



(51) International Patent Classification:
C12Q 1/68 (2006.01)

(21) International Application Number:
PCT/US2016/060369

(22) International Filing Date:
3 November 2016 (03.11.2016)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
62/250,565 4 November 2015 (04.11.2015) US

(71) Applicants: **THE TRANSLATIONAL GENOMICS RESEARCH INSTITUTE** [US/US]; 445 N. Fifth Street, Suite 600, Phoenix, Arizona 85004 (US). **ARIZONA BOARD OF REGENTS ON BEHALF OF NORTHERN ARIZONA UNIVERSITY** [US/US]; 1298 South Knoles Drive, ARD Building #56, Suite 240, Flagstaff, Arizona 86011-4087 (US).

(72) Inventors: **DRIEBE, Elizabeth**; 5120 Meadow Vista Court, Flagstaff, Arizona 86001 (US). **BOWERS, Jolene**; 6225 N. Dodge Avenue, Flagstaff, Arizona 86004 (US). **ENGELTHALER, David**; 1255 N. Rockridge, Flagstaff, Arizona 86001 (US). **KEIM, Paul**; 3115 W. Tina Lane, Flagstaff, Arizona 86001 (US).

(74) Agent: **FULLER, Rodney J.**; Booth Udall Fuller, PLC, 1255 West Rio Salado Parkway, Suite 215, Tempe, Arizona 85281 (US).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

- without international search report and to be republished upon receipt of that report (Rule 48.2(g))
- with sequence listing part of description (Rule 5.2(a))

(54) Title: SYSTEMS AND METHODS OF DIAGNOSING AND CHARACTERIZING INFECTIONS

(57) Abstract: Embodiments of the invention include methods of identifying microorganisms and/or diagnosing infections in subjects cause by microorganisms. Embodiments of the invention may also include further characterizing (e.g., determining the presence of one or more antibiotic resistance markers) the microorganisms and determining a strain identity of the microorganisms.



SYSTEMS AND METHODS OF DIAGNOSING AND CHARACTERIZING INFECTIONS

CROSS REFERENCE TO RELATED APPLICATIONS

[0001] This application claims priority to and the benefit of U.S. Provisional Application Serial No. 62/250,565, filed on November 4, 2015, the contents of which are hereby incorporated by reference in their entirety.

INCORPORATION-BY-REFERENCE OF MATERIAL ELECTRONICALLY FILED

[0002] Incorporated by reference in its entirety herein is a computer-readable nucleotide/amino acid sequence listing submitted concurrently herewith and identified as follows: One 77,824 byte ASCII (text) file named "91482_196_Sequence_Listing" created on October 29, 2016.

FIELD OF INVENTION

[0003] The present invention is generally related to systems and methods of sequencing biological molecules for determinations regarding the presence and/or absence of infectious agents, and particularly related to systems and methods for the diagnosis of infections and determination of the presence of antimicrobial-resistance markers using nucleic-acid sequencing (e.g., amplicon-based sequencing).

BACKGROUND OF THE INVENTION

[0004] Once highly treatable, common bacterial infections are becoming increasingly more difficult to characterize and manage. Gram-negative infections, including *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa*, *Enterobacter* species and *Escherichia coli* and Gram-positive infections, including *Staphylococcus aureus*, are all considered high-priority target pathogens for which new therapeutics are urgently needed. These are the components of the widely recognized ESKAPE pathogens, which have become a focal point for the development of new antibiotics. The lack of new drugs makes it critical that clinicians receive rapid and informative data on their patients' infections and their likely response to therapy. This clinical and public health crisis is coming at a time when we have greater scientific understanding of the mechanisms of resistance and better access to molecular technology to detect and characterize those mechanisms than

ever before; however, the barriers between genomic research and the clinical lab also remains large and therefore the separation between genotype and phenotype remains largely un-bridged.

[0005] Some of these ESKAPE pathogens (e.g., *K. pneumoniae*, *A. baumannii*, *P. aeruginosa*, *Enterobacter spp.* and ExPEC *E. coli*) are all widely distributed Gram-negative bacteria demonstrating high levels of multi-drug resistance, including resistance to carbapenems and other last-line drugs. While similar in many respects, including all being primary causes of healthcare-associated infections (HAI), individually, these organisms present many unique problems to clinical medicine and public health. *P. aeruginosa* is a major cause of pneumonia in cystic fibrosis patients and outbreaks in neonatal intensive care units, and is frequently shown to colonize water networks in hospitals. Its high rates of intrinsic resistance frequently confound clinical susceptibility testing. *A. baumannii* is responsible for up to 80% of all intensive care unit (ICU) bloodstream infections; its high-rate of transmission in the healthcare setting can be attributed to high desiccation tolerance, enabling it to survive for long periods on fomites. Extra-intestinal Pathogenic *E. coli* (ExPEC) has become a global cause of antimicrobial resistance (AR) bacteremia and urinary tract infections, particularly specific strains, such as the well-studied ST131 strain. *Enterobacter* has been seen to have comparable prevalence of ESBLs and other plasmid-mediated resistance as *Klebsiella* and *E. coli*. In addition, at least some of the aforementioned pathogens have developed resistance to multiple antibiotics typically through several different mechanisms. The mechanisms for resistance are variable, can occur in combinations and in some cases are synergistic.

[0006] The rapid depletion of available antibiotics combined with an even more rapid increase in bacterial mechanisms of multi-drug resistance is a call to action for molecular researchers, medical and public health officials and the biomedical industry to join forces to develop and translate more informative diagnostic tools for the clinician, to better characterize and respond to AR in bacterial infections. While bacterial culturing may remain the gold standard for diagnosis and characterizing resistance, standard culturing processes are time consuming, potentially hazardous, and require specific microbiological skills and facilities.

[0007] In addition, antibiotic susceptibility analysis techniques may vary for different species, as there is no single approach that can be used to rapidly diagnose

and characterize infections with all target agents. Clearly a new approach is required – the detection of established and emerging multi-drug resistant bacteria directly from clinical specimens demands a sensitive and comprehensive diagnostic approach that is capable of rapidly identifying trace amounts of genotypic targets in human samples and using that information to develop phenotypic profiles. The scientific and clinical research community now has more knowledge and cutting edge technology than ever before, allowing them to address this health crisis with such an approach.

[0008] Given these public-health issues, there is a demonstrated need to develop a novel, multiplexed, targeted sequence analysis system for the detection and characterization of AR markers in critical ESKAPE bacteria, typically associated with HAIs, directly from clinical samples. As described herein, the inventors have developed an approach that is capable of identifying all known classes of AR mechanisms, including mutations, SNPs, mobile elements, gene amplification, and expression modulation.

DETAILED DESCRIPTION

[0009] As used herein, the verb “comprise” as is used in this description and in the claims and its conjugations are used in its non-limiting sense to mean that items following the word are included, but items not specifically mentioned are not excluded. Plural encompasses singular and vice versa; e.g., the singular forms “a,” “an,” and “the” include plural referents unless expressly and unequivocally limited to one referent.

[00010] Some embodiments of the invention provide systems and methods for diagnosing and/or identifying infectious organisms and thereafter making one or more genotypic determinations. Moreover, in addition to the diagnostic and genotypic capabilities of the instant invention, some embodiments may also provide methodologies of making genetic / typing determinations of the infectious organisms to provide those of ordinary skill in the art with epidemiological information about the infectious organisms (e.g., strain type). In addition, some embodiments of the invention can be employed with samples that do not require isolation and culturing of the infectious organisms (i.e., embodiments include the capability to directly analyze samples from a patient).

[00011] Some embodiments of the invention may provide methods of (i) diagnosing and/or identifying an infectious organism; (ii) determining the presence of one or more antibiotic resistance targets of the infectious organism; and/or (iii) determining a strain identity of the infectious organism. In some aspects, these three steps of some embodiments of the invention may be performed substantially or completely simultaneously, or, in other aspects, may be performed sequentially (e.g., in any sequence). Moreover, in some embodiments, one or more of these three steps may be omitted. For example, in some embodiments, the first two steps may be performed according to some aspects of the invention and the user may omit the third step in the event that the user does not desire strain information about the infectious organism. In other embodiments, one or more of the first and/or second steps recited above may be omitted.

[00012] With respect to the three steps / methodologies, some or all of these steps may include an analysis of one or more markers at each step. For example, pre-selected markers can be used to make determinations regarding the presence and/or absence of one or more infectious microorganisms (e.g., ESKAPE pathogens). In particular, a series of pre-selected markers may be chosen to provide relevant differential information regarding the identification. For example, markers can be selected that are specific for a plurality of potentially pathogenic microorganisms so that presence or absence of certain markers provides highly specific relevant information. Moreover, the same and/or different pre-selected markers can be used to make predictions about a phenotype of the infectious microorganisms. In some embodiments, as described in greater detail below, the phenotypic predictions can be generally or specifically directed to determining the presence of one or more AR markers. In addition, some of the same and/or different pre-selected markers as recited above can be used to provide users with information regarding a strain identity of the infectious microorganism.

[00013] As described in greater detail herein, some embodiments of the invention may include amplicon-based sequencing of the one or more markers to make the aforementioned determinations. Some embodiments of the invention include systems and methods of preparing samples for one or more downstream processes that can be used for assessing one or more markers for any of the previously mentioned purposes. Some embodiments of the invention may comprise a universal indexing sequencing strategy for use in downstream sequencing platform

processes. By way of example only, some embodiments of the invention comprise a universal indexing sequencing strategy that can be used to amplify multiple genomic regions (e.g., markers, as described below) from a DNA sample simultaneously in a single reaction for the sequencing of one or more amplicons. One or more embodiments of the invention can be used with any desired sequencing platform, such as the ILLUMINA® Next Generation Sequencing (e.g., MiSEQ) platform, Life Technologies' Ion Torrent System, or any other sequencing system now known or developed in the future.

[00014] Some embodiments may be configured to enable relatively simple, rapid (e.g., microorganism-culture independent), inexpensive, and efficient preparation of samples for use on, in, and/or with downstream sequencing platforms. For example, some embodiments may use a sequence coupled to one or more oligonucleotides / primers (as used herein, oligonucleotides and primers are used interchangeably). More specifically, one or more amplicons per sample can be generated using a hybrid oligonucleotide that is designed for amplification of a marker and incorporation of at least one universal tail sequence into the resulting amplicon. As a result, additional steps that may be conventionally required to prepare samples for sequencing can be limited or removed entirely. Further information regarding the universal tail, amplicon-based sequencing strategy can be found in PCT/US2014/064890, which is hereby incorporated by reference in its entirety for all purposes.

[00015] In some embodiments, the methodology may include performing downstream sequencing on one or more amplicons. For example, in order to minimize and/or eliminate the need for cultures of microorganisms or large inputs of nucleic acids, methodologies of the instant invention may include an initial PCR step to create amplicons that correspond to the one or more pre-selected markers. As such, some embodiments require only limited amounts of starting material are necessary and the starting material need not be of high quality (e.g., genomic DNA, crude DNA extracts, single stranded DNA, RNA, cDNA, etc.). In contrast, many conventional sample preparation systems may require relatively large amounts of starting material of relatively high quality, which can limit the use of some conventional systems.

[00016] Some embodiments of the invention can be used for and/or in complement with high-throughput amplicon sequencing of markers, which can be

very useful for a variety of molecular genetic genotyping / predicted-phenotyping applications, including clinical sample analysis. For example, use of the systems and methods of the invention can be employed with sequencing platforms to provide rapid, high-yield sequence data, which can enable the sequencing of multiple markers/amplicons from many samples in a relatively short period of time. Specifically, in some embodiments, amplicons can be selected and PCR reactions can be designed to provide information that can be used to make clinically relevant determinations after sequencing of the amplicons.

[00017] Overall, embodiments of the invention may generally include systems and methods that provide multiplexed marker / sequence amplification, sequencing, and analysis for the identification / detection and characterization of AR markers in critical ESKAPE pathogens and HAI-related pathogens from culture-independent samples. In other words, embodiments of the invention provide diagnostic methodologies with the ability to provide phenotypic and predicted-genotypic information without the need for separately culturing a sample from a subject. In conventional diagnostic procedures, the need for culturing can add time and complexity, which is avoided by the instantly recited invention.

