNA comprise circulating cfNA in a subject's bloodstream. In some embodiments, the nucleic acids comprise circulating cfDNA, circulating cfRNA, cfDNA, cfRNA, circulating DNA, circulating RNA, or any combination thereof. In some embodiments, a sample of a body fluid can comprise a cfNA.

[00126] The cfNA described herein are, in some embodiments, nucleic acids that, when floating within a body fluid of a subject, are not encapsulated by a cell. In some embodiments, a cfNA comprises nucleic acids that are not encapsulated by a human cell, not encapsulated by a microbial cell, or not encompassed by either a human cell or a microbial cell. The cfNAs can be free-floating, such as cfDNA fragments in plasma. In some embodiments, cfNA includes vesicle-associated cfNA, such as cfNA associated with exosomes, extracellular vesicles, microvesicles, apoptotic bodies, or any combination thereof. In some embodiments, cfNA do not comprise vesicle-associated cfNA, such as cfNA associated with exosomes, extracellular vesicles,

microvesicles, apoptotic bodies, or any combination thereof. cfNAs can also be associated with proteins or other cellular constituents, outside of an intact cell. For example, cfNA can comprise free-floating nucleosome-associated cfNA. cfNAs can arise from various biological processes, including cell death (apoptosis, necrosis) or active secretion.

[00127] In some embodiments, cfNA can comprise viral nucleic acids that are not encapsulated by a capsid, that are fragmented, or a combination thereof. Of note, the methods provided herein can, in some embodiments, be practiced using viral nucleic acids derived from whole viruses floating in a bodily fluid.

[00128] In some embodiments, cfNA can be alternatively referred to as free-circulating nucleic acids. In some embodiments, a cfNA can originate from cell death and other processes that release fragments of nucleic acids into a bloodstream or other biological fluid. In some embodiments, a cfNA can be derived from any source of nucleic acids provided herein. In some embodiments, cfNA present in a raw biological sample can be isolated from genomic nucleic acid in the raw biological sample by processing the raw biological sample into an initial sample by removing intact cells. In some embodiments, removing intact cells can comprise centrifuging or filtering a raw biological sample to produce a cell-free fraction of a biological fluid comprising cfNA. In some cases, the removal of intact cells may use a technique targeting a specific cell type. For example, centrifugation at a relative low RPM can remove human or mammalian cells. In some cases, centrifugation at a higher RPM can be used to remove microbial cells such as bacterial cells. In some cases, centrifugation can be performed at an even higher speed (e.g, via ultracentrifugation) in order to remove viral particles. In some cases, afaster spin can be performed after an initial removal (e.g., by centrifugation) of mammalian cells. In some embodiments, a sample can comprise non-host nucleic acids. In some embodiments, non-host nucleic acids can comprise microbial nucleic acids. In some embodiments, microbial nucleic acids can comprise microbial cell-free nucleic acid (mcfNA). In some embodiments, the phrase "target nucleic acids" as used herein can refer to target cfNA. In some embodiments, the phrase "target nucleic acids" as used herein can refer to target mcfNA. In some embodiments, an mcfNA can be derived from one or more kingdoms, divisions, classes, orders, families, genuses, species and/or strain of microbe. In some embodiments, a mcfNA can be derived from a prokaryotic or a eukaryotic microbe. In some embodiments, an mcfNA can comprise a bacterial cfNA, a fungal cfNA, a viral cfNA, a protozoan cfNA, an archaeal cfNA, an algal cfNA, or any combination thereof. In some embodiments, a sample can comprise a non-microbial nucleic acid (e.g., a non-microbial cell-free nucleic acid). In some embodiments, a sample can comprise mcfNAs from one or more species of microbes. In some embodiments, a sample can comprise mcfNAs from at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least

11, at least 12, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, at least 21, at least 22, at least 23, at least 24, at least 25, at least 26, at least 27, at least 28, at least 29, or at least 30 species of microbes.

[00129] In some embodiments, a sample can comprise a mixture of nucleic acids. In some embodiments, a sample can comprise target nucleic acids (e.g., target cfNAs) and/or non-target nucleic acids. In some cases, the cfNAs comprise cfNA derived from a subject or microbial genome. In some cases, the cfNAs are derived from a housekeeping gene, e.g., a human or mammalian housekeeping gene, a microbial housekeeping gene, a microbe-specific housekeeping gene, and/or a microbe non-specific housekeeping gene.

[00130] In some embodiments, a sample can further comprise contaminant nucleic acids. In some embodiments, contaminant nucleic acids can comprise nucleic acids from a general environment (e.g., a sample collection site). In some embodiments, contaminant nucleic acids are introduced during any step of sample processing. In some embodiments, the contaminant nucleic acids comprise contaminating microbial nucleic acids, contaminating nucleic acids from a different sample, contaminating host nucleic acids, or any combination thereof.

[00131] In some embodiments, cell-free nucleic acids (cfNAs) can comprise a mixture of cfNAs. In some embodiments, a mixture of cfNAs can comprise cfNAs originated from one or more organisms. In some embodiments, a mixture of cfNAs can comprise microbial nucleic acids (e.g., mcfNAs) originated from one or more species of microbes. In some embodiments, an mcfNA can comprise a bacterial-derived cfNA, a fungal-derived cfNA, a viral-derived cfNA, a protozoan-derived cfNA, an archaeal-derived cfNA, an algal-derived cfNA, or any combination thereof.

[00132] In some embodiments, a cfNA can comprise a double-stranded nucleic acid (dsNA), a single-stranded nucleic acid (ssNA), nicked double-stranded, or a combination thereof. In some embodiments, a cfNA can comprise a cell-free DNA (cfDNA), a cell-free RNA (cfRNA), a cell-free DNA-RNA hybrid (cfDNA-RNA), or a combination thereof. In some embodiments, hcfNA can comprise host cell-free DNA (hcfDNA), host cell-free RNA (hcfRNA), host cell-free DNA-RNA hybrid (hcfDNA-RNA), or a combination thereof. In some embodiments, mcfNA can comprise microbial cell-free DNA (mcfDNA), microbial cell-free RNA (mcfRNA), microbial cell-free DNA-RNA hybrid (mcfDNA-RNA), or any combination thereof.

[00133] As used herein, a cell-free samples is generally a sample devoid, or almost devoid, of cells. In some instances, the cell-free sample is devoid of human cells (e.g., intact human cells). In some instances, the cell-free sample is devoid of microbial cells (e.g., intact microbial cells). In some instances, the cell-free sample is devoid of microbial and human cells. In some instances, the cell-free sample is devoid of a particular type of microbial cell or virus, while comprising a

different type of microbial cell or virus. For example, the cell-free sample can, in some embodiments, be devoid of intact bacterial cells while containing intact viruses. In some instances, the cell-free sample is devoid of all types of microbes, including eukaryotic or prokaryotic cells. The cell-free sample can be obtained from a biological sample provided herein. In some instances, the cell-free sample is a plasma, which has been processed in order to remove blood cells, subject cells, and/or intact microbes, or fragments thereof. In some instances, the cell-free sample can be obtained by a sample preparation process, such as centrifuging, ultracentrifuging, or filtering a biological sample.

[00134] In some embodiments, a cfNA can comprise a host nucleic acid, a non-host nucleic acid, a target nucleic acid, or a combination thereof. In some embodiments, a cfNA can be derived from a host (host cell free nucleic acids or "hcfNA") or a non-host. In some embodiments, a hcfNA can be derived from nuclear nucleic acids, mitochondria nucleic acids, exosomal nucleic acids, fetal nucleic acids, or any combination thereof. In some embodiments, a sample can comprise a host nucleic acid (e.g., a host cell-free nucleic acid). In some embodiments, a host is any subject provided herein.

[00135] In some embodiments, the nucleic acids from a sample can be extracted and/or enriched to generate enriched nucleic acids. In some embodiments, the enriched nucleic acids comprise degraded nucleic acids, ultra-short nucleic acids, single stranded nucleic acids, double stranded nucleic acids, nicked nucleic acids, microbial cell-free nucleic acids (mcfNA), subject's cell-free nucleic acids, circulating tumor nucleic acids (ctNA), mitochondrial nucleic acids (mtNA), or any combination thereof. In some embodiments, the methods comprise enriching for degraded nucleic acids, ultra-short nucleic acids, single stranded nucleic acids, or nicked double stranded nucleic acids. As used herein, "degraded nucleic acid" refers to fragments of DNA and RNA that are released into circulation due to cell death, turnover, or pathological processes. Degraded nucleic acids can originate from various sources, including natural sources or artificial sources. In some embodiments, the degraded nucleic acids originate from normal cellular apoptosis, necrosis, or disease-related processes such as cancer or infections. In some embodiments, the degraded nucleic acids originate from factors during laboratory handling, including, but not limited to, sample degradation due to prolonged storage, or temperature fluctuations. In some embodiments, ultrashort nucleic acids comprise nucleic acids less than 100 bp, less than 90 bp, less than 80 bp, less than 70 bp, less than 60 bp, less than 50 bp, less than 40 bp, or less than 30 bp in length. In some embodiments, at least 70%, 75%, 80%, 85%, 90%, or 95% of the mcfNA are degraded nucleic acids.

Sizes of cfNAs

[00136] In some embodiments, a cfNA as disclosed herein or fragments thereof can be approximately less than about 10 bp, less than about 15 bp, less than about 20 bp, less than about 25 bp, less than about 30 bp, less than about 35 bp, less than about 40 bp, less than about 45 bp, less than about 50 bp, less than about 55 bp, less than about 60 bp, less than about 65 bp, less than about 70 bp, less than about 75 bp, less than about 80 bp, less than about 85 bp, less than about 90 bp, less than about 95 bp, less than about 100 bp, less than about 105 bp, less than about 110 bp, less than about 115 bp, less than about 120 bp, less than about 125 bp, less than about 130 bp, less than about 135 bp, less than about 140 bp, less than about 145 bp, less than about 150 bp, less than about 155 bp, less than about 160 bp, less than about 165 bp, less than about 170 bp, less than about 175 bp, less than about 180 bp, less than about 185 bp, less than about 190 bp, less than about 195 bp, or less than about 200 bp long.

[00137] In some embodiments, cfNAs provided herein or fragments thereof can be approximately about 10 bp, about 15 bp, about 20 bp, about 25 bp, about 30 bp, about 35 bp, about 40 bp, about 45 bp, about 50 bp, about 55 bp, about 60 bp, about 65 bp, about 70 bp, about 75 bp, about 80 bp, about 85 bp, about 90 bp, about 95 bp, about 100 bp, about 105 bp, about 110 bp, about 115 bp, about 120 bp, about 125 bp, about 130 bp, about 135 bp, about 140 bp, about 145 bp, about 150 bp, about 155 bp, about 160 bp, about 165 bp, about 170 bp, about 175 bp, about 180 bp, bout 185 bp, about 190 bp, about 195 bp, or about 200 bp long. In some embodiments, the cfNAs provided herein comprise ultra short cfNAs. As used herein, "ultra short nucleic acids" refer to subnucleosomal nucleic acids or fragments shorter than subnucleosomal nucleic acids. In some embodiments, ultra short cfNAs can be from about 10 bp to about 100 bp long. In some embodiments, the ultra short cfNAs can be from about 30 bp to about 80 bp long. In some embodiments, the ultra short cfNAs can be from about 40 bp to about 60 bp long. In some embodiments, the ultra short cfNAs are about 50 bases long.

[00138] In some embodiments, mcfNA can be present at higher concentrations relative to hcfNA at lengths that fall outside a nucleosomal interval. In some embodiments, mcfNA can be enriched relative to hcfNA by enriching for cfNA of less than 180bp, less than 170bp, less than 160bp, less than 150bp, less than 140bp, less than 130bp, less than 120bp, less than 110bp, less than 100bp, less than 90bp, less than 80bp, less than 70bp, less than 60bp, less than 50bp, less than 40bp, less than 30bp, or less than 20bp. In some embodiments, enriching for mcfNA can comprise enriching for cfNA between 10-180 bp.

[00139] In some embodiments, a cfNA can comprise any nucleic acid that is not encapsulated by a cell (e.g., a eukaryotic or microbial cell). In some embodiments, a cfNA can originate from any

nucleic acids. In some embodiments, a cfNA can comprise a plurality of chemical forms of deoxyribonucleic acid (DNA), ribonucleic acid (RNA), or DNA/RNA hybrid. In some embodiments, nucleic acids can comprise a plurality of structural forms of DNA, RNA, or DNA/RNA hybrid. In some embodiments, a cfNA can comprise linear nucleic acids or circular nucleic acids. In some embodiments, a cfNA can comprise single stranded nucleic acids (ssNA), double strand nucleic acids (dsNA) or hybrid nucleic acids. In some embodiments, nucleic acids can be from a genome of an organism or an organelle of a cell (e.g., an exosome or a mitochondria). In some embodiments, a cfNA can comprise a mitochondrial DNA, an intercellular signal nucleic acid, an exogenous nucleic acid, a DNA enzyme, a RNA enzyme, a food-derived nucleic acid, any metabolic form of nucleic acid-based therapeutic, or any combination thereof. In some embodiments, a cfNA can be derived from a member selected from the group consisting of genomic DNA, cDNA, mRNA, cRNA, tRNA, ribosomal RNA, miRNA, siRNA, nuclear DNA, nuclear RNA, mitochondrial DNA, mitochondrial RNA, exosomal DNA, exosomal RNA, fetal DNA, fetal RNA, plasmids, vectors, and any combination thereof.

[00140] In some embodiments, nucleic acids can comprise a mixture of nucleic acids from various sources. In some embodiments, nucleic acids can be derived from a plurality of biological fluids. In some embodiments, nucleic acids can be from a plurality of organisms. In some embodiments, nucleic acids can be from a subject. In some embodiments, nucleic acids can be from one or more species of microbes. In some embodiments, nucleic acids can comprise environmental nucleic acids. In some embodiments, environmental nucleic acids can comprise any nucleic acid at or near a sample collection site, or any nucleic acid introduced by personnel, equipment or a reagent used in collecting and/or processing a sample from a subject.

Process control molecules

[00141] As used herein, the phrase "process control molecules" refers to molecules that are added to a sample before or during nucleic acid library generation to aid in the identification or quantification of nucleic acids in a sample. In some embodiments, process control molecules can comprise nucleic acids. In some embodiments, process control molecules can comprise synthetic nucleic acids. In some embodiments, process control molecules are separate from and not integrated in the target molecules. In some embodiments, process control molecules can have special features such as specific sequences, lengths, GC content, degrees of degeneracy, degrees of sequence diversity, different secondary, tertiary, or quaternary structures, and/or known starting concentrations. In some embodiments, process control molecules can be used for normalizing the signal in a sample to account for variations in sample processing or to control process performance. In some embodiments, process control molecules can include whole assay

internal control (WINC) molecules. In some embodiments, at least 10,000, at least 15,000, at least 20,000, at least 25,000, at least 30,000, at least 35,000, at least 40,000, at least 45,000, or at least 50,000 unique WINC molecules are spike in the sample. In some embodiments, process control molecules can include sample identifiers. In some embodiments, process control molecules can comprise dephosphorylation control molecules, denaturation control molecules, and/or ligation control molecules. In some embodiments, multiple different types or sets of control molecules can be added to a sample.

[00142] As used herein, the phrase "adapter attachment control molecule" refers to a control molecule that allows monitoring of the efficiency of an adapter attachment reaction. An adapter attachment reaction can be ligation-based, TdT-based, template-switching-based, primer-extension-based, amplification-based, or a combination thereof.

[00143] As used herein, the phrase "degradation assessment molecules" refers to a control molecule used to evaluate sample and spiked sample integrity during processing.

[00144] As used herein, the phrase "spiked initial sample" refers to an initial sample to which process control molecules (or synthetic spike-ins) have been added prior to the start of generating a sequencing library.

[00145] As used herein, "sequence diversity controls" refers to degenerate pools, or pools of nucleic acids with diverse sequences, which degenerate pools can often be used for diversity assessment, abundance calculation, and/or determination of information transfer efficiency/

[00146] As used herein, "size controls," "length controls," "GC Spike-in Panel" or "GC size/length controls" refers to nucleic acids that are size or length or GC-content markers, which can be used for abundance normalization, development, and/or analysis purposes and other purposes.

[00147] As used herein, "ID Spike(s)" refers to identification spikes that can be used, for example without limitation, for sample identification tracking, cross-contamination detection, reagent tracking, and/or reagent lot tracking.

Sample Preparation and Processing

[00148] In some embodiments, a sample comprising nucleic acids can be prepared prior to a sequencing assay. In some embodiments, a raw biological sample comprising whole blood can be processed by centrifugation to generate an initial sample of plasma.

[00149] In some embodiments, whole blood can be collected in a K2-EDTA tube. In some embodiments, whole blood draws are not pooled. In some embodiments, a tube can be gently inverted multiple times after draw. In some embodiments, a tube can be centrifuged for about 1200 RCF (g), about 1400 RCF (g), about 1600 RCF (g), or more after draw in order to separate

plasma from the blood. The centrifugation can occur at ambient temperature. In some cases, the centrifugation occurs for greater than 5 minutes, 7 minutes, 10 minutes, 15 minutes or 20 minutes. In some embodiments, for tubes containing less than 4 mL a tube manufacturer's instruction and centrifugation speed and time can be used. In some embodiments, the plasma fraction can be transferred into a new tube. In some embodiments, the plasma is subjected to centrifugation a second time to remove residual cells (e.g., mammalian cells and microbial cells). The additional centrifugation can be conducted at, e.g., about 1400 RCF (g), about 1600 RCF (g), about 1800 RCF (g), about 2000 RCF (g), or more.

[00150] In some embodiments, at least 0.4 ml, at least 0.5 ml, at least 0.7 ml, or at least 1.0 ml of plasma can be transferred into a sterile polypropylene tube, with care taken to not disturb a buffy coat when transferring. In some embodiments, a tube can be labelled with a patient's first and last name, a unique identifier (DOB or MRN), and/or a date and time of specimen collection. In some embodiments, if a specimen is unlikely to reach a testing facility within 96 hours of collection it can be frozen directly in K2-EDTA after centrifugation. In some embodiments, if a gel plug does not rise to separate cells from plasma, then a tube can be re-centrifuged at a higher speed. In some embodiments, a specimen tube can be shipped to a testing facility.

[00151] In some embodiments, the methods comprise analyzing a cell-free sample. In some embodiments, the methods comprise performing a sequencing assay on the cell-free sample. In some embodiments, the methods comprise preparing a cell-free sample from a biological sample for the sequencing assay. In some instances, the methods comprise centrifuging a biological sample to generate a cell-free sample. In some embodiments, the cell free sample is plasma. In some instances, the methods comprise centrifuging a biological sample to generate a cell-free sample devoid of or almost devoid of human cells. In some instances, the methods comprise centrifuging a biological sample to generate a cell-free sample devoid of or almost devoid of non-human cells, such as a microbe. In some embodiments, the methods described herein comprise processing the sample comprising nucleic acids to maximize collection of target nucleic acids. In some embodiments, the methods can enrich for or maximize the collection of degraded nucleic acids, ultra-short nucleic acids, single stranded nucleic acids, double stranded nucleic acids, nicked nucleic acids, or rare nucleic acids.

[00152] In some embodiments, a nucleic acid can be extracted from a sample. In some embodiments, an extraction can comprise separating nucleic acids from other cellular components and contaminants that can be present in a sample. In some embodiments, a nucleic acid can be extracted from a sample using a liquid extraction (e.g., a Trizol, a DNAzol) technique. In some embodiments, an extraction can be performed by phenol chloroform extraction or precipitation by organic solvents (e.g., ethanol, or isopropanol). In some embodiments, an

extraction can be performed using a nucleic acid-binding column, a nucleic acid-binding spin column, or a combination thereof. In some cases, an extraction of a cell-free nucleic acid can involve filtration or ultra-filtration. In some embodiments, a nucleic acid can be extracted or purified by use of magnetic beads that bind nucleic acids. In some embodiments, compositions of a binding buffer can be adjusted to control a strength of bonds between functional groups and a nucleic acid, allowing for controlled and reversible binding. In some embodiments, a nucleic acid can be released from a magnetic particle with an elution buffer.

[00153] In some embodiments, the methods described herein comprise enriching a population of cfNA. In some embodiments, enriching a population of cfNA comprises bioinformatically enriching or physically enriching. In some embodiments, enriching a population of cfNA comprises isolating, extracting, or selectively amplifying a desired population of cfNA (e.g., target cfNA) from the initial sample. In some embodiments, enriching a population of cfNA does not comprise isolating or extracting the desired population of cfNA from the initial sample. In some embodiments, enriching a population of cfNA does not comprise amplifying the desired population of cfNA. In some embodiments, enriching a population of cfNA comprises removing the undesired cfNA population (e.g., non-target cfNA or contaminants) from the initial sample. In some embodiments, enriching cfNA can comprise differentiating and/or selecting the cfNA by one or more characteristics comprising size, sequence, GC content, secondary structure, biological source, or protein-binding.

[00154] In some embodiments, the methods comprise enriching microbial cfNA (mcfNA) in the biological sample. In some embodiments, the methods provided herein comprise enriching for at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, or at least 95% of the mcfNA in the biological sample. In some embodiments, enriching mcfNA comprises enriching cfNA that are less than about 20 bp, 30 bp, 40 bp, 50 bp, 60 bp, 70 bp, 80 bp, 90 bp, 100 bp, 110 bp, 120 bp, 130 bp, 140 bp, 150 bp, 160 bp, 170 bp, 180 bp, 190 bp, 200 bp, 210 bp, 220 bp, 230 bp, 240 bp, or 250 bp in length. In some embodiments, enriching mcfNA comprises amplifying the nucleic acids in the initial sample with primers containing non-human nucleic acid sequences. In some embodiments, enriching mcfNA comprises removing non-microbial cfNA. In some embodiments, enriching mcfNA comprises removing nucleosome-bound cfNA. In some embodiments, the methods comprise enriching non-microbial cfNA (e.g., host cfNA) in the biological sample. In some embodiments, the methods comprise enriching mammalian cfNA in the biological sample.

