

CLAIMS

We Claim:

1. A system for characterization of a microorganism-related condition, the system comprising:
 - a sample handling system comprising a sequencing system operable to determine microorganism genetic sequences based on samples associated with a set of subjects, wherein the samples comprise microorganism nucleic acids associated with the microorganism-related condition from site-diverse samples corresponding to a plurality of collection sites comprising at least two of gut, genitals, mouth, skin, and nose;
 - a set of microbiome characterization modules operable to apply a set of analytical techniques comprising at least two of a statistical test, a dimensionality reduction technique, and an artificial intelligence approach, and wherein the set of microbiome characterization modules comprises:
 - a first microbiome characterization module operable to apply a first analytical technique, of the set of analytical techniques, to determine a set of microbiome features based on the microorganism genetic sequences, wherein the set of microbiome features is associated with the microorganism-related condition;
 - and
 - a second microbiome characterization module operable to apply a second analytical technique, of the set of analytical techniques, to determine a processed microbiome feature set based on the set of microbiome features, wherein the processed microbiome feature set is adapted to improve the characterization of the microorganism-related condition; and
 - a microorganism-related condition model generated based on the processed microbiome feature set, wherein the microorganism-related condition model is operable to determine a characterization of the microorganism-related condition for a user, the characterisation comprising a set of site-wise disease propensity metrics based on the set of site-diverse samples and the microorganism-related condition model, wherein each site-wise disease propensity metric, of the set of site-wise disease propensity metrics, corresponds to a different collection site of the plurality of collection sites and is associated with the microorganism-related condition.
2. The system of Claim 1, wherein the first analytical technique comprises a statistical test comprising at least one of a t-test, a Kolmogorov-Smirnov test, and a regression model, and wherein the first microbiome characterization module is operable to apply the statistical test to determine the set of microbiome features based on the microorganism genetic sequences.
3. The system as in any one of Claims 1 and 2,
 - wherein the samples comprise site-diverse samples collected from a plurality of collection sites comprising at least two of gut, genitals, mouth, skin, and nose, and

wherein the first microbiome characterization module is operable to apply the statistical test to determine first subsets of microbiome features of the set microbiome features based on the site-diverse samples, wherein each subset of microbiome features from the first subsets of microbiome features corresponds to a different collection site from the plurality of collection sites.

4. The system of Claim 3,
wherein the second microbiome characterization module is operable to apply an additional statistical test to determine second subsets of microbiome features of the set of microbiome features based on the site-diverse samples, and
wherein the microorganism-related condition model is generated based on the first subsets and the second subsets of microbiome features.
5. The system as in any one of Claims 1 and 2, wherein the second microbiome characterization module is operable to apply the second analytical technique to perform at least one of feature selection, feature weighting, and warm start, for processing the set of microbiome features into the processed microbiome feature set.
6. The system of Claim 1, wherein the microorganism-related condition model comprises a skin-related characterization model generated based on the processed microbiome feature set, wherein the skin-related characterization model is operable to determine the characterization of a photosensitivity-associated condition for the user, and wherein the set of microbiome features comprises features associated with at least one of: *Alloprevotella* (genus), *Prevotella* sp. WAL 2039G (species), *Corynebacterium mastitidis* (species), *Bacteroidaceae* (family), *Blautia* (genus), *Bacteroides* (genus), *Desulfovibrio* (genus), *Clostridium* (genus), *Bacteroides vulgatus* (species), *Faecalibacterium prausnitzii* (species), *Blautia faecis* (species), *Alistipes putredinis* (species), *Bacteroides* sp. AR20 (species), *Bacteroides* sp. AR29 (species), *Bacteroides acidifaciens* (species), *Dielma* (genus), *Slackia* (genus), *Eggerthella* (genus), *Adlercreutzia* (genus), *Paraprevotella* (genus), *Alistipes* (genus), *Holdemania* (genus), *Eisenbergiella* (genus), *Enterorhabdus* (genus), *Adlercreutzia equolifaciens* (species), *Phascolarctobacterium succinatutens* (species), *Roseburia inulinivorans* (species), *Phascolarctobacterium* sp. 377 (species), *Desulfovibrio piger* (species), *Eggerthella* sp. HGA1 (species), *Lactonifactor longoviformis* (species), *Alistipes* sp. HGB5 (species), *Holdemania filiformis* (species), *Collinsella intestinalis* (species), *Neisseria macacae* (species), *Clostridiaceae* (family), *Gemella sanguinis* (species), *Bacteroides fragilis* (species), *Enterobacteriaceae* (family), *Lachnospiraceae* (family), *Pasteurellaceae* (family), *Pasteurellales* (order), *Enterobacteriales* (order), *Sphingobacteriales* (order), *Haemophilus* (genus), *Leuconostoc* (genus), *Brevundimonas* (genus), *Prevotella oris* (species), *Odoribacter* (genus), *Capnocytophaga* (genus), *Flavobacterium* (genus), *Pseudomonas brenneri* (species), *Flavobacterium ceti* (species), *Brevundimonas* sp. FXJ8.080 (species), *Ruminococcaceae* (family), *Vibrionaceae* (family),

Flavobacteriaceae (family), Fusobacteriaceae (family), Porphyromonadaceae (family), Brevibacteriaceae (family), Rhodobacteraceae (family), Intrasporangiaceae (family), Bifidobacteriaceae (family), Sphingobacteriaceae (family), Caulobacteraceae (family), Campylobacteraceae (family), Bacteroidia (class), Fusobacteriia (class), Flavobacteriia (class), Bifidobacteriales (order), Neisseriales (order), Bacteroidales (order), Rhodobacterales (order), Flavobacteriales (order), Vibrionales (order), Fusobacteriales (order), Caulobacterales (order), Fusobacteria (phylum), Actinobaculum (genus), Varibaculum (genus), Fusicatenibacter (genus), Brevibacterium (genus), Faecalibacterium (genus), Campylobacter (genus), Actinobacillus (genus), Porphyromonas (genus), Fusobacterium (genus), Chryseobacterium (genus), Megasphaera (genus), Rothia (genus), Neisseria (genus), Lactobacillus sp. BL302 (species), Bacteroides plebeius (species), Corynebacterium ulcerans (species), Varibaculum cambriense (species), Blautia wexlerae (species), Staphylococcus sp. WB18-16 (species), Streptococcus sp. oral taxon G63 (species), Propionibacterium acnes (species), Anaerococcus sp. 9401487 (species), Haemophilus parainfluenzae (species), Staphylococcus epidermidis (species), Campylobacter ureolyticus (species), Janibacter sp. M3-5 (species), Prevotella timonensis (species), Peptoniphilus sp. DNFO0840 (species), Finegoldia sp. S8 F7 (species), Prevotella disiens (species), Porphyromonas catoniae (species), Fusobacterium periodonticum (species), Infectious Diseases (KEGG2), Poorly Characterized (KEGG2), Metabolic Diseases (KEGG2), Immune System Diseases (KEGG2), Cellular Processes and Signaling (KEGG2), Restriction enzyme (KEGG3), Nucleotide excision repair (KEGG3).