[00018] In some preferred aspects, the methodology may include creating a series of oligonucleotides designed to provide multiplexed amplification of one or more markers to produce the desired amplicons. In particular, the one or more markers and amplicons thereof can be selected / amplified to provide users with clinically relevant information related to identification of one or more potentially infectious microorganisms and phenotypic and genotypic information about the microorganisms (e.g., AR status and strain identity). After production of the amplicons (e.g., via PCR amplification), which may include the universal tail sequences, as detailed above and below, the method may include processing the resulting amplicons for downstream sequencing and thereafter sequencing the processed amplicons. After processing and analysis of the resulting sequencing data, one of skill in the art can make any necessary determinations regarding the identification of one or more microorganisms that may have been contained within the sample and predicted-phenotypic and/or genotypic information revealed.

[00019] Generally, some embodiments of the present invention can be used to detect, identify, assess, sequence, or otherwise evaluate a marker. A marker may be any molecular structure produced by a cell, expressed inside the cell, accessible

on the cell surface, or secreted by the cell. A marker may be any protein, carbohydrate, fatty acid, nucleic acid, catalytic site, or any combination of these such as an enzyme, glycoprotein, cell membrane, virus, a particular cell, or other uni- or multimolecular structure. A marker may be represented by a sequence of a nucleic acid or any other molecules derived from the nucleic acid. Examples of such nucleic acids include miRNA, tRNA, siRNA, mRNA, cDNA, genomic DNA sequences, single-stranded DNA, or complementary sequences thereof. Alternatively, a marker may be represented by a protein sequence. The concept of a marker is not limited to the exact nucleic acid sequence or protein sequence or products thereof, rather it encompasses all molecules that may be detected by a method of assessing the marker. Without being limited by the theory, the detection, identification, assessment, sequencing, or any other evaluation of the marker may encompass an assessment of a change in copy number (e.g., copy number of a gene or other forms of nucleic acid) or in the detection of one or more translocations. Moreover, in some embodiments, the marker may be relevant to a particular phenotype or genotype. By way of example only, in some embodiments, the marker may be related to phenotypes including antibiotic resistance, virulence, or any other phenotype.

[00020] Therefore, examples of molecules encompassed by a marker represented by a particular sequence further include alleles of the gene used as a marker. An allele includes any form of a particular nucleic acid that may be recognized as a form of the particular nucleic acid on account of its location, sequence, or any other characteristic that may identify it as being a form of the particular gene. Alleles include but need not be limited to forms of a gene that include point mutations, silent mutations, deletions, frameshift mutations, single nucleotide polymorphisms (SNPs), inversions, translocations, heterochromatic insertions, and differentially methylated sequences relative to a reference gene, whether alone or in combination. An allele of a gene may or may not produce a functional protein; may produce a protein with altered function, localization, stability, dimerization, or protein-protein interaction; may have overexpression, underexpression or no expression; may have altered temporal or spatial expression specificity; or may have altered copy number (e.g., greater or less numbers of copies of the allele). An allele may also be called a mutation or a mutant. An allele may be compared to another allele that may be termed a wild type form of an allele. In some cases, the wild type allele is more common than the mutant.

[00021] In some aspects, the markers may include one or more sets of amplifiable nucleic acids that can provide diagnostic information about the microorganisms. For example, the markers may include amplifiable nucleic acid sequences that can be used to assess the presence and/or absence of one or more microorganism that may have the potential to cause a diseased state in the subject. In some embodiments, the markers may include amplifiable nucleic acid sequences that can be used to identify one or more of the following exemplary microorganisms: *Klebsiella pneumoniae*, *Serratia marcescens*, *Proteus mirabilis*, *Providencia alcalifaciens*, *Providencia stuartii*, *Enterococcus faecalis*, *Enterococcus faecium*, *Klebsiella oxytoca*, *Pseudomonas aeruginosa*, *Staphylococcus aureus*, *Escherichia coli*, *Acinetobacter baumannii*, *Acinetobacter calcoaceticus* (including the Acb complex), *Streptococcus pneumoniae*, *Streptococcus agalactiae*, *Enterobacter cloacal*, *Enterobacter aerogenes*, *Streptococcus pyogenes*, *Streptococcus dysgalactia*, and *Streptococcus equi*. In some embodiments, the methods may include the use of one or more than one marker per microorganism. Moreover, in some embodiments, one or more of the microorganisms may not be considered pathogenic to certain subjects, but the methodology employed herein can still rely on detection of pathogenic and non-pathogenic microorganisms for differential diagnoses / diagnostics. In some embodiments, the oligonucleotides (with or without the universal tail sequences detailed herein) listed in **Table 1** can be used with embodiments of the invention to amplify one or more markers from the microorganisms to provide diagnostic / identification information to the user.

[00022] Moreover, in some embodiments, one or more the markers associated with the plurality of microorganisms can be amplified in a multiplex manner. For example, in some aspects, nucleic acids can be obtained from the sample and the oligonucleotides used to amplify one or more of the markers used to identify / diagnose can be added to a single mixture to produce a plurality of amplicons in a single reaction mixture. In other aspects, the oligonucleotides can be added to multiple mixtures to provide for the creation of multiple amplicons in multiple mixtures.

[00023] In some aspects, the markers may include one or more sets of amplifiable nucleic acids that can provide characterization information about the microorganisms. For example, the markers may include amplifiable nucleic acid sequences that can be used to assess the presence and/or absence of one or more

antibiotic resistance (AR) mechanisms within the microorganisms. In some embodiments, the markers may include amplifiable nucleic acid sequences that can be used to identify the presence of one or more of the following AR markers: bla_{tem}, bla_{shv}, bla_{rob}, bla_{oxa}, bla_Z, aadB, aacC1, aacC2, aacC3, aac6'-IIa, aacA4, aad(6'), vanA, vanB, vanC, msrA, sarA, aac(6') aph(2''), vat, vga, ermA, ermB, ermC, mecA, int, sul, mecA, aac2ia, aac2ib, aac2ic, aac2id, aac2i, aac3ia, aac3iia, aac3iib, aac3iii, aac3iv, aac3ix, aac3vi, aac3viii, aac3vii, aac3x, aac6i, aac6ia, aac6ib, aac6ic, aac6ie, aac6if, aac6ig, aac6iia, aac6iib, aad9, aad9ib, aadd, acra, acrb, adea, adeb, adec, amra, amrb, ant2ia, ant2ib, ant3ia, ant4iia, ant6ia, aph33ia, aph33ib, aph3ia, aph3ib, aph3ic, aph3iia, aph3iva, aph3va, aph3vb, aph3via, aph3vii, aph4ib, aph6ia, aph6ib, aph6ic, aph6id, arna, baca, bcra, bcrc, bl1_acc, bl1_ampc, bl1_asba, bl1_ceps, bl1_cmy2, bl1_ec, bl1_fox, bl1_mox, bl1_och, bl1_pao, bl1_pse, bl1_sm, bl2a_1, bl2a_exo, bl2a_iii2, bl2a_iii, bl2a_kcc, bl2a_nps, bl2a_okp, bl2a_pc, bl2be_ctxm, bl2be_oxy1, bl2be_per, , bl2be_shv2, bl2b_rob, bl2b_tem1, bl2b_tem2, bl2b_tem, bl2b_tle, bl2b_ula, bl2c_bro, bl2c_pse1, bl2c_pse3, bl2d_lcr1, bl2d_moxa, bl2d_oxa10, bl2d_oxa1, bl2d_oxa2, bl2d_oxa5, bl2d_oxa9, bl2d_r39, bl2e_cbla, bl2e_cepa, bl2e_cfxa, bl2e_fpm, bl2e_y56, bl2f_nmca, bl2f_sme1, bl2_ges, bl2_kpc, bl2_len, bl2_veb, bl3_ccra, bl3_cit, bl3_cpha, bl3_gim, bl3_imp, bl3_l, bl3_shw, bl3_sim, bl3_vim, ble, blt, bmr, cara, cata10, cata11, cata12, cata13, cata14, cata15, cata16, cata1, cata2, cata3, cata4, cata5, cata6, cata7, cata8, cata9, catb1, catb2, catb3, catb4, catb5, ceoa, ceob, cml_e1, cml_e2, cml_e3, cml_e4, cml_e5, cml_e6, cml_e7, cml_e8, dfra10, dfra12, dfra13, dfra14, dfra15, dfra16, dfra17, dfra19, dfra1, dfra20, dfra21, dfra22, dfra23, dfra24, dfra25, dfra25, dfra25, dfra26, dfra5, dfra7, dfrb1, dfrb2, dfrb3, dfrb6, emea, emrd, emre, erea, ereb, erma, ermb, ermc, ermd, erme, ermf, ermg, ermh, ermn, ermo, ermq, ermr, erms, ermt, ermu, ermv, ermw, ermx, emy, fosa, fosb, fosc, fosx, fusb, fush, ksga, lmra, lmrb, lnua, lnub, lsa, maca, macb, mdte, mdtf, mdtg, mdth, mdtk, mdtl, mdtm, mdtn, mdto, mdtp, meca, mecrl, mefa, mepa, mexa, mexb, mexc, mexd, mexe, mexf, mexh, mexi, mexw, mexx, mexy, mfpa, mpha, mphb, mphc, msra, norm, oleb, opcm, oprd, oprj, opm, oprn, otra, otrb, pbp1a, pbp1b, pbp2b, pbp2, pbp2x, pmra, qac, qaca, qacb, qnra, qnrb, qnrs, rosa, rosb, smea, smeb, smec, smed, smee, smef, srm, sta, str, sul1, sul2, sul3, tcma, tcr3, tet30, tet31, tet32, tet33, tet34, tet36, tet37, tet38, tet39, tet40, teta, tetb, tetc, teth, tete, tetg, teth, tetj, tetk, tetl, tetm, teto, tetpa, tetpb, tet, tetq, tets, tett, tetu, tetv, tetw, text, tety, tetz, tlrc, tmrb, tolc, tsnr, vana, vanb, vanc, vand, vane,

vang, vanha, vanhb, vanhd, vanra, vanrb, vanrc, vanrd, vanre, vanrg, vansa, vansb, vansc, vands, vanse, vansg, vant, vante, vantg, vanug, vanwb, vanwg, vanxa, vanxb, vanxd, vanxyc, vanxye, vanxyg, vanya, vanyb, vanyd, vanyg, vanz, vata, vatb, vatc, vatd, vate, vgaa, vgab, vgba, vgbb, vph, ykkc, and ykkd.

[00024] In some specific embodiments of the methodology, the method may include an analysis of the following AR markers: blaKPC (beta lactam resistance), blaNDM (beta lactam resistance), rmtA (aminoglycoside resistance), rmtB (aminoglycoside resistance), rmtC (aminoglycoside resistance), rmtD1 and rmtD2 (aminoglycoside resistance), npmA (aminoglycoside resistance), armA (aminoglycoside resistance), gyrA (fluoroquinolone resistance), parC (fluoroquinolone resistance), qnrA (fluoroquinolone resistance), qnrC (fluoroquinolone resistance), qnrD (fluoroquinolone resistance), qnrS (fluoroquinolone resistance), qnrB (fluoroquinolone resistance), blaCTX-M (beta lactam resistance), dfrA markers (dihydrofolate reductase production affecting resistance to trimethoprim), dfrB markers (dihydrofolate reductase production affecting resistance to trimethoprim), sul markers (dihydropteroate synthase production affecting resistance to sulfonamides) *Staphylococcus* rpoB (vancomycin intermediate resistance), and known *Staphylococcus* AR markers, including linA, mecA, vanA, aacA, blaZ, ermA, ermC, tetK, tetM, and msrA. In some embodiments, the oligonucleotides (with or without the universal tail sequences detailed herein) listed in **Table 2** can be used with embodiments of the invention to amplify one or more markers from the microorganisms to provide AR information to the user.