[00155] In some embodiments, the methods comprise enriching cfNA by size selection. In some embodiments, size selection can comprise removing nucleic acids not in the desired size range. In some embodiments, the desired size range comprises an artificial or engineered threshold. In

some embodiments, size selection comprises separating nucleic acids by size via chromatography (e.g., size-exclusion chromatography), electrophoresis (e.g., gel or capillary electrophoresis), centrifugation (e.g., density-gradient centrifugation), filtration (e.g., membrane ultrafiltration), magnetic bead-based methods (e.g., SPRI beads), affinity-based methods (e.g., streptavidin beads), or any combination thereof. In some embodiments, the methods described herein can comprise discriminating or differentiating cfNA by size and/or selectively isolating or extracting the cfNA of the desired size range. In some embodiments, the methods comprise differentiating microbial cfNA from the non-microbial cfNA (e.g., host cfNA) in the sample and selectively removing the non-microbial cfNA in the sample by size. In some embodiments, the non-microbial cfNA comprises mammalian cfNA (e.g., human or animal cfNA).

[00156] In some embodiments, the methods can comprise selectively removing nucleic acid fragments greater than about 500 bp, about 450 bp, about 400 bp, about 350 bp, about 300 bp, about 250 bp, about 200 bp, about 150 bp, about 140 bp, about 130 bp, about 120 bp, about 110 bp, about 100 bp, about 90 bp, about 80 bp, about 70 bp, or about 60 bp in length. In some embodiments, the methods can comprise selectively enriching nucleic acid fragments at most about 20 bp, about 30 bp, about 40 bp, about 50 bp, about 60 bp, about 70 bp, about 80 bp, about 90 bp, about 100 bp, about 110 bp, about 120 bp, about 130 bp, about 140 bp, about 150 bp, about 160 bp, about 170 bp, about 180 bp, about 190 bp, about 200 bp, about 210 bp, about 220 bp, about 230 bp, about 240 bp, or about 250 bp in length.

[00157] In some embodiments, the methods can comprise selectively enriching nucleic acid fragments of about 10 bp to about 20 bp, about 10 bp to about 30 bp, about 10 bp to about 40 bp, about 10 bp to about 50 bp, about 10 bp to about 60 bp, about 10 bp to about 70 bp, about 10 bp to about 80 bp, about 10 bp to about 90 bp, about 10 bp to about 100 bp, about 10 bp to about 110 bp, about 10 bp to about 120 bp, about 10 bp to about 130 bp, about 10 bp to about 140 bp, about 10 bp to about 150 bp, about 10 bp to about 160 bp, about 10 bp to about 170 bp, about 10 bp to about 180 bp, about 10 bp to about 190 bp, about 10 bp to about 200 bp, about 10 bp to about 210 bp, about 10 bp to about 220 bp, about 10 bp to about 230 bp, about 10 bp to about 240 bp, or about 10 bp to about 250 bp in length. In some embodiments, the methods can comprise selectively enriching nucleic acid fragments of about 20 bp to about 250 bp, about 20 bp to about 200 bp, about 20 bp to about 150 bp, about 20 bp to about 100 bp, about 20 bp to about 90 bp, about 20 bp to about 80 bp, about 20 bp to about 70 bp, about 20 bp to about 60 bp, about 20 bp to about 50 bp, about 30 bp to about 250 bp, about 30 bp to about 200 bp, about 30 bp to about 150 bp, about 30 bp to about 100 bp, about 30 bp to about 90 bp, about 30 bp to about 80 bp, about 30 bp to about 70 bp, about 30 bp to about 60 bp, about 30 bp to about 50 bp, about 40 bp to about 250 bp, about 40 bp to about 200 bp, about 40 bp to about 150 bp, about 40 bp to

about 100 bp, about 40 bp to about 90 bp, about 40 bp to about 80 bp, about 40 bp to about 70 bp, about 40 bp to about 60 bp, or about 40 bp to about 50 bp in length.

Exemplary Process 1 – double-stranded cfDNA

[00158] In some embodiments, the methods provided herein comprise performing process 1 to prepare a sample for high throughput sequencing assay. Process 1 provides an example of preparing a sequencing library from the double-stranded cfDNA in the original sample. In some embodiments, a control molecule can be added to an initial sample of plasma to generate a spiked plasma sample. In some embodiments, nucleic acid extraction can be performed on a spiked plasma sample to generate purified and concentrated cfDNA.

[00159] In some embodiments, a library preparation process can be performed on a purified and concentrated cfDNA sample. In some embodiments, the library preparation can comprise attaching (e.g., by ligation) double-stranded adapters to double-stranded cfDNA. In some embodiments, library preparation can comprise performing unbiased amplification on the sample.

[00160] In some embodiments, a sample preparation method does not comprise extracting nucleic acids from a raw or initial sample. For example, in some cases, nucleic acids can be extracted during or following library preparation, if at all.

[00161] In some embodiments, an adapter pair can be attached to cfDNA fragments in a sample such as by ligation or PCR amplification. In some embodiments, a pair of adapters can comprise a p5 adapter that is attached to a 5' end of a molecule and a p7 adapter that is attached to a 3' end of a molecule. In some embodiments, a p5 and p7 sequence can allow a nucleic acid library to bind and generate clusters on a flow cell.

[00162] In some cases, the cfDNA can be attached to adapters comprising identifier sequences that can differentiate between multiple samples. In some embodiments, the multiple samples comprise a plurality of patient samples and/or control samples (e.g., positive control, negative control). In some embodiments, samples can be pooled after barcoding, then sequenced, then de-multiplexed to assign each cluster to its sample.

Exemplary Process 2– double-stranded cfDNA and single-stranded cfDNA

[00163] In some embodiments, the methods provided herein comprise performing process 2 to prepare a sample for high throughput sequencing assay. Process 2 provides an example of preparing a sequencing library from the double-stranded and single-stranded cfDNA in the original sample. In some embodiments, a control molecule can be added to an initial sample of plasma to generate a spiked plasma sample. In some embodiments, nucleic acid extraction can be performed on a spiked plasma sample to generate purified and concentrated cfDNA. In some embodiments, a sample preparation method does not comprise extracting nucleic acids from a

raw or initial sample. For example, in some cases, nucleic acids can be extracted during or following library preparation, if at all.

[00164] In some embodiments, Process 2 comprises physically enriching for degraded nucleic acids, ultra-short nucleic acids, single stranded nucleic acids, double stranded nucleic acids, nicked nucleic acids, or rare nucleic acids.

[00165] In some embodiments, the cfDNA can be denatured in order to separate strands of double-stranded cfDNA, converting the double-stranded cfDNA into single-stranded cfDNA. In some embodiments, the resulting sample can comprise single-stranded cfDNA derived from both double-stranded and single-stranded cfDNA in the original sample.

[00166] In some embodiments, the methods described herein can comprise attaching a splint adapter (e.g., a splint oligonucleotide) to the single-stranded cfDNA in the sample. In some embodiments, the splint adapter can comprise dsDNA with a ssDNA overhang. In some cases, the ssDNA overhang can comprise random or degenerate nucleotides (e.g., NNNNNN). In some embodiments, the ssDNA overhang can comprise a specific sequence capable of hybridizing to a target molecule.

[00167] In some embodiments, the splint adapter can randomly hybridize to cfDNA via the single-stranded DNA overhang within the splint adapter. In some embodiments, attaching the splint adapter can further comprise ligating the hybridized adapter to cfDNA, e.g., using an enzyme (e.g., ligase). In some embodiments, the ligase comprises T4 ligase, CircLigase II, CircLigase ssDNA Ligase, Splint ligase, any engineered variants thereof, any natural variants thereof, or any combination thereof.

[00168] In some embodiments, an adapter pair can be attached to cfDNA fragments in a sample such as by ligation or PCR amplification. In some embodiments, a pair of adapters can comprise a p5 adapter that is attached to a 5' end of a molecule and a p7 adapter that is attached to a 3' end of a molecule. In some embodiments, a p5 and p7 sequence can allow a nucleic acid library to bind and generate clusters on a flow cell.

[00169] In some cases, the cfDNA can be attached to adapters comprising identifier sequences that can differentiate between multiple samples. In some cases, the sample identifiers attached to the cfDNA via ligation or PCR amplification. In some embodiments, the sample identifiers are incorporated into the splint oligonucleotides. In some embodiments, the multiple samples comprise a plurality of patient samples and/or control samples (e.g., positive control, negative control). In some embodiments, multiple samples can be pooled after barcoding, then sequenced, then de-multiplexed to assign each cluster to its sample. Sequencing Methods

[00170] Disclosed herein are methods of sequencing a sample comprising a nucleic acid as disclosed herein. In some embodiments, disclosed herein are methods for sequencing nucleic

acids, cell-free nucleic acids (cfNAs), or a combination thereof present in a sample. In some embodiments, a cfNA as disclosed herein can be prepared into a nucleic acid library for use in a sequencing assay. As used herein, a "nucleic acid library" refers to a collection of nucleic acid fragments. In some embodiments, a collection of nucleic acid fragments can be used, for example, for sequencing. Disclosed herein in some embodiments are methods comprising subjecting nucleic acids in a nucleic acid library derived from a sample to a sequencing assay to generate sequence reads. In some embodiments, the methods further comprise performing any necessary steps, methods, or techniques to prepare the sample for the sequencing assay. For example, the methods may comprise adding process control molecules (or synthetic spike-in molecules) to the sample, ligating adapters to the nucleic acids, or generating a nucleic acid library. In some embodiments, the methods further comprise analyzing the sequence reads generated from the sequencing assay. In some embodiments, analyzing the sequence reads comprises performing a bioinformatic analysis. In some embodiments, a bioinformatic analysis can comprise calculating an abundance of sequence reads, generating fragment length profiles of sequence reads, identifying and mapping sequence reads to known references to identify a source organism from which a sequenced nucleic acid was derived, or any combination thereof. In some embodiments, a method can further comprise generating a report of a result of a sequencing assay.

[00171] In some embodiments, a sequencing assay provided herein can be performed by any sequencing methods suitable for sequencing a nucleic acid provided herein. In some embodiments, the sequencing assay herein can comprise massively parallel sequencing. In some embodiments, the massively parallel sequencing can comprise whole genome sequencing. In some embodiments, a massively parallel sequencing can comprise Next Generation Sequencing or a Next Next Generation sequencing. In some embodiments, the methods provided herein comprise determining a concentration or quantity of a mcfDNA. In some embodiments, the methods comprise monitoring a concentration or quantity of a mcfDNA over time. In some embodiments, the method comprises identifying fragments of mcfDNA that vary during a course of treatment. In some embodiments, the sequencing assay or the sequencing method comprises sequencing-by-synthesis. In some embodiments, sequencing methods provided herein comprise Maxam-Gilbert sequencing-based techniques, chain-termination-based techniques, shotgun sequencing, bridge PCR sequencing, single-molecule real-time sequencing, ion semiconductor sequencing, nanopore sequencing, pyrosequencing, sequencing by synthesis, sequencing by ligation, sequencing by electron microscopy, dideoxy sequencing reactions (Sanger method), massively parallel sequencing, polony sequencing, DNA nanoball sequencing and any variation thereof. The term "Next Generation Sequencing (NGS)" herein refers to sequencing methods that allow for massively parallel sequencing of nucleic acid molecules during which a plurality,

e.g., millions, of nucleic acid fragments from a single sample or from multiple different samples are sequenced simultaneously. Non-limiting examples of NGS include sequencing-by-synthesis, sequencing-by-ligation, real-time sequencing, and nanopore sequencing. In some embodiments, sequencing involves hybridizing a primer to the template to form a template/primer duplex, contacting the duplex with a polymerase enzyme in the presence of detectably labeled or unlabeled nucleotides under conditions that permit the polymerase to add labeled or unlabeled nucleotides to the primer in a template-dependent manner, detecting a signal from the incorporated labeled nucleotide or detecting a signal resulting from the process of incorporating labeled or unlabeled nucleotide (e.g., proton release), and sequentially repeating the contacting and/or detecting steps at least once, wherein sequential detection of incorporated labeled or unlabeled nucleotide determines the sequence of the nucleic acid. In some embodiments, exemplary detectable labels include radiolabels, fluorescent labels, protein labels, dye labels, enzymatic labels, etc. In some embodiments, the detectable label can be an optically detectable label, such as a fluorescent label. Exemplary fluorescent labels include cyanine, rhodamine, fluorescein, coumarin, BODIPY®, Alexa Fluor®, or conjugated multi-dyes.

[00172] In some embodiments, a method disclosed herein can comprise calculating an abundance of sequence reads generated from a nucleic acid present in a sample. In some embodiments, a method can comprise calculating an abundance of nucleic acids (e.g., cell-free nucleic acids (cfNA)) in a sample. In some embodiments, an abundance of nucleic acids in a sample can be calculated based on an abundance of sequence reads generated from a nucleic acid in a sample. In some embodiments, an abundance can comprise an absolute abundance, a relative abundance, or a normalized abundance. In some embodiments, a relative abundance or a normalized abundance can be calculated based on a reference value. In some embodiments, a reference value can comprise an abundance of sequence reads generated from any nucleic acids in the sample. For example, the reference value may be an abundance of sequence reads from microbial cell-free DNA (mcfDNA) or of sequence reads from a synthetic spike-in molecule (or process control molecule). In some embodiments, the methods further comprise calculating the relative abundance of a sequence read at least in part by comparing the abundance of the sequence read to the abundance of another sequence read. In some embodiments, the methods further comprise calculating the normalized abundance of a sequence read at least in part by comparing the abundance of the sequence read from mcfDNA in the sample to the abundance of sequence reads from synthetic spike-in molecules.

[00173] In some embodiments, the methods further comprise mapping a sequence read to a reference sequence. In some embodiments, the reference sequence can comprise a microbial sequence. In some embodiments, the reference sequence comprises a non-microbial sequence. In

some embodiments, the reference sequence comprises a genomic sequence. In some embodiments, the method further comprise mapping a sequence read to a reference genome. In some embodiments, a method disclosed herein can comprise identifying a species of microbe by mapping a sequence read to a reference genome. In some embodiments, the methods further comprise identifying the source of nucleic acids in the sample from which the sequence read originated. In some embodiments, the methods further comprise performing a bioinformatic analysis. In some embodiments, performing the bioinformatics analysis comprises assembling sequence data, detecting and quantifying sequence reads, distinguish populations of nucleic acids, detecting the presence and measuring the abundance of microbial nucleic acids, comparing sequence reads, comparing abundances of sequence reads, identifying contaminant nucleic acids from the sample collection site, identifying target nucleic acids (e.g., cell-free nucleic acids), identifying host nucleic acids, generating fragment lengths profiles of microbial nucleic acids, generating fragment lengths profiles of control process molecules, comparing fragment lengths profiles of the microbial nucleic acids, detecting site of infection, detecting the state of infection, detecting the risk of organ rejection in a transplant patient, determining the eligibility of a subject for a transplant, detecting potential for drug resistance, or any combination thereof.

[00174] Disclosed herein in some embodiments are methods for identifying sequence reads obtained through sequencing as host or non-host. In some embodiments, a host can comprise a subject. In some embodiments, sequence reads identified as non-host can then be aligned to a nucleotide database. In some embodiments, the nucleotide database comprises microbial reference sequences. In some embodiments, the database can be selected for those microbial sequences known to be associated with the host, e.g., the set of commensal and pathogenic microorganisms of the subject (e.g., animal or human). In some embodiments, the microbial database can be optimized to mask or remove contaminating sequences. For example, many public database entries include artifactual sequences not derived from the microorganism, e.g., primer sequences, host sequences, and other contaminants. In some embodiments, sequence reads can be aligned to a reference sequence comprising artifactual sequences. In some embodiments, regions that show irregularities in read coverage when multiple samples are aligned can be masked or removed as an artifact. In some embodiments, the detection of such irregular coverage can be done by various metrics, such as the ratio between coverage of a specific nucleotide and the average coverage of the entire contig within which this nucleotide is found. In some embodiments, a sequence that is represented as greater than about 5×, about 10×, about 25×, about 50×, about 100× the average coverage of the reference sequence comprising artifactual sequences can be artifactual. In some embodiments, a binomial test can be applied to provide a per-base likelihood of coverage given the overall coverage of the contig. In some

embodiments, each high confidence read can align to multiple organisms in the given microbial database. In some embodiments, to correctly assign organism abundance based upon this possible mapping redundancy, an algorithm can be used to compute the most likely organism (for example, see Lindner et al. Nucl. Acids Res. (2013) 41 (1): e10, which is referenced herein in its entirety). For example, GRAMMy or GASiC algorithms can be used to compute the most likely organism that a given read came from. In some embodiments, alignments and assignment to a host sequence or to a non-host (e.g., microbial) sequence can be performed in accordance with art-recognized methods. For example, a read of 50 nt. can be assigned as matching a given genome if there is not more than 1 mismatch, not more than 2 mismatches, not more than 3 mismatches, not more than 4 mismatches, not more than 5 mismatches, etc. over the length of the read. In some embodiments, publicly available algorithms can be used for alignments and identification. A non-limiting example of such an alignment algorithm is the bowtie2 program (Johns Hopkins University). In some embodiments, these assignments of reads to an organism (e.g., host organism, non-host organism, microbe, pathogen, etc.) can then be totaled and used to compute the estimated number of reads assigned to each organism in a given sample, in a determination of the prevalence of the organism in the sample (for example, a cell-free nucleic acid sample). In some embodiments, this information can be used to determine an origin of a pathogen or contaminant. In some embodiments, the analysis can be used to normalize the counts for the size of the microbial genome to provide a calculation of coverage for a microbe. In some embodiments, the normalized coverage for each microbe can be compared to the host sequence coverage in the same sample to account for differences in sequencing depth between samples. In some embodiments, a dataset of microbial organisms represented by sequences in the sample, and the prevalence of those microorganisms can be optionally aggregated and displayed for ready visualization, e.g., in the form of a report.

[00175] In some embodiments, the methods provided herein comprise generating sequence reads (or sequencing reads) from the nucleic acids in the sample. "Sequence read", "sequenced read" and "sequencing read" are used interchangeably herein. In some embodiments, the methods further comprise detecting, identifying, analyzing, processing, comparing, aligning, or mapping the sequence reads generated from the nucleic acids in the sample. In some embodiments, the sequence reads are generated from the microbial cell-free nucleic acids (mcfNA) in the sample. In some embodiments, the methods comprise mapping the sequence reads generated from the nucleic acids in the sample to a reference sequence. In some embodiments, the reference sequence can comprise a microbial sequence. In some embodiments, the microbial sequence comprises a region of a microbial genome. In some embodiments, the reference sequence comprises an artifactual sequence.

[00176] In some embodiments, a sequencing can generate at least 100, at least 250, at least 500, at least 750, at least 1000, at least 1500, at least 2500, at least 3000, at least 3500, at least 4000, at least 4500, at least 5000, at least 5500, at least 6000, at least 7000, at least 8000, at least 9000, at least 10,000, at least 12,500, at least 15,000, at least 17,500, at least 20,000, at least 30,000, at least 40,000, at least 50,000, at least 60,000, at least 70,000, at least 80,000, at least 90,000, at least 100,000, at least 200,000, at least 300,000, at least 400,000, at least 500,000, at least 600,000, at least 700,000, at least 800,000, at least 900,000, at least 1,000,000, at least 2,000,000, at least 3,000,000, at least 4,000,000, at least 5,000,000, at least 6,000,000, at least 7,000,000, at least 8,000,000, at least 9,000,000, or at least 10,000,000 sequence reads from mcfNA in a sample.

Hardware and Software

[00177] Disclosed herein in some embodiments are kits and systems configured to perform a method disclosed herein. In some embodiments, a kit or system can comprise a nucleic acid sequencer for generating DNA or RNA sequence information. In some embodiments, a kit or system can further comprise a computer comprising software that performs a bioinformatics analysis on the DNA or RNA sequence information. In some embodiments, a bioinformatics analysis can include, without limitation, assembling sequence data, detecting and quantifying sequence reads, distinguish populations of nucleic acids, detecting the presence and measuring the abundance of microbial nucleic acids, comparing sequence reads, comparing abundances of sequence reads, identifying contaminant nucleic acids from the sample collection site, identifying target nucleic acids (e.g., cell-free nucleic acids), identifying host nucleic acids, generating fragment lengths profiles of microbial nucleic acids, generating fragment lengths profiles of control process molecules, comparing fragment lengths profiles of the microbial nucleic acids, detecting site of infection, detecting the state of infection, detecting the risk of organ rejection in a transplant patient, determining the eligibility of a subject for a transplant, and/or detecting potential for drug resistance.

[00178] The kit or system can also include computer control systems with machine-executable instructions (e.g., software) to implement the methods. **FIG. 6** shows a computer system 601 that is programmed or otherwise configured to implement methods of the present disclosure The computer system 601 includes a central processing unit (CPU, also "processor" and "computer processor" herein) 605, which can be a single core or multi core processor, or a plurality of processors for parallel processing. The computer system 601 also includes memory or memory location 610 *(e.g.,* random-access memory, read-only memory, flash memory), electronic storage unit 615 *(e.g.,* hard disk), communication interface 620 *(e.g.,* network adapter) for

communicating with one or more other systems, and peripheral devices 625, such as cache, other memory, data storage and/or electronic display adapters.