7. The system of Claim 1, wherein the microorganism-related condition model comprises a skin-related characterization model generated based on the processed microbiome feature set, wherein the skin-related characterization model is operable to determine the characterization of a dry skin-associated condition for the user, and wherein the set of microbiome features comprises features associated with at least one of: Corynebacteriaceae (family), Bacilli (class), Lactobacillales (order), Actinomycetales (order), Firmicutes (phylum), Corynebacterium (genus), Dermabacteraceae (family), Lactobacillaceae (family), Propionibacteriaceae (family), Actinobacteria (class), Dermabacter (genus), Dialister (genus), Facklamia (genus), Lactobacillus (genus), Propionibacterium (genus), Corynebacterium ulcerans (species), Facklamia hominis (species), Corynebacterium sp. (species), Propionibacterium sp. MSP09A (species), Facklamia sp. 1440-97 (species), Staphylococcus sp. C9I2 (species), Anaerococcus sp. 9402080 (species), Corynebacterium glucuronolyticum (species), Dermabacter hominis (species), Enterobacteriaceae (family), Pseudomonadaceae (family), Staphylococcaceae (family), Gammaproteobacteria (class), Bacillales (order), Enterobacteriales (order), Bifidobacterium (genus), Pseudomonas (genus), Anaeroglobus (genus), Kluyvera (genus), Atopobium (genus), Staphylococcus (genus), Lactobacillus sp. BL302 (species), Corynebacterium mastitidis (species),

Bifidobacterium longum (species), *Anaeroglobus geminatus* (species), *Anaerococcus* sp. S9 PR-16 (species), *Prevotella timonensis* (species), *Kluyvera georgiana* (species), *Actinobaculum* (genus), *Finegoldia* (genus), *Cronobacter* (genus), *Acinetobacter* sp. WB22-23 (species), *Anaerococcus octavius* (species), *Finegoldia* sp. S9 AA1-5 (species), *Staphylococcus* sp. C-D-MA2 (species), *Peptoniphilus* sp. 7-2 (species), *Cronobacter sakazakii* (species), *Pasteurellaceae* (family), *Acidobacteriia* (class), *Sphingobacteriia* (class), *Sphingobacteriales* (order), *Acidobacteria* (phylum), *Porphyromonas* (genus), *Haemophilus* (genus), *Acinetobacter* (genus), *Anaerococcus* sp. 8405254 (species), *Sphingomonadaceae* (family), *Sphingomonadales* (order), *Kocuria* (genus), *Gemella* (genus), *Veillonella* sp. CM60 (species), *Lactobacillus* sp. 7_1_47FAA (species), *Gemella* sp. 933-88 (species), *Porphyromonas catoniae* (species), *Haemophilus parainfluenzae* (species), *Bacteroides* sp. AR20 (species), *Bacteroides vulgatus* (species), *Bacteroides* sp. D22 (species), *Dorea longicatena* (species), *Parabacteroides merdae* (species), *Bacteroides* sp. AR29 (species), *Dorea* (genus), *Collinsella* (genus), *Bacteroides* (genus), *Oscillospiraceae* (family), *Ruminococcaceae* (family), *Bacteroidaceae* (family), *Verrucomicrobiaceae* (family), *Coriobacteriaceae* (family), *Clostridiales* (order), *Bacteroidales* (order), *Verrucomicrobiales* (order), *Coriobacteriales* (order), *Thermoanaerobacterales* (order), *Clostridia* (class), *Bacteroidia* (class), *Verrucomicrobiae* (class), *Verrucomicrobia* (phylum), *Bacteroidetes* (phylum), Translation (KEGG2), Cellular Processes and Signaling (KEGG2), Amino Acid Metabolism (KEGG2), Cell Growth and Death (KEGG2), Replication and Repair (KEGG2), Metabolism of Other Amino Acids (KEGG2), Neurodegenerative Diseases (KEGG2), Metabolism of Cofactors and Vitamins (KEGG2), Transport and Catabolism (KEGG2), Endocrine System (KEGG2), Immune System Diseases (KEGG2), Excretory System (KEGG2), Enzyme Families (KEGG2), Membrane Transport (KEGG2), Carbohydrate Metabolism (KEGG2), Biosynthesis of Other Secondary Metabolites (KEGG2), Metabolism of Terpenoids and Polyketides (KEGG2), Infectious Diseases (KEGG2), Genetic Information Processing (KEGG2), Nervous System (KEGG2), Environmental Adaptation (KEGG2), Nucleotide Metabolism (KEGG2), Signaling Molecules and Interaction (KEGG2), Signal Transduction (KEGG2), Inorganic ion transport and metabolism (KEGG3), Chromosome (KEGG3), Cell cycle - *Caulobacter* (KEGG3), Ribosome Biogenesis (KEGG3), DNA replication proteins (KEGG3), Translation factors (KEGG3), Glycine, serine and threonine metabolism (KEGG3), Sulfur metabolism (KEGG3), Other ion-coupled transporters (KEGG3), Valine, leucine and isoleucine biosynthesis (KEGG3), Nitrogen metabolism (KEGG3), Peptidoglycan biosynthesis (KEGG3), Homologous recombination (KEGG3), Peroxisome (KEGG3), Sulfur relay system (KEGG3), Peptidases (KEGG3), Protein kinases (KEGG3), Mismatch repair (KEGG3), Xylene degradation (KEGG3), Ribosome (KEGG3), RNA polymerase (KEGG3), Tryptophan metabolism (KEGG3), Histidine metabolism (KEGG3), Vitamin metabolism (KEGG3), Cell motility and secretion (KEGG3), Pyrimidine metabolism (KEGG3), Cytoskeleton proteins (KEGG3), DNA replication (KEGG3), Amino sugar and