[00025] Moreover, in some embodiments, one or more the AR markers associated can be amplified in a multiplex manner. For example, in some aspects, nucleic acids can be obtained from the sample and the oligonucleotides used to amplify one or more of the AR markers used to identify / characterize the AR potential of the microorganisms can be added to a single mixture to produce a plurality of amplicons in a single reaction mixture. In other aspects, the oligonucleotides can be added to multiple mixtures to provide for the creation of multiple amplicons in multiple mixtures. In some aspects, amplification of the AR markers and the markers used to identify microorganisms / diagnose an infection can also occur in a multiplex manner such that some or all of the amplicons are generated in a single reaction for a particular sample. In other aspects, amplification

of the AR markers and the markers used to identify microorganisms / diagnose an infection can occur in multiple reaction vessels.

[00026] In some aspects, the markers may include one or more sets of amplifiable nucleic acids that can provide strain identity information about the microorganisms. For example, the markers may include amplifiable nucleic acid sequences that can be used to assess the presence and/or absence of specific strains of microorganisms. In some embodiments, the markers may include amplifiable nucleic acid sequences that can be used to identify the presence of one or more of the following strains of microorganisms: methicillin-resistant *Staphylococcus aureus* (orfX junction), methicillin-resistant *Staphylococcus epidermidis* (orfX junction or cfr), *S. aureus* strain USA300, *S. aureus* strain CC8, *S. aureus* strain USA300/500, *S. aureus* strain Archaic Iberian, *S. aureus* strain True Iberian, *S. aureus* strain ST239, *S. aureus* strain CC5, *Staphylococcus* strains comprising the *tuf* gene, SCCmec typing, including the following *S. aureus* strains, *ccrB2*, *mecI*, IS1272J, *ccrC*, *ccrB1*, *ccrB3*, *ccrB4*, *mecC2*, SCCmecIVa-J1, and *sccmecIVa*. The strain identification markers may also include further *S. aureus* typing markers MRSA sea and MRSA seb and *E. coli* pathovar typing, including the following markers: Shiga toxin genes, *E. coli* aggR, T3SS, *E. coli* virulence markers, *E. coli* and *Shigella* invasion antigen, *E. coli* enterotoxin, *E. coli* UPEC, *E. coli* H30-Rx, *trpA*, *pabB*, *uidA*, *rfbO16*, and *rfbO25b*. In some aspects, the strain identification markers may further include markers for strain identification of *K. pneumoniae* strains, including ST258-strain specific SNP 3576137-Y, *rpoB*, *gapA*, *mdh*, phosphoglucose isomerase, *phoE*, *infB*, and *tonB*. The markers for strain identification may also include the CG258 ID assay, a marker for *Klebsiella* capsule type (*wzi* gene of *cps* locus), markers for *E. coli* strain identification (*adk*, *fumC*, *gyrB*, *icd*, *mdh*, *purA*, and *recA*), markers for *S. aureus* strain identification (*arc*, *aro*, *glp*, *gmk*, *pta*, *tpi*, and *yqi*), and markers for *Streptococcus pyogenes* identification. Moreover, in some embodiments, one or more control nucleotide sequences can be added to any reactions to ensure proper amplification. In some embodiments, the oligonucleotides (with or without the universal tail sequences detailed herein) listed in **Table 3** can be used with embodiments of the invention to amplify one or more markers from the microorganisms to provide AR information to the user. In other embodiments, the oligonucleotides (with or without the universal tail sequences detailed herein) listed in **Table 9** can be used with embodiments of the invention to

amplify one or more markers from the microorganisms to provide lineage identification information to the user.

[00027] Moreover, in some embodiments, one or more the strain identification markers associated can be amplified in a multiplex manner. For example, in some aspects, nucleic acids can be obtained from the sample and the oligonucleotides used to amplify one or more of the strain identification markers used to identify the strain of the microorganisms can be added to a single mixture to produce a plurality of amplicons in a single reaction mixture. In other aspects, the oligonucleotides can be added to multiple mixtures to provide for the creation of multiple amplicons in multiple mixtures. In some aspects, amplification of the strain identification markers, the AR markers, and/or the markers used to identify microorganisms / diagnose an infection can also occur in a multiplex manner such that some or all of the amplicons are generated in a single reaction for a particular sample. In other aspects, amplification of the strain identification markers, the AR markers, and/or the markers used to identify microorganisms / diagnose an infection can occur in multiple reaction vessels. Overall, as described in greater detail below, regardless of the multiplex nature of some embodiments of the invention, after amplification of the markers detailed above, the method may include processing and sequencing the resulting amplicons to provide information related to the identification, characterization, and strain identity of one or more microorganisms that may be present within the sample.

[00028] Some embodiments of the invention may comprise the use of one or more methods of amplifying a nucleic acid-based starting material (i.e., a template, including genomic DNA, crude DNA extract, single-stranded DNA, double-stranded DNA, cDNA, RNA, or any other single-stranded or double-stranded nucleic acids). Nucleic acids may be selectively and specifically amplified from a template nucleic acid contained in a sample. In some nucleic acid amplification methods, the copies are generated exponentially. Examples of nucleic acid amplification methods known in the art include: polymerase chain reaction (PCR), ligase chain reaction (LCR), self-sustained sequence replication (3SR), nucleic acid sequence based amplification (NASBA), strand displacement amplification (SDA), amplification with Q β replicase, whole genome amplification with enzymes such as ϕ 29, whole genome PCR, in vitro transcription with T7 RNA polymerase or any other RNA

polymerase, or any other method by which copies of a desired sequence are generated.

[00029] In addition to genomic DNA, any polynucleotide sequence can be amplified with an appropriate set of primer molecules. In particular, the amplified segments created by the PCR process itself are, themselves, efficient templates for subsequent PCR amplifications.

[00030] PCR generally involves the mixing of a nucleic acid sample, two or more primers or oligonucleotides (primers and oligonucleotides are used interchangeably herein) that are designed to recognize the template DNA, a DNA polymerase, which may be a thermostable DNA polymerase such as Taq or Pfu, and deoxyribose nucleoside triphosphates (dNTP's). In some embodiments, the DNA polymerase used can comprise a high fidelity Taq polymerase such that the error rate of incorrect incorporation of dNTPs is less than one per 1,000 base pairs. Reverse transcription PCR, quantitative reverse transcription PCR, and quantitative real time reverse transcription PCR are other specific examples of PCR. In general, the reaction mixture is subjected to temperature cycles comprising a denaturation stage (typically 80-100° C), an annealing stage with a temperature that is selected based on the melting temperature (T_m) of the primers and the degeneracy of the primers, and an extension stage (for example 40-75° C). In real-time PCR analysis, additional reagents, methods, optical detection systems, and devices known in the art are used that allow a measurement of the magnitude of fluorescence in proportion to concentration of amplified template. In such analyses, incorporation of fluorescent dye into the amplified strands may be detected or measured.

[00031] Either primers or primers along with probes allow a quantification of the amount of specific template DNA present in the initial sample. In addition, RNA may be detected by PCR analysis by first creating a DNA template from RNA through a reverse transcriptase enzyme (i.e., the creation of cDNA). The marker expression may be detected by quantitative PCR analysis facilitating genotyping analysis of the samples.

[00032] "Amplification" is a special case of nucleic acid replication involving template specificity. Amplification may be a template-specific replication or a non-template-specific replication (i.e., replication may be specific template-dependent or not). Template specificity is here distinguished from fidelity of replication (synthesis of the proper polynucleotide sequence) and nucleotide (ribo- or deoxyribo-)

specificity. Template specificity is frequently described in terms of “target” specificity. Target sequences are “targets” in the sense that they are sought to be sorted out from other nucleic acid. Amplification techniques have been designed primarily for this sorting out. The amplification process may result in the production of one or more amplicons.

[00033] The term “template” refers to nucleic acid originating from a sample that is analyzed for the presence of one or more markers. In contrast, “background template” or “control” is used in reference to nucleic acid other than sample template that may or may not be present in a sample. Background template is most often inadvertent. It may be the result of carryover, or it may be due to the presence of nucleic acid contaminants sought to be purified out of the sample. For example, nucleic acids from organisms other than those to be detected may be present as background in a test sample.

[00034] In addition to primers and probes, template specificity is also achieved in some amplification techniques by the choice of enzyme. Amplification enzymes are enzymes that, under the conditions in which they are used, will process only specific sequences of nucleic acid in a heterogeneous mixture of nucleic acid. Other nucleic acid sequences will not be replicated by this amplification enzyme. Similarly, in the case of T7 RNA polymerase, this amplification enzyme has a stringent specificity for its own promoters (Chamberlin et al. (1970) *Nature* (228):227). In the case of T4 DNA ligase, the enzyme will not ligate the two oligonucleotides or polynucleotides, where there is a mismatch between the oligonucleotide or polynucleotide substrate and the template at the ligation junction (Wu and Wallace (1989) *Genomics* (4):560). Finally, Taq and Pfu polymerases, by virtue of their ability to function at high temperature, are found to display high specificity for the sequences bounded and thus defined by the primers; the high temperature results in thermodynamic conditions that favor primer hybridization with the target sequences and not hybridization with non-target sequences (H. A. Erlich (ed.) (1989) *PCR Technology*, Stockton Press).

[00035] The term “amplifiable nucleic acid” refers to nucleic acids that may be amplified by any amplification method. It is contemplated that “amplifiable nucleic acid” will usually comprise “sample template.” The terms “PCR product,” “PCR fragment,” “amplification product,” and “amplicon” refer to the resultant mixture of compounds after two or more cycles of the PCR steps of denaturation, annealing

and extension. These terms encompass the case where there has been amplification of one or more segments of one or more target sequences.

[00036] In some forms of PCR assays, quantification of a target in an unknown sample is often required. Such quantification may be determined in reference to the quantity of a control sample. The control sample starting material / template may be co-amplified in the same tube in a multiplex assay or may be amplified in a separate tube. Generally, the control sample contains template at a known concentration. The control sample template may be a plasmid construct comprising only one copy of the amplification region to be used as quantification reference. To calculate the quantity of a target in an unknown sample, various mathematical models are established. Calculations are based on the comparison of the distinct cycle determined by various methods, e.g., crossing points (CP) and cycle threshold values (Ct) at a constant level of fluorescence; or CP acquisition according to established mathematic algorithm.

[00037] Some embodiments of the invention may comprise a multiplex assay. As used herein, the term "multiplex" refers to the production of more than one amplicon, PCR product, PCR fragment, amplification product, etc. in a single reaction vessel. In other words, multiplex is to be construed as the amplification of more than one marker-specific sequences within a PCR reaction or assay within the same PCR assay mixture (e.g., more than one amplicon is produced within a single vessel that contains all of the reagents necessary to perform a PCR reaction). In some embodiments, a step prior to performing the PCR (or RT-PCR, quantitative RT-PCR, etc.) reaction can occur such that sets of primers and/or primers and probes are designed, produced, and optimized within a given set of reaction conditions to ensure proper amplicon production during the performance of the PCR.

[00038] The algorithm for Ct values in real time-PCR calculates the cycle at which each PCR amplification reaches a significant threshold. The calculated Ct value is proportional to the number of marker copies present in the sample, and the Ct value is a precise quantitative measurement of the copies of the marker found in any sample. In other words, Ct values represent the presence of respective marker that the primer sets are designed to recognize. If the marker is missing in a sample, there should be no amplification in the Real Time-PCR reaction.

[00039] Alternatively, the Cp value may be utilized. A Cp value represents the cycle at which the increase of fluorescence is highest and where the logarithmic

phase of a PCR begins. The LightCycler® 480 Software calculates the second derivatives of entire amplification curves and determines where this value is at its maximum. By using the second-derivative algorithm, data obtained are more reliable and reproducible, even if fluorescence is relatively low.

[00040] The various and non-limiting embodiments of the PCR-based method detecting marker expression level as described herein may comprise one or more probes and/or primers. Generally, the probe or primer contains a sequence complementary to a sequence specific to a region of the nucleic acid of the marker gene. A sequence having less than 60% 70%, 80%, 90%, 95%, 99% or 100% identity to the identified gene sequence may also be used for probe or primer design if it is capable of binding to its complementary sequence of the desired target sequence in marker nucleic acid.