[00179] The memory 610, storage unit 615, interface 620, and peripheral devices 625 are in communication with the CPU 605 through a communication bus (solid lines), such as a motherboard. The storage unit 615 can be a data storage unit (or data repository) for storing data.

[00180] The computer system 601 can be operatively coupled to a computer network ("network") 630 with the aid of the communication interface 620. The network 630 can be the Internet, an internet and/or extranet, or an intranet and/or extranet that is in communication with the Internet. The network 630 in some embodiments is a telecommunication and/or data network. The network 630 can include one or more computer servers, which can enable distributed computing, such as cloud computing.

[00181] The network 630, in some embodiments with the aid of the computer system 601, can implement a peer-to-peer network, which can enable devices coupled to the computer system 601 to behave as a client or a server.

[00182] The CPU 605 can execute a sequence of machine-readable instructions, which can be embodied in a program or software. The instructions can be stored in a memory location, such as the memory 610. The instructions can be directed to the CPU 605, which can subsequently program or otherwise configure the CPU 605 to implement methods of the present disclosure. Examples of operations performed by the CPU 605 can include fetch, decode, execute, and writeback. The CPU 605 can be part of a circuit, such as an integrated circuit. One or more other components of the system 601 can be included in the circuit. In some embodiments, the circuit is an application specific integrated circuit (ASIC).

[00183] The storage unit 615 can store files, such as drivers, libraries, and saved programs. The storage unit 615 can store user data, *e.g.,* user preferences and user programs. The computer system 601 in some embodiments can include one or more additional data storage units that are external to the computer system 601, such as located on a remote server that is in communication with the computer system 601 through an intranet or the Internet.

[00184] The computer system 601 can communicate with one or more remote computer systems through the network 630. For instance, the computer system 601 can communicate with a remote computer system of a user. Examples of remote computer systems include personal computers, slate or tablet PC's, telephones, smart phones, or personal digital assistants. The user can access the computer system 601 via the network 630.

[00185] The kit or system can be implemented by way of machine *(e.g.,* computer processor) executable code stored on an electronic storage location of the computer system 601, such as, for example, on the memory 610 or electronic storage unit 615. The machine executable or machine-

readable code can be provided in the form of software. During use, the code can be executed by the processor 605. In some embodiments, the code can be retrieved from the storage unit 615 and stored on the memory 610 for ready access by the processor 605. In some situations, the electronic storage unit 615 can be precluded, and machine-executable instructions are stored on memory 610. The code can be pre-compiled and configured for use with a machine having a processor adapted to execute the code or can be compiled during runtime. The code can be supplied in a programming language that can be selected to enable the code to execute in a pre-compiled or as-compiled fashion.

[00186] Parts of the kits and systems, such as the computer system 601, can be embodied in programming. Various aspects of the technology can be thought of as "products" or "articles of manufacture" typically in the form of machine (or processor) executable code and/or associated data that is carried on or embodied in a type of machine readable medium. Machine-executable code can be stored on an electronic storage unit, such as memory *(e.g.,* read-only memory, random-access memory, flash memory) or a hard disk. "Storage" type media can include any or all of the tangible memory of the computers, processors or the like, or associated modules thereof, such as various semiconductor memories, tape drives, disk drives and the like, which can provide non-transitory storage at any time for the software programming. All or portions of the software can at times be communicated through the Internet or various other telecommunication networks. Such communications, for example, can enable loading of the software from one computer or processor into another, for example, from a management server or host computer into the computer platform of an application server. Thus, another type of media that can bear the software elements includes optical, electrical, and electromagnetic waves, such as used across physical interfaces between local devices, through wired and optical landline networks and over various air-links. The physical elements that carry such waves, such as wired or wireless links, optical links, or the like, also can be considered as media bearing the software. As used herein, unless restricted to non-transitory, tangible "storage" media, terms such as computer or machine "readable medium" refer to any medium that participates in providing instructions to a processor for execution.

[00187] Hence, a machine readable medium, such as computer-executable code, can take many forms, including but not limited to, a tangible storage medium, a carrier wave medium or physical transmission medium. Non-volatile storage media include, for example, optical or magnetic disks, such as any of the storage devices in any computer(s) or the like, such as can be used to implement the databases, *etc.* shown in the drawings. Volatile storage media include dynamic memory, such as main memory of such a computer platform. Tangible transmission media include coaxial cables; copper wire and fiber optics, including the wires that comprise a bus within a computer

system. Carrier-wave transmission media can take the form of electric or electromagnetic signals, or acoustic or light waves such as those generated during radio frequency (RF) and infrared (IR) data communications. Common forms of computer-readable media therefore include for example: a floppy disk, a flexible disk, hard disk, magnetic tape, any other magnetic medium, a CD- ROM, DVD or DVD-ROM, any other optical medium, punch cards paper tape, any other physical storage medium with patterns of holes, a RAM, a ROM, a PROM and EPROM, a FLASH-EPROM, any other memory chip or cartridge, a carrier wave transporting data or instructions, cables or links transporting such a carrier wave, or any other medium from which a computer can read programming code and/or data. Many of these forms of computer readable media can be involved in carrying one or more sequences of one or more instructions to a processor for execution.

[00188] The computer system 601 can include or be in communication with an electronic display 635 that comprises a user interface (UI) 640 for providing, an output of a report, which can include a diagnosis of a subject or a therapeutic intervention for the subject. Examples of UI's include, without limitation, a graphical user interface (GUI) and web-based user interface. The analysis can be provided as a report. The report can be provided to a subject, to a health care professional, a lab-worker, or other individual.

[00189] Methods and systems of the present disclosure can be implemented by way of one or more algorithms. An algorithm can be implemented by way of software upon execution by the central processing unit 605. The algorithm can, for example, facilitate the enrichment, sequencing and/or detection of pathogen or microbe or other target nucleic acids.

[00190] Information about a patient or subject can be entered into a computer system, for example, patient background, patient medical history, or medical scans. The computer system can be used to analyze results from a method described herein, report results to a patient or doctor, or come up with a treatment plan.

Applications

[00191] In some embodiments, the methods provided comprise detecting a clinically-relevant genetic marker. In some embodiments, the genetic marker is associated with an infection or a non-communicable disease. In some embodiments, the methods are useful for detecting an anti-microbial resistance (AMR) marker. In some embodiments, the methods are useful for detecting a cancer marker. In some embodiments, the methods are useful for detecting a disease site.

[00192] In some embodiments, the methods provided herein are useful for informing diagnostic and treatment decisions for infectious diseases in clinical settings. The methods provided herein can be useful for rapid detection of the microbe causing the infection without performing

microbial culture. In some embodiments, the methods can be useful for detecting an AMR genetic marker, an AMR gene cassette, and/or a carrier microbe harboring the AMR genetic marker. In some embodiments, the methods comprise generating a report listing the microbes detected by the methods. In some embodiments, the methods comprise generating a report listing the microbes and probabilities of the microbes being resistant or susceptible to antimicrobial treatment for clinical use. In some embodiments, the methods comprise generating a report on the microbes and their resistance or susceptibility to antimicrobial treatment.

[00193] In some embodiments, the methods provided herein are useful for linking the AMR genetic marker to carrier microbes harboring the AMR genetic marker. In some embodiments, by detection of the AMR genetic marker can provides additional phenotypic characteristics of the carrier microbe. In some embodiments, a phenotypic characteristic is resistance to a particular AMR drug or class of AMR drugs. In some embodiments, the methods comprise performing a first sequencing assay to detect one or more carrier microbes harboring the AMR genetic marker and based on this detection, perform a second sequencing assay to link the AMR genetic marker to a specific carrier microbe. For example, many carrier microbes can harbor the same AMR gene, making it important to identify the exact microbes in the infection that carry the gene to better inform treatment. For instance, the mecA gene can be carried by multiple Staphylococcus species, such as S. aureus and S. epidermidis, so mecA could be present in either or both microbes in the infection. The methods provided herein can be useful for determining the likely etiology of the infection and establishing the microbial species containing the mecA gene. In some embodiments, the methods further comprise administering a therapeutic or developing a treatment plan based on the identification of a microbial species, such as a microbial species carrying an AMR marker.

[00194] In some embodiments, the methods are useful for guiding the selection of antimicrobial treatment. For example, these methods can help determine the appropriate class of antimicrobial drugs for an infection based on the results of AMR detection. Once a link is established between an AMR genetic marker and the detected carrier microbe, the treatment plan for the infection can be adjusted accordingly based on the results generated by the methods.

[00195] In some embodiments, the methods comprise modifying the antimicrobial drug class administered to a subject when microbes harboring resistance genes are detected. When such resistance genes are detected, in some embodiments, the methods comprise administering an alternative regimen to the subject. For example, if the methods detect that Staphylococcus aureus infecting the patient harbors the SCCmec cassette, mecA gene, and/or mecC gene, this indicates that the infection is caused by methicillin-resistant Staphylococcus aureus (MRSA), which is resistant to all β-lactams except ceftaroline and ceftobiprole. As a result, the clinical treatment

can shift from β-lactams to alternative primary regimens such as vancomycin, linezolid, daptomycin, or ceftobiprole. For example, if the methods detect that Enterococcus infecting the patient harbors the VanA or VanB gene, this indicates that the infection may be caused by vancomycin-resistant Enterococcus (VRE). VRE is often resistant to penicillins, aminoglycosides, and glycopeptides. As a result, the clinical treatment would shift from vancomycin or penicillin-based therapies to alternative regimens such as linezolid or daptomycin. For example, if the methods detect that the carrier microbes harbor the blaCTX-M gene, this indicates the presence of an extended-spectrum β-lactamase (ESBL)-producing strain. ESBL-producing bacteria are generally resistant to penicillins, penicillin-BLI combinations, most cephalosporins (except cephamycins), and aztreonam. As a result, the clinical treatment would shift from β-lactam antibiotics to alternative regimens such as ceftolozane-tazobactam, ertapenem, imipenem-cilastatin, meropenem, or aminoglycosides. For example, if the methods detect that the infecting bacteria harbor the blaKPC gene, this indicates the presence of Klebsiella pneumoniae carbapenemase (KPC)-producing bacteria. KPC producers are resistant to penicillins, penicillin-BLI combinations, cephalosporins, carbapenems, and aztreonam. As a result, the clinical treatment would shift from traditional β-lactam therapies to alternative regimens such as ceftazidime-avibactam, meropenem-vaborbactam, imipenem-cilastatin-relebactam, cefiderocol, or certain aminoglycosides.

[00196] In some embodiments, the methods provided herein are useful for identification of multiple species involved in an infection or a co-infection. In some embodiments, the methods comprise non-biased sequencing of microbial nucleic acids and can identify many microbes present in an infection. This is particularly useful for diagnosing and treating patients susceptible to multiple infections, including co-infections and complex infections, such as those occurring in immunocompromised individuals. In some embodiments, the methods comprise identifying microbes from different kingdoms within an infection, such as a viral, bacterial, and fungal infection occurring simultaneously. For example, immunocompromised patients, such as those with human immunodeficiency virus (HIV), may experience co-infections involving multiple pathogens. In some embodiments, the methods identify multiple resistant bacteria in difficult-to-treat infections. For instance, chronic infections in cystic fibrosis patients are particularly challenging to manage due to the persistence of multidrug-resistant bacteria. In some embodiments, the methods are useful for identifying resistance markers across a broad range of microbes. For example, the methods can detect clinically relevant AMR genes in Cytomegalovirus (CMV) and other DNA viruses, the rpoB gene in Mycobacterium tuberculosis, additional classes of carbapenemase and ESBL genes in Gram-negative bacteria, and antifungal resistance genes in Aspergillus spp. and other molds.

[00197] In some embodiments, the methods provided herein are useful for rapid diagnosis of life-threatening infections, particularly in immunocompromised or critically ill patients, such as those undergoing chemotherapy, organ transplants, or long-term immunosuppressive therapy. For example, these methods can be used to diagnose endocarditis, pneumonia in immunocompromised hosts (ICH), invasive fungal infections, febrile neutropenia, and/or fever of unknown origin (FUO). The methods described herein offer a rapid and comprehensive approach to identifying causative pathogens, improving targeted treatment, and minimizing the use of broad-spectrum antimicrobials.

[00198] In some embodiments, the methods provided herein can be used independently of microbial culturing or can be used to confirm the results of a microbial culture. The methods can offer various advantages over traditional culturing, including faster turnaround times and the ability to detect microbes that are difficult to culture. In some embodiments, the methods described herein are more sensitive and/or more accurate than microbial culturing. For example, even when a patient's blood culture fails to show any organism, the methods provided herein can detect microbial nucleic acids from pathogens causing the infection in the patient's blood sample, thereby identifying the etiology of the infection. In some embodiments, the methods provided herein can be used independently of other detection methods, such as method that depend solely on PCR. The methods provided herein can offer various advantages over methods that solely use PCR. For example, the methods provided herein can detect multiple microbial nucleic acids in a biological sample in a high-throughput manner, providing unique insights into the patient's microbiome, infectome, and resistome that may be unavailable with other detection methods. In some embodiments, the methods described herein more accurately detect the microbe infecting the patient and more accurately link the detection of an AMR genetic marker to its carrier microbe.

[00199] In some embodiments, the methods provided herein are useful for monitoring the development of AMR genetic markers in an infected patient or a population of patients over time, providing longitudinal data on microbial resistance. In some embodiments, the methods monitor AMR genetic marker development at clinical locations, such as hospital wards. The methods may be useful for preventing the spread of antimicrobial-resistant infections in clinical settings. In some instances, the methods can help reduce the acquisition and spread of antimicrobial resistance due to the inappropriate use of antimicrobials, both in individual patients and among populations.

[00200] In some embodiments, the methods provided herein are useful for treating a patient who has a cancer, wherein the cancer harbors a cancer marker or multiple markers for cancer. For example, a cancer marker can be detected by high throughput sequencing; followed by a targeted assay from a different aliquot of a sample in order to detect additional cancer markers. In some

embodiments, the methods comprise determining, adjusting, or administering a cancer therapy provided herein to the patient.

[00201] In some embodiments, an mcfNA in a subject's bodily fluid can originate from a microbe living in or on a subject. In some embodiments, a mcfNA can be detected in a subject that has been exposed to an infectious disease. In some embodiments, an mcfNA can be detected in a healthy individual. In some embodiments, detection of mcfNA by the methods disclosed herein can be performed as a biomarker of infection. In some embodiments, mcfNAs can be associated with communicable and/or non-communicable diseases. In some embodiments, mcfNAs can be associated with a range of diseases and conditions of a subject described herein, including an infection, an inflammatory bowel disease (IBD), a Kawasaki disease (KD), a human immunodeficiency virus (HIV), a cardiovascular disease (CVD), a cystic fibrosis (CF), a pneumonia, a sepsis, a cancer, a gastric cancer (GC), a hepatocellular carcinoma (HCC), a melanoma, or any combination thereof.

[00202] In some embodiments, the methods provided herein comprise detecting, diagnosing, treating, monitoring, staging, or prognosing a disease or a disorder. In some embodiments, the disease or disorder is an infection or another medical indication related to an infection. In some embodiments, the methods are for determining an infection site or identifying a source of an infection. In some embodiments, the methods are for determining the biological relationship between a microbe and a host (e.g., the subject herein). In some embodiments, the methods are for detecting or identifying a commensal microbe of the subject. In some cases, a commensal microbe of a first subject may be a potential pathogen to a second subject. For example, a commensal microbe of an animal may be a potential pathogen to a human. In some embodiments, methods disclosed herein can be used in conjunction with one or more medical tests.

[00203] In some embodiments, the methods described herein are for determining an eligibility of a subject in transplantation. The subject described herein can be a donor or a recipient of the transplantation. In some embodiments, the transplantation comprises organ transplantation, tissue transplantation, composite tissue transplantation, living donor transplantation, or xenotransplantation. In some embodiments, the methods are for determining the eligibility of an animal donor for xenotransplant. In some embodiments, the methods are for determining the eligibility of a human donor for organ transplant. In some embodiments, the methods are for determining the eligibility of a human transplant recipient.

[00204] In some embodiments, the methods described herein are for individualized treatment for an infected subject or a subject who is susceptible or at risk for infections (e.g., immunosuppressed, immunocompromised, living conditions, or genetic variations resulting in increased susceptibility for infection). In some embodiments, individualized treatment can

include predicting if an infection will progress to an invasive disease stage, monitoring the efficacy of a therapy in a subject, modifying a therapeutic regimen depending on the subject's response to the therapy, and determining a pathogen's resistance to a particular therapeutic.

[00205] In some embodiments, the methods disclosed herein can be used to detect, diagnose, predict, or prognose a pathogen's resistance to a particular therapeutic. In some cases, the methods disclosed herein can further comprise sequencing of the subject's DNA for genetic variations that are associated with therapeutic resistance to therapeutics or to a particular therapeutic.

[00206] The methods provided herein can comprise, in some embodiments, determining a subject's response to a particular treatment. In some cases, samples can be collected serially at various times before or during the course of the infection to determine the pathogen's and subject's response to a treatment. In some cases, a treatment plan is individually tailored to the subject. In some embodiments, samples can be collected at various timepoints following administration of a treatment. In some embodiments, samples can be collected at various timepoints following discontinuation of a treatment or adjustment of a treatment. In some cases, serially collected samples are compared to each other to determine whether the infection is improving or worsening in the subject. In some embodiments, after a response to treatment is identified, the methods can comprise maintaining, discontinuing, modifying, adjusting the dose of a particular treatment. In some cases, modifying the treatment may comprise replacing a therapeutic drug with a different or alternative therapeutic drug, particularly if an AMR marker is detected following treatment. In some cases, the methods comprise discontinuing an antimicrobial treatment if an AMR gene or cassette is identified. In some cases, if an anti-microbial resistance is detected, the methods comprise changing a treatment to a different therapy, or alternative regimen. In some cases, the methods can comprise adjusting a dose of current treatment, either up or down, based on detection of an AMR marker.

[00207] In some embodiments, the methods described herein can be used to adjust a therapeutic regimen. For example, the subject can be administered a drug to treat an infection. In some embodiments, methods provided herein can be used to track or monitor the efficacy of the drug treatment. In some cases, the therapeutic regimen can be adjusted, depending on upward or downward course of the infection. For example, if the methods provided herein indicate that an infection is not improving with drug treatment, the therapeutic regimen can be adjusted by changing the type of drug or treatment, discontinuing the use of the drug, continuing the use of the drug, increasing the dose of the drug, or adding a new drug or treatment to the subject's therapeutic regimen.

[00208] In some embodiments, the methods provided herein are useful for treating a patient who has an infection by one or more carrier microbes harboring an AMR marker. In some embodiments, the methods comprise determining, adjusting, changing, or administering an antimicrobial treatment of a subject based on identification of a multiple microbes.

[00209] In some embodiments, a treatment can involve administering a drug or other therapy to reduce or eliminate the colonization or invasive disease associated with an infection. In some cases, the subject can be treated prophylactically to prevent the development of an infection. The methods provided herein can comprise performing procedures or administering a treatment to improve or reduce the symptoms of an infection. Some nonlimiting exemplary drugs that can be used are antibiotics (such as ampicillin, sulbactam, penicillin, vancomycin, gentamycin, aminoglycosides, clindamycin, cephalosporin, metronidazole, timentin, ticarcillin, clavulanic acid, cefoxitin), antiretroviral drugs (e.g., highly active antiretroviral therapy (HAART), reverse transcriptase inhibitors, nucleoside/nucleotide reverse transcriptase inhibitors (NRTIs), Non-nucleoside RT inhibitors, and/or protease inhibitors), immunoglobulins, or any variant or combination thereof.

[00210] In some embodiments, the methods described herein are for detection, monitoring, diagnosis, prognosis, treatment, prediction, or prevention of colonization by the microbes disclosed herein. The disclosure also provides methods to detect, monitor, diagnose, prognose, treat, predict, or prevent invasive disease caused by the microbes described herein. The methods of the disclosure can be applied to any pathogen that has various stages of infection, including early stages. The methods can be especially useful for pathogens that have a colonization stage and an invasive disease stage. In some cases, the invasive disease stage can be caused by the pathogen infection. In some cases, the invasive disease stage can be associated with the pathogen infection.

[00211] In some embodiments, the methods can be used for to distinguish populations of nucleic acids or for detecting a microbe in a subject. In some embodiments, the methods provide a more comprehensive view of the state and diversity of the infection or symbiotic microbes in a subject. For example, the identification of both RNA and DNA in a sample can be useful to detect RNA and DNA type viruses, or to detect bacterial, protist, parasitic or fungal genomic DNA and/or gene expression products, e.g., mRNA. Such process can also be able to differentiate between latent infection (e.g., which might be indicated by the presence of integrated retroviral DNA) versus active infection (e.g., which might be indicated by the presence of viral RNA from intact viral particles). Such processes can also be able to detect drug resistance and/or the origin of infection. Such processes can also be used to analyze host response. Such analyses can include analysis of cell-free, circulating nucleic acids, e.g., for microbial or viral infection identification.

Subjects

[00212] In some embodiments, the sample described herein has been obtained from a subject. In some embodiments, the subject can comprise a human or a non-human animal. In some embodiments, a subject can comprise a male or a female. In some embodiments, a subject can be of any age. In some embodiments, a subject can be a child. In some embodiments, a subject can comprise an embryo or a fetus.