nucleotide sugar metabolism (KEGG3), Folate biosynthesis (KEGG3), Carbon fixation in photosynthetic organisms (KEGG3), Phosphatidylinositol signaling system (KEGG3), Lysine degradation (KEGG3), Selenocompound metabolism (KEGG3), Fructose and mannose metabolism (KEGG3), Inositol phosphate metabolism (KEGG3), Protein folding and associated processing (KEGG3), PPAR signaling pathway (KEGG3), Lipid metabolism (KEGG3), Valine, leucine and isoleucine degradation (KEGG3), Glyoxylate and dicarboxylate metabolism (KEGG3), Arginine and proline metabolism (KEGG3), Limonene and pinene degradation (KEGG3), D-Alanine metabolism (KEGG3), Porphyrin and chlorophyll metabolism (KEGG3), C5-Branched dibasic acid metabolism (KEGG3), Chaperones and folding catalysts (KEGG3), Fatty acid metabolism (KEGG3), Glutathione metabolism (KEGG3), Pentose phosphate pathway (KEGG3), Phosphotransferase system (PTS) (KEGG3), Pantothenate and CoA biosynthesis (KEGG3), Proximal tubule bicarbonate reclamation (KEGG3), Galactose metabolism (KEGG3), Starch and sucrose metabolism (KEGG3), Primary immunodeficiency (KEGG3), Cysteine and methionine metabolism (KEGG3), Ubiquinone and other terpenoid-quinone biosynthesis (KEGG3), DNA repair and recombination proteins (KEGG3), Tyrosine metabolism (KEGG3), Phenylalanine, tyrosine and tryptophan biosynthesis (KEGG3), Aminoacyl-tRNA biosynthesis (KEGG3), Alanine, aspartate and glutamate metabolism (KEGG3), Photosynthesis (KEGG3), Other transporters (KEGG3), Butanoate metabolism (KEGG3), Bacterial secretion system (KEGG3), Glycerophospholipid metabolism (KEGG3), Oxidative phosphorylation (KEGG3), Type I diabetes mellitus (KEGG3), Glycolysis / Gluconeogenesis (KEGG3), Photosynthesis proteins (KEGG3), Transporters (KEGG3), Terpenoid backbone biosynthesis (KEGG3), Biosynthesis of unsaturated fatty acids (KEGG3), Signal transduction mechanisms (KEGG3), Synthesis and degradation of ketone bodies (KEGG3), Nucleotide excision repair (KEGG3), Secretion system (KEGG3), Alzheimer's disease (KEGG3), Zeatin biosynthesis (KEGG3), Type II diabetes mellitus (KEGG3), D-Glutamine and D-glutamate metabolism (KEGG3), Taurine and hypotaurine metabolism (KEGG3), Glutamatergic synapse (KEGG3), Plant-pathogen interaction (KEGG3), Vitamin B6 metabolism (KEGG3), Citrate cycle (TCA cycle) (KEGG3), Ethylbenzene degradation (KEGG3), Base excision repair (KEGG3), Replication, recombination and repair proteins (KEGG3), Ribosome biogenesis in eukaryotes (KEGG3), Aminobenzoate degradation (KEGG3), Bacterial motility proteins (KEGG3), Biosynthesis of ansamycins (KEGG3), Ion channels (KEGG3), Metabolism (KEGG2), Poorly Characterized (KEGG2), Biosynthesis and biodegradation of secondary metabolites (KEGG3), Lipoic acid metabolism (KEGG3), Amino acid related enzymes (KEGG3), Translation proteins (KEGG3), Ascorbate and aldarate metabolism (KEGG3), Thiamine metabolism (KEGG3), Function unknown (KEGG3), Glycosaminoglycan degradation (KEGG3), Others (KEGG3), Pentose and glucuronate interconversions (KEGG3), Biotin metabolism (KEGG3), Phenylalanine metabolism (KEGG3), Glycosphingolipid biosynthesis - ganglio series (KEGG3), Pores ion channels

(KEGG3), Membrane and intracellular structural molecules (KEGG3), Purine metabolism (KEGG3), One carbon pool by folate (KEGG3), Phosphonate and phosphinate metabolism (KEGG3), Lysosome (KEGG3), Drug metabolism - other enzymes (KEGG3), Penicillin and cephalosporin biosynthesis (KEGG3), Huntington's disease (KEGG3), Nicotinate and nicotinamide metabolism (KEGG3), Drug metabolism - cytochrome P450 (KEGG3), Lipopolysaccharide biosynthesis proteins (KEGG3), Metabolism of xenobiotics by cytochrome P450 (KEGG3), Tuberculosis (KEGG3), and Polycyclic aromatic hydrocarbon degradation (KEGG3).

8. The system of Claim 1, wherein the microorganism-related condition model comprises a skin-related characterization model generated based on the processed microbiome feature set, wherein the skin-related characterization model is operable to determine the characterization of a scalp-related condition for the user, and wherein the set of microbiome features comprises features associated with at least one of: Actinobacteria (class), Lactobacillales (order), Actinomycetales (order), Firmicutes (phylum), Dermabacteraceae (family), Lactobacillaceae (family), Propionibacteriaceae (family), Corynebacteriaceae (family), Lactobacillus (genus), Corynebacterium (genus), Propionibacterium (genus), Dermabacter (genus), Eremococcus (genus), Corynebacterium freiburgense (species), Eremoc(KEGG3)occus coleocola (species), Corynebacterium sp. (species), Staphylococcus sp. C9I2 (species), Anaerococcus sp. 8405254 (species), Corynebacterium glucuronolyticum (species), Dermabacter hominis (species), Coriobacteriaceae (family), Enterobacteriaceae (family), Staphylococcaceae (family), Enterobacteriales (order), Bacillales (order), Bifidobacterium (genus), Staphylococcus (genus), Atopobium (genus), Megasphaera (genus), Corynebacterium mastitidis (species), Streptococcus sp. BS35a (species), Finegoldia magna (species), Staphylococcus aureus (species), Haemophilus influenzae (species), Corynebacterium sp. NML 97-0186 (species), Streptococcus sp. oral taxon G59 (species), Dorea (genus), Roseburia sp. 11SE39 (species), Dorea longicatena (species), Prevotellaceae (family), Veillonellaceae (family), Oscillospiraceae (family), Negativicutes class, Selenomonadales (order), Finegoldia (genus), Oscillospira (genus), Intestinimonas (genus), Flavonifractor (genus), Prevotella (genus), Moryella (genus), Catenibacterium mitsuokai (species), Collinsella aerofaciens (species), Peptoniphilus sp. 2002-2300004 (species), Corynebacterium canis (species), Finegoldia sp. S9 AA1-5 (species), Prevotella buccalis (species), Dialister invisus (species), Moraxella (genus), Neisseria (genus), Neisseria mucosa (species), Rikenellaceae (family), Metabolism of Cofactors and Vitamins (KEGG2), Enzyme Families (KEGG2), Lipid Metabolism (KEGG2), Immune System Diseases (KEGG2), Glycolysis / Gluconeogenesis (KEGG3), Primary immunodeficiency (KEGG3), Pyruvate metabolism (KEGG3), Transport and Catabolism (KEGG2), Neurodegenerative Diseases (KEGG2), Endocrine System (KEGG2), Amino Acid Metabolism (KEGG2), Cellular Processes and Signaling (KEGG2), Signaling Molecules and Interaction (KEGG2), Metabolism of Other Amino Acids (KEGG2), Replication