[00041] Some embodiments of the invention may include a method of comparing a marker in a sample relative to one or more control samples. A control may be any sample with a previously determined level of expression. A control may comprise material within the sample or material from sources other than the sample. Alternatively, the expression of a marker in a sample may be compared to a control that has a level of expression predetermined to signal or not signal a cellular or physiological characteristic. This level of expression may be derived from a single source of material including the sample itself or from a set of sources.

[00042] The sample in this method is preferably a biological sample from a subject. The term "sample" or "biological sample" is used in its broadest sense. Depending upon the embodiment of the invention, for example, a sample may comprise a bodily fluid including whole blood, serum, plasma, urine, saliva, cerebral spinal fluid, semen, vaginal fluid, pulmonary fluid, tears, perspiration, mucus and the like; an extract from a cell, chromosome, organelle, or membrane isolated from a cell; a cell; genomic DNA, RNA, or cDNA, in solution or bound to a substrate; a tissue; a tissue print, or any other material isolated in whole or in part from a living subject or organism. Biological samples may also include sections of tissues such as biopsy and autopsy samples, and frozen sections taken for histologic purposes such as blood, plasma, serum, sputum, stool, tears, mucus, hair, skin, and the like. Biological samples also include explants and primary and/or transformed cell cultures derived from patient tissues. In some embodiments, sample may comprise

a portion of a non-animal organism, such as a plant (e.g., castor beans or derivatives thereof).

[00043] In some embodiments, sample or biological sample may include a bodily tissue, fluid, or any other specimen that may be obtained from a living organism that may comprise additional living organisms. By way of example only, in some embodiments, sample or biological sample may include a specimen from a first organism (e.g., a human) that may further comprise an additional organism (e.g., bacteria, including pathogenic or non-pathogenic / commensal bacteria, viruses, parasites, fungi, including pathogenic or non-pathogenic fungi, etc.). In some embodiments of the invention, the additional organism may be separately cultured after isolation of the sample to provide additional starting materials for downstream analyses. In some embodiments, the sample or biological sample may comprise a direct portion of the additional, non-human organism and the host organism (e.g., a biopsy or sputum sample that contains human cells and bacteria).

[00044] With respect to use of the sample or biological sample, embodiments of the claimed methodology provide improvements compared to conventional methodologies. Specifically, conventional methodologies of identifying and characterizing microorganisms include the need for morphological identification and culture growth. As such, conventional methodologies may take an extended period of time to identify the microorganism and may then require further time to identify whether the microorganism possesses and AR markers. Some embodiments of the invention can provide a user with information about any microorganisms present in a sample without the need for additional culturing because of the reliance of nucleic acid amplification and sequencing. In other words, direct extraction of nucleic acids coupled with amplification of the desired markers and downstream sequencing can reduce significantly the time required to obtain diagnostic, AR, and strain identifying information.

[00045] The invention may further comprise the step of sequencing the amplicon. Methods of sequencing include but need not be limited to any form of DNA sequencing including Sanger, next-generation sequencing, pyrosequencing, SOLiD sequencing, massively parallel sequencing, pooled, and barcoded DNA sequencing or any other sequencing method now known or yet to be disclosed.

[00046] In Sanger Sequencing, a single-stranded DNA template, a primer, a DNA polymerase, nucleotides and a label such as a radioactive label conjugated with the nucleotide base or a fluorescent label conjugated to the primer, and one chain terminator base comprising a dideoxynucleotide (ddATP, ddGTP, ddCTP, or ddTTP, are added to each of four reaction (one reaction for each of the chain terminator bases). The sequence may be determined by electrophoresis of the resulting strands. In dye terminator sequencing, each of the chain termination bases is labeled with a fluorescent label of a different wavelength that allows the sequencing to be performed in a single reaction.

[00047] In pyrosequencing, the addition of a base to a single-stranded template to be sequenced by a polymerase results in the release of a pyrophosphate upon nucleotide incorporation. An ATP sulfurylase enzyme converts pyrophosphate into ATP that in turn catalyzes the conversion of luciferin to oxyluciferin which results in the generation of visible light that is then detected by a camera or other sensor capable of capturing visible light.

[00048] In SOLiD sequencing, the molecule to be sequenced is fragmented and used to prepare a population of clonal magnetic beads (in which each bead is conjugated to a plurality of copies of a single fragment) with an adaptor sequence and alternatively a barcode sequence. The beads are bound to a glass surface. Sequencing is then performed through 2-base encoding.

[00049] In massively parallel sequencing, randomly fragmented targeted nucleic acids and/or amplicons are attached to a surface. The fragments / amplicons are extended and bridge amplified to create a flow cell with clusters, each with a plurality of copies of a single fragment sequence. The templates are sequenced by synthesizing the fragments in parallel. Bases are indicated by the release of a fluorescent dye correlating to the addition of the particular base to the fragment.

[00050] Nucleic acid sequences may be identified by the IUAPC letter code which is as follows: A - Adenine base; C- Cytosine base; G - guanine base; T or U - thymine or uracil base. M - A or C; R - A or G; W - A or T; S - C or G; Y - C or T; K - G or T; V - A or C or G; H - A or C or T; D - A or G or T; B - C or G or T; N or X - A or C or G or T. Note that T or U may be used interchangeably depending on whether the nucleic acid is DNA or RNA. A sequence having less than 60%, 70%, 80%, 90%, 95%, 99% or 100% identity to the identifying sequence may still be encompassed by the invention if it is able of binding to its complimentary sequence

and/or facilitating nucleic acid amplification of a desired target sequence. In some embodiments, as previously mentioned, the method may include the use of massively parallel sequencing, as detailed in U.S. Patent Nos. 8,431,348 and 7,754,429, which are hereby incorporated by reference in their entirety.

[00051] Some embodiments of the invention comprise multiple steps and/or processes that are carried out to execute the universal tail indexing strategy to prepare amplicons corresponding to desired markers for sequencing. In some embodiments, one or more markers for a given sample or template can be selected, as described above. Some embodiments of the invention can be used in conjunction with an analysis of one or more markers (e.g., genes / alleles) associated with a particular phenotype (e.g., resistance to one or more pharmaceuticals, such as antibiotics). By way of example only, some embodiments of the invention can be used to detect and/or quantify the development of antibiotic resistance in populations of patients infected with an organism. As such, in some aspects, prior to performing additional steps, an investigator can assess the markers present within the genome of the organism to determine which markers are implicated in the development of antibiotic resistance. For example, markers can be selected that may contain a SNP or other change or alteration that can confer at least partial antibiotic resistance. In other aspects of the invention, markers can be selected that are not implicated in antibiotic resistance, but are associated with other phenotypes / genotypes that are desirable for further analysis.

[00052] After selection of the markers, marker-specific primers / oligonucleotides can be designed for the amplification of the markers to produce the desired amplicons, as detailed above. As is known in the art, a forward and a reverse marker-specific primer can be designed to amplify the marker from a nucleic acid sample. In some embodiments, the forward and reverse primers can be designed to produce an amplicon (e.g., some or all of the sequence of the marker) of a desired length. For example, the length of the amplicon may comprise approximately 50 base pairs (bp), 100 bp, 150 bp, 200 bp, 250 bp, 300 bp, 350 bp, 400 bp, 450 bp, 500 bp, 1,000 bp, or any size amplicon greater in size or therebetween.

[00053] As previously mentioned, some embodiments of the invention may include a multiplex PCR reaction. For example, marker-specific primers can be designed for multiple markers or multiple regions of the same marker such that

multiple amplicons of between about 50 bp and 1,000 bp are being produced within a single PCR reaction vessel. In other words, the forward and reverse primers can be designed to function within a given set of temperature parameters such that more than one amplicon can be successfully amplified from a given template within a single PCR reaction mixture. As such, multiple amplicons can be prepared using the universal tail indexing strategy for sequencing preparation.

[00054] In some embodiments, the forward and reverse primers that have been designed for each of the markers can be modified to include a universal tail. For example, the universal tail sequences can be relatively or completely unique sequences of nucleotides that are coupled to the 5' ends of some or all of the forward and reverse marker-specific primers. In some aspects, the universal tail sequences can be selected such that there is little to no overlap in sequence between portions of the markers that are being amplified and the universal tail sequences. Moreover, the universal tail sequences can comprise a length between ten and twenty nucleotides in length. In some embodiments, the universal tail sequences can be any other length, as desired by the user to meet the needs and requirements of the reaction. As such, the universal tail sequences can exhibit a relatively negligible impact on binding of the forward and reverse marker-specific primers to the template sequence to enable amplification. Moreover, as a result of being included on the 5' end of the forward and reverse marker-specific primers, the universal tail sequences will form a portion of the resulting amplicons. In addition, in some aspects of the invention, the sequences selected for the universal tail sequences can be at least partially correlated with the chemical composition of the template nucleic acids. For example, in some aspects, the sequences selected for the universal tail sequences can be at least partially correlated with the G-C content of the organism from which the template is isolated.

[00055] In some aspects, some or all of the universal tail sequences can be at least partially unique. In some embodiments, each of the 5' ends of all of the forward marker-specific primers within a given PCR assay mixture can comprise the same or a similar universal tail sequence (e.g., a first universal tail sequence or UT1). Similarly, each of the 5' ends of all of the reverse marker-specific primers within the same PCR assay mixture can comprise a second universal tail sequence (UT2) that differs from the first universal tail sequence. As such, each respective sample from which a template sequence is used in the multiplex PCR assay will

have two unique universal tail sequences. Accordingly, each forward and reverse marker-specific primer within a multiplex PCR mixture will include a unique universal tail sequence. For example, if the PCR includes 35 different samples, 35 universal tail sequences can be employed for the forward primers in each of the 35 unique reactions (i.e., not including technical replicates) and 35 universal tail sequences can be employed for the reverse primers in each of the 35 unique reactions (i.e., not including technical replicates). Overall, the forward and reverse marker-specific primers that each comprise the universal tail sequences can comprise a generally short length (e.g., 25-50 bp), which can facilitate simultaneous amplification of multiple targets in a single reaction.

[00056] In addition, some embodiments of the invention may comprise performing quantitative PCR to optimize the multiplex PCR assay. For example, after design of the forward and reverse marker-specific primers that each include a universal tail sequence, the contemplated multiplex PCR assays can be performed using quantitative PCR (e.g., using DNA as a template) to assess relative quantities of the amplicons produced. Accordingly, the sequence coverage of each amplicon is considered to be equal if the quantities of the amplicons produced by the multiplex quantitative PCR appear to be equal. If the quantities of the amplicons produced by the multiplex quantitative PCR do not appear to be equal, the forward and/or reverse marker-specific primers can be altered and re-optimized until adequate quantities of amplicons are produced.

[00057] After design and adequate optimization of the multiplex PCR assay comprising multiple forward and reverse marker-specific primers that each include universal tail sequences, the multiplex PCR can be performed to obtain the amplicons associated with the above-described markers. In some embodiments, template that has been previously isolated from a sample can be used for the amplification of the amplicons. In some aspects, multiple PCR reaction replicates can be performed for each sample template and one or more control templates.

[00058] In some embodiments, after successful production of the amplicons during the multiplex PCR assay, the resulting amplicons can be further processed to provide sequencing-ready amplicons. For example, some embodiments of the invention may comprise an indexing extension step. In some aspects, the indexing extension step may comprise extending the optimized multiplex amplicons using a set of indexing and common primers that recognize the respective universal tail

sequences used for the particular group of amplicons in a minimal cycle PCR assay (e.g., 5-10 total cycles). In particular, each multiplex set of amplicons to be sequenced can be extended with a different set of index oligonucleotides and common oligonucleotides that recognize UT1 and UT2, respectively. In some aspects, the index sequence of the index oligonucleotides can be custom designed to allow for the selection of an index sequence from potentially thousands of different index sequences.