[00213] In some embodiments, a subject can comprise a healthy subject. In some embodiments, a subject can have, be suspected of having, or be at risk of having a disease or a disorder. In some embodiments, the subject has an infection by a microbe. In some embodiments, the subject is infected by a pathogenic microbe, a commensal microbe, or a combination thereof. In some embodiments, the subject is infected by a carrier microbe harboring one or more AMR genetic marker provided herein. In some embodiments, the subject has a medical indication related to an infection or an abnormal immune response to an infection. In some embodiments, the subject is wholly or partially immunocompromised or abnormally susceptible to infections. In some embodiments, the subject has a reoccurring infection by a microbe, a secondary infection by another microbe, or a co-infection by two or more microbes. In some embodiments, the subject has a disorder or a disease comprising a cancer, transplantation, a surgery, a burn, an infection, a malnourishment, a chronic kidney disease, diabetes mellitus, an autoimmune disease or disorder, or an immune disorder (e.g., an acquired immunodeficiency syndrome (AIDS)). In some embodiments, a subject can have a medical condition. In some embodiments, a medical condition can comprise pregnancy, lactation, menopause, frailty, malnutrition, graft or organ rejection, chronic fatigue, or a combination thereof.

[00214] In some embodiments, the subject is taking an immunosuppressant agent. The immunosuppressant agent may comprise chemotherapy, radiation, corticosteroids, transplant medications, or certain biologics.

[00215] In some embodiments, the subject is taking an anti-infective agent provided herein. In some embodiments, the anti-infective agent comprises an antibacterial drug (e.g., an antibiotic), an antiviral drug, an antifungal drug, an antiparasitic drug, or any combination thereof. In some embodiments, the subject is taking one anti-infective agent without improvement of symptoms and needs a different anti-infective regimen. In some embodiments, the subject is taking one anti-infective agent but has developed an infection by the one or more carrier microbes harboring an AMR genetic marker. In some embodiments, the subject has a co-infection of one or more microbes, wherein the one or more microbes comprise one or more carrier microbes harboring an AMR genetic marker.

[00216] In some embodiments, a subject can have or be at an elevated risk for developing an infection. In some embodiments, a subject can have or be at an elevated risk for developing a cancer. In some embodiments, a subject can be eligible as a recipient of transplantation or is an actual recipient of a transplanted organ or graft. In some embodiments, a subject can be an organ donor or preparing to be an organ donor. In some embodiments, a subject can comprise an animal organ donor for use in a xenotransplant or an animal being prepared for organ donation in a xenotransplant.

[00217] The subject can also be referred to as a "host." As used herein, "host" refers to an organism that harbors another organism or microbe. For example, a living thing e.g., a mammal such as a human being can be a host that harbors a microbe, the microbe being the non-host. As used herein, "host nucleic acids" and all derivative terms such as "host cell-free nucleic acids", "host cell-free DNA", etc. refer to nucleic acids derived from the host genome. In some embodiments, a host genome can comprise nucleic acids derived from a nucleus, a mitochondrion, a cytoplasm, an exosome, cell-free nucleic acids derived from any of these, or any combination thereof. In some embodiments, the target nucleic acids comprise host nucleic acids. In some embodiments, the host nucleic acids comprise a genetic marker associated with a disease or disorder of the host. In some embodiments, the disease or disorder comprises an infectious disease or a non-communicable disease. For example, the host nucleic acid can comprise a genetic marker associated with a cancer.

[00218] In some embodiments, a subject can comprise an animal. In some embodiments, an animal can comprise a vector for disease transmission from which a sample is being tested to determine a presence or absence of a pathogen in the animal. In some embodiments, a disease vector can comprise an animal that has come into contact with a human subject. In some embodiments, an animal coming into contact with a human subject can comprise an animal biting a human, a human ingesting an animal or a secretion of an animal, or a combination thereof. In some embodiments, an animal can comprise a mammal, a bird, a reptile, an amphibian, a fish, an insect, or an arachnid. In some embodiments, an animal can comprise a research animal, an animal for medical use (e.g., xenotransplant donor), a companion animal, a farm animal, a working animal, a performance animal, or a wild animal. In some embodiments, a mammal can comprise a non-human primate (e.g., a macaque or rhesus monkey), a rodent, a carnivore (e.g., a canine or a feline), a bat, a cetacean (e.g., a dolphin), an ungulate, or an insectivore (e.g., a hedgehog). In some embodiments, an ungulate can comprise a swine, a sheep, a cow, a deer, or a horse.

Diagnosis and Treatment

[00219] In some embodiments, the methods are for detecting, diagnosing, treating, monitoring, staging, or prognosing an infectious disease or an infection. The infection may be caused by any microbe, including but not limited to pathogenic or commensal microbes. In some embodiments, the infection is caused by a carrier microbe harboring the one or more AMR genetic markers.

[00220] In some embodiments, the methods are for detecting, diagnosing, treating, monitoring, staging or prognosing a non-communicable disease or disorder in a subject. In some embodiments, non-communicable disease or disorder is associated with altered or abnormal gene expression. In some embodiments, the non-communicable disease or disorder is associated with a genetic alteration (e.g., genetic marker). In some embodiments, the non-communicable disease is associated with a cancer genetic marker. In some embodiments, a non-communicable disease or disorder comprises a cancer, an autoimmune disease, a neurodegenerative disease, diabetes, arthritis, a heart disease, hepatitis, a kidney disease, multiple sclerosis, a disease of the integumentary system, an anemia, Duchenne muscular dystrophy, hemophilia, Down syndrome, Angelman syndrome, Prader-Willi syndrome, Rett syndrome, Fragile X syndrome, Li-Fraumeni syndrome, a cardiomyopathy, obesity, an inflammatory bowel disease, celiac disease, an autism spectrum disorder, a prion disease, asthma, a genetic defect, osteoporosis, neurofibromatosis, beta-thalassemia, Marfan syndrome, and cystic fibrosis.

[00221] In some embodiments, the methods are for detecting, diagnosing, treating, monitoring, staging, or prognosing a medical indication related to an infection. In some embodiments, a medical indication related to an infection comprises any disease, disorder or procedure (e.g., medical or surgical) that renders a subject immunocompromised or susceptible to infections. In some embodiments, the medical indication related to an infection may cause a new or reoccurring infection in a subject. For example, the medical indication related to an infection can comprise any disease for which an immunosuppressant (e.g., chemotherapy, radiation, corticosteroids, transplant medications, or certain biologics) or an anti-infective agent is required as a treatment. In some embodiments, a medical indication related to an infection may comprise cancers, transplantation, surgeries, burns, infections, malnourishment, chronic kidney diseases, diabetes mellitus, autoimmune diseases, or immune disorders (e.g., acquired immunodeficiency syndrome (AIDS)).

[00222] In some embodiments, the methods further comprise administering a treatment to the subject. The treatment can comprise a pharmacological treatment or a non-pharmacological treatment for the disease or disorder that the subject has. In some embodiments, the pharmacological treatments may comprise a small molecule drug or a biologic drug. In some

embodiments, the biologic drug comprises a peptide, a protein, an antibody or an antigen binding fragment thereof, nucleic acids, a cell-based therapy, or any combination thereof. The drug treatment may comprise medications to be administered via a plurality of routes. In some embodiments, the drug treatment is administered orally, topically, or by injections. Topical administration may comprise intranasal administration, intravaginal administration, transdermal administration, or inhalation. Injections may comprise intravenous, subcutaneous, intramuscular, or intrathecal administration. In some embodiments, non-pharmacological treatment comprises a medical device (e.g., laser or photodynamic therapy, radiofrequency or thermal ablation, or cryotherapy), a surgical intervention (e.g., invasive, minimally invasive, laparoscopic, or others), radiotherapy, radioisotopic/nuclear therapy, physical therapy, occupational therapy, phonoaudiological therapy, a rehabilitation intervention, a digital health intervention, a behavioral intervention, a psychological intervention, or any combination thereof.

[00223] In some embodiments, the treatment comprises an antimicrobial or an anti-infective agent. In some embodiments, the antimicrobial agent comprises an antibacterial drug (e.g., an antibiotic), an antiviral drug, an antifungal drug, an antiparasitic drug, or any combination thereof.

[00224] In some embodiments, the antibiotic drug comprises a β-Lactam drug (e.g., carbapenems, methicillin), an aminoglycoside drug (e.g., gentamicin), a Macrolide-Lincosamide-Streptogramin (MLS) drug (e.g., clindamycin), a tetracycline drug (e.g., doxycycline), a fluoroquinolone drug (e.g., ciprofloxacin), a glycopeptide drug (e.g., vancomycin), a polymyxin drug (e.g., colistin), rifamycin (e.g., rifampin), a sulfonamide drug (e.g., trimethoprim-sulfamethoxazole), or any combination thereof. In some embodiments, the antibiotic drug is a broad-spectrum antibiotic drug. In some embodiments, the antibiotic drug is a narrow-spectrum or targeted antibiotic drug.

[00225] In some embodiments, the antiviral drug comprises a nucleoside/nucleotide analogue (e.g., Acyclovir, Tenofovir, Sofosbuvir), a reverse transcriptase inhibitor (RTI) (e.g., Zidovudine, Efavirenz), a protease inhibitor (PI) (e.g., Ritonavir, Glecaprevir), an integrase inhibitor (e.g., Raltegravir, Dolutegravir), an entry inhibitor (e.g., Maraviroc, Enfuvirtide), a neuraminidase inhibitor (e.g., Oseltamivir, Zanamivir), an RNA polymerase inhibitor (e.g., Remdesivir, Sofosbuvir), a broad-spectrum antiviral drug (e.g., Ribavirin, Nitazoxanide), an antibody (e.g., Palivizumab, Inmazeb), or any combination thereof. In some embodiments, the antiviral drug comprises remdesivir, favipiravir, galidesivir, nirmatrelvir, or molnupiravir.

[00226] In some embodiments, the antifungal drug comprises a polyene (e.g., Amphotericin B, Nystatin), an azole (e.g., Fluconazole, Voriconazole, Posaconazole), an echinocandin (e.g., Caspofungin, Micafungin), a pyrimidine analog (e.g., flucytosine), an allylamine (e.g., terbinafine), Griseofulvin, Ciclopirox, Tavaborole, or any combination thereof.

[00227] In some embodiments, the methods described herein further comprise adjusting a treatment that the subject has received at the time of sample collection.

[00228] In some embodiments, the treatment comprises a cancer treatment. In some embodiments, the cancer treatment comprises surgery, radiation, chemotherapy, targeted therapy, immunotherapy, hormone therapy, gene therapy, or any combination thereof.

Definitions

[00229] Unless defined otherwise, all technical and scientific terms used herein have the meaning commonly understood by a person skilled in the art to which this disclosure belongs.

[00230] The articles "A," "an," and "the", as used herein, can each include singular or plural references unless expressly limited to one reference, or unless otherwise indicated by context.

[00231] As used herein, the term "or" is used to refer to a nonexclusive "or"; as such, "A or B" includes "A but not B," "B but not A," and "A and B," unless otherwise indicated.

[00232] The term "about" as used herein generally means plus or minus ten percent (10%) of a value, inclusive of the value, unless otherwise indicated by the context of the usage. For example, "about 100" refers to any number from 90 to 110 and includes the number 100, unless otherwise indicated by the context in which the term is used. The term "about" a range refers to that range minus 10% of its lowest value and plus 10% of its greatest value. When the term "about" is used before a non-numerical term that is a stand-in for a numerical value (e.g., horizontal, perpendicular, aligned), the term "about" refers to the value of the non-numerical term (e.g., 90 degrees, 1800 degrees) plus or minus 10% of that value. Numeric ranges are inclusive of the numbers defining the range.

[00233] As used herein, "abundance" refers to the quantity of something, such as, for example, the quantity or number of molecules, such as nucleic acids. As used herein, "relative abundance" is the abundance of a molecule or molecules of interest per abundance of a reference molecule or molecules of interest. For example, relative abundance of target nucleic acid molecules *(e.g.,* microbial cell-free nucleic acids*)* refers to abundance per reference nucleic acids (*e.g.,* host nucleic acids, synthetic nucleic acid added to the sample, *etc.).* As used herein, "absolute abundance" is the abundance of molecules per a defined unit of initial sample or sample quantity. For example, absolute abundance of target nucleic acid molecules *(e.g.,* microbial cell-free nucleic acids*)* refers to the abundance per defined unit of sample quantity *(e.g.,* sample volume, sample mass *etc.).*

[00234] Whenever the term "at least," "greater than," "greater than or equal to," "no more than," "less than," or "less than or equal to" precedes the first numerical value in a series of two or more numerical values, the term applies to each of the numerical values in that series of numerical

values, unless otherwise specified. For example, greater than or equal to 1, 2, or 3 is equivalent to greater than or equal to 1, greater than or equal to 2, or greater than or equal to 3.

[00235] Whenever the term "no more than," "less than," or "less than or equal to" precedes the first numerical value in a series of two or more numerical values, the term "no more than," "less than," or "less than or equal to" applies to each of the numerical values in that series of numerical values. For example, less than or equal to 3, 2, or 1 is equivalent to less than or equal to 3, less than or equal to 2, or less than or equal to 1.

[00236] The terms ″increased″ , ″increasing″ , or ″increase″ are used herein to generally mean an increase by a statically significant amount. In some cases, the terms ″increased,″ or ″increase,″ mean an increase of at least 10% as compared to a reference level, for example an increase of at least about 10%, at least about 20%, or at least about 30%, or at least about 40%, or at least about 50%, or at least about 60%, or at least about 70%, or at least about 80%, or at least about 90% or up to and including a 100% increase or any increase between 10-100% as compared to a reference level, standard, or control. Other examples of ″increase″ include an increase of at least 2-fold, at least 5-fold, at least 10-fold, at least 20-fold, at least 50-fold, at least 100-fold, at least 1000-fold or more as compared to a reference level.

[00237] The terms, “decreased”, “decreasing”, or “decrease” are used herein generally to mean a decrease by a statistically significant amount. In some cases, “decreased” or “decrease” means a reduction by at least 10% as compared to a reference level, for example a decrease by at least about 20%, or at least about 30%, or at least about 40%, or at least about 50%, or at least about 60%, or at least about 70%, or at least about 80%, or at least about 90% or up to and including a 100% decrease (e.g., absent level or non-detectable level as compared to a reference level), or any decrease between 10-100% as compared to a reference level. In the context of a marker or symptom, by these terms is meant a statistically significant decrease in such level. The decrease can be, for example, at least 10%, at least 20%, at least 30%, at least 40% or more, and is preferably down to a level accepted as within the range of normal for an individual without a given disease.

[00238] Throughout this application, various cases may be presented in a range format. It should be understood that the description in range format is merely for convenience and brevity and should not be construed as an inflexible limitation on the scope of the disclosure. Accordingly, the description of a range should be considered to have specifically disclosed all the possible subranges as well as individual numerical values within that range. For example, description of a range such as from 1 to 6 should be considered to have specifically disclosed subranges such

as from 1 to 3, from 1 to 4, from 1 to 5, from 2 to 4, from 2 to 6, from 3 to 6 etc., as well as individual numbers within that range, for example, 1, 2, 3, 4, 5, and 6. This applies regardless of the breadth of the range.

[00239] As used herein, "adapter" or "portions of an adapter" refers to a chemically synthesized, single-stranded, or double-stranded oligonucleotide that can be attached, *e.g.,* covalently *(e.g.,* ligation, primer extension) or non-covalently *(e.g.,* hybridization), to the ends of nucleic acid molecules, such as DNA or RNA molecules. Adapter sequences can be of any length.

Adapter can refer to either a full-length adapter or a portion of the adapter, *e.g.,* partial adapters can be attached in some embodiments before the full-lengths are introduced by e.g., indexing primers in amplification steps. 3'-end adapters and 5'-end adapters can be full-length or a portion of an adapter sequence that are attached to the opposite ends of a target nucleic acid, a copy of a target nucleic acid, or a target nucleic acid complement. 3'-end adapters and 5'-end adapters sequences end up being attached to the opposite ends of e.g., a template that can be sequenced that comprises target nucleic acid, a copy of a target nucleic acid, and/or a target nucleic acid complement. The 3'-end adapter and 5'-end adapter sequences can be the same or they can be different.

[00240] As used herein, "control" refers to a standard of comparison. A "negative control" refers to a standard of comparison that is used to identify contaminants from samples, or to identify the nature of a signal in the absence of a sample (e.g., a background signal). A "positive control" refers to a standard of comparison that is designed to produce a positive result or signal. Generally, the presence of a substance (e.g., nucleic acid) is detected in a positive control that is run during an assay. Some embodiments of the disclosure comprise a positive and/or negative control. Some embodiments of the disclosure comprise an initial sample or samples without a positive and/or negative control. Some embodiments of the disclosure comprise an initial sample or samples without a positive control. Some embodiments of the disclosure comprise an initial sample or samples without a negative control.

[00241] As used herein, "denaturing" refers to a process in which biomolecules, such as proteins or nucleic acids, lose their native or higher order structure. Native and higher order structure can include, for example, without limitation, quaternary structure, tertiary structure, or secondary structure. For example, a double-stranded nucleic acid molecule can be denatured into two single-stranded molecules.

[00242] As used herein, the length of a nucleic acid sequence is often expressed in "bases" or "bases in length." For single-stranded nucleic acids, this indicates the number of nucleotides (nt). For double-stranded nucleic acids, this indicates the number of base pairs (bp). Sometimes,

"bases" is used interchangeably with "nucleotides (nt)" or "base pairs (bp)" depending on the context.

[00243] As used herein, "detect" refers to quantitative or qualitative detection, including, without limitation, detection by identifying the presence, absence, quantity, frequency, concentration, sequence, form, structure, origin, or amount of an analyte.

[00244] As used herein, "pathogen" refers to a microbe that can cause a disease, ailment, or an infection.

[00245] As used herein, "microbe," or "microbial," generally refers to archaea, bacteria, fungi, protists, parasites, viruses, or other entities that are usually detectable using a microscope (e.g., an optical microscope or electron microscopy). As used herein, the term "microorganism" refers to a uni- or multi- cellular organism, such as, for example, a microscopic organism or macroscopic organism including but not limited to bacteria, fungi, protists, and parasites. Microbes herein can be a prokaryote or a eukaryote. Microbes are often pathogens responsible for disease, but can also exist in a non-pathogenic, symbiotic, commensalistic, mutualistic, or amensalistic relationship with a host, such as a human.

[00246] The term "sequencing," as used herein, generally refers to methods and technologies for determining the sequence of nucleotide bases in one or more polynucleotides. Sequencing can involve basic methods including Maxam-Gilbert sequencing and chain-termination methods, or *de novo* sequencing methods including shotgun sequencing and bridge PCR**,** or next-generation sequencing (NGS) methods (or massively-parallel sequencing method) including but not limited to polony sequencing, pyrosequencing, sequencing-by-synthesis, sequencing by ligation, ion semiconductor sequencing, single molecule sequencing, single-molecule real-time sequencing, nanopore sequencing, and others. Sequencing can be performed by various systems currently available, such as, without limitation, a sequencing system by Illumina®, Pacific Biosciences®, Oxford Nanopore®, Genia Technologies®, or Life Technologies® and others. Such devices can provide a plurality of raw genetic data corresponding to the genetic information of a host *(e.g.,* human), a non-host *(e.g.,* a pathogen, an organ donor), a host-derived variant genetic sequence *(e.g.,* a single nucleotide polymorphism), and/or combinations thereof as generated by the device from a sample provided by the subject.

[00247] The term "derived from" encompasses the terms "originated from," "obtained from," "obtainable from" and "created from," and generally indicates that one specified material finds its origin in another specified material or has features that can be described with reference to the specified material. For example, a sample can be derived from a blood draw, a nucleic acid can be derived from a sample, a sequence read can be derived from sequencing a nucleic acid, or any combination thereof.

[00248] As used herein, the phrase "uniformly distributed" refers to a distribution that is continuous or uniform between members of a family such that for each member of a family there is a predictable or symmetric interval between them. The term "non-uniformly distributed" refers to a distribution of members of a family that does not have a predictable or symmetric interval between them.

[00249] As used herein, "copy number" refers to the number of times a particular gene or genomic region is present in the genome of an organism. In some embodiments, "copy number" refers to the number of times a microbial gene or genomic region (an AMR gene) is present in a microbe. As used herein, "copy number" can also refer to "copy number variations (CNVs)" in eukaryotes.

[00250] As used herein, "cell-free nucleic acids (cfNAs)" refers to nucleic acids that exist freely outside of cells in biological samples such as blood, urine, cerebrospinal fluid, and synovial fluid. cfNAs can be free-floating, such as cfDNA fragments in plasma. cfNAs can also be associated with cells but not contained within intact cells, as seen in vesicle-associated cfNA or nucleosome-associated cfNA. cfNAs can arise from various biological processes, including cell death (apoptosis, necrosis), active secretion, or viral shedding.

[00251] As used herein, "cell-free sample" refers to a sample devoid, or almost devoid, of cells. In some instances, the cell-free sample is devoid of any human cells. In some instances, the cell-free sample is devoid of any microbial cells. In some instances, the cell-free sample is devoid of all types of cells, including eukaryotic or prokaryotic cells. The cell-free sample can be obtained from a biological sample provided herein. In some instances, the cell-free sample is a plasma, which is almost devoid of blood cells. In some instances, the cell-free sample can be obtained by the sample preparation process described herein, such as centrifuging a biological sample.

EXAMPLES

[00252] The following examples are provided to further illustrate some embodiments of the present disclosure, but are not intended to limit the scope of the disclosure; it will be understood by their exemplary nature that other procedures, methodologies, or techniques known to those skilled in the art may alternatively be used.