and Repair (KEGG2), Translation (KEGG2), Cell Growth and Death (KEGG2), Membrane Transport (KEGG2), Biosynthesis of Other Secondary Metabolites (KEGG2), Metabolism of Terpenoids and Polyketides (KEGG2), Inorganic ion transport and metabolism (KEGG3), Vitamin metabolism (KEGG3), Valine, leucine and isoleucine biosynthesis (KEGG3), Peroxisome (KEGG3), Ribosome Biogenesis (KEGG3), Selenocompound metabolism (KEGG3), Histidine metabolism (KEGG3), Chromosome (KEGG3), Sulfur metabolism (KEGG3), PPAR signaling pathway (KEGG3), Porphyrin and chlorophyll metabolism (KEGG3), Phosphatidylinositol signaling system (KEGG3), Inositol phosphate metabolism (KEGG3), Sulfur relay system (KEGG3), Glycine, serine and threonine metabolism (KEGG3), DNA replication proteins (KEGG3), Pantothenate and CoA biosynthesis (KEGG3), Translation factors (KEGG3), Protein folding and associated processing (KEGG3), Type II diabetes mellitus (KEGG3), Protein kinases (KEGG3), Folate biosynthesis (KEGG3), Lysine degradation (KEGG3), RNA polymerase (KEGG3), D-Alanine metabolism (KEGG3), Carbon fixation in photosynthetic organisms (KEGG3), Nitrogen metabolism (KEGG3), Glycerophospholipid metabolism (KEGG3), Biosynthesis of ansamycins (KEGG3), Valine, leucine and isoleucine degradation (KEGG3), Cytoskeleton proteins (KEGG3), Peptidases (KEGG3), Fatty acid metabolism (KEGG3), Cell cycle - *Caulobacter* (KEGG3), Phosphotransferase system (PTS) (KEGG3), Pyrimidine metabolism (KEGG3), Alzheimer's disease (KEGG3), Butanoate metabolism (KEGG3), Tryptophan metabolism (KEGG3), Signal transduction mechanisms (KEGG3), Pentose phosphate pathway (KEGG3), Other ion-coupled transporters (KEGG3), Homologous recombination (KEGG3), Replication, recombination and repair proteins (KEGG3), Xylene degradation (KEGG3), Mismatch repair (KEGG3), Glyoxylate and dicarboxylate metabolism (KEGG3), Arginine and proline metabolism (KEGG3), Peptidoglycan biosynthesis (KEGG3), Chaperones and folding catalysts (KEGG3), Type I diabetes mellitus (KEGG3), DNA replication (KEGG3), Bacterial secretion system (KEGG3), Tyrosine metabolism (KEGG3), Citrate cycle (TCA cycle) (KEGG3), Amino sugar and nucleotide sugar metabolism (KEGG3), Ribosome (KEGG3), Limonene and pinene degradation (KEGG3), Cell motility and secretion (KEGG3), Taurine and hypotaurine metabolism (KEGG3), Oxidative phosphorylation (KEGG3), Fructose and mannose metabolism (KEGG3), Vitamin B6 metabolism (KEGG3), Ion channels (KEGG3), Synthesis and degradation of ketone bodies (KEGG3), Other transporters (KEGG3), Galactose metabolism (KEGG3), Polycyclic aromatic hydrocarbon degradation (KEGG3), Transporters (KEGG3), DNA repair and recombination proteins (KEGG3), Starch and sucrose metabolism (KEGG3), Alanine, aspartate and glutamate metabolism (KEGG3), Ribosome biogenesis in eukaryotes (KEGG3), Secretion system (KEGG3), Biosynthesis of unsaturated fatty acids (KEGG3), Cysteine and methionine metabolism (KEGG3), Base excision repair (KEGG3), Aminobenzoate degradation (KEGG3), Photosynthesis (KEGG3), Photosynthesis proteins (KEGG3), Pores ion

channels (KEGG3), Lipid biosynthesis proteins (KEGG3), and D-Glutamine and D-glutamate metabolism (KEGG3).

9. A method for characterizing a microorganism-related condition, the method comprising:
 - determining a microorganism sequence dataset for a user based on microorganism nucleic acids from a sample associated with the user, the sample comprising a set of site-diverse samples corresponding to a plurality of collection sites comprising at least two of gut, genitals, mouth, skin, and nose; and
 - determining a characterization of the microorganism-related condition for the user based on the microorganism sequence dataset and a microorganism-related condition model generated based on the application, with a set of microbiome characterization modules, of a set of analytical techniques to determine a set of microbiome features, wherein the set of analytical techniques comprises at least one of a statistical test, a dimensionality reduction technique, and an artificial intelligence approach, wherein the set of microbiome characterization modules comprises:
 - a first microbiome characterization module operable to apply a first analytical technique of the set of analytical techniques, and
 - a second microbiome characterization module operable to apply a second analytical technique of the set of analytical techniques. Andwherein the characterisation comprises a set of site-wise disease propensity metrics based on the set of site-diverse samples and the microorganism-related condition model, wherein each site-wise disease propensity metric, of the set of site-wise disease propensity metrics, corresponds to a different collection site of the plurality of collection sites and is associated with the microorganism-related condition.
10. The method of Claim 9, wherein the application of the set of analytical techniques to determine the set of microbiome features comprises:
 - determination of an initial set of microbiome features based on the microorganism sequence dataset; and
 - application, with the first microbiome characterization module of the set of microbiome characterization modules, of the dimensionality reduction technique on the initial set of microbiome features to determine the set of microbiome features, wherein the dimensionality reduction technique comprises at least one of missing values ratio, principal component analysis, probabilistic principal component analysis, matrix factorization techniques, compositional mixture models, and feature embedding techniques.
11. The method of Claim 10, wherein the determination of the initial set of microbiome features comprises application, with the second microbiome characterization module of the set of microbiome characterization modules, of the statistical test with the microorganism sequence dataset to determine the initial set of microbiome features,

wherein the statistical test comprises at least one of a t-test, a Kolmogorov-Smirnov test, and a regression model.