[00059] After this step, the resulting products include a set of amplicons for each sample / template that comprise the same index and any necessary sequences that may be required for a particular sequencing platform (e.g., platform sequences associated with the ILLUMINA® Next Generation sequencing platform). Thereafter, the resulting extension-reaction products can be quantified, pooled, and sequenced using a desired platform. In some aspects, the inclusion of the universal tail sequences on the index and common primers can coincide with the use of genomic and index read primers in the mixture of sequencing primer reagents. For example, some embodiments of the invention are capable of pooling multiple amplicons with multiple indices in a single sequencing run to provide 40,000x-95,000x coverage across the amplicons. In other embodiments, the systems and methods associated with the invention can be configured to provide any level of sequencing coverage that is desirable to the user (e.g., higher or lower than the coverage levels discussed above). In some embodiments, after sequencing and generation of the sequence data, the resulting data can be demultiplexed and the sequence files can be aligned to the appropriate reference sequences for subsequent sequence analyses.

[00060] Embodiments of the invention offer additional advantages relative to conventional systems. For example, some embodiments of the invention comprise the use of PCR before sequencing such that only limited amounts of starting material are necessary and the starting material need not be of high quality (e.g., genomic DNA, crude DNA extracts, single stranded DNA, RNA, cDNA, etc.). In contrast, many conventional sample preparation systems may require relatively large amounts of starting material of relatively high quality, which can limit the use of these systems. Moreover, the inclusion of non-desirable template materials can also interfere in one or more downstream processes in conventional systems and methods. For example, if an investigation is being conducted that focuses on one or more organisms that may be associated with another organism (e.g., bacteria associated

with a human); the sampling of the target organism may result in template contamination from the host organism.

[00061] In particular, in some aspects, obtaining samples of pathogenic or commensal bacteria from, on, or within a human may also result in the collection of human tissue. As such, when isolating the template, human nucleic acids may contaminate the bacterial template. Some embodiments of the invention are configured such that the contaminating template (e.g., from a human) would not interfere with downstream processes, including sequencing. For example, some embodiments of the invention operate such that only a limited amount of starting template (e.g., 500 femtograms or greater) can be used. Moreover, some embodiments are also configured such that the starting material (e.g., template contaminated with foreign nucleic acids) can still produce the required amplicons for sequencing in the presence of more than a 1,000-fold excess of contaminating template with no discernible inhibition of the multiplex PCR.

[00062] In certain aspects, the present invention provides an assay that works with as little as about 1pg, about 900fg, about 800fg, about 700fg, about 600fg, about 500fg, about 400fg, about 300fg, about 200fg, or about 100fg of genomic DNA.

EXAMPLES

[00063] The inventors employed embodiments of the methods detailed herein to assess the presence of microorganisms in control samples from respiratory, wound, nasal swab, and urine specimens using oligonucleotides / primers detailed in **Tables 1** and **2** to amplify desired markers. Thereafter, the resulting amplicons were sequenced using next-generation sequencing technology to obtain information on the identity and AR characteristics of the microorganisms in the samples.

[00064] As illustrated in **Table 4**, the inventive methodology produced extremely reliable results, compared to standard / conventional culture methods. Moreover, as illustrated by the samples that include the *** indication, the claimed methodology was also able to reveal the presence of an AR marker, which was not detected using the traditional culture methodology. Moreover, as illustrated in **Table 5**, by detecting organisms that were not previously detected using traditional culture-

based methods, the method further shows increased sensitivity compared to the culture-dependent methodology.

[00065] As illustrated in **Tables 6, 7, and 8**, the inventive methodology was used to identify microorganisms from different sample types. For example, in **Table 6**, the inventive methodology was used to identify the listed species and the listed antibiotic resistances in 27 respiratory specimen samples from subjects with cystic fibrosis. Similarly, in **Table 7**, the inventive methodology was used to identify the listed species and the listed antibiotic resistances in 24 respiratory specimen samples from subjects with chronic rhinosinusitis. Finally, in **Table 8**, the inventive methodology was used to identify the listed species and the listed antibiotic resistances in 38 respiratory specimen samples from subjects with various respiratory diseases.

[00066] It should be understood from the foregoing that, while particular embodiments have been illustrated and described, various modifications can be made thereto without departing from the spirit and scope of the invention as will be apparent to those skilled in the art. Such changes and modifications are within the scope and teachings of this invention as defined in the claims appended hereto.

Table 1. Oligonucleotides / Primers for Use in Identifying Bacterial Species

Target purpose	Target ID	Assay	Sequence	SEQ ID NO:
Species ID	Klebsiella pneumoniae	Kp_M1_UT1v (Cluster 1245)	CGTTTCACAACTGCGGATG	1
			TGGCGGTCATGTTCTTACTC	2
		Kp_M1_UT2v (Cluster 1245)	TTCACACCGCCATCGTC	3
			CAGCATCTCCACCGACAG	4
		Kp_M1_UT1nv (Cluster 4240)	CGTTCCTCACCGTAGTGG	5
			TCCAGCGTGACATAATCGG	6
		Kp_M1_UT2nv (Cluster 4240)	TTTCCACCTACGCCGATAAAG	7
			CCAGCGGAATAACCCAGGAC	8
		Kp_M2_UT1v (Cluster 1825)	GGCGCGATGAGTTTGT	9
			ATGTTGCGGATAGCCTGATAG	10
		Kp_M2_UT2v (Cluster 1825)	CGTCCACTATCAGGCTATC	11
			TTCGGCGATGGGAATAAACA	12
		Kp_M2_UT1nv (Cluster 2858)	TGCCTATCGCCACTTTATTGA	13
			CGGTCGTTAATCGCCTTCT	14
		Kp_M2_UT2nv (Cluster 2858)	CCTCAGGTACGCTCATTCG	15
			TCGTCACATCCCCTCCTCT	16
Species ID	Serratia marcescens	Smar_UT1	GATCAGCAGCAACCGTGAC	17
			GGCGATGTAATCCTGCGAGA	18
Species ID	Proteus mirabilis	Pmbs_UT1	GGGATATCCGTGGAATGCGT	19
			TAATGTGATCACCGCTCCCG	20
Species ID	Providencia alcalifaciens	Pal_UT1	ACGAGCCAGCCTGATGAAAA	21
			CATCGCAACTGCTGCATCTC	22
Species ID	Providencia stuartii	Pst_UT1	GCCTTGCGCCTTAGTTTGT	23
			GGCTATCCATTTTCCAGCCG	24
Species ID	Enterococcus faecalis	Efs_UT1	TTCTTTTCCAGGAGCAACGC	25

Target purpose	Target ID	Assay	Sequence	SEQ ID NO:
			AGCAAGACAGAAATAAGTAAAAAGA	26
		Efs_UT2	TCTTCGTCATGGTSGTTTC	27
			AACGGAACATGGTGAGCAAC	28
Species ID	Enterococcus faecium	Efm_UT1	GKTTTGATGGCTGGGTCAGT	29
			TGTTTGGCTTTTTCAGTTG	30
			TACTTRGCTTTTTCAGTTG	31
		Efm_UT2	TCAGTGAAAACRACCCACGA	32
			ATYGCTGCTGTCCGAGTTCT	33
Species ID	Klebsiella oxytoca	Koxy_UT2	CCGTCGCCGTATTACTGAT	34
			TCTCTACAACACGCTACCCCTA	35
Species ID	Klebsiella oxytoca	Koxy_UT3	GGTGAGAACGATGTGATTGTG	36
			TGACCCAAAGGCGATTCTG	37
Species ID	Pseudomonas aeruginosa	PSAR_28621	GAAGCGGTTGAGGAACAG	38
			ACCCACGACATGATGTAC	39
		PSAR_27569	GGACGATCCATTCCATACCG	40
			GACGAAGCGGTGATCCAC	41
Species ID	Staphylococcus aureus	Sa_M1_UT1	ATAGGTGTGGCTTTTGTAGGG	42
			CTGCTGCTAATAACGCTTGC	43
		Sa_M2_UT1	ATGGTTATGAAATTGTTGGTATGAA	44
			CCGTTGCTTGATATTCCTTCGTAT	45
		Sa_M4_UT1	TACGAAATCAGAAAGTGGCTCAA	46
			ACGTTAAATCCTGCATTTTCCAA	47
		Sa_M4_UT2	TAGCGTTGGTATTAAAGTGGTTGT	48
			GTCATAGCATAGTTCGGGTCA	49
Genus and species ID	Staphylococcus	tuf_UT1	ACWGGCCGCTGTTGAACGTG	50

Target purpose	Target ID	Assay	Sequence	SEQ ID NO:
		tuf_UT2	CACCAGCTTCAGCGTAGTCTAAT	51
			CTCAACWGGCATCATGAATGGTTT	52
			AGGTGACGATGTRCCTGTAAATC	53
Species ID	Escherichia coli	Ec_M1_UT	CGTTGTTTGGACCCYGTATA	54
			AATGTAAATTCGCTGRGATCGT	55
		Ec_M2_UT	CGCCTGGCTCTGTTTCA	56
			AGCAAGTGCATTACGAGTCT	57
Species ID	Acinetobacter baumannii	AB_marker1	GTCGTGTTGAATGATGATGATGTCG	58
			ATCCCATTTTACAAATGGCAAAACA	59
			TGCAGAACTAGCCCATTTTACA	60
			TGCWGAACATAKCCCATTTTACA	61
Species ID	Acinetobacter baumannii	Ab_M2_UT1	TTTAGCGRACAGATTGTGTTTACA	62
			CGGTTGCCCTTGCTGAGTT	63
		Ab_M2_UT2	TCCGYAAACCTCAACTCATTT	64
			GCCAGTGAACCATCTAATACAGT	65
		Ab_M3_UT1	TAGAGCAGCAACAATAACAAGA	66
			TTATTATGAGTCGCCCTTGGCAATT	67
		Ab_M3_UT2	TAGAGCAGCAACAATAACAAGA	68
			TCGCATGGCTTCTGACTCA	69
	Acinetobacter calcoaceticus-baumannii complex	Acb_M2_UT1	AAATTGCYAAAGCACAAAGGTTT	70
			CGGGAATTGCYTCATGGTAAG	71
		Acb_M3_UT1	GATAGCGATTACTCGGTTCCATC	72
			CCTGTGCCTAAAAACCAGTGAAA	73
Species ID	Streptococcus pneumoniae	Sp_M1_UT2	TCCTGATATAATCGGTGTCACAAG	74
			AGAACAGAGGATATGTCATCTAACTT	75

Target purpose	Target ID	Assay	Sequence	SEQ ID NO:
		Sp_M2_UT1	CTGATGGATATGGAGACATTAAGGAT	76
			GTCCAATTATATGCGGATCTATACCT	77
		Sp_M2_UT2	CTGATGGATATGGAGACATTAAGGAT	78
			AATTAGTTGAGTAGCCATGACAGT	79
		Sp_M3_UT1	TACGGTGTAAAGTGCTGATAAAGG	80
			GCTTCAAAGTTAAGTCCATCCAAAT	81
Species ID	Streptococcus agalactiae	GBS_UT1	GCCAACCGTGGTATGTCA	82
			TTCTGCCGGGTTTCITTTG	83
		GBS_UT2	CTCGTGCTGCTACTTCTC	84
			CGTGGATGTGCTYGTAAACAAC	85
Species ID	Enterobacter cloacae group	Encl_M1_UT	AACTRTCTTGTGCTTGC	86
			ACATAGTCGTTGGCATAACAGA	87
			ATTGTTGATGACTGTTCCCTCTC	88
			CGACGAYGATGACGAATG	89
			CGCAATTAAGTCGGTCTCTG	90
Species ID	Enterobacter cloacae group	Encl_M2_UT	CCTTCAACGTGRATGTTTCA	91
			AATWCCYACTTCGCCCTTCA	92
			TGGCGTTCTGGAYGTCAA	93
			AGTTGTTTATACCGCCAAAGC	94
			AGTTGTCTACACTGCCAAAGC	95
Species ID	Enterobacter aerogenes	Eaer_M1_UT	ACCCAAAAGTATTCTGATTGTTGAG	96
			ATGAGGTCCTTATTATCGGTCAACTG	97
Species ID	Enterobacter aerogenes	Eaer_M2_UT	AATCGGCAATACCAGCAGTT	98
			CCTTCACGGAATAACATATCCAGAT	99
Species ID	Streptococcus pyogenes	Spyo_UT1	GCTTATTGCTAATGACTGTGCTTAT	100
			TGCCGACATTGACAACCATA	101