Example 1

[00253] A bioinformatic enhancement to high throughput sequencing of mfDNA that determined presence or absence of an SCC*mec* element in *Staphylococcus aureus*, a genetic element harboring *mecA* and *mecC* resistance genes was described in S. Venkatasubrahmanyam, et al. 2020 A-3136 ASM Microbe. While the analytical and clinical accuracy of this feature was very high, the diagnostic yield was relatively low, at approximately 50%, because there was no

enrichment for SCC*mec*. However, the results were available on the same day as identification of the potential pathogen.

[00254] To increase the diagnostic yield of this feature for *staphylococci* spp. and to add additional antimicrobial resistance (AMR) markers to the test, a more sensitive assay was that directly targeted a panel of 6 additional common AMR genes in 18 selected bacteria encoding resistance to methicillin (*mec*A and *mec*C), vancomycin (*van*A and *van*B), oxyimino-cephalosporins and aztreonam ($bla_{\text{CTX-M}}$), and carbapenems (bla_{KPC}). Consistent with good laboratory practice all the AMR markers were linked to the organism detected and the report provided interpretations to assist in guiding antimicrobial management.

[00255] These new features represent a novel application of plasma mcfDNA sequencing to detect AMR genes by enriching for ultra short, ultra rare target fragments in plasma prior to sequencing without the need for a positive culture or intact microbial genomes. Disclosed herein are the analytical and clinical validation of mcfDNA sequencing for detection of AMR genetic markers that have been added to the workflow.

1.1 Material and methods

1.1.1 High-throughput sequencing test with AMR

[00256] Microbial cell-free DNA (mcfDNA) was extracted from 250 µL of the plasma, converted to DNA libraries, and sequenced on an Illumina NextSeq®500 or NovaSeq® at a CAP-accredited and CLIA-certified laboratory according to previously validated methods (Karius®, Redwood City, CA). Any of the over 1,000 organisms included in the Karius® clinical reportable range found to be present above a predefined statistical threshold were reported as previously described. The quantity for each organism identified was expressed in molecules per microliter (MPM), representing the number of cfDNA molecules from the reported microorganism present per microliter of plasma (absolute quantity) and estimated deduplicated reads (EDR), representing the number of unique DNA sequencing reads from the reported microorganism present in the sequenced library (relative quantity). For monitoring the yield of the laboratory workflow of the method disclosed herein, whole assay internal control (WINC) molecules are synthetic oligonucleotides spiked into each sample at a known concentration as a control for sequencing depth or human background levels. A quality control minimum of 25,000 unique WINC molecules was set for reporting microorganism absolute quantities in MPM.

[00257] For samples in which *S. aureus* was reported, the Test analytical pipeline further reported:

- Methicillin-resistant *S. aureus* (MRSA), when there was strong evidence consistent with the presence of a *mecA* or *mecC* containing SCC*mec* cassette;
- Methicillin-susceptible *S. aureus* (MSSA), when there was strong evidence for the lack of an SCC*mec* cassette; or

- Indeterminate, when the evidence was ambiguous or insufficient.

Broadly, the determination was based on the number of reads that align to a collection of SCC*mec* (44 SCC*mec* variants and 4 non-*mec* SCC variants), a 20-70 kb genetic element known to harbor the *mecA* or *mecC* genes. The number of reads was compared to the number of reads mapping to the *S. aureus* genome and a statistical model was used to determine the likelihood of either the presence or absence of methicillin resistance. To reduce false-positive MRSA calls, indeterminate was also reported if the SCC in the clinical sample was likely to be a non-*mec* SCC or if SCC fragments in the clinical sample may have originated from a microbe other than *S. aureus*, as determined by comparing the abundances of a set of "interfering" species, known to harbor sequences highly homologous to SCC*mec,* to that of *S. aureus*.

[00258] Test results that reported one or more of the 18 targeted bacteria (**Table 1**) triggered the AMR gene detection workflow (**Fig.1**), provided additional requirements were met:

- sufficient plasma volume remaining (>500 μL);
- for $bla_{\text{CTX-M}}$ and bla_{KPC} testing, the target bacterial species was reported at >1000 MPM;
- for *S. aureus*, the SCC*mec* caller reported "indeterminate"; and
- for *S. aureus* with MPM < 300, the abundance of interfering coagulase-negative staphylococci was <5 MPM.

[00259] In the wet-lab workflow, 500 μL of plasma was spiked with two synthetic control DNA molecules, an ID spike, and an AMR WINC. mcfDNA was extracted, followed by targeted amplification using more than 300 primers to amplify mcfDNA fragments from *mecA*, *mecC*, *vanA*, *vanB*, $bla_{\text{CTX-M}}$, bla_{KPC}, several species-specific housekeeping genes, and the synthetic control molecules. Following sequencing, libraries were demultiplexed and amplicons aligned to AMR markers, human reference, and synthetic sequence spike-ins. The number of unique template molecules (estimated deduplicated transcripts or templates (EDT)) was estimated from the amplicon alignments for each marker and housekeeping gene and from the unique molecular identifiers for the quality control oligonucleotides. Background EDT levels of each marker and housekeeping gene were estimated per sequencing batch using the batch controls. An upper bound estimate of background signal was subtracted from the observed EDT for each gene to generate a corrected EDT estimate.

[00260] Corrected AMR marker genes and WINC EDT were used as input into the AMR caller statistical model. For each AMR marker gene, the AMR caller identifies all of the pathogens detected by the test known to carry that gene, referred to here as "carriers", and their respective MPM values. The statistical model of the AMR caller evaluates the likelihood of all the possible combinations of carriers and AMR gene copy number given an established prior to output a

probability of linkage of AMR to each carrier. For carriers that qualified for the follow-on AMR testing, the test AMR caller reported:

- "detected" if there was strong evidence for linkage of AMR presence to the pathogen;
- "detected-unassigned" if there was strong evidence for AMR presence, but linkage cannot be assigned to one of multiple qualified pathogens;
- "not detected if there was strong evidence for linkage of absence of AMR to the pathogen; or
- "indeterminate" if there was insufficient evidence to support linkage of AMR presence or absence to a qualified pathogen.

For *S. aureus*, *S. epidermidis*, and *E. faecium* an additional check of consistency of housekeeping gene EDT to pathogen MPM and AMR EDT was performed. Failure of this check resulted in an "indeterminate" report.

[00261] Several sample and sequencing batch quality control checks were implemented. EDT of ID Spikes, oligonucleotides unique to each sample, was used to monitor for sample mix-ups and sample cross-contamination. The reaction efficiency of the enrichment process was monitored by the WINC EDT - below 100 WINC EDT, the sample was requeued, excluding if there was an AMR detection or if an AMR "absence" call is supported by an internal check using EDT of housekeeping genes for *S. aureus, S. epidermidis*, or *E. faecium*. Failure of these sample internal quality control checks led to an "indeterminate" report for the affected sample and pathogen. Sequencing batch quality control checks included a check that the correct AMR genes were detected/not detected in the positive control samples and that the background EDT for the genes present in the batch controls did not exceed the upper bound of expectations established for the reagent lot. Failure of the sequencing batch quality control checks led to requeues and "indeterminate" AMR reports for all the affected genes and/or pathogens in the sequencing batch.

1.1.2 Limits of detection SCCmec

[00262] The limit of detection (LoD) of the MRSA/MSSA calls made by the SCC*mec* caller was established by measuring sensitivity using *in-silico* modeling based on an established dilution series of a *S. aureus* genome spiked into a human cfDNA background representing the 10th percentile of human cfDNA concentration observed in commercial samples. The 10 half-log dilutions ranged from 316,000 MPM down to 10 MPM, in addition to un-spiked samples. Ten genome assemblies were chosen (5 MRSA and 5 MSSA) that were geographically and genotypically diverse, publicly available, and tested for phenotypic susceptibility (Table 1). For each such genome, an *in-silico* dilution series was created from the established dilution series by swapping existing reads with simulated reads from the genome. Using these simulated data, a probit analysis was performed to establish a 95% LoD for each genome after subsampling to the

unique WINC QC minimum (25,000) as well as the typical level (300,000) observed in production samples. The final LoD values presented in *S. aureus* MPM and EDR were averaged across the genomes.

1.1.3 AMR gene panel

[00263] The LoD for calling AMR presence and absence was calculated using contrived laboratory samples containing multiplexed mixtures of sheared genomes microbes carrying the target AMR genes from the panel (Supplemental Table 2). The multiplexed mixtures were spiked in 8 half-log dilutions into a low human cfDNA plasma background (Supplemental Table 3). The *vanB* variant carried by the ATCC strain used in the genome mix shares 95% nucleotide identity with the *vanB* used to design targeting primers; 9 of 16 primer-binding sites contain a mutation. To compensate for the loss of signal due to imperfect primer binding sites, *vanB* concentration in the genome mixes was selectively increased five-fold. Therefore, LoD values for presence of *vanB* were reported as a function of this increased concentration. To establish the LoD for calling AMR(-) genomes, e.g. absence of *mec*A, samples from the dilution series of a "partner" genome were used: *mecA*(-)/*S. aureus mecC*(+); *vanA*(-)/*E. faecalis vanB*(+); *vanB*(-)/*E. faecium vanA*(+); bla_{KPC}(-)/*E. coli* $bla_{CTX\text{-}M}$(+); and $bla_{CTX\text{-}M}$(-)/*K. pneumoniae* bla_{KPC}(+). Replicates were tested over 10 assay runs across 10 days to incorporate variance in typical laboratory conditions. The LoD was estimated using a probit analysis and reported in units of pathogen MPM.

[00264] To examine how sequencing depth affects the LoD, each read set was subsampled to 300,000, 3.4 million, and 7.8 million reads, representing the 10th, 25th, and 50th percentile of typical clinical sample sequencing depths (Supplemental Fig. 1), and the LoD was recalculated. Lower sequencing depth is expected to increase the LoD. To best simulate a realistic reduced-depth sequencing scenario for an individual sample, the calibrated background subtraction method was applied based on non-down sampled batch controls to down sampled LoD samples. After evaluating all sample quality control metrics, two replicates were removed from all subsequent analyses that use LoD samples due to failure of the cross-contamination quality control metrics.

1.15 Inclusivity and exclusivity SCCmec

[00265] To assess inclusivity, 360 *in-silico* samples were simulated by spiking simulated reads from 36 MRSA-labeled genome assemblies at a particular read count into a healthy human plasma matrix subsampled to the unique WINC QC minimum. The read count was drawn from distribution matched to *S. aureus* calls observed in commercial samples run with the Test, and the database was blinded to the assembly from which the reads were simulated. To assess exclusivity, the same procedure was performed for 11 MSSA-labeled genome assemblies,

generating 110 *in-silico* samples. The MSSA-labeled genome assemblies included genomes containing a non-*mec* SCC element that lacks a *mec* resistance gene.

1.1.6 AMR gene panel

[00266] To test the robustness of the primers and bioinformatics protocols to sequence diversity, *in-silico* samples were created using available AMR alleles from curated NCBI® nucleotide sequences available in the Comprehensive Antibiotic Resistance Database (CARD). To generate an *in-silico sample*, an allele from the desired on-panel AMR gene was chosen from the set of available nucleotide sequences at random. From the LoD experiment, the empirically tested concentration closest to, but not below, twice the LoD for calling the presence of the gene at median sequencing depth was selected.

[00267] In the technical replicate sequencing reads, the AMR target amplicons, and their corresponding AMR target primers, which are referred to to as "activated" primers, i.e., primers that successfully amplified a template molecule, were identified. The nucleotide sequences of the amplicons were then swapped with nucleotide sequences corresponding to the chosen AMR allele. As a conservative modeling of primer binding dynamics, only amplicons whose activated primers had a perfect match on the allele were swapped; otherwise, they were discarded from the *in-silico* sample. The inclusivity of housekeeping genes was considered (Supplemental Methods) and found that housekeeping genes have high percent identity matches to all the target species assemblies tested. Given high inclusivity of housekeeping genes for their target species, the "read swapping" protocol was not carried out for them. For each on-panel AMR gene, 50 such *in-silico* samples were created.

[00268] To evaluate robustness of primers and bioinformatics protocols to cross-reactive sequences, AMR genes and primers were mapped to microbe, fungi, and viral reference genomes from the Test pathogen reference genome database and the human reference genome to generate a list of cross-reactive sequences for each on-panel AMR gene. High identity matches (>95% identity, >99% coverage) to the AMR target genes or alleles of the AMR target genes are excluded.

[00269] In addition, genomes of clinical isolates of organisms in Table 1 with presumed absence of the associated AMR gene(s) were queried using NCBI®'s Pathogen Detection database. Absence of AMR genes was judged using NCBI®'s AMRFinderPlus annotation for the clinical isolate genome. Only complete genomes were queried in order to have high quality sequence data.

[00270] To generate a single *in-silico* sample for an AMR gene, an AMR (-) genome from the appropriate target species or a cross-reactive sequence from the above lists was chosen. Then,

the empirically tested concentration closest to, but not below, twice the LoD for calling the absence of the AMR gene at median sequencing depth from the LoD experiment was selected. Using relaxed parameters that allow less-than-perfect primer matches, the activated primers were mapped to the chosen AMR (-) genome or cross-reactive sequence to identify putative primer-binding regions and a "read swapping" protocol was performed as described above. This is a conservative approach that does not fully account for the selectivity of primers in a reaction. 100 such *in-silico* samples were created per AMR gene (50 with AMR-negative genomes and 50 with putative cross-reactive sequences).

1.1.7 Clinical validation.

[00271] A total of 115 residual patient plasma samples which had one of the 18 target pathogens detected in the Test, concordant culture results, applicable orthogonal AST results, and adequate residual volume were included in the clinical validation study. The study sites, number of samples with orthogonal data by site, and the data or sample collection protocols that met the inclusion criteria are shown in Table 2. The orthogonal culture results were obtained from a wide variety of sample types including 58 blood, 8 sputum, 7 BAL, 7 wound, 5 tissue/aspirate, and 2 urine. The samples contained 12/18 (66.7%) of the target bacteria including 58 *Staphylococcus aureus*, 16 *Pseudomonas aeruginosa*, 14, *Escherichia coli*, 6 *Klebsiella pneumoniae*, 5 *Serratia marcescens*, 5 *Enterobacter cloacae* complex, 4 *Enterococcus faecium*, 2 *S. epidermidis*, 2 *E. faecalis*, 1 *Proteus mirabilis*, 1 *K. oxytoca* and 1 *Salmonella enterica*.

[00272] The following rules were used to assign labels to the plasma samples included in the analysis: Positive culture for one of the panel staphylococci and reported resistant or susceptible to oxacillin for SCC*mec* and *mecA* or *mecC*. Positive culture for *E. faecalis* or *E. faecium* and reported resistant or susceptible to vancomycin for *vanA* or *vanB*. Positive culture for one of the panel gram-negative bacilli and reported resistant to any one or more of the following antibiotics tested: cefotaxime, ceftazidime, ceftriaxone, cefepime or aztreonam for putative $bla_{\text{CTX-M}}$ positives and reported susceptible or intermediate to all key indicator drugs for putative $bla_{\text{CTX-M}}$ negatives. Not all drugs were tested for each pathogen. Positive culture for one of the panel gram-negative bacilli and reported resistant to any one or more of the carbapenems tested: doripenem, ertapenem, imipenem, or meropenem for putative bla_{KPC} positives and reported susceptible to all carbapenems tested for putative bla_{KPC} negatives. As above, not all drugs were tested for each pathogen. Overall, for those plasma samples associated with concordant cultures that revealed strains with different AST results, the most resistant phenotype was used as the label.

[00273] The orthogonal phenotypic AST was performed by Vitek2 for 46.1% and by MicroScan WalkAway system for 47.8% of samples included in the clinical validation. No phenotypic AST method results were recorded for 7 (6.1%) of the Indiana University MSKI study samples.

[00274] Orthogonal genotypic results were also reported for 11 of the AMORTISE Study samples from positive blood cultures including 2 Verigene, 5 GeneMark and 4 Biofire BCID panel results. All the Duke MRSA/MSSA and the Indiana MSKI study samples had associated BCID panel results reported for positive blood cultures for *S. aureus*.

1.1.8 Statistical Analyses.

[00275] All statistical analyses were performed in Python 3.10. The LoD was calculated using a maximum-likelihood probit model fit to the data. This model assumed that the number of positives at each concentration followed a binomial distribution and the success rate of the binomial distribution at each concentration was determined by the value of the cumulative normal distribution at that concentration. The parameters of the cumulative normal distribution (probit fit) associating concentration with sensitivity were fitted to maximize the likelihood of the dilution series. The LoD concentration was determined using the probit (inverse of cumulative normal distribution) at 95% sensitivity. The 95% confidence intervals on proportions were calculated using the Clopper-Pearson interval. Descriptive statistics used to assess test performance characteristics, positive percent agreement (PPA), negative percent agreement (NPA), and categorical agreement (CA) were calculated by accepted methods. Diagnostic yield (DY) was defined as the percentage of tests that yielded an actionable result either detected or not detected.

1.1.9 Ethical considerations

[00276] Institutional review board approvals were obtained locally by all clinical centers and conducted in compliance with good clinical practice guidelines set forth by the International Conference on Harmonization (ICH-E6). Advarra IRB approved Karius® - KAR-0002, AMR On Remnant TISsuE (AMORTISE) and granted waiver of consent for collection of additional clinical data to support this study.

1.10 Robustness of limit of detection to AMR gene copy number

[00277] To examine the robustness of the limit of detection (LoD) to the copy number of the AMR gene, the copy number of the AMR genes was estimated using the Test sequencing data (without targeting) for 100,000 MPM of contrived sheared pathogen genome carrying the AMR gene. The copy number was estimated using known ratios between the AMR gene and the pathogen genome length, the Test estimated pathogen read count, and the AMR read counts from

a custom alignment to the AMR genes (Supplemental Table 2). The observed AMR gene EDT was scaled by the copy number, e.g., with five gene copies, 20 EDT becomes 4 EDT, and recalculated the LoD. Each gene copy was assumed to contribute the same EDT signal as all the others, which is a conservative assumption, given that it does not allow for attribution of most of the EDT signal to a single gene copy.

1.11 Precision SCCmec

[00278] To assess the precision of the SCC*mec* detection capabilities, four commercial (2 MRSA and 2 MSSA) samples were chosen for which there was an *S. aureus* call with a concentration that exceeded the minimum LoD at the QC minimum for unique WINC molecules (25,000). The expectation was that the call in all reruns of these samples would be identical to that of the original result. Each sample was requeued 3 times, either within the same sequencing batch (repeatability) or on different batches (reproducibility). AMR gene panel. To assess repeatability and within-lab reproducibility, samples from the LoD experiment were used. To assess precision of presence calls per gene, replicates containing the gene at an empirically tested concentration closest to, but not below, twice the presence LoD were chosen for a maximum total of 20 replicates examined per gene. The same procedure was repeated to assess precision of absence calls using empirically tested concentration closest to, but not below, twice the absence LoD. To evaluate repeatability and within-lab reproducibility in clinical samples, four samples were identified as MRSA by the Test SCC*mec* caller where the *S. aureus* call is above twice the LoD for *mecA* and *mecC* detection. Six samples were identified following the same protocol for MSSA, with the *S. aureus* call above twice the LoD for determining *mecA* and *mecC* absence. The MRSA and MSSA plasma samples were made into two standardized clinical plasma pools. Three aliquots each from the MRSA and MSSA pools were tested within a sequencing run to determine repeatability. Three aliquots each from the MRSA and MSSA pools were tested across three sequencing runs to determine within-lab reproducibility.

1.12 Housekeeping gene (HK) inclusivity/exclusivity.

[00279] Housekeeping gene targets for *S. aureus*, *S. epidermidis*, and *E. faecium* were used in the targeted AMR gene assay as a secondary check for species presence and judgment of AMR presence/absence. If the ratio of AMR EDT to housekeeping gene EDT for one of the three species departed from the expected relationship, the AMR call was "indeterminate". For example, lack of adequate housekeeping gene EDT would lower confidence in an AMR absence call, making it indeterminate. Housekeeping genes were chosen from species-specific core protein coding genes and filtered to ones where the nucleotide sequence would not cross-react with genomes of other pathogens. Only housekeeping genes for *S. aureus*, *S. epidermidis*, and *E.*

faecium showed high species specificity and were then incorporated into the targeted AMR gene assay.

[00280] Housekeeping genes were checked against the Test pathogen reference genome database and evaluated for the percent identity, alignment length, and number of species reference genome assemblies they aligned to. An ideal species-specific housekeeping gene would align fully at 100% identity to all of the species reference genomes in the database, but no other genomes.

[00281] Except for C-SA, all the chosen housekeeping genes align fully with almost 100% identity to all the available target species assemblies in the database (Supplemental Table 5). C-SA is one of 9 housekeeping genes for *S. aureus* and its presence or absence does not strongly bias the total housekeeping EDT recovery.

[00282] The cross-reactivity of the housekeeping genes with other species is very limited (Supplemental Table 6), the matches are either low percent identity (<80%), which would lead to a poor match of primers to the cross-reactive sequence or filtering via the analytical pipeline or the matches are very short DNA sequences (<10% of housekeeping gene length), which are unlikely to be bound by primers or contribute enough EDT signal to systematically bias housekeeping gene recovery.

1.2 RESULTS

[00283] Table 1. Target Bacteria and associated AMR markers.

Bacterium	AMR marker						
	SCC*mec*	*mecA*	*mecC*	*vanA*	*vanB*	bla_{CTX-M}	bla_{KPC}
Methicillin-resistant staphylococci							
Staphylococcus aureus	+	+	+				
S. epidermidis		+	+				
S. lugdunensis		+	+				
Vancomycin-resistant enterococci							
Enterococcus faecalis				+	+		
E. faecium				+	+		
ESBL and carbapenemase producing gram-negative bacteria							
Enterobacter cloacae complex						+	+
Escherichia coli						+	+
Klebsiella aerogenes						+	+
K. pneumoniae						+	+
K. oxytoca						+	+
Proteus mirabilis						+	+
P. vulgaris						+	+
Salmonella bongori						+	+
S. enterica						+	+
Serratia marcescens						+	+
Pseudomonas aeruginosa						+	+
Acinetobacter baumannii						+	+
A. calcoaceticus						+	+

[00284] Table 2. Study sites providing data or specimens for the clinical validation of the Test AMR markers.