12. The method of Claim 10, wherein the application of the set of analytical techniques comprises, with the second microbiome characterization module of the set of microbiome characterization modules, application of a machine learning approach to determine relevance scores for the set of microbiome features, wherein the microorganism-related condition is generated based on the set of microbiome features and the relevance scores.
13. The method of Claim 10, wherein determining the characterization comprises determining a drug metabolism characterization associated with the microorganism-related condition based on the microorganism-related condition model, the sample from the user, and known associations between the set of microbiome features and drug metabolization.
14. The method of Claim 14, further comprising determining a microorganism dataset associated with the plurality of collection sites based on the set of site-diverse samples, wherein determining the overall disease propensity metric comprises:
determining at least one of a covariance metric and a correlation metric, based on the microorganism dataset, wherein the at least one of the covariance metric and the correlation metric is associated with the plurality of collection sites; and
determining the overall disease propensity metric for the user based on the set of site-wise disease propensity metrics and the at least one of the covariance metric and the correlation metric.
15. The method of Claim 9, wherein the microorganism-related condition model is generated based on, prior to the application of the set of analytical techniques to determine the microbiome features, filtering of the microorganism sequence dataset by at least one of: a) removing first sample data corresponding to first sample outliers of the set of samples, wherein the first sample outliers are determined by at least one of principal component analysis, a dimensionality reduction technique, and a multivariate methodology; b) removing second sample data corresponding to second sample outliers of the set of samples, wherein the second sample outliers are determined based on corresponding data quality for the set of microbiome features; and c) removing a microbiome feature from the set of microbiome features based on a sample number for the microbiome feature failing to satisfy a threshold sample number condition, wherein the sample number corresponds to a number of samples associated with high quality data for the microbiome feature.

16. The method of Claim 9, wherein determining the characterization of the microorganism-related condition for the user comprises determining a skin-related characterization of a photosensitivity-associated condition for the user based on a set of user microbiome features and the microorganism-related condition model, wherein the set of user microbiome features comprises features associated with at least one of: *Alloprevotella* (genus), *Prevotella* sp. WAL 2039G (species), *Corynebacterium mastitidis* (species), *Bacteroidaceae* (family), *Blautia* (genus), *Bacteroides* (genus), *Desulfovibrio* (genus), *Clostridium* (genus), *Bacteroides vulgatus* (species), *Faecalibacterium prausnitzii* (species), *Blautia faecis* (species), *Alistipes putredinis* (species), *Bacteroides* sp. AR20 (species), *Bacteroides* sp. AR29 (species), *Bacteroides acidifaciens* (species), *Dielma* (genus), *Slackia* (genus), *Eggerthella* (genus), *Adlercreutzia* (genus), *Paraprevotella* (genus), *Alistipes* (genus), *Holdemania* (genus), *Eisenbergiella* (genus), *Enterorhabdus* (genus), *Adlercreutzia equolifaciens* (species), *Phascolarctobacterium succinatutens* (species), *Roseburia inulinivorans* (species), *Phascolarctobacterium* sp. 377 (species), *Desulfovibrio piger* (species), *Eggerthella* sp. HGA1 (species), *Lactonifactor longoviformis* (species), *Alistipes* sp. HGB5 (species), *Holdemania filiformis* (species), *Collinsella intestinalis* (species), *Neisseria macacae* (species), *Clostridiaceae* (family), *Gemella sanguinis* (species), *Bacteroides fragilis* (species), *Enterobacteriaceae* (family), *Lachnospiraceae* (family), *Pasteurellaceae* (family), *Pasteurellales* (order), *Enterobacteriales* (order), *Sphingobacteriales* (order), *Haemophilus* (genus), *Leuconostoc* (genus), *Brevundimonas* (genus), *Prevotella oris* (species), *Odoribacter* (genus), *Capnocytophaga* (genus), *Flavobacterium* (genus), *Pseudomonas brenneri* (species), *Flavobacterium ceti* (species), *Brevundimonas* sp. FXJ8.080 (species), *Ruminococcaceae* (family), *Vibrionaceae* (family), *Flavobacteriaceae* (family), *Fusobacteriaceae* (family), *Porphyromonadaceae* (family), *Brevibacteriaceae* (family), *Rhodobacteraceae* (family), *Intrasporangiaceae* (family), *Bifidobacteriaceae* (family), *Sphingobacteriaceae* (family), *Caulobacteraceae* (family), *Campylobacteraceae* (family), *Bacteroidia* (class), *Fusobacteriia* (class), *Flavobacteriia* (class), *Bifidobacteriales* (order), *Neisseriales* (order), *Bacteroidales* (order), *Rhodobacterales* (order), *Flavobacteriales* (order), *Vibrionales* (order), *Fusobacteriales* (order), *Caulobacterales* (order), *Fusobacteria* (phylum), *Actinobaculum* (genus), *Varibaculum* (genus), *Fusicatenibacter* (genus), *Brevibacterium* (genus), *Faecalibacterium* (genus), *Campylobacter* (genus), *Actinobacillus* (genus), *Porphyromonas* (genus), *Fusobacterium* (genus), *Chryseobacterium* (genus), *Megasphaera* (genus), *Rothia* (genus), *Neisseria* (genus), *Lactobacillus* sp. BL302 (species), *Bacteroides plebeius* (species), *Corynebacterium ulcerans* (species), *Varibaculum cambriense* (species), *Blautia wexlerae* (species), *Staphylococcus* sp. WB18-16 (species), *Streptococcus* sp. oral taxon G63 (species), *Propionibacterium acnes* (species), *Anaerococcus* sp. 9401487 (species), *Haemophilus parainfluenzae* (species), *Staphylococcus epidermidis* (species), *Campylobacter ureolyticus* (species), *Janibacter* sp. M3-5 (species), *Prevotella timonensis* (species), *Peptoniphilus* sp. DNFoo840 (species), *Finegoldia* sp. S8 F7 (species), *Prevotella*

disiens (species), Porphyromonas catoniae (species), Fusobacterium periodonticum (species), Infectious Diseases (KEGG2), Poorly Characterized (KEGG2), Metabolic Diseases (KEGG2), Immune System Diseases (KEGG2), Cellular Processes and Signaling (KEGG2), Restriction enzyme (KEGG3), Nucleotide excision repair (KEGG3).