Target purpose	Target ID	Assay	Sequence	SEQ ID NO:
Species ID	Streptococcus dysgalactiae	Sdys_UT1	CATCTGCTTGACCCTTTATGAAAC	102
			GTGAGGCTAACACATCAAACTAG	103
Species ID	Streptococcus dysgalactiae	Sdys_UT2	ATGCGGGAGGCTCTGTTTCTAC	104
			CCAGCCGGTAAAGAAACTC	105
Species ID	Streptococcus equi	Sequi_UT1	GTCAAGCACCTTAYCACCATCTAG	106
			GGCATGCTTGTGAACAACA	107
Species ID	Streptococcus equi	Sequi_UT2	GTGCCRATTGAGGATCTGGTTTC	108
			CYGTGGACCACCTTCTGTCAT	109

Table 2. Oligonucleotides / Primers for Use in Identifying Antibiotic Resistance in Bacteria

Target purpose	Target ID	Assay	Sequence	SEQ ID NO:
B-lactam resistance	blaKPC	KPC_UT1	CGTCTAGTTCTGCTGCTTGT	110
			ACCGTCATGCCGTGTTGTC	111
		KPC_UT2	TTGTTGATTGGCTAAAGGGAAC	112
			CAGACGACGGCATAGTCATT	113
B-lactam resistance	blaNDM	NDM_UT1	GGACAAGATGGCGGTATG	114
			CGGCGTAGTCTCAGTG	115
		NDM_UT2	CAACTGGATCAAGCAGGAGAT	116
			CGACAACGCATTGGCATAAG	117
Aminoglycoside resistance	rmtA	rmtA_UT1	GAATTGGACTGCCTCTACGATT	118
			GCACGCCCATACAGATGT	119
Aminoglycoside resistance	rmtB	rmtB_UT1	AAGGCATGGAGGCGAAC	120
			AAGTATATAAGTTCTGTTCCGATGGT	121
Aminoglycoside resistance	rmtC	rmtC_UT1	AATACTCCACACTTTATCCACCAA	122
			TTCTTGCGAACCTCCTTCTC	123
Aminoglycoside resistance	rmtD1 and rmtD2	rmtD1_ & D2	AACGATGCGACGATCCATT	124
			GCGATTTGCTGTGCGAAA	125
Aminoglycoside resistance	npmA	npmA	CCGCTTGCTGGTACATATCTA	126
			CCTATCTCGTCCGCTATCTG	127

Target purpose	Target ID	Assay	Sequence	SEQ ID NO:
Aminoglycoside resistance	armA	armA	ACTATTCTGCCTATCCTAATTGGG	128
			TCATTTAATGTTGCGACTCTTTCA	129
fluoroquinolone resistance	gyrA	gyrA_ACBA_UT1	ATGACTATAACAAAAGCYTACAAAGAAATC	130
			ACAATGGTTTCATAAACAGCTAAGTC	131
			AACAAGCCCTACAAGAAATCC	132
			CGCACTTCGGGTATCG	133
			AACAAGCCCTACAAGAAATCC	134
			CTCGTGCCATCGTAGT	135
			CATTGCCAGATGTTTCGTGAC	136
			GCCATCAACAAGCGGATAAC	137
			AAATCTGCCCGTGTCGTT	138
			ACCATCCACCAGCATGTAAC	139
			CTCCTATCTGGATTATGCGATGT	140
			CGTCAATAGAACCGAAGTTACC	141
			CAATGACTGGAAACAAAGCCT	142
			CGATGGAACCAAGTTACCC	143
fluoroquinolone resistance	parC	parC_PSAR_UT1	TTGTAGTGTTCAGGTCGTC	144
			CCGTCTGCTATTGTTCAAGG	145
			GAAATTCAAAAAGTCCGCC	146
			GGATAGCGGTAAGAGAACGG	147
			CAGCGCGAAATTCAAAAAGT	148
			GCGAAAGATTGGGATCGTC	149
			GTTCTCTTACCGTTATCCGC	150
			TATTCGACAACCGGATTTC	151
			TGTGTATGCGATGTCTGAAC	152

Target purpose	Target ID	Assay	Sequence	SEQ ID NO:
			ATATTTGACAAACCGGATTC	153
		parC_ACBA_UT1	GAGCGAGCTAGGCTTAAAAA	154
			TCCCCTGACCTTCGATTAA	155
		parC_ACBA_UT2	GTACGTCATTATGACCGTG	156
			CTATAAGCCGAGAGTTTGGC	157
		parC_staph_UT1	TCCGTAAAAGTCCGAAACA	158
			AGCTTCAGTGTAACGCATTG	159
		parC_staph_UT2	CCAGATGTTTCGTGATGGTT	160
			CTTAGCTTCAGTGTAACGCA	161
Fluoroquinolone resistance	QnrA	qnrA_UT1	TAGAGTTCAGGGAGTGCGAT	162
			CCGAGCAGAAGTACATCTTATGG	163
		qnrA_UT2	GATTTGAGYGACAGCCGTTT	164
			GCAGAAGTACATCTTATGGCTGA	165
Fluoroquinolone resistance	QnrC	qnrC_UT1	GCTAATTTCTCACAGGCAAACTTT	166
			ACAACCCGTAATGTAAGCAGAG	167
	QnrD	qnrD_UT1	GTTTGATTGGTCTTTGGCTGATT	168
			CCATCCAACTTCACTCCATCT	169
		qnrD_UT2	CGACAGGAATAGCTTGGAAAG	170
			CCAGTTATCACAGTGCCATTC	171
	QnrS	qnrS_UT1	CAATTTATGCCACGCCGAAC	172
			TAATTTATGTCACGCCGAAC	173
			GATAAACCAACAATACCCAGTGCTT	174
		qnrS_UT2	GTGCTAACTTGCGTGATACGA	175
			TCCATATTGGCATAGGAAAGATTACA	176
			TCCATATTGGCATAAGACAGGTTACA	177

Target purpose	Target ID	Assay	Sequence	SEQ ID NO:
	QnrB	qnrB-C1_UT1	GACGTTCAAGTGGTTCRGATCT	178
			KGCTCGCCAGTCGAAAGT	179
		qnrB-C1_UT2	CGACGTTCAAGTGGTTCRGATCT	180
			GCKGCTCGCCAGTCGAAA	181
		qnrB-C2_UT1	ACCAATCTAAGCTACGCCAACTT	182
			CCTGAGTTCCCATCCAGCG	183
		qnrB-C3_UT1	TACGCACTGTGATTTGACCAAT	184
			GGAKCAACGATGCCCTGGTAG	185
		qnrB-C3_UT2	CGCATATATCACCATAACCAACTT	186
			GTCCAGGAKCAACGATGCC	187
B-lactamase production	blaCTX-M genes	CTX-M-G1_64_UT1	CCGTCACGCTGTTGTTAGG	188
			CGCTCATCAGCACGATAAAGT	189
		CTX-M-G1_64_UT2	CGCTGATTCTGGTCATTACTTC	190
			ACGGCTTTCTGCCCTTAGGT	191
		CTX-M-G1_UT1	CGATGTGCAGCACCCAGTAA	192
			TCGGTTCGCTTTCACTTTTCT	193
		CTX-M-G1_UT2	GACGATGTCACTGGCTGAG	194
			CCACAACCCAGGAAGCAG	195
		CTX-M-G2_74-75_UT1	TGGCGCAGACCCTGAAAA	196
			ATATCGTTGGTGGTGCCATAA	197
		CTX-M-G2_74-75_UT2	ATGGCGCAGACCCTGAAA	198
			CCGCTGCCCGGTTTATCG	199
		CTX-M-G8_G25_UT1	GAGCCGACGCTCAACACC	200
			CCCGACAACCCACGATGT	201

Target purpose	Target ID	Assay	Sequence	SEQ ID NO:
		CTX-M-8_G25_UT2	GCTCAACACCGCGATCCC	202
			CCCGACAACCCACGATGT	203
		CTX-M-9_UT1	TTCGTCGGATCGCACTGA	204
			GATGATTCTCGCCGCTGAAG	205
		CTX-M-9_UT2	CGCTGGTTCTGGTGACCTA	206
			GATGATTCTCGCCGCTGAAG	207
Dihydrofolate reductase production (DNA synthesis) uninhibited by trimethoprim	dfrA genes	dfrA1_UT1	AATGGCTGTGGTTGGACG	208
			CATACTTTCGGTTGGTAATGCT	
		dfrA15_UT1	AATATGCCGTTGTAACGTTCA	
			ACACAATCACATGATCCGTTATCG	
		dfrA16_UT1	AGTATGCAGTTGTAACTCGCTCTA	
			CACCACCACAGAAACGATAAC	
		dfrA14-30_UT2	GTGATTGGTTGCGGTCCA	
			CCCGCCACCAGACACTAT	
		dfrA6-31_UT1	YGAGAAATGGAGTAATTGGCTCT	
			WATTTCAACCAACCACAGAAACAAA	
		dfrA26-13_UT1	GGWGCCAATCGGGTTAT	
			CTCAGTGAGTCTGCGAAA	
			CTCGGTGAGCCTGCGAAA	
		dfrA8_UT1	AAAGACTACGAGCAGAAATGGC	
			ACGGTAAGTGAAGTAAGTGAAG	
		dfrA3b_UT1	AACGCTGCCATTGTTACCA	
			AAGCCTTGAAGTGTCTCGAG	

Target purpose	Target ID	Assay	Sequence	SEQ ID NO:
		dfrA9_UT1	AAGACAGGAGGTATCGGATTGA	225
			CGTAGGCAGCTAAGTTCTCGTA	226
		dfrA24_UT1	AAGACCGCATCAATATCGTCATC	227
			CATAGCAAGCCGTCCAAGAA	228
		dfrA27-28_UT1	AAGACTCTTACGAACCATGTTGTT	229
			CCTCTGGCTCGGAATCTATTG	230
		dfrA25_UT1	AAGCACTGACCTATAACCAATGG	231
			CCCAGGAATGTTCCGGAAGAAA	232
		dfrA10_UT1	AAGCATTGAGAGACACAAACCAA	233
			AACCAACACCAATGACAT	234
		dfrA32_UT1	AAGGTGAGCAGCTAATCTTTAAGG	235
			TGACCCCTGAAATTCATTCTTTGA	236
		dfrA20_UT1	AAGTCGCACAACATCTTGAAGG	237
			AGATTGAGCACCACCAATAATGA	238
		dfrA23_UT1	AATCAATATCACGACAGCGATCAA	239
			CTTCACGGGATGGGCTCA	240
		dfrA7_UT1	AATCAGTGGCTCCTTGTGG	241
			GGAAGAACACCCATAGAGTCAAAT	242
		dfrA29_UT1	AATCAGTGGCTTCTTGTCGG	243
			GTGGATGATAGATAAGTGGATGGT	244
		dfrA17_UT1	AATGGCGTAATCGGTAGTGG	245
			GCTTGAAATCCGTTCTTTGACA	246
		dfrA18_UT1	ACGCATTGCTGTCATTGGT	247
			CTCGCTGGCACTGGAATC	248
		dfrA3_UT1	ACTCTATGCCGAGGCTCTG	249
			CGCTGACGACTCAAGGTAAC	250