Study sites	**No. samples**	**Protocol**
AMORTISE: Orlando Health Carle Foundation Advent Health Orlando OSF Health City of Hope	53	Data collection for orthogonal AST results (Vitek 2, bioMérieux) for any of the 18 target bacteria from a variety of specimen types that were performed within 7 d of the Test sample collection.
Duke BSI[1]	12	Data collection for orthogonal AST results (MicroScan WalkAway, Beckman Coulter) for any of the 18 target bacteria from blood cultures that were performed ≤48 h of the Test sample collection.
Indiana University MSKI[2]	7	Data collection for Biofire BCID2, (bioMérieux) panel results for *S. aureus* from blood cultures from patients with musculoskeletal infections that were performed ≤96 h of the Test sample collection.
Duke MRSA/MSSA	43	Plasma samples collected 48-72 h of positive blood cultures with *S. aureus* labeled as MRSA or MSSA by orthogonal methods (MicroScan WalkAway and Biofire BCID2).

[00285] Table 3. Limit of detection characterization for SCC*mec*.

	Average Limit of Detection	
Sequencing depth	**SCC*mec* presence (MRSA)**	**SCC*mec* absence (MSSA)**
QC minimum (25,000 unique WINC)	54,803 MPM / 4,567 EDR	10,304 MPM / 858 EDR
Standard (300,000 unique WINC)	3,916 MPM / 3,916 EDR	683 MPM / 683 EDR

[00286] Table 4. LoD by target and sequencing depth for the AMR gene panel.

Target	**Original sequencing depth**	**50th percentile sequencing depth**	**25th percentile sequencing depth**	**10th percentile sequencing depth**
mecA presence	662 MPM	674 MPM	966 MPM	1,712 MPM
mecA absence	2,316 MPM	2,316 MPM	2,177 MPM	6,694 MPM
mecC presence	808 MPM	856 MPM	1,180 MPM	1,803 MPM
mecC absence	472 MPM	472 MPM	472 MPM	1,000 MPM
vanA presence	425 MPM	372 MPM	437 MPM	646 MPM
vanA absence	1,044 MPM	1,044 MPM	2,617 MPM	1,000 MPM
*vanB presence**	4,985 MPM	4,996 MPM	5,000 MPM	6,730 MPM
vanB absence	1,074 MPM	1,105 MPM	1,105 MPM	1,180 MPM
$bla_{\text{CTX-M}}$ presence	6,107 MPM	6,633 MPM	5,793 MPM	6,971 MPM
$bla_{\text{CTX-M}}$ absence	22,680 MPM	19,640 MPM	–	3,471 MPM
bla_{KPC} presence	1,346 MPM	1,346 MPM	1,380 MPM	1,769 MPM
bla_{KPC} absence	1,555 MPM	1,555 MPM	1,555 MPM	4,672 MPM

[00287] Table 5. Inclusivity by AMR marker.

Target	**Inclusivity**	**95% CI**
SCC*mec*	100% (36/36)	90%-100%
mecA	100% (50/50)	93%-100%
mecC	100% (50/50)	93%-100%
vanA	100% (50/50)	93%-100%
vanB	100% (50/50)	93%-100%
$bla_{\text{CTX-M}}$	54% (27/50)	39%-68%
bla_{KPC}	100% (50/50)	93%-100%

[00288] Table 7. Correlation of SCC*mec* results with phenotype for 58 *Staphylococcus aureus* detections. The 95% confidence intervals for performance metric estimates are in brackets.

SCC*mec* result	**Phenotype**	
	Methicillin resistant	Methicillin susceptible
Detected	13	1[a]
Not detected	0	17
Indeterminate	12	15
Agreement	PPA: 13/13 (100%) [73.5-100%]	NPA: 17/18 (94.4%) [72.7-99.8%]
Categorical Agreement	30/31 (96.8%) [83.3-99.9%]	
Diagnostic Yield	31/58 (53.4%) [39.8-66.6%]	

[a]Both SCC*mec* and *mecA* were detected.

[00289] Table 8. Correlation of *mecA* results with phenotype for 58 *Staphylococcus aureus* detections. The 95% confidence intervals for performance metric estimates are in brackets.

***mecA* result**	**Phenotype**	
	Methicillin resistant	Methicillin susceptible
Detected	14	1[a]
Not detected	1	19
Indeterminate	10	13
Agreement	PPA: 14/15 (93.3%) [68.0-99.8%]	NPA: 19/20 (95.0%) [75.1-99.8%]
Categorical Agreement	33/35 (94.3%) [80.8-99.3%]	
Diagnostic Yield	35/58 (60.3%) [46.6-72.9%]	

Table 9. Correlation or results when SCC*mec* and *mecA* tests are combined for 58 *Staphylococcus aureus* and 2 *S. epidermidis* detections in the decision workflow for methicillin resistance in staphylococci. The 95% confidence intervals for performance metric estimates are in brackets.

Result	**Phenotype**	
	Methicillin resistant	Methicillin susceptible
SCC*mec* or *mecA* detected	19	1[a]
SCC*mec* and *mecA* not detected	1	21
SCC*mec* and *mecA* Indeterminate	6	12
Agreement	PPA: 19/20 (95.0%) [75.1-99.9%]	NPA: 21/22 (95.4%) [77.1-99.9%]
Categorical Agreement	40/42 (95.2%) [83.8-99.4%]	
Diagnostic Yield	42/60 (70.0%) [56.8-81.1%]	

[a]Both SCC*mec* and *mecA* were detected.

[00290] Table 10. Correlation of *vanA* results with phenotype for 6 *Enterococcus spp.* detections. The 95% confidence intervals for performance metric estimates are in brackets.

***vanA* result**	**Phenotype**	
	Vancomycin resistant	Vancomycin susceptible
Detected	3[a]	0
Not detected	0	2[b]
Indeterminate	1[c]	0
Agreement	PPA: 3/3 (100%) [29.2-100%]	NPA: 2/2 (100%) [15.8-100%]
Categorical Agreement	5/5 (100%) [47.8-100%]	
Diagnostic Yield	5/6 (83.3%) [35.8-99.5%]	

[a]All *E. faecium;* [b]Both *E. faecalis'* [c]*E. faecium*

[00291] Table 11. Correlation of *bla*$_{CTX-M}$ results with ESBL phenotype for 49 Gram-negative bacilli detections. The 95% confidence intervals for performance metric estimates are in brackets.

***bla*$_{CTX-M}$ result**	**ESBL Phenotype[a]**	
	Yes	No
Detected	5[b]	0
Not detected	4[c]	26
Indeterminate	0	14
Agreement	PPA: 5/9 (55.5%); [21.2-86.3%]	NPA: 26/26 (100%); [86.7-100%]
Categorical agreement	31/35 (88.5%) [73.2-96.8%]	
Diagnostic Yield	35/49 (71.4%) [56.7-83.4%]	

[a]Yes indicate resistance to at least one of following antimicrobials: Cefotaxime, ceftazidime, ceftriaxone, cefepime (oxyimino-beta-lactams), or aztreonam (monobactam) and no indicates susceptible or intermediate results for all key indicator drugs. Not all drugs were tested for each sample.

[b]Includes 4 detections of *Escherichia coli* and 1 detection of *Klebsiella pneumoniae*.

[c]Includes 3 detections of *Enterobacter cloacae* complex and 1 detection of *E. coli*.

[00292] Table 12. Resolution of *bla*$_{CTX-M}$ false negative detections[a].

Lab no.	**Category**	**Bacterium**	**Phenotype[b]**	**Genotype**	**Resolution**
DK-27435	FN	*E. cloacae* complex	Cefepime - S Ceftazidime - R Ceftriaxone - R	Not available	TN
LB-70865	FN	*E. cloacae* complex	Aztreonam - R Cefepime - R Ceftriaxone - R	Not available	FN

Lab no.	Category	Bacterium	Phenotype[b]	Genotype	Resolution
LB-75775	FN	*E. cloacae* complex	Cefepime - S Ceftriaxone - R	*bla*$_{CTX-M}$ not detected by Verigene® BC-GN test	TN
LB-77110	FN	*E. coli*	Cefepime - S Ceftriaxone - R	Not available	TN

[a]Abbreviations: FN, false negative; TN, true negative; S, susceptible; and R, resistant.

[b]ESBL and AmpC β-lactamases can be distinguished phenotypically based on differential susceptibility to cefepime.

[00293] Table 13. Correlation of *bla*$_{KPC}$ results with carbapenemase phenotype for 44 Gram-negative bacilli detections. The 95% confidence intervals for performance metric estimates are in brackets.

***bla*$_{KPC}$ result**	**Carbapenemase Phenotype[a]**	
	Yes	No
Detected	0	0
Not detected	2[b]	23
Indeterminate	3	16
Agreement	PPA: 0/2 (0%); [0-84.2%]	NPA: 23/23 (100%); [85.1-100%]
Accuracy	23/25 (92.3%) [73.9-99.0%]	
Diagnostic Yield	25/44 (56.8%) [41.0-71.6%]	

[a]Yes indicate resistance to doripenem, ertapenem, imipenem, or meropenem and no indicates susceptibility to all key indicator drugs. Not all drugs were tested for each sample.

[b]Includes one detection of *Pseudomonas aeruginosa* by the Test in a patient who had two cultures positive for *P. aeruginosa* strains with different carbapenem phenotypes (BAL meropenem susceptible and sputum meropenem resistant) and one *Escherichia coli* detection that was that was resistant to both ertapenem and meropenem.

[00294] Supplemental Table 1. *Staphylococcus aureus* genomes used for SCC*mec* LoD assessment.

Assembly accession	**MRSA/ MSSA**	**Source**	**Geolocation**	**Collection Year**	**Strain**	**SCC*mec* subtype[a]**
GCF_001018905.2	MRSA	Wound	USA:Washington	-	NRS387; USA800; 1045 (FDAARGOS_26)	IVc(2B) / IV
GCF_001018875.2	MRSA	Sputum	USA:Tennessee	-	NRS127 (FDAARGOS_22)	II(2A) / II
GCF_001046135.1	MRSA	Blood	USA: New York	2009	C1440 t008	IVa(2B) / -
GCF_001046615.1	MRSA	Nares	USA: Florida	2009	C1555, t008	IVg(2B) / -
GCF_001019565.2	MRSA	Blood	Denmark	1964	E2125 (FDAARGOS_46)	I(1B) / I
GCF_001018685.2	MSSA	Buttock abscess	USA:Texas	2002	TCH 959 (USA300-HOU-MS); ATCC BAA-1718 (FDAARGOS_4)	None
GCF_001018975.2	MSSA	-	-	-	NRS143 (FDAARGOS_30)	None

Assembly accession	MRSA/ MSSA	Source	Geolocation	Collection Year	Strain	SCC*mec* subtype[a]
GCF_001018715.2	MSSA	-	-	-	NRS128 (FDAARGOS_11)	None
GCF_001019375.2	MSSA	Vaginal tampon	-	-	NRS156 (FDAARGOS_17)	None
GCF_001018725.2	MSSA	-	French Guiana	1996	NRS129 (FDAARGOS_10)	None

[a]SCCmecFinder inferred / BioSample metadata)

[00295] Supplemental Table 2. Organisms and genomes used to establish LoD.

Organism	Resistance Genotype	Estimated gene copy number	ATCC Strain ID (Strain Name)	Mix
Staphylococcus aureus	*mecA*(+)	1	700699 (*S. aureus subsp. aureus* Mu50)	P3.2
S. aureus	*mecC*(+)	1	BAA-2313 (*S. aureus* M10/0148)	P3.3
Enterococcus faecium	*vanA*(+)	6	700221 (*E. faecium* VRE)	P3.2
E. faecalis	*vanB*(+)	2	700802 (*E. faecalis* V583)	P3.3
Klebsiella pneumoniae	bla_{KPC}(+)	3	BAA-1705 (*K. pneumoniae* ART 2008133)	P3.3
Escherichia coli	bla_{CTX-M}(+)	2	BAA-2326 (*E. coli* TY-2482)	P3.2

[00296] Supplemental Table 3. Limit of Detection experimental design (N = 260).

Dilution	**Pathogen Input Concentration in MPM* (No. replicates)**
1 (P3.2, P3.3) N=20	100,000 (x10) **vanB*: 500,000
2 (P3.2, P3.3) N=20	31,600 (x10) **vanB*: 158,000
3 (P3.2, P3.3) N=40	10,000 (x20) **vanB*: 50,000
4 (P3.2, P3.3) N=40	3,160 (x20) **vanB*: 15,800
5 (P3.2, P3.3) N=40	1,000 (x20) **vanB*: 5,000
6 (P3.2, P3.3) N=40	316 (x20) **vanB*: 1,580
7 (P3.2, P3.3) N=40	100 (x20) **vanB*: 500
8 (P3.2, P3.3) N=20	32 (x10) **vanB*: 160
9 (plasma background)	0 (x40 or 4 per sequencing run)

[00297] Supplemental Table 4. LoD per target and copy number for the AMR gene panel.

Target	**Original LoD**	**LoD at 1 gene copy**
mecA presence	662 MPM	662 MPM
mecC presence	808 MPM	808 MPM
vanA presence	425 MPM	3,296 MPM
vanB presence	*4,985 MPM*	*6,012 MPM*
$bla_{\text{CTX-M}}$ presence	6,107 MPM	1,458 MPM
bla_{KPC} presence	1,346 MPM	3,169 MPM

[00298] Supplemental Table 5. HK gene inclusivity

Species	**HK gene name**	**% matching assemblies**	**Median alignment coverage**	**Median percent identity**
S. aureus	*A-SA*	22/22 (100%)	1.000000	99.4445
	B-SA	22/22 (100%)	1.000000	97.8210
	C-SA	13/22 (59%)	1.000000	99.4410
	D-SA	22/22 (100%)	1.000000	98.6670
	E-SA	22/22 (100%)	1.000000	99.0820
	glpF	22/22 (100%)	1.000000	99.6775
	orfX	22/22 (100%)	1.000000	98.5420
	arcC	22/22 (100%)	1.000000	99.3420
	aroE	22/22 (100%)	1.000000	98.9040
S. epidermidis	*A-SE*	8/8 (100%)	1.000000	98.9820
	B-SE	8/8 (100%)	1.000000	98.2330
	C-SE	8/8 (100%)	1.000000	97.8850
	D-SE	8/8 (100%)	1.000000	96.9095
	yqiL	8/8 (100%)	1.000000	99.1545
	pyrR	8/8 (100%)	1.000000	97.1960
	aroE	8/8 (100%)	0.998960	97.8145
	tpiA	8/8 (100%)	1.000000	99.4090
E. faecium	*atpA*	20/20 (100%)	1.000000	97.8420
	purK	20/20 (100%)	0.998984	99.7965
	ddl	20/20 (100%)	1.000000	98.7100

[00299] Supplemental Table 6. HK gene exclusivity

Species	**HK gene name**	**no. matching assemblies of non-target species**	**Median alignment coverage**	**Median percent identity**
S. aureus	*A-SA*	–	–	–
	B-SA	–	–	–
	C-SA	1	0.061453	93.9390
	D-SA	4	0.104000	89.7440
	E-SA	–	–	–
	glpF	52	1.000000	78.3550
	orfX	110	0.995833	79.7920
	arcC	45	0.574561	72.5000
	aroE	36	0.995614	74.8565
S. epidermidis	*A-SE*	–	–	–
	B-SE	1	0.037691	93.7500
	C-SE	1	0.069374	87.8050
	D-SE	1	0.033435	93.9390
	ymaiL	8	0.350242	72.2390
	pyrR	57	0.271028	75.7510
	aroE	42	0.979210	72.4470
	tpiA	887	0.891253	72.4600
E. faecium	*atpA*	1185	0.953237	75.7125
	purK	20	0.887195	75.9730
	ddl	74	0.969892	79.3790

1.21 Limit of Detection and Precision SCC*mec*

[00300] To estimate LoD, multiple genome assemblies and two levels of sequencing depths, representing quality control minimum for absolute quantification and typical production depth, were simulated in an *in-silico* dilution derived from an original dilution series of a *S. aureus* sheared genome in a human plasma background. An *in-silico* approach allowed for incorporation of multiple variables, e.g., different MRSA and MSSA genomes and sequencing depths, that the original dilution series did not address. The LoD values at different sequencing depths and SCC*mec* presence or absence were averaged across the genome assemblies assessed (Table 3). Sequencing depth was measured using the number of unique WINC molecules in the sample as a proxy, with the minimum depth 10-fold lower than the typical depth. The presence and absence LoD in MPM were sensitive to sequencing depth, with the LoD increasing approximately 10-fold at the minimum depth. This is expected, as higher pathogen abundance is needed to make a determination in a shallowly sequenced sample. In a typical production sample, the presence LoD is 3,916 MPM and absence LoD is 683 MPM. For precision assessment, replicates of clinical samples with a MRSA or MSSA determination by the SCCmec caller were reprocessed in the same sequencing batch for repeatability or across sequencing batches for reproducibility (Supplemental Methods). 100% repeatability and reproducibility were observed in the MRSA and MSSA clinical samples (n=6 per sample type).

1.22 AMR gene panel

[00301] To estimate LoD for each AMR gene, a dilution series was created using sheared microbial genomes carrying AMR genes spiked into human plasma background. This dilution series was computationally down sampled to assess LoD at different sequencing depths. Table 4 shows the LoD based on probit model fits for non-down sampled data, and then the data down sampled to each of the predefined sequencing depths. Overall, the presence LoDs range from 425-6,107 MPM and the absence LoDs range from 472-22,680 MPM. LoD was not strongly sensitive to lower sequencing depths - this is because the targeted amplification method creates many amplicons from the original fragment making “surplus” signal. Presence LoD is largely influenced by the length of the targeted region on the gene, as well as the gene copy number and gene variant. The *vanB* presence LoD is a conservative estimate, a limitation imposed by the *vanB* gene variant in the genome mixture (see Methods). Since only one microbial genome was assessed to determine the gene LoD, LoD was also characterized assuming one copy of the AMR gene (Table 4), to give a conservative upper bound for the LoD as several of the panel AMR genes were carried in multiple copies by the microbes (Table 2). For AMR genes $bla_{\text{CTX-M}}$ that were carried in multiple copies, except $bla_{\text{CTX-M}}$, LoD increased when assuming one copy.

Decrease in *bla*$_{CTX-M}$ LoD at one copy number is due to elimination of positive calls at low pathogen MPM due to reagent background (Fig. 2) once the *bla*$_{CTX-M}$ signal was scaled to simulate one gene copy, leading to a different probit model fit (Fig. 3).

[00302] For *bla*$_{CTX-M}$, LoD was also impacted by the presence of reagent background signal (Supplemental Fig. 2), especially the absence LoD. For *bla*$_{CTX-M}$, down sampling to lower sequencing depths decreased the *bla*$_{CTX-M}$ reagent background signal, leading to more effective background subtraction and improved probit fits to the observed data. *bla*$_{CTX-M}$ absence LoD was unable to be estimated using a probit model for samples down sampled to the 25th percentile of sequencing depth because there was no concentration at which there was a ≥95% absence call rate.

[00303] Repeatability and reproducibility of the AMR assay was assessed with the contrived samples from the dilution series and confirmed with MRSA and MSSA clinical samples as labeled by the SCC*mec* caller. 100% repeatability and reproducibility were observed when evaluating precision in the contrived samples (n=10 to 20 per target), as well as in the clinical MRSA and MSSA samples (n=6 per sample type).

1.23 Inclusivity and exclusivity.

[00304] Different alleles of AMR genes may exhibit different sensitivities in the AMR assay due to primer-template mismatches or failure to successfully align to the reference sequences in the bioinformatics pipeline. To test a wide array of alleles representative of the range of variation present in the population, *in-silico* simulations were performed, derived from the LoD dilution series contrived samples. Across the different alleles assessed, 100% PPA was observed for all targets except *bla*$_{CTX-M}$ (Table 5). Inclusivity for *bla*$_{CTX-M}$ was limited to 54% in this analysis (based on a single sampling of 50 alleles) because the targeted approach has primers with perfect homology only to a single clade of *bla*$_{CTX-M}$ variants (*bla*$_{CTX-M-3}$ clade), but not the other highly divergent clades (Fig. 2). In general, based on the prevalences derived from CARD resistomes dataset and literature, the sensitivity for *bla*$_{CTX-M}$ variants is expected to be 67-75%. Common *bla*$_{CTX-M}$ variants captured by AMR gene enrichment include *bla*$_{CTX-M-3}$, *bla*$_{CTX-M-15}$, *bla*$_{CTX-M-32}$, *bla*$_{CTX-M-55}$, and *bla*$_{CTX-M-1}$. Common CTX-M variants poorly captured or missed by AMR gene enrichment include *bla*$_{CTX-M-2}$, *bla*$_{CTX-M-9}$, *bla*$_{CTX-M-14}$, *bla*$_{CTX-M-27}$, and *bla*$_{CTX-M-65}$. Except for the highly divergent *bla*$_{CTX-M}$, the AMR assay is largely insensitive to variation in gene sequences.