17. The method of Claim 9, wherein determining the characterization of the microorganism-related condition for the user comprises determining a skin-related characterization of a dry skin-associated condition for the user based on a set of user microbiome features and the microorganism-related condition model, wherein the set of user microbiome features comprises features associated with at least one of: Corynebacteriaceae (family), Bacilli (class), Lactobacillales (order), Actinomycetales (order), Firmicutes (phylum), Corynebacterium (genus), Dermabacteraceae (family), Lactobacillaceae (family), Propionibacteriaceae (family), Actinobacteria (class), Dermabacter (genus), Dialister (genus), Facklamia (genus), Lactobacillus (genus), Propionibacterium (genus), Corynebacterium ulcerans (species), Facklamia hominis (species), Corynebacterium sp. (species), Propionibacterium sp. MSP09A (species), Facklamia sp. 1440-97 (species), Staphylococcus sp. C9I2 (species), Anaerococcus sp. 9402080 (species), Corynebacterium glucuronolyticum (species), Dermabacter hominis (species), Enterobacteriaceae (family), Pseudomonadaceae (family), Staphylococcaceae (family), Gammaproteobacteria (class), Bacillales (order), Enterobacteriales (order), Bifidobacterium (genus), Pseudomonas (genus), Anaeroglobus (genus), Kluyvera (genus), Atopobium (genus), Staphylococcus (genus), Lactobacillus sp. BL302 (species), Corynebacterium mastitidis (species), Bifidobacterium longum (species), Anaeroglobus geminatus (species), Anaerococcus sp. S9 PR-16 (species), Prevotella timonensis (species), Kluyvera georgiana (species), Actinobaculum (genus), Finegoldia (genus), Cronobacter (genus), Acinetobacter sp. WB22-23 (species), Anaerococcus octavius (species), Finegoldia sp. S9 AA1-5 (species), Staphylococcus sp. C-D-MA2 (species), Peptoniphilus sp. 7-2 (species), Cronobacter sakazakii (species), Pasteurellaceae (family), Acidobacteriia (class), Sphingobacteriia (class), Sphingobacteriales (order), Acidobacteria (phylum), Porphyromonas (genus), Haemophilus (genus), Acinetobacter (genus), Anaerococcus sp. 8405254 (species), Sphingomonadaceae (family), Sphingomonadales (order), Kocuria (genus), Gemella (genus), Veillonella sp. CM60 (species), Lactobacillus sp. 7_1_47FAA (species), Gemella sp. 933-88 (species), Porphyromonas catoniae (species), Haemophilus parainfluenzae (species), Bacteroides sp. AR20 (species), Bacteroides vulgatus (species), Bacteroides sp. D22 (species), Dorea longicatena (species), Parabacteroides merdae (species), Bacteroides sp. AR29 (species), Dorea (genus), Collinsella (genus), Bacteroides (genus), Oscillospiraceae (family), Ruminococcaceae (family), Bacteroidaceae (family), Verrucomicrobiaceae (family), Coriobacteriaceae (family), Clostridiales (order), Bacteroidales (order), Verrucomicrobiales (order), Coriobacteriales (order), Thermoanaerobacterales

(order), Clostridia (class), Bacteroidia (class), Verrucomicrobiae (class), Verrucomicrobia (phylum), Bacteroidetes (phylum), Translation (KEGG2), Cellular Processes and Signaling (KEGG2), Amino Acid Metabolism (KEGG2), Cell Growth and Death (KEGG2), Replication and Repair (KEGG2), Metabolism of Other Amino Acids (KEGG2), Neurodegenerative Diseases (KEGG2), Metabolism of Cofactors and Vitamins (KEGG2), Transport and Catabolism (KEGG2), Endocrine System (KEGG2), Immune System Diseases (KEGG2), Excretory System (KEGG2), Enzyme Families (KEGG2), Membrane Transport (KEGG2), Carbohydrate Metabolism (KEGG2), Biosynthesis of Other Secondary Metabolites (KEGG2), Metabolism of Terpenoids and Polyketides (KEGG2), Infectious Diseases (KEGG2), Genetic Information Processing (KEGG2), Nervous System (KEGG2), Environmental Adaptation (KEGG2), Nucleotide Metabolism (KEGG2), Signaling Molecules and Interaction (KEGG2), Signal Transduction (KEGG2), Inorganic ion transport and metabolism (KEGG3), Chromosome (KEGG3), Cell cycle - Caulobacter (KEGG3), Ribosome Biogenesis (KEGG3), DNA replication proteins (KEGG3), Translation factors (KEGG3), Glycine, serine and threonine metabolism (KEGG3), Sulfur metabolism (KEGG3), Other ion-coupled transporters (KEGG3), Valine, leucine and isoleucine biosynthesis (KEGG3), Nitrogen metabolism (KEGG3), Peptidoglycan biosynthesis (KEGG3), Homologous recombination (KEGG3), Peroxisome (KEGG3), Sulfur relay system (KEGG3), Peptidases (KEGG3), Protein kinases (KEGG3), Mismatch repair (KEGG3), Xylene degradation (KEGG3), Ribosome (KEGG3), RNA polymerase (KEGG3), Tryptophan metabolism (KEGG3), Histidine metabolism (KEGG3), Vitamin metabolism (KEGG3), Cell motility and secretion (KEGG3), Pyrimidine metabolism (KEGG3), Cytoskeleton proteins (KEGG3), DNA replication (KEGG3), Amino sugar and nucleotide sugar metabolism (KEGG3), Folate biosynthesis (KEGG3), Carbon fixation in photosynthetic organisms (KEGG3), Phosphatidylinositol signaling system (KEGG3), Lysine degradation (KEGG3), Selenocompound metabolism (KEGG3), Fructose and mannose metabolism (KEGG3), Inositol phosphate metabolism (KEGG3), Protein folding and associated processing (KEGG3), PPAR signaling pathway (KEGG3), Lipid metabolism (KEGG3), Valine, leucine and isoleucine degradation (KEGG3), Glyoxylate and dicarboxylate metabolism (KEGG3), Arginine and proline metabolism (KEGG3), Limonene and pinene degradation (KEGG3), D-Alanine metabolism (KEGG3), Porphyrin and chlorophyll metabolism (KEGG3), C5-Branched dibasic acid metabolism (KEGG3), Chaperones and folding catalysts (KEGG3), Fatty acid metabolism (KEGG3), Glutathione metabolism (KEGG3), Pentose phosphate pathway (KEGG3), Phosphotransferase system (PTS) (KEGG3), Pantothenate and CoA biosynthesis (KEGG3), Proximal tubule bicarbonate reclamation (KEGG3), Galactose metabolism (KEGG3), Starch and sucrose metabolism (KEGG3), Primary immunodeficiency (KEGG3), Cysteine and methionine metabolism (KEGG3), Ubiquinone and other terpenoid-quinone biosynthesis (KEGG3), DNA repair and recombination proteins (KEGG3), Tyrosine metabolism (KEGG3), Phenylalanine, tyrosine and tryptophan biosynthesis (KEGG3), Aminoacyl-