Target purpose	Target ID	Assay	Sequence	SEQ ID NO:
Dihydrofolate reductase production (DNA synthesis) uninhibited by trimethoprim	dfrB genes	dfrB1-8_UT1		251
			WATGGGAGATCGCGTGGC	
			GCWGTACCAACCGACAATCT	
			GCAGGGTCAAGTYGTCGG	
			TCGGACTCGACSGCATAG	
			ACCCGACAACCTTGACCCCT	
			ACCAACACACAATGGAGTCA	
			AATCTACCCAGGCTCAGT	
dihydropteroate synthase (DNA synthesis) uninhibited by sulfonamides	sul genes	sul1_UT1	CCGTTCAAGCGCAGTCAT	252
				253
				254
				255
				256
				257
				258
				259
	sul genes	sul1_UT1	GCTGGTGGTTATGCACTCAG	260
			CGCCCAAGAAGGATTTCGG	261
			CGTGCTGTCGAACCTTCAA	262
			GCTGGACCCAGATCCTTTACA	263
			CATCATTTTCGGCATCGTCAAC	264
			GCGACAAGGCATAGGCTT	265
			GTCACATAACCTCGGACAGTT	266
			CTCGGCCATCAGCTTACG	267
			AAAGCCTTAATGACAGGTTTGAGT	268
			GAAGATGGAGCAGATGTGATTGAT	269
			GGCAAAGTCAGATTGCCAACTTG	

Target purpose	Target ID	Assay	Sequence	SEQ ID NO:
Vancomycin-intermediate S. aureus (VISA)	Staphylococcus rpoB (VISA)	VISA_rpoB-481_UT1	CCTCTCCGGATTCTGTTCAAC	270
			CCAGGTCCTAATGCTGATAGACG	271
			ACCGTCGTTTACGTTCTGTAGGTG	272
			CCAGGTCCTAATGCTGATAGACG	273
			GGACCAAGCAAYCCATTAGCT	274
Staphylococcus antibiotic resistance	linA	linA_UT1		275
			GTGAAGGCATCCAATSAACCTT	
			AAACMACAAAGAGAACACAGAGAT	276
			GGAACGATGCCATCTCATATGCT	277
			ATAGCGTCATTATCCAGGAATGCA	278
			CGGCTCGACTTCCTGATGA	279
			TGTGCGGTATTGGGAAACAG	280
			GCCACACTATCATATAACCACCTACCGA	281
			TCCAAGAGCAATAAGGGCATAACCAA	282
			ACACTCTTGCGCGTTTCACT	283
			CCTAAGGGCCAATCTGAACCTATT	284
			CAACCAATTGATTTCAAAGAAGGACTAC	285
			TCAAAGCCTGTCGGAATTGGT	286
			ATTTAATCGTGGAAATACGAGTTTGCTAA	287
			CGTCAATTCTCGCATGTTTTAAGG	288
			AGTTTGAGCTGTCTTGTTTCATTG	289
			TGCAGCAGATCCTACTCCTTGCTAC	290
			CTTTCTGGGCTTCCATTGGTTTATC	291
			CGAGCTCTCATACTGCATTCCA	292
			CTTCTTCCAAATGTTCCTCATCTTTT	293

Target purpose	Target ID	Assay	Sequence	SEQ ID NO:
			ACCAGATCGTTTAAGTGCAATCAAA	294
	cfr	cfr_UT1	AACGAAGGGCAGGTAGAAGC	295
			TGACCAACAAGCAGCGTCAAT	296
		cfr_UT2	GTGAGGAACGCAGCAAAATTGA	297
			GCTTCTACCTGCCCTTCGTT	298

Table 3. Oligonucleotides / Primers for Use in Genetic Characterization of Bacteria

Target purpose	Target ID	Assay	Sequence	SEQ ID NO:
MRSA ID – orfX junction	MRSA-orfX	MRSA-orfX_UT1	GGATCAAAACGGCCTGCACA	299
	MRSE ID – orfX junction	MRSE-orfX_UT1	TGGCTCCAATGGTYTACACCAA	300
			GTCAAAAATCATGAACCTCATTACTTATG	301
			ATTTCATATATGTAATTCCTCCACATCTC	302
			CAATATTATCTCGTAATTTACCTTGTTT	303
			CTCTGCTTTATATTATAAAAATTACGGCTG	304
			CACTTTTTATCTTCAAAGATTGAGC	305
SCCmec typing	S. aureus ccrB2	SCCmec_ccrB2_UT	CTCATGTTACARATACTTGCG	306
			CCTTGATAATAGCCTTCTTGG	307
	S. aureus mecI	SCCmec_mecI_UT	CGTTATAAGTGACGAATGGTTTTTG	308
			TCATCTGCAGAATGGGAAGTT	309
	S. aureus IS1272J	SCCmec_IS1272J_UT	GAAGCTTTGGGCGATAAAGA	310
			GCACTGTCTCGTTTAGACCAATC	311
	S. aureus ccrC	SCCmec_ccrC_UT	TCCAGTCTATAAAGGSTATGTCAG	312
			ACTTATAATGGCTTCATGCTTACC	313
	S. aureus ccrB1	SCCmec_ccrB1_UT	ACCACAAACACACTTAAAGATG	314
			CAATTTCAAGTATTTGGTCCATAAC	315
	S. aureus ccrB3	SCCmec_ccrB3_UT	AACACAACGAACACATTGAAAG	316
			CGTATTTCTCAATCACATCAGC	317
	S. aureus ccrB4	SCCmec_ccrB4_UT	CGAAGTATAGACACTGGAGCGATA	318
			GCGACTCTCTTGGCGTTTA	319
	S. aureus mecC2	SCCmec_mecC2_UT	TCAGTTCAATTGCTCACGATATG	320
			GCCAACGGCTACAGTGATAA	321
	SCCmecIVa -J1	SCCmecIVa-J1_UT	CAGGATATTTTTTCAAACCTCCTTCA	322

Target purpose	Target ID	Assay	Sequence	SEQ ID NO:
			TGGAGGACCAAGGATATTCG	323
	SCCmecIVa	SCCmecIVa_UT	CCTTTGAATGCCCTCCATGAATAAAAT	324
			GCATATAGAAAAGATAGAAAGTTCGAAAGA	325
S. aureus virulence	MRSA sea	sea_UT	TATGGCTAGACGGTAAACAAAATACAG	326
			CTCTGAACCTTCCCATCAAAAACATC	327
	MRSA seb	seb_UT	TTACTGTTMGGGTATTTGAAGATGG	328
			CCGTTTCATAAGGYGAGTTGT	329
E. coli virulence	Shiga toxin genes	stx1_UT	ATGTCAGAGGGATAGATCCA	330
			TATAGCTACTGTCACCAGACAAT	331
		stx2_UT	AGTTCCTGCGTTTTGTCACTGTC	332
			CGAAGCACATTGCTGATT	333
	E. coli aggR	EAEC_aggR_UT	GATACATTAAAGACGCCCTAAAGGATGC	334
			CCTTTGACCAATTCGGACAACT	335
	T3SS in E. coli	escV_UT	TGGAATAATCATCATCATATTACAGCCA	336
			GAGTCATATCACGATCTTATTCTGGC	337
	E. coli virulence	bfpA_UT	TACCAGTCTGCGTCTGATTCC	338
			ACGTTGCGCTCATTACTTCTG	339
	E. coli (and Shigella) invasion antigen	EIEC_lpaH3_UT	TGAGTTACCTGAATCACTGGAAG	340
			TCGAGGATGATAGTGCAGGTC	341
	E. coli enterotoxins	ETEC_eltA_UT	CTGCGTTAGGTGGAATACCAT	342
			CTGGGTCTCCTCATTACAAGTATC	343
		ETEC_estA_UT	GTTCCAGCCTGCCATCTG	344
			TCGCCAGCAACACATTCAG	345
	E. coli UPEC	UPEC_fyuA_UT	AAGCACGCTGGTGGTTAC	346
			CTGGATGGTGTGGTGGA	347
Klebsiella capsule type	wzi gene of cps locus	wzi_UT	CGCGAGYGCCTTTCTATCTTG	348

Target purpose	Target ID	Assay	Sequence	SEQ ID NO:
			GAGASCCACTGGTTCCAGAA	349
CONTROL PCR	IPSC	IPSC_UT1	ACCCAACTGAATGGAGCGGGCGGA	350
			CGAAAACCCCTTGACACAG	
			ACGCACCTTGACTTGCTTCGCCGGG	351
			ATGCCCTTACCTAGACGCAATGA	
		IPSC_UT2	ACCCAACTGAATGGAGCGGGCAGC	352
			CGTTGAGGGCAAAAGTGATAC	
			ACGCACCTTGACTTGCTTCGGAGTTC	353
			CGTCCGGTTAAGCGTGACAGTC	

Table 4 – Organisms Detected Correctly Compared to Culture Methods

Specimen Type	Organisms Detected Correctly Compared to Culture Methods/Total Tested					
	<i>K. pneumoniae</i>	<i>S. aureus</i>	<i>S. pneumoniae</i>	<i>P. aeruginosa</i>	<i>E. coli</i>	<i>A. baumannii</i>
Respiratory	4/4***	4/4	2/2	4/4	2/2	1/1
Urine	9/9***				1/1	
Wound	1/1	3/3				
Nasal Swab		5/5	3/3	1/1		

** None, Negative culture results or culture detected other organisms

Bold-face type indicates antibiotic resistance genes detected by amplicon sequencing

*** In one sample, the inventors found a fluoroquinolone-conferring mutation in the *gyrA* gene.

Table 5 – False Positives

Specimen Type	False Positives*					
	<i>K. pneumoniae</i>	<i>S. aureus</i>	<i>S. pneumoniae</i>	<i>P. aeruginosa</i>	<i>E. coli</i>	<i>A. baumannii</i>
Respiratory	1	3	0	0	1	1
Urine					1	
Wound						
Nasal Swab		2	0	0		

* Detected by methodology when culture was negative or detected other organisms.

Table 6. Identification of Bacterial Species and Antibiotic Resistances in 27 Respiratory Specimen Samples from Subjects with Cystic Fibrosis

Bacterial Species Identified in Respiratory Specimen Samples	
<i>Bacteria Identified</i>	<i>Number of Identifications</i>
<i>P. aeruginosa</i>	22
<i>E. coli</i>	2 (UPEC)
MRSA	15 (9 CC5, SCCmecII)
MRSE	11
MSSA	1
<i>S. pneumoniae</i>	2
<i>A. baumannii</i>	2
GBS	5
<i>E. faecalis</i>	2
Other <i>Staphylococcus</i> species	3
<i>P. mirabilis</i>	2
Negative	2
Antibiotic Resistance	
<i>Antibiotic Resistance</i>	<i>Number of Identifications</i>
Macrolides	10
Fluoroquinolones	3
Trimethoprim-sulfamethoxazole	5

Clindamycin	1
-------------	---

Table 7. Identification of Bacterial Species and Antibiotic Resistances in 24 Respiratory Specimen Samples from Subjects with Chronic Rhinosinusitis

Bacterial Species Identified in Respiratory Specimen Samples	
<i>Bacteria Identified</i>	<i>Number of Identifications</i>
<i>P. aeruginosa</i>	9
MRSA	7 (1 CC5/SCCmecII, several SCCmecIVa, 1 USA300, 5 SCCmecV)
MRSE	19 (several SCCmecIVa)
MSSA	5 (3 CC8)
<i>E. faecalis</i>	1
<i>E. faecium</i>	1
Other <i>Staphylococcus</i> species	2
<i>P. mirabilis</i>	6
<i>E. aerogenes</i>	2
<i>K. pneumoniae</i>	1 (CG258)
Antibiotic Resistance Found	
<i>Antibiotic Resistance</i>	<i>Number of Identifications</i>
Aminoglycosides	11

ESBL	9
Macrolides	31
Fluoroquinolones	15
Trimethoprim-sulfamethoxazole	12
Clindamycin	4
Tetracycline	8
Vancomycin	2

Table 8. Identification of Bacterial Species and Antibiotic Resistances in 38 Respiratory Specimen Samples from Subjects with Various Respiratory Diseases