[00305] The ability of the Test AMR detection assays to call absence may be impeded by the presence of non-AMR sequences that are a match for the primers and similar in sequence to the AMR gene. Like the inclusivity assessment above, an *in-silico* approach was chosen to assess a wide array of possible cross-reactive sequences. An exclusivity of 100% was observed (table in

FIG. 7), showing a remarkably high assay specificity for the AMR markers tested. From a literature and database search, allotypes were identified of *mecC*, *mecC1* and *mecC2*, which have 93.7% and 96.3% sequence identity to *mecC*, respectively. When simulating samples using these allotypes, *mecC* was called in all simulations, due to the very high sequence similarity. Only *mecC1* is an allotype known to be phenotypically susceptible to methicillin. This allotype has the potential to affect exclusivity, but it is expected to be extremely rare in the patient population, given that *mecC* itself has a low prevalence in patients (0-3%). Overall, exclusivity is highly unlikely to be impacted by any cross-reactive sequence with <95% sequence identity to the AMR markers.

1.24 Clinical validation.

[00306] The positive percent agreement (PPA), negative percent agreement NPA), categorical agreement (CA), and diagnostic yield (DY) and 95% confidence intervals in brackets were estimated for each maker. All data for clinical validation samples are shown in Supplemental **Table 7**.

[00307] Correlation of SCC*mec* results with phenotype for 58 *S. aureus* detections are shown in Table 7. The PPA, NPA, CA and DY were 14/15 (93.3%) [68.0-99.8%], 19/20 (95.0%) [75.1-99.8%], 33/35 (94.3%) [80.8-99.3%] and 35/58 (60.3%) [46.6-72.9%], respectively. In the one false positive both SCC*mec* and *mecA* were detected and likely represents a true positive. The single false negative remained a false negative after review.

[00308] Correlation of *mecA* results with phenotype for the same 58 *S. aureus* detections are shown in Table 8. The PPA, NPA, CA, and DY were 14/15 (93.3%) [68.0-99.8%], 19/20 (95.0%) [75.1-99.8%], 33/35 (94.3%) [80.8-99.3%] and 35/58 (60.3%) [46.6-72.9%], respectively. The single false positive and false negative occurred with the same two strains described above for SCC*mec*. No *mecC*-positive strains were detected.

[00309] The results when the SCC*mec* and *mecA* tests were combined for 58 *Staphylococcus aureus* and 2 *S. epidermidis* detections in the decision workflow for methicillin resistance in staphylococci are shown in Table 9. As expected, the PPA, NPA and CA were nearly identical to what was observed when each target was assessed separately. However, the DY increased to 70% [56.8-81.1%].

[00310] Correlation of *vanA* results with phenotype for 6 *Enterococcus spp.* detections (4 *E. faecium* and 2 *E. faecalis*) are shown in Table 10. The PPA, NPA, CA, and DY were 3/3 (100%) [29.2-100%], 2/2 (100%) [15.8-100%], 5/5 (100%) [47.8-100%] and 5/6 (83.3%) [35.8-99.5%], respectively. No *vanB*-positive strains were detected.

[00311] Correlation of *bla*$_{\text{CTX-M}}$ results with ESBL phenotype for 49 Gram-negative bacilli detections are shown in Table 11. The PPA, NPA, CA, and DY were 5/9 (55.5%) [21.2-86.3%], 26/26 (100%) [86.7-100%], 31/35 (88.5%) [73.2-96.8%] and 35/49 (71.4%) [56.7-83.4%], respectively. The 5 true positive *bla*$_{\text{CTX-M}}$ detections included 4 *E. coli* and 1 *K.pneumoniae*. The 4 false negative *bla*$_{\text{CTX-M}}$ detections included 3 *E. cloacae* complex detections and 1 detection of *E. coli*.

[00312] ESBLs are not readily distinguishable phenotypically from each other and de-repressed or plasmid-encoded AmpC β-lactamases. Susceptibility to cephamycins can help distinguish between them but data were unavailable for these strains. However, ESBL and AmpC β-lactamases can be distinguished phenotypically based on differential susceptibility to cefepime. Cefepime is usually not hydrolyzed by AmpC β-lactamases whereas ESBL producers most often have elevated cefepime MICs and are interpreted as susceptible dose-dependent or resistant to cefepime.

[00313] Two of the three *E. cloacae* complex and the single *E. coli* strains were susceptible to cefepime. In one of the *E. cloacae* complex strains *bla*$_{\text{CTX-M}}$ was also not detected by the Verigene BC-GN test. The remaining *E. cloacae* complex detection was resistant to cefepime, ceftriaxone and aztreonam. The resolution of the four false negatives using these revised criteria is shown in Table 12. Following resolution of false negative result discrepancies, the PPA of *bla*$_{\text{CTX-M}}$ increased from 55.5% to 83.3% and the CA increased from 88.5% to 97.1%. The other performance metrics remained the same. As a check on the revised criteria for labeling samples it was observed that all five true positive *bla*$_{\text{CTX-M}}$ detections were from strains that were resistant to cefepime.

[00314] Correlation of *bla*$_{\text{KPC}}$ results with carbapenemase phenotype for 44 Gram-negative bacilli detections are shown in Table 13. The PPA, NPA, CA, and DY were 0/2 (0%) [0-84.2%], 23/23 (100%) [85.1-100%], 23/25 (92.3%) [73.9-99.0%] and 25/44 (56.8%) [41.0-71.6%], respectively. There were no true positive results. The 2 false negative results included 1 detection of *P. aeruginosa* by the Test in a patient who had two cultures positive for strains with different carbapenem phenotypes (BAL meropenem susceptible and sputum meropenem resistant) and one *E. coli* detection that was that was resistant to both ertapenem and meropenem. The most resistant *P. aeruginosa* strain was selected to label the specimen, but the strain from the BAL was deemed more likely to be associated with disease and therefore most likely a true negative. The false negative detection for the *E. coli* could not be resolved.

1.3 Discussion

[00315] Metagenomic sequencing is rapidly evolving and has the potential for detection of microbial genomes directly in clinical samples, particularly in the context of culture-negative samples. However, sequencing of mcfDNA from plasma has the distinct advantage of not requiring samples of infected body fluids or tissues which often require invasive procedures to collect. There are currently no such FDA-cleared assays for detection of bacterial pathogens and resistance genes/determinants; however, this field is expected to advance quickly, given the rapid advances in technology and the clinical need. This is the first disclosure on the use of plasma mcfDNA sequencing to identify AMR genetic markers for prediction of AMR. mcfDNA comprises < 0.001% of the total cell free DNA in plasma and consists of short fragments of <100 bp in length most with fragment lengths of approximately 40 bp.

[00316] The rarity and short fragment lengths of mcfDNA in plasma present significant technical challenges for detection of AMR markers. Another challenge is the small size of the AMR markers in relation to bacterial genome sizes. The SCC*mec* marker ranges in size from 21 to 60 kb and the individual AMR target genes range from approximately 900 to 2000 bp. As an example, *S. aureus* has a genome size of 2.8 mb so SCC*mec* and *mecA* represent approximately only 2.1% and 0.07% of its genome, respectively. While the analytical and clinical accuracy of the SCC*mec* feature was very high, the diagnostic yield was relatively low, at approximately 50%, because there is no enrichment for SCC*mec*. In contrast, the single AMR gene tests required PCR amplification prior to sequencing to enrich target sequences. Other mNGS assays have also used selective enrichment of AMR genes prior to sequencing to enhance their detection.

[00317] Despite substantial technical challenges it was possible to develop and incorporate robust assays for AMR markers into the Test. The LoDs for both the presence and absence of all AMR markers were determined. In a typical production sample (unique WINC 300,000), the presence LoD is 3,916 MPM and absence LoD is 683 MPM for the SCC*mec*. The higher presence LoD is likely due to interfering species in the human background which cause no-calls for MRSA assemblies when SCC*mec* and *S. aureus* are low. In addition, the probit fits did not fit the observed data patterns very well for MSSA because there was a sharp boundary between no absence calls and >95% absence calls.

[00318] Overall, the presence LoDs ranged from 425 to 6,107 MPM and the absence LoDs ranged from 472 to 22,680 MPM for the single AMR gene assays. The *vanB* and bla_{CTX} assays represented the biggest LoD challenges for the presence and absence of these genes at 6,107 and 22,680 MPM, respectively. The *vanB* variant carried by the ATCC strain used in the genome mix shares 95% nucleotide identity with *vanB* used to design targeting primers. It was compensated

for the loss of signal due to imperfect primer binding sites by increasing the *vanB* concentration in the genome five-fold. The high LoD for *bla*$_{CTX}$ was a result of elevated reagent background signal for this gene. All of the AMR marker assays had within- and between-run precision of 100%.

[00319] The inclusivity of the assays was 100% for all alleles except *bla*$_{CTX}$. Inclusivity for *bla*$_{CTX}$ was limited to 54% because the assay has primers with perfect homology only to a single clade of *bla*$_{CTX-M}$ variants (*bla*$_{CTX-M-3}$ clade), but not the other highly divergent clades. However, based on the prevalence of *bla*$_{CTX}$ variants in North America the capture of 67-75% of these variants in the assay was estimated. The exclusivity was estimated to be 100% for all markers.

[00320] Aside from good laboratory practice, it is now a CAP requirement (CAP checklist item MIC.21855) to link AMR determinants to a specific organism in the final patient report. Linkage of the AMR marker with a specific bacterium has been challenging for mNGS assays. For each AMR marker gene, the assays identify all the bacteria detected by the Test known to carry that gene, referred to here as "carriers", and their respective MPM values. A statistical model evaluates the likelihood of all the possible combinations of carriers and AMR gene copy number given an established prior to output a probability of linkage of AMR to each carrier. If the linkage cannot be established the result is reported as indeterminate.

[00321] Unlike many molecular and phenotypic assays, the assays not only detect *vanA*, *vanB*, *mecA*, and *mecC* but also differentiate between them. While *vanA* encodes resistance to vancomycin and other glycopeptides, *vanB* encodes resistance to vancomycin only. Although both *mecA* and *mecC* encode resistance to methicillin, *mecC* is much less common and more difficult to detect phenotypically, and as a result may be missed and reported as MSSA. In addition, this feature may increase the understanding of the prevalence of *mecC*. However, neither *vanB* nor *mecC* genes were detected in the clinical validation samples.

[00322] The Test AMR marker detection covers 8 of 10 bacterial species driving global AMR-attributable deaths and all of the ESKAPE pathogens which are important drivers of multidrug resistance, and the panel of 7 markers detect resistance to 5 different classes of antibiotics, including anti-staphylococcal semisynthetic penicillins, glycopeptides, oxyimino-cephalosporins, aztreonam and carbapenems. Resistance to these drugs is among CDC's top urgent and serious AMR threats in the US. In addition, it was estimated from a commercial sample database that the target bacteria comprise approximately 18% of all reported detections and thus could be eligible for AMR marker detection.

[00323] Clinical accuracy was assessed with 115 plasma samples obtained from patients at 7 study sites with concordant culture results for any of the 18 target bacteria isolated from a variety of specimen types and correlated with available orthogonal phenotypic antimicrobial

susceptibility test results in all cases and genotypic results when available. The performance characteristics of the SCC*mec* and *mecA* assays were assessed individually with the same 58 *S. aureus* detections and were found to be virtually identical except for DY which increased from 53.4% to 60.3% with the addition of the *mecA* assay. The performance for the combination of SCC*mec* and *mecA* for all staphylococci as they were intended to be used together in the AMR workflow (Fig. 1) showed excellent PPA, NPA, CA, all >95%, and a DY of 70.0%. Although the number of samples that contained vancomycin-resistant enterococci was small, the PPA, NPA, and CA were 100%, and the DY was 83.3%.

[00324] The estimate of the PPA for the $bla_{\text{CTX-M}}$ was only 55.5% based on initial criteria of resistance to any oxyimino-beta-lactams, or aztreonam to label samples as putative ESBL producers. However, ESBLs are not readily distinguishable phenotypically from each other and from de-repressed or plasmid-encoded AmpC β-lactamase*s*. ESBL and AmpC β-lactamases can be distinguished phenotypically based on differential susceptibility to cefepime. Cefepime is usually not hydrolyzed by AmpC β-lactamases whereas ESBL producers most often have elevated cefepime MICs and are interpreted as susceptible dose-dependent or resistant to cefepime. Using susceptibility to cefepime as the sole criterion for labeling samples as putative ESBL producers, the PPA and DY improved to 83.3.% and 97.1%, respectively. The remaining false negative was an *E. cloacae* complex detection that was resistant to cefepime, ceftriaxone and aztreonam.

[00325] Our clinical validation samples contained only two bacteria that were phenotypically resistant to carbapenems, *E. coli* and *P. aeruginosa*, and in both bla_{KPC} was not detected. In a recent study of prevalence of carbapenem non-susceptible *Enterobacterales* (CNSE), 34,623 clinically significant isolates collected from 86 US hospitals from 2016 to 2020 were submitted to a central laboratory for analysis. Only 450 (1.3%) were CNSE. Of those, KPC production was the most common resistance mechanism among CNSE isolates: 281/450 (62.4%) carried bla_{KPC}, including three new variants. OXA-48-like and metallo-β-lactamase (MBL) encoding genes were detected among seven (1.7%) and 12 (2.7%) isolates, respectively. Among MBL genes, $bla_{\text{NDM-1}}$ was the most common, but $bla_{\text{NDM-5}}$, $bla_{\text{VIM-1}}$ and $bla_{\text{IMP-27}}$ were also identified. One hundred and sixty-nine (37.6%) isolates did not produce carbapenemases.

[00326] Membrane permeability, together with derepression of the intrinsic beta-lactamase gene, is the global prevailing mechanism of carbapenem resistance in *P. aeruginosa*, but the acquired genes for carbapenemases are of increasing concern. Production of serine carbapenemases (e.g., KPC) by *Pseudomonas aeruginosa* are rare but MBLs may be present. The failure to detect a bla_{KPC} positive strain is likely due to the small sample size, rarity of the gene, and the diversity of mechanisms of carbapenem resistance in gram-negative bacilli.

[00327] A limitation of the study is the relatively small number of samples available for clinical validation. This was due to the stringent criteria used to qualify the samples: residual patient plasma samples which had one of target pathogens detected in the Test, concordant culture results, applicable orthogonal phenotypic AST results, and adequate residual plasma volume. However, the supporting data and samples were collected from 7 study sites across the US and included a wide variety of sample types. Taken together, the results with this sample set suggest that they are generalizable and demonstrate that Test can detect AMR markers in bacteria causing infections throughout the body.

[00328] In some embodiments, orthogonal phenotypic AST was performed with the Vitek2 and MicroScan Walkaway systems. In some embodiments, these systems are used by clinical microbiology laboratories. In some embodiments, these systems provide a better estimate of the performance of these assays in clinical practice, such as compared to other clinical methods.

[00329] DY was defined for each AMR marker as the percentage of tests that yielded a clinically actionable result (detected/not detected) and ranged from 56.8% for bla_{KPC} to 83.3% for *vanA*. While some might consider the DY low, the assays were optimized for accuracy. This is best demonstrated by the results for methicillin resistance in staphylococci and cefepime resistance in gram-negative bacilli in which the CA were estimated to be 95.2% and 97.1%, respectively.

[00330] Based on earlier identification of the cause of infection and increased diagnostic yield, mcfDNA sequencing results led to a change in antimicrobial management in 32%-72% of the reported cases. The addition of specific AMR markers to the Test will provide clinicians more opportunities for optimization of therapy and anticipate the greatest clinical impact in those patients with infections in which the microbe is detected only by mcfDNA sequencing and in immunocompromised hosts who are particularly vulnerable to infections with multidrug-resistant bacteria. Faster identification of drug-resistant organisms has the potential to prevent the spread of multidrug resistance, improve patient outcomes by enabling appropriate therapy to get patients on the right targeted therapy faster, and potentially reducing mortality and hospitalization costs. In the AMR workflow that is described the results can be reported within a clinically relevant time frame with results being available 1 to 3 days after receipt of the sample in the laboratory.

[00331] In summary, the validation of a novel application of mcfDNA sequencing for detection of AMR markers that has been incorporated into the Test workflow without the need for a positive culture or intact microbial genomes is reported herein. The key to success was enrichment for the ultrashort and rare mcfDNA fragments prior to sequencing for identification of AMR markers. The workflow could readily be adapted to expand the number of target bacteria

and AMR genes as needed. Future studies will be focused on real-world evidence assessing the clinical impact of the Test with expanded AMR marker detection.

[00332] The subject matter defined in the appended claims is not necessarily limited to the specific implementations described above. Various alternatives to the embodiments of the invention described herein may be employed in practicing the invention. The specific implementations described above are disclosed as examples only. It is intended that the following claims define the scope of the invention and that methods and structures within the scope of these claims and their equivalents be covered thereby.

CLAIMS

WHAT IS CLAIMED IS:

1. A method for performing a multi-step assay for detecting nucleic acids in a sample from a subject, the method comprising:

   (a) providing a first aliquot of the sample from the subject, wherein the sample comprises nucleic acids comprising a first target nucleic acid and a second target nucleic acid that is different from the first target nucleic acid;

   (b) performing a sequencing assay on the nucleic acids to produce sequence reads comprising sequence reads associated with the first target nucleic acid;

   (c) analyzing the sequence reads, thereby obtaining an identification of the first target nucleic acid;

   (d) after the identification of the first target nucleic acid is obtained, providing a second aliquot of the sample comprising nucleic acids comprising the first target nucleic acid and the second target nucleic acid;

   (e) introducing primers into the second aliquot or to nucleic acids derived from the second aliquot wherein the primers specifically target the second target nucleic acid and do not target the first target nucleic acid; and

   (f) conducting an amplification reaction on the second target nucleic acid with the primers, thereby amplifying the second target nucleic acid and producing amplicons associated with the second target nucleic acid.

2. The method of claim 1, further comprising performing a high-throughput sequencing assay on the amplicons associated with the second target nucleic acid.

3. The method of any one of the preceding claims, wherein the second target nucleic acid is not detected by the sequencing assay in (b).

4. The method of any one of the preceding claims, wherein the first nucleic acid is associated with a genome of an organism but not a phenotype of interest of the organism and the second nucleic acid is associated with the phenotype of interest of the organism.

5. The method of any one of the preceding claims wherein the first target nucleic acid comprises at least two, at least three, at least four, at least five, at least ten, at least 15, at least 20, or at least 25 first target nucleic acids; or the second target nucleic acid comprises at least two, at least three, at least four, at least five, at least ten, at least 15, at least 20, or at least 25 second target nucleic acids.

6. The method of any one of the preceding claims wherein the first target nucleic acid or the second target nucleic acid are not negative or positive controls for any step of the multi-step assay.

7. The method of any one of the preceding claims, wherein the primers comprise multiple primers targeting multiple target nucleic acids.

8. The method of any one of the preceding claims, wherein the first target nucleic acid comprises nucleic acids from an animal, a virus, a bacterium, a protozoa, a fungus, an archaea, or an algae.

9. The method of any one of the preceding claims, wherein the second target nucleic acid comprises nucleic acids from an animal, a virus, a bacterium, a protozoa, a fungus, an archaea, or an algae.

10. The method of any one of the preceding claims, wherein the second target nucleic acid comprises a cancer marker.

11. The method of any one of the preceding claims, wherein the first target nucleic acid is associated with a carrier microbe harboring a target genetic marker and the second target nucleic acid comprises a sequence associated with the target genetic marker.

12. The method of claim 11, where the carrier microbe comprises a virus, a bacterium, a protozoa, a fungus, an archaea, or an algae.

13. The method of any one of the preceding claims, wherein the target genetic marker comprises an antimicrobial resistance (AMR) genetic marker.

14. A method for performing a multi-step sample processing procedure for detecting nucleic acids in a sample from a subject:

(a) providing a first aliquot of the sample from the subject, wherein the sample comprises one or more first nucleic acids, wherein the one or more first nucleic acids comprise one or more microbial nucleic acids derived from one or more carrier microbes harboring an antimicrobial resistance (AMR) genetic marker;

(b) performing a sequencing assay on the one or more first nucleic acids to produce sequence reads comprising microbial nucleic acid sequence reads from the one or more carrier microbes;

(c) detecting the one or more carrier microbes by analyzing the microbial nucleic acid sequence reads;

(d) after at least one carrier microbe is detected, providing a second aliquot of the sample from the subject, wherein the second aliquot comprises second nucleic acids from at least one AMR genetic marker;

(e) annealing primers targeting at least one AMR genetic marker to the second nucleic acids; and

(f) amplifying the AMR genetic marker in an amplification reaction to produce amplicons associated with the at least one AMR genetic marker.

15. A method for assaying nucleic acids in a sample from a subject comprising:

(a) providing a first aliquot of the sample from the subject, wherein the sample comprises one or more first nucleic acids, wherein the one or more first nucleic acids comprise one or more microbial nucleic acids derived from one or more carrier microbes harboring an antimicrobial resistance (AMR) genetic marker;

(b) performing a sequencing assay on the one or more first nucleic acids to produce sequence reads comprising microbial nucleic acid sequence reads from the one or more carrier microbes;

(c) detecting the one or more carrier microbes by analyzing the microbial sequence reads;

(d) providing a second aliquot of the sample from the subject, wherein the second aliquot comprises second nucleic acids, wherein the second aliquot comprises second nucleic acids from at least one AMR genetic marker;

(e) annealing primers to the second nucleic acids that target the at least one AMR genetic marker, wherein the primers that target the at least one AMR genetic marker are annealed prior to amplification of nucleic acids in the sample; and

(f) conducting an amplification reaction on the second nucleic acids using the primers targeting the at least one AMR genetic marker, thereby amplifying at least one AMR genetic marker and producing amplicons associated with the at least one AMR genetic marker.