tRNA biosynthesis (KEGG3), Alanine, aspartate and glutamate metabolism (KEGG3), Photosynthesis (KEGG3), Other transporters (KEGG3), Butanoate metabolism (KEGG3), Bacterial secretion system (KEGG3), Glycerophospholipid metabolism (KEGG3), Oxidative phosphorylation (KEGG3), Type I diabetes mellitus (KEGG3), Glycolysis / Gluconeogenesis (KEGG3), Photosynthesis proteins (KEGG3), Transporters (KEGG3), Terpenoid backbone biosynthesis (KEGG3), Biosynthesis of unsaturated fatty acids (KEGG3), Signal transduction mechanisms (KEGG3), Synthesis and degradation of ketone bodies (KEGG3), Nucleotide excision repair (KEGG3), Secretion system (KEGG3), Alzheimer's disease (KEGG3), Zeatin biosynthesis (KEGG3), Type II diabetes mellitus (KEGG3), D-Glutamine and D-glutamate metabolism (KEGG3), Taurine and hypotaurine metabolism (KEGG3), Glutamatergic synapse (KEGG3), Plant-pathogen interaction (KEGG3), Vitamin B6 metabolism (KEGG3), Citrate cycle (TCA cycle) (KEGG3), Ethylbenzene degradation (KEGG3), Base excision repair (KEGG3), Replication, recombination and repair proteins (KEGG3), Ribosome biogenesis in eukaryotes (KEGG3), Aminobenzoate degradation (KEGG3), Bacterial motility proteins (KEGG3), Biosynthesis of ansamycins (KEGG3), Ion channels (KEGG3), Metabolism (KEGG2), Poorly Characterized (KEGG2), Biosynthesis and biodegradation of secondary metabolites (KEGG3), Lipoic acid metabolism (KEGG3), Amino acid related enzymes (KEGG3), Translation proteins (KEGG3), Ascorbate and aldarate metabolism (KEGG3), Thiamine metabolism (KEGG3), Function unknown (KEGG3), Glycosaminoglycan degradation (KEGG3), Others (KEGG3), Pentose and glucuronate interconversions (KEGG3), Biotin metabolism (KEGG3), Phenylalanine metabolism (KEGG3), Glycosphingolipid biosynthesis - ganglio series (KEGG3), Pores ion channels (KEGG3), Membrane and intracellular structural molecules (KEGG3), Purine metabolism (KEGG3), One carbon pool by folate (KEGG3), Phosphonate and phosphinate metabolism (KEGG3), Lysosome (KEGG3), Drug metabolism - other enzymes (KEGG3), Penicillin and cephalosporin biosynthesis (KEGG3), Huntington's disease (KEGG3), Nicotinate and nicotinamide metabolism (KEGG3), Drug metabolism - cytochrome P450 (KEGG3), Lipopolysaccharide biosynthesis proteins (KEGG3), Metabolism of xenobiotics by cytochrome P450 (KEGG3), Tuberculosis (KEGG3), and Polycyclic aromatic hydrocarbon degradation (KEGG3).

18. The method of Claim 9, wherein determining the characterization of the microorganism-related condition for the user comprises determining a skin-related characterization of a scalp-related condition for the user based on a set of user microbiome features and the microorganism-related condition model, wherein the set of user microbiome features comprises features associated with at least one of: Actinobacteria (class), Lactobacillales (order), Actinomycetales (order), Firmicutes (phylum), Dermabacteraceae (family), Lactobacillaceae (family), Propionibacteriaceae (family), Corynebacteriaceae (family), Lactobacillus (genus), Corynebacterium (genus), Propionibacterium (genus), Dermabacter (genus), Eremococcus (genus),

Corynebacterium freiburgense (species), *Eremococcus coleocola* (species), *Corynebacterium* sp. (species), *Staphylococcus* sp. C9I2 (species), *Anaerococcus* sp. 8405254 (species), *Corynebacterium glucuronolyticum* (species), *Dermabacter hominis* (species), *Coriobacteriaceae* (family), *Enterobacteriaceae* (family), *Staphylococcaceae* (family), *Enterobacteriales* (order), *Bacillales* (order), *Bifidobacterium* (genus), *Staphylococcus* (genus), *Atopobium* (genus), *Megasphaera* (genus), *Corynebacterium mastitidis* (species), *Streptococcus* sp. BS35a (species), *Finegoldia magna* (species), *Staphylococcus aureus* (species), *Haemophilus influenzae* (species), *Corynebacterium* sp. NML 97-0186 (species), *Streptococcus* sp. oral taxon G59 (species), *Dorea* (genus), *Roseburia* sp. 11SE39 (species), *Dorea longicatena* (species), *Prevotellaceae* (family), *Veillonellaceae* (family), *Oscillospiraceae* (family), *Negativicutes* class, *Selenomonadales* (order), *Finegoldia* (genus), *Oscillospira* (genus), *Intestinimonas* (genus), *Flavonifractor* (genus), *Prevotella* (genus), *Moryella* (genus), *Catenibacterium mitsuokai* (species), *Collinsella aerofaciens* (species), *Peptoniphilus* sp. 2002-2300004 (species), *Corynebacterium canis* (species), *Finegoldia* sp. S9 AA1-5 (species), *Prevotella buccalis* (species), *Dialister invisus* (species), *Moraxella* (genus), *Neisseria* (genus), *Neisseria mucosa* (species), *Rikenellaceae* (family), Metabolism of Cofactors and Vitamins (KEGG2), Enzyme Families (KEGG2), Lipid Metabolism (KEGG2), Immune System Diseases (KEGG2), Glycolysis / Gluconeogenesis (KEGG3), Primary immunodeficiency (KEGG3), Pyruvate metabolism (KEGG3), Transport and Catabolism (KEGG2), Neurodegenerative Diseases (KEGG2), Endocrine System (KEGG2), Amino Acid Metabolism (KEGG2), Cellular Processes and Signaling (KEGG2), Signaling Molecules and Interaction (KEGG2), Metabolism of Other Amino Acids (KEGG2), Replication and Repair (KEGG2), Translation (KEGG2), Cell Growth and Death (KEGG2), Membrane Transport (KEGG2), Biosynthesis of Other Secondary Metabolites (KEGG2), Metabolism of Terpenoids and Polyketides (KEGG2), Inorganic ion transport and metabolism (KEGG3), Vitamin metabolism (KEGG3), Valine, leucine and isoleucine biosynthesis (KEGG3), Peroxisome (KEGG3), Ribosome Biogenesis (KEGG3), Selenocompound metabolism (KEGG3), Histidine metabolism (KEGG3), Chromosome (KEGG3), Sulfur metabolism (KEGG3), PPAR signaling pathway (KEGG3), Porphyrin and chlorophyll metabolism (KEGG3), Phosphatidylinositol signaling system (KEGG3), Inositol phosphate metabolism (KEGG3), Sulfur relay system (KEGG3), Glycine, serine and threonine metabolism (KEGG3), DNA replication proteins (KEGG3), Pantothenate and CoA biosynthesis (KEGG3), Translation factors (KEGG3), Protein folding and associated processing (KEGG3), Type II diabetes mellitus (KEGG3), Protein kinases (KEGG3), Folate biosynthesis (KEGG3), Lysine degradation (KEGG3), RNA polymerase (KEGG3), D-Alanine metabolism (KEGG3), Carbon fixation in photosynthetic organisms (KEGG3), Nitrogen metabolism (KEGG3), Glycerophospholipid metabolism (KEGG3), Biosynthesis of ansamycins (KEGG3), Valine, leucine and isoleucine degradation (KEGG3), Cytoskeleton proteins (KEGG3), Peptidases (KEGG3), Fatty acid

metabolism (KEGG3), Cell cycle - *Caulobacter* (KEGG3), Phosphotransferase system (PTS) (KEGG3), Pyrimidine metabolism (KEGG3), Alzheimer's disease (KEGG3), Butanoate metabolism (KEGG3), Tryptophan metabolism (KEGG3), Signal transduction mechanisms (KEGG3), Pentose phosphate pathway (KEGG3), Other ion-coupled transporters (KEGG3), Homologous recombination (KEGG3), Replication, recombination and repair proteins (KEGG3), Xylene degradation (KEGG3), Mismatch repair (KEGG3), Glyoxylate and dicarboxylate metabolism (KEGG3), Arginine and proline metabolism (KEGG3), Peptidoglycan biosynthesis (KEGG3), Chaperones and folding catalysts (KEGG3), Type I diabetes mellitus (KEGG3), DNA replication (KEGG3), Bacterial secretion system (KEGG3), Tyrosine metabolism (KEGG3), Citrate cycle (TCA cycle) (KEGG3), Amino sugar and nucleotide sugar metabolism (KEGG3), Ribosome (KEGG3), Limonene and pinene degradation (KEGG3), Cell motility and secretion (KEGG3), Taurine and hypotaurine metabolism (KEGG3), Oxidative phosphorylation (KEGG3), Fructose and mannose metabolism (KEGG3), Vitamin B6 metabolism (KEGG3), Ion channels (KEGG3), Synthesis and degradation of ketone bodies (KEGG3), Other transporters (KEGG3), Galactose metabolism (KEGG3), Polycyclic aromatic hydrocarbon degradation (KEGG3), Transporters (KEGG3), DNA repair and recombination proteins (KEGG3), Starch and sucrose metabolism (KEGG3), Alanine, aspartate and glutamate metabolism (KEGG3), Ribosome biogenesis in eukaryotes (KEGG3), Secretion system (KEGG3), Biosynthesis of unsaturated fatty acids (KEGG3), Cysteine and methionine metabolism (KEGG3), Base excision repair (KEGG3), Aminobenzoate degradation (KEGG3), Photosynthesis (KEGG3), Photosynthesis proteins (KEGG3), Pores ion channels (KEGG3), Lipid biosynthesis proteins (KEGG3), and D-Glutamine and D-glutamate metabolism (KEGG3).

19. The method of Claim 9, wherein the microorganism-related condition comprises a skin-related condition, wherein the method further comprises promoting a probiotic therapy to the user for the skin-related condition based on the characterization, and wherein the probiotic therapy is associated with microorganisms associated with any one of the following: *Corynebacterium ulcerans*, *Facklamia hominis*, *Corynebacterium* sp., *Propionibacterium* sp. MSP09A, *Facklamia* sp. 1440-97, *Staphylococcus* sp. C9I2, *Anaerococcus* sp. 9402080, *Corynebacterium glucuronolyticum*, *Dermabacter hominis*, *Lactobacillus* sp. BL302, *Corynebacterium mastitidis*, *Bifidobacterium longum*, *Anaeroglobus geminatus*, *Anaerococcus* sp. S9 PR-16, *Prevotella timonensis*, *Kluyvera georgiana*, *Acinetobacter* sp. WB22-23, *Anaerococcus octavius*, *Finegoldia* sp. S9 AA1-5, *Staphylococcus* sp. C-D-MA2, *Peptoniphilus* sp. 7-2, *Cronobacter sakazakii*, *Anaerococcus* sp. 8405254, *Veillonella* sp. CM60, *Lactobacillus* sp. 7_1_47FAA, *Gemella* sp. 933-88, *Porphyromonas catoniae*, *Haemophilus parainfluenzae*, *Bacteroides* sp. AR20, *Bacteroides vulgatus*, *Bacteroides* sp. D22, *Dorea longicatena*, *Parabacteroides merdae*, *Bacteroides* sp. AR29, *Prevotella* sp. WAL 2039G, *Faecalibacterium prausnitzii*, *Blautia faecis*, *Alistipes putredinis*,

Bacteroides acidifaciens, *Adlercreutzia equolifaciens*, *Phascolarctobacterium succinatutens*, *Roseburia inulinivorans*, *Phascolarctobacterium* sp. 377, *Desulfovibrio piger*, *Eggerthella* sp. HGA1, *Lactonifactor longoviformis*, *Alistipes* sp. HGB5, *Holdemania filiformis*, *Collinsella intestinalis*, *Neisseria macacae*, *Gemella sanguinis*, *Bacteroides fragilis*, *Prevotella oris*, *Pseudomonas brenneri*, *Flavobacterium ceti*, *