Bacterial Species Identified in Respiratory Specimen Samples	
<i>Bacteria Identified</i>	<i>Number of Identifications</i>
P aeruginosa	8
MRSA	6 (4 SCCmecIVa, 3 USA300)
MRSE	4 (1 SCCmecIVa)
MSSA	6 (1 CC5)
E. faecalis	8
E. faecium	1
Other Staphylococcus species	8 (2 SCCmecIVa)

S. pneumoniae	4
A. baumannii	3
P. mirabilis	1
E. aerogenes	1
E. cloacae	3
K. pneumoniae	4
E. coli	4
K. oxytoca	2
GBS	1
Negative	2
Antibiotic Resistance Found	
<i>Antibiotic Resistance</i>	<i>Number of Identifications</i>
Aminoglycosides	1
Macrolides	11
Fluoroquinolones	7
Trimethoprim-sulfamethoxazole	4
Clindamycin	1
Tetracycline	3

Table 9. Oligonucleotides of Lineage Identification Assays

Target purpose	Target ID	Assay	Sequence	SEQ ID NO:
Streptococcus pyogenes M59	M59	M59_UT1	AGAGGAGGAGATAGAGTAGGAGAT	354
			TACCAGCCAGTATGATAAGAAGAGA	355
Staphylococcus aureus USA300	SNP	USA300_UT1	GCACGTTGATGACTTCTGACA	356
			GCCCAACAACAATCCAACCTTAC	357
S. aureus CC8 (excluding ST239)	SNP	CC8_UT1	TGCCCCATAACACATTTTGACACTTT	358
			TTCGGCCACAGCTAAACTCG	359
S. aureus USA300/500	SNP	USA3-500_UT1	ACCTTATACGGAACATAGCAGACG	360
			TCGATGCGCTTCTATCACCTC	361
S. aureus Archaic Iberian	SNP	Archlber_UT_B+	CGCCAAATGACTCGCATTGT	362
			GCAATGTGCCITTTCCGAARTAAA	363
S. aureus True Iberian	SNP	Trulber_UT1	GCGCAACAGGGAAGCAA	364
			TGCGGATGTCCCTATGTCTGAAAG	365
S. aureus ST239	SNP	ST239_UT1	CATGACCGCCACTATAACCAGA	366
			ATGCAACATTAGCAGGAGGATG	367
S. aureus CC5	SNP	CC5_UT1	CTGGCGCTCTCCAGCA	368
			TTGCAATTAGTGTGTAGGTGGTAA	369
E. coli strain ID	E. coli H30-Rx	H30-Rx-snp200_UT	GACACCATGCGTTTTTGCTTC	370
			TCGTACCGGCAACAATTGAC	371
	E. coli H30-Rx	H30-Rx-snp264_UT	GTGGCGATTTACGCTGTTA	372
			TATCCAGCACGTTCCAGGTG	373
	trpA	trpA_ST131_O16_UT	AAAACCGCGCGCGTTACCT	374

Target purpose	Target ID	Assay	Sequence	SEQ ID NO:
			CCAGAAATCGCGCCCGCATT	375
	pabB	pabB_ST131_O25b	TCCAGCAGGTGCTGGATCGT	376
			GCGAAATTTTCGCCGTACTGT	377
	uidA	uidA_1_UT	CATTACGGCAAGTGTGGTCAAT	378
			CCATCAGCACGTTATCGAATCCTT	379
	uidA	uidA_2_UT	CGTATCACHGTTTGTGTGAACAA	380
			GGATTCACYACTTGCAAAGTCC	381
	rfbO16	rfbO16_UT	ATACCGACGACGCCGATCTG	382
			GGATCATTTATGCTGGTACG	383
	rfbO25b	rfbO25b_UT	ATACCGACGACGCCGATCTG	384
			TGCTATTCAATTATGCGCAGC	385
K. pneumo MLST	rpoB gene	KpST-rpoB_UT	GGCGAAATGGCWGAGAAACCA	386
			GAGTCTTCGAAGTTGTAACC	387
	gapA	KpST-gapA_UT	TGAAATATGACTCCACTCACGG	388
			CTTCAGAAGCGGCTTTGATGGCTT	389
	mdh	KpST-mdh_UT	CCCAACTCGCTTCAGGTTTCAG	390
			CCGTTTTTCCCCCAGCAGCAG	391
	phosphoglucose isomerase	KpST-pgi_UT	GAGAAAAACCTGCCTGTACTGCTGGC	392
			CGCGCCACGCTTTATAGCGGTTAAT	393
	phoE	KpST-phoE_UT	ACCTACCGCAACACCCGACTTCTTCGG	394
			TGATCAGAACTGTAGGTGAT	395
	infB	KpST-infB_UT	CTCGCTGCTGGACTATATTTCG	396
			CGCTTTCAGCTCAAGAACTTC	397
	tonB	KpST-tonB_UT	CTTTATACCTCGGTACATCAGGTT	398
			ATTCGCCGGCTGRGCRGAGAG	399
ST258 ID assay / Kp-species ID	SNP 3576137-Y	ST258_UT_FB	ATGGTGGTGCGCCAGTG	400

Target purpose	Target ID	Assay	Sequence	SEQ ID NO:
CG258 ID assay			GCTGACCGAGACGTTGTC	401
		CG258_UT_FB	ACGGCAGCGGATTTGATTTAACG	402
E. coli MLST			AGCTGCGTGATCGAGACCTATC	403
	adk	EcST-adk_UT	TCATCATCTGCACCTTTCCGC	404
			CCAGATCAGCGCGAACTTCA	405
	fumC	EcST-fumC_UT	TCACAGGTGCGCAGCGCTTC	406
			TCCCGCAGATAAGCTGTGG	407
	gyrB	EcST-gyrB_UT	TCGGCGACACGGATGACGGC	408
			GTCCATGTAGGCGTTCAGGG	409
	icd	EcST-icd_UT	ATGGAAGTAAAGTAGTTGTTTC CGGCACA	410
			GGACGCAGCAGGATCTGTT	411
	mdh	EcST-mdh_UT	AGCGGTTCTGTTCAAATGC	412
			CAGGTTCAGAACTCTCTCTGT	413
	purA	EcST-purA_UT	TCGGTAACGGTGTGTGTCTG	414
			CATACGGTAAGCCACGCAGA	415
	recA	EcST-recA_UT	ACCTTTGTAGCTGTACCACG	416
			AGCGTGAAGGTAAACCTGTG	417
S. aureus MLST	arc	SaST-arc_UT	TTGATTCAACGCGCGTATTGTC	418
			AGGTATCTGCTTCAATCAGCG	419
	aro	SaST-aro_UT	ATCGGAAATCCTATTTACACATTC	420
			GGTGTGTGTAATAAACGATATC	421
	glp	SaST-glp_UT	CTAGGAACTGCAATCTTAATCC	422
			TGGTAAATCGCATGTCCAATTC	423
	gmk	SaST-gmk_UT	ATCGTTTTATCGGGACCATC	424
			TCATTAACTACAAACGTAATCGTA	425
	pta	SaST-pta_UT	GTTAAATCGTATTACCTGAAGG	426
			GACCCCTTTGTTGAAAAGCTTAA	427

Target purpose	Target ID	Assay	Sequence	SEQ ID NO:
	tpi	SaST-tpi_UT	TCGTTCAATTCTGAACGTCGTGAA	428
			TTTGACACCTTCTAACAATTGTAC	429
	yqi	SaST-yqi_UT	CAGCATACAGGACACACCTATTGGC	430
			CGTTGAGGGAATCGATACTACTGGAAC	431

CLAIMS

What is claimed is:

1. A method of detecting and characterizing one or more microorganisms within a sample, the method comprising the steps of:

receiving the sample;

extracting a template from the sample;

performing a multiplex polymerase chain reaction assay comprising the steps of:

amplifying a first marker from the template to create a first amplicon, wherein the first amplicon comprises a first universal tail sequence and a second universal tail sequence, and further wherein the first marker is specific for at least a genus of microorganism;

amplifying a second marker from the template to create a second amplicon, wherein the second amplicon comprises the first universal tail sequence and the second universal tail sequence, and further wherein the second marker is specific for at least one antibiotic resistance gene;

sequencing the first and second amplicons to detect and characterize the one or more microorganisms.

2. The method of claim 1 and further comprising amplifying a third marker from the template to create a third amplicon, wherein the third amplicon comprises a first universal tail sequence and a second universal tail sequence, and further wherein the third marker is specific for strain identity and/or lineage of the one or more microorganisms.

3. The method of claim 1 or 2, wherein the first amplicon comprises a sequence selected from SEQ ID NOs: 1-109.

4. The method according to any one of claims 1 to 3, wherein the second amplicon comprises a sequence selected from SEQ ID NOs: 110-298.

5. The method of claim 2, wherein the third amplicon comprises a sequence selected from SEQ ID NOs: 299-431.
6. The method of claim 2, wherein the amplification of the first, second, and third markers occurs within a single reaction vessel.
7. The method of claim 2 further comprising sequencing the third amplicon.
8. The method according to any one of claims 1 to 7 further comprising adding an index to the first amplicon and the second amplicon using at least one indexing oligonucleotide, wherein the at least one indexing oligonucleotide comprises a complimentary sequence that recognizes at least one of the first universal tail sequence and the second universal tail sequence.
9. The method according to any one of claims 1 to 8, wherein at least one of the one or more microorganisms is a pathogenic microorganism.
10. The method of claim 9, wherein the pathogenic microorganism is an ESKAPE pathogen selected from the group consisting of *Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa*, and *Enterobacter* spp.
11. The method according to any one of claims 1 to 10, wherein the template comprises nucleic acids from a subject and the one or more microorganisms.
12. The method of claim 11, wherein the template is DNA.
13. A method of detecting and characterizing a microorganism in a sample from a subject, the method comprising the steps of:
 - receiving the sample;
 - performing a polymerase chain reaction comprising the steps of:
 - amplifying a first marker within the sample, wherein amplifying the first marker comprises mixing the sample with a first oligonucleotide comprising a first universal tail sequence and a

second oligonucleotide comprising a second universal tail, and further wherein the first marker is specific for at least a genus of microorganism;

amplifying a second marker within the sample, wherein amplifying the second marker comprises mixing the sample with a third oligonucleotide comprising the first universal tail sequence and a fourth oligonucleotide comprising the second universal tail, and further wherein the second marker comprises an antibiotic resistance gene;

wherein amplifying the first marker and the second marker results in a first amplicon and a second amplicon; and

sequencing the first amplicon and the second amplicon to detect and characterize the microorganism.

14. The method of claim 13, wherein the subject is an animal.
15. The method of claim 14, wherein the animal is a human.
16. The method according to any one of claims 13 to 15, wherein the sample is selected from at least one of a respiratory sample, a nasal swab sample, a urine sample, and a blood sample.
17. The method according to any one of claims 13 to 16, wherein the first oligonucleotide and the second oligonucleotide independently comprise a sequence selected from SEQ ID NOs: 1-109.
18. The method according to any one of claims 13 to 17, wherein the third oligonucleotide and the fourth oligonucleotide independently comprise a sequence selected from SEQ ID NOs: 110-298.
19. The method according to any one of claims 13 to 18, wherein the amplification of the first amplicon and the second amplicon occurs in a single reaction vessel.

20. The method according to any one of claims 13 to 19, further comprising amplifying a third marker within the sample, wherein amplifying the third marker comprises mixing the sample with a fifth oligonucleotide comprising the first universal tail sequence and a sixth oligonucleotide comprising the second universal tail, and further wherein the third marker is specific for strain identity and/or lineage of the one or more microorganisms, wherein amplifying the third marker results in a third amplicon.

21. The method of claim 20, wherein the fifth oligonucleotide and the sixth oligonucleotide independently comprise a sequence selected from SEQ ID NOs: 299-431.

22. The method of claim 20 or 21, wherein the amplification of the first amplicon, the second amplicon, and the third amplicon occurs in a single reaction vessel.

23. The method according to any one of claims 13 to 22, wherein the microorganism is a pathogenic microorganism.

24. The method of claim 23, wherein the pathogenic microorganism is an ESKAPE pathogen selected from the group consisting of *Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa*, and *Enterobacter* spp.