16. A method for performing a multi-step sample processing procedure for detecting nucleic acids in a sample from a subject:

(a) providing a first aliquot of the sample from the subject, wherein the sample comprises one or more first nucleic acids, wherein the one or more first nucleic acids comprise one or more microbial nucleic acids derived from one or more carrier microbes harboring an antimicrobial resistance (AMR) genetic marker;

(b) performing a sequencing assay on the one or more first nucleic acids to produce sequence reads comprising microbial nucleic acid sequence reads from the one or more carrier microbes;

(c) detecting the one or more carrier microbes by analyzing the microbial nucleic acid sequence reads;

(d) after at least one carrier microbe is detected, providing a second aliquot of the sample from the subject, wherein the second aliquot comprises second nucleic acids from at least one AMR genetic marker;

(e) annealing primers to the second nucleic acids that target at least one AMR genetic marker; and

(f) amplifying the AMR genetic marker and producing amplicons associated with the at least one AMR genetic marker.

17. A method for assaying nucleic acids in a sample from a subject comprising:

(a) providing a first aliquot of the sample from the subject, wherein the first aliquot comprises first nucleic acids, and the first nucleic acids comprise an AMR genetic marker and a gene cassette comprising the AMR genetic marker derived from one or more carrier microbes harboring the AMR genetic marker;

(b) performing a sequencing assay on the first nucleic acids to produce sequence reads comprising microbial nucleic acid sequence reads from the one or more carrier microbes;

(c) detecting the AMR genetic marker or the gene cassette comprising the AMR genetic marker in the sample and quantifying microbial sequence reads from the AMR genetic marker and/or microbial sequence reads from the gene cassette;

(d) after and based on the quantification, providing a second aliquot of the sample from the subject, wherein the second aliquot comprises second nucleic acids, wherein the second nucleic acids comprise the AMR genetic marker;

(e) introducing primers targeting the AMR genetic marker to the second nucleic acids, wherein the primers targeting the AMR genetic marker are introduced prior to amplification of nucleic acids in the second aliquot; and

(f) conducting an amplification reaction on the second nucleic acids using the primers targeting the AMR genetic marker, thereby amplifying the AMR genetic marker and producing amplicons associated with the AMR genetic marker.

18. The method of any one of the preceding claims, further comprising in (d) comparing an abundance of the microbial sequence reads from the AMR genetic marker to a threshold value.

19. The method of any one of the preceding claims, further comprising in (d) comparing an abundance of the microbial sequence reads from the gene cassette to a threshold value.

20. The method of any one of the preceding claims, further comprising providing a second aliquot for the amplification reaction when the abundance of the microbial sequence reads from the AMR genetic marker is below the threshold.

21. The method of any one of the preceding claims, further comprising providing a second aliquot for the amplification reaction when the abundance of the microbial sequence reads from the gene cassette is below the threshold.

22. The method of any one of the preceding claims, further comprising in (d) quantifying mcfNA sequencing reads from the one or more microbes.
23. The method of any one of the preceding claims, further comprising comparing the abundance of the mcfNA sequencing reads from the one or more microbes to a threshold.
24. The method of any one of the preceding claims, further comprising providing a second aliquot for the amplification reaction when the abundance of the mcfNA sequencing reads from the one or more microbes is below the threshold.
25. The method of any one of the preceding claims, further comprising performing high-throughput sequencing on the amplicons associated with the AMR genetic marker.
26. The method of any one of the preceding claims, wherein the sequencing assay on the first nucleic acids comprises a high-throughput sequencing assay.
27. The method of any one of the preceding claims, further comprising conducting a polymerase chain reaction (PCR) to amplify the AMR genetic marker, thereby producing amplicons associated with the AMR genetic marker.
28. The method of any one of the preceding claims, wherein the PCR comprises multiplex PCR, random PCR (rPCR), non-biased PCR, Nested PCR, Hot Start PCR, or Assembly PCR.
29. The method of any one of the preceding claims, further comprising attaching an adapter sequence to the second nucleic acids.
30. The method of any one of the preceding claims, wherein the primers comprise an adapter sequence.
31. The method of any one of the preceding claims, further comprising physically manipulating the sample to produce a fraction of cfNA enriched for degraded cfNA, wherein the fraction of cfNA comprises the AMR genetic marker.
32. The method of any one of the preceding claims, wherein the degraded cfNA comprises ultra short cfNA, single stranded cfNA, or nicked double stranded cfNA.
33. The method of any one of the preceding claims, wherein the ultra short cfNA comprises cfNA fragments that are less than 100 nucleotides in length.
34. The method of any one of the preceding claims, wherein the ultra short cfNA comprises cfNA fragments that are from 30 nucleotides to 70 nucleotides in length.
35. The method of any one of the preceding claims, wherein physically manipulating the sample comprises performing size selection of the nucleic acids in the sample.
36. The method of any one of the preceding claims, wherein performing size selection using a method selected from the group consisting of: chromatography, size-exclusion chromatography, electrophoresis, gel electrophoresis, automated electrophoresis,

capillary electrophoresis, high-throughput size selection, centrifugation, density-gradient centrifugation, filtration, membrane ultrafiltration, magnetic beads, and affinity-based beads.

37. The method of any one of the preceding claims, wherein the second aliquot comprises at least 500 μl of plasma.
38. The method of any one of the preceding claims, wherein detecting the carrier microbe comprises determining an abundance of the carrier microbe.
39. The method of any one of the preceding claims, wherein the abundance is expressed as molecules of mcfNA per microliter of sample (MPM).
40. The method of any one of the preceding claims, further comprising calculating an AMR gene copy number from the amplification or sequencing of the AMR genetic marker.
41. The method of any one of the preceding claims, further comprising linking the AMR genetic marker to the carrier microbe using the abundance of the mcfNA from the carrier microbe and the AMR gene copy number.
42. The method of any one of the preceding claims, where the AMR gene copy number is an episomal gene copy number.
43. The method of any one of the preceding claims, comprising introducing at least 200 primers targeting the plurality of AMR genetic markers into the second aliquot of the sample.
44. The method of any one of the preceding claims, further comprising introducing primers targeting housekeeping genes into the second aliquot of the sample.
45. The method of any one of the preceding claims, further comprising determining antimicrobial resistance of the microbe infecting the subject.
46. The method of any one of the preceding claims, wherein the sample comprises plasma.
47. The method of any one of the preceding claims, wherein the first nucleic acids are DNA.
48. The method of any one of the preceding claims, wherein the second nucleic acids are DNA.
49. The method of any one of the preceding claims, wherein the sample is not subjected to a process that primarily causes cell lysis.
50. The method of any one of the preceding claims, further comprising preparing a library from the amplicons associated with the AMR genetic marker .
51. The method of any one of the preceding claims, wherein the primers are added directly to the second aliquot.
52. The method of any one of the preceding claims, wherein the sample comprises cell-free nucleic acids.

53. The method of any one of the preceding claims, wherein the sample comprises cell-free DNA.

54. The method of any one of the preceding claims, wherein the first nucleic acids comprise microbial cell-free nucleic acids.

55. The method of any one of the preceding claims, wherein the second nucleic acids comprise microbial cell-free nucleic acids.

56. The method of any one of the preceding claims, wherein intact microbes are not actively lysed prior to performing the sequencing assay.

57. A method for preparing a cell-free DNA fraction from a subject infected with a microbe useful for analyzing an antimicrobial resistance (AMR) genetic marker, comprising:

(a) providing a sample from the subject, wherein:

(i) the sample comprises cell-free nucleic acids (cfNA) from the subject and microbial nucleic acids from the microbe infecting the subject;

(ii) the cfNA comprises degraded cfNA, single-stranded cfNA and intact cfNA; and

(iiii) the microbe harbors an AMR genetic marker;

(b) physically manipulating the sample in order to produce a fraction of cfNA enriched for degraded cfNA and single-stranded cfNA, wherein the fraction of cfNA comprises microbial nucleic acids comprising a genetic locus associated with the AMR genetic marker;

(c) performing a sequencing assay on the fraction of cfNA enriched for the degraded cfNA and single-stranded cfNA; and

(d) analyzing the genetic locus associated with the AMR genetic marker.

58. The method of any one of the preceding claims, wherein the microbial nucleic acids comprise microbial cell-free nucleic acids (mcfNA) from the microbe.

59. The method of any one of the preceding claims, wherein the microbial nucleic acids comprise nucleic acids associated with a microbial cell.

60. The method of any one of the preceding claims, wherein (c) comprises enriching for at least at least 75%, at least 80%, at least 85%, or at least 90% of the mcfNA in the sample.

61. The method of any one of the preceding claims, wherein the degraded cfNA comprise ultra short cfNA, single stranded cfNA, nicked double stranded cfNA, or any combination thereof.

62. The method of any one of the preceding claims, wherein the degraded cfNA comprise cfNA fragments that are less than 100 nucleotides in length.

63. The method of any one of the preceding claims, wherein the degraded cfNA comprise cfNA fragments that are from 30 nucleotides to 70 nucleotides in length.

64. The method of any one of the preceding claims, wherein physically manipulating the sample comprises performing size selection of the nucleic acids in the sample.

65. The method of claim 64, wherein performing size selection using a method selected from the group consisting of: chromatography, size-exclusion chromatography, electrophoresis, gel electrophoresis, automated electrophoresis, capillary electrophoresis, high-throughput size selection, centrifugation, density-gradient centrifugation, filtration, membrane ultrafiltration, magnetic beads, and affinity-based beads.

66. The method of any one of the preceding claims, further comprising generating sequence reads from the cfNA from the subject, wherein the sequence reads comprise microbial sequencing reads derived from the microbe infecting the subject.

67. The method of any one of the preceding claims, wherein the sequence reads further comprise microbial sequencing reads from one or more carrier microbe.

68. The method of any one of the preceding claims, further comprising calculating an abundance of the mcfNA from the microbe in the sample and an abundance of the mcfNA from the one or more carrier microbe.

69. The method of any one of the preceding claims, wherein the abundance is expressed as molecules of mcfNA per microliter of sample (MPM).

70. The method of any one of the preceding claims, wherein the abundance is expressed as molecules of mcfNA per 100 nanoliters of sample.

71. The method of any one of the preceding claims, further comprising calculating an AMR gene copy number for each of the one or more carrier microbe.

72. The method of any one of the preceding claims, where the AMR gene copy number is an episomal gene copy number.

73. The method of any one of the preceding claims, further comprising detecting one or more carrier microbe from the sequencing assay.

74. The method of any one of the preceding claims, further comprising identifying the one or more carrier microbe as the microbe infecting the subject.

75. The method of any one of the preceding claims, further comprising linking the AMR genetic marker to one of the carrier microbes or the microbe infecting the subject.

76. The method of any one of the preceding claims, further comprising determining antimicrobial resistance of the microbe infecting the subject.

77. A method for determining antimicrobial resistance of a microbe infecting a subject, comprising:

(a) providing a sample comprising nucleic acids from the subject, wherein

(b) the nucleic acids comprise microbial nucleic acids from one or more potential carrier microbes harboring an antimicrobial resistance (AMR) genetic marker;

(c) performing a sequencing assay on the nucleic acids from the subject to generate sequence reads and detecting the one or more potential carrier microbe based on the sequence reads;

(d) if a carrier microbe is detected in (c), performing an amplification reaction on an aliquot of the sample, wherein the amplification reaction targets one or more AMR genetic markers in order to provide a copy number for the one or more AMR genetic markers; and

(e) conducting a statistical analysis to link one or more of the AMR genetic markers to one of the one or more potential carrier microbe in the subject, thereby determining that a potential carrier microbe comprises an AMR genetic marker.

78. The method of any one of the preceding claims, further comprising obtaining an abundance of the one or more carrier microbes based on an abundance of carrier microbe sequences in the sample.

79. The method of any one of the preceding claims, wherein conducting a statistical analysis comprises using the abundance of the microbial nucleic acids of the one or more potential carrier microbes.

80. The method of any one of the preceding claims, wherein conducting a statistical analysis comprises using the AMR gene copy number.

81. The method of any one of the preceding claims, where the AMR gene copy number is an episomal gene copy number.

82. The method of any one of the preceding claims, wherein the abundance is expressed as molecules of mcfNA per microliter of sample (MPM).

83. The method of any one of the preceding claims, wherein the abundance is expressed as molecules of mcfNA per 100 nanoliters of sample.

84. The method of any one of the preceding claims, further comprising calculating a probability of each of the one or more potential carrier microbes being the microbe harboring the AMR genetic marker, thereby identifying the potential carrier microbe as the microbe infecting the subject.

85. The method of any one of the preceding claims, further comprising determining the antimicrobial resistance of the microbe infecting the subject based on the calculated probability.

86. A method for determining antimicrobial resistance of a microbe infecting a subject, comprising:

(a) providing a first aliquot of the sample comprising first nucleic acids, wherein the first nucleic acids comprise microbial nucleic acids derived from a potential carrier microbe harboring an antimicrobial resistance (AMR) genetic marker;

(b) detecting the potential carrier microbe by analyzing the first nucleic acids;

(c) after the potential carrier microbe is detected, providing a second aliquot of the sample comprising second nucleic acids, wherein the second nucleic acids comprise microbial nucleic acids from the potential carrier microbe harboring an AMR genetic marker;

(d) conducting an amplification reaction on the second nucleic acids to detect the AMR genetic marker; and

(e) determining the antimicrobial resistance of the microbe infecting the subject based on the AMR genetic marker detection.

87. The method of any one of the preceding claims, wherein the detecting comprises performing a sequencing assay.

88. The method of any one of the preceding claims, wherein the sequencing assay comprises a high-throughput sequencing assay.

89. The method of any one of the preceding claims, wherein the amplification reaction comprises introducing primers targeting a plurality of AMR genetic markers.

90. The method of any one of the preceding claims, wherein the amplification reaction produces amplicons associated with the AMR genetic marker.

91. The method of any one of the preceding claims, further comprising performing high-throughput sequencing on the amplicons associated with the AMR genetic marker.

92. The method of any one of the preceding claims, wherein the sample comprises plasma.

93. The method of any one of the preceding claims, wherein the first nucleic acids are DNA.

94. The method of any one of the preceding claims, wherein the second nucleic acids are DNA.

95. The method of any one of the preceding claims, wherein the sample does not undergo cell lysis.

96. The method of any one of the preceding claims, wherein intact microbes are not actively lysed.

97. The method of any one of the preceding claims, wherein the amplicons associated with the AMR genetic marker undergoes direct library preparation.

98. The method of any one of the preceding claims, wherein the sample comprises cell-free nucleic acids.

99. The method of any one of the preceding claims, wherein the sample comprises cell-free DNA.

100. The method of any one of the preceding claims, wherein the first nucleic acids comprise cell-free nucleic acids.

101. The method of any one of the preceding claims, wherein the second nucleic acids comprise cell-free nucleic acids.

102. The method of any one of the preceding claims, wherein (d) comprises adding primers targeting a plurality of AMR genetic markers into the second aliquot of the sample.

103. The method of any one of the preceding claims, comprising adding at least 200 primers targeting a plurality of AMR genetic markers into the second aliquot of the sample.

104. The method of any one of the preceding claims, further comprising introducing primers targeting housekeeping genes into the second aliquot of the sample.

105. The method of any one of the preceding claims, wherein the second aliquot of the sample comprises at least 500 µl of plasma.

106. The method of any one of the preceding claims, wherein detecting the pathogen comprises detecting the abundance of mcfNA from the pathogen in the sample over a threshold.

107. The method of any one of the preceding claims, further comprising calculating a positive percent agreement (PPA), negative percent agreement (NPA), diagnostic yield (DY), or any combination thereof.

108. The method of any one of the preceding claims, further comprising detecting cfDNA from one or more housekeeping genes.

109. The method of any one of the preceding claims, further comprising spiking one or more control molecules into the sample at a known concentration.

110. The method of any one of the preceding claims, wherein the one or more control molecules are synthetic oligonucleotides.

111. The method of any one of the preceding claims, wherein the control molecules comprise whole assay internal control (WINC) molecules.

112. The method of any one of the preceding claims, further comprising spiking at least 25,000 unique WINC molecules at known concentrations.

113. The method of any one of the preceding claims, further comprising generating a report listing the carrier microbes or pathogen detected in the subject.

114. The method of any one of the preceding claims, wherein the report further comprises the abundance of microbial cell-free DNA (mcfDNA) from microbes detected in the sample or the antimicrobial resistance of the microbes infecting the subject.

115. The method of any one of the preceding claims, wherein the sample comprises a biological sample obtained from the subject.

116. The method of any one of the preceding claims, wherein the biological sample is a whole blood sample, a plasma sample, a serum sample, a cerebrospinal fluid sample, a synovial fluid sample, a urine sample, a stool sample, a bronchoalveolar lavage sample, or any combination thereof.

117. The method of any one of the preceding claims, wherein the sample is a plasma sample.

118. The method of any one of the preceding claims, where the AMR genetic marker comprises a gene, a genetic element, a genetic cassette, an allele, a mutation, or any combination thereof.

119. The method of any one of the preceding claims, where the AMR genetic marker is associated with one or more genes selected from the group consisting of: *SCCmec, mecA, mecC, vanA, vanB,* $bla_{CTX\text{-}M}$, bla_{KPC}, *OXA-48-like*, *OXA-23, NDM, VIM, IMP*, or *mcr-1*.

120. The method of any one of the preceding claims, where the AMR genetic marker provides resistance to an anti-microbial agent selected from the group consisting of: methicillin, vancomycin, cephalosporin, carbapenem, and oxyimino-cephalosporin/aztreonam resistance.

121. The method of any one of the preceding claims, where the microbe infecting the subject or the carrier microbe comprises a virus, a bacterium, a protozoa, a fungus, an archaea, or an algae.

122. The method of any one of the preceding claims, where the microbe infecting the subject or the carrier microbe comprises a gram-positive bacterium or a fungus.

123. The method of any one of the preceding claims, where the microbe infecting the subject or the carrier microbe is a microbe listed in Table 1.

124. The method of any one of the preceding claims, where the microbe infecting the subject or the carrier microbe harbors at last two AMR genetic markers.

125. The method of any one of the preceding claims, wherein the antimicrobial resistance is phenotypic antimicrobial resistance.

126. The method of any one of the preceding claims, wherein the sequencing comprises next-generation sequencing or a sequencing method beyond next generation sequencing.

127. The method of any one of the preceding claims, wherein the sequencing comprises sequencing by synthesis.

128. The method of any one of the preceding claims, wherein the subject is an animal.

129. The method of any one of the preceding claims, wherein the subject is a human.

130. The method of any one of the preceding claims, wherein the subject has an infection by the microbe infecting the subject harboring the AMR genetic marker.

131. The method of any one of the preceding claims, further comprising administering an anti-infective agent to the subject.

132. The method of any one of the preceding claims, wherein the subject has been treated with an anti-infective agent for an infection.

133. The method of any one of the preceding claims, further comprising adjusting the anti-microbial agent received by the subject at least in part based on the antimicrobial resistance of the microbe infecting the subject.

134. The method of any one of the preceding claims, further comprising determining antimicrobial resistance of a microbe.

135. The method of claim 134, wherein the microbe comprises a microbe infecting the subject.

136. The method of claim 135, wherein the microbe comprises a carrier microbe.

137. The method of any one of the preceding claims wherein the method comprises determining whether an AMR gene is carried by a microbe infecting the subject or by a carrier microbe.

138. The method of any one of claims 134-137, wherein the determining comprises linking an AMR genetic marker to one of the carrier microbes or the microbe infecting the subject.

139. The method of any one of claims 134-138, wherein the determining comprises determining a copy number of the AMR genetic marker in the microbe.

140. The method of claim 138 or 139, wherein the linking comprises performing a statistical analysis on the sequence reads.

141. The method of claim 140, wherein the statistical analysis comprises a generalized linear model, a maximum-likelihood estimation, a probit model, a logistic regression, a linear probability, a linear regression, a complimentary log-log, a Poisson regression, a support vector machine, a decision tree, a random forest, a neural network, a gradient boosted model, a Bayesian model, a hidden Markov model, or any combination thereof.

142. The method of any one of claims 134-141, wherein the method further comprises comparing the sequence read data to the amplification data.

143. The method of any one of claims 134-142, wherein the method further comprises calculating estimated deduplicated templates (EDT) for an AMR gene, wherein calculating EDT comprises calculating a number of unique DNA sequencing reads from a microorganism present in a sequenced library.

144. The method of any one of claims 134-143, wherein the method further comprises calculating an EDT for a housekeeping gene, wherein calculating EDT comprises calculating a number of unique DNA sequencing reads from a microorganism present in a sequenced library.

145. The method of claim 143 or 144, further comprising determining a presence of an organism in a sample by calculating a ratio of an AMR estimated deduplicated templates (EDT) to a housekeeping gene EDT.

146. The method of any one of the preceding claims, wherein the method further comprises administering an alternative therapy to the subject if an AMR marker is detected.

147. The method of any one of the preceding claims, wherein the method further comprises detecting methicillin-resistant Staph aureus by detecting SCCmec, mecA, or mecC.

148. The method of any one of the preceding claims, wherein the method further comprises detecting methicillin-resistant Staph aureus and administering ceftaroline, ceftobiprole, vancomycin, linezolid, daptomycin, ceftobipole or combination thereof to the subject in order to treat the methicillin-resistant staph aureus.

149. The method of any one of the preceding claims, wherein the method further comprises detecting vancomycin resistant enterococcus by detecting VanA or VanB.

150. The method of any one of the preceding claims, wherein the method further comprises detecting vancomycin resistant enterococcus and administering linezolid or daptomycin to the subject in order to treat the vancomycin resistant enterococcus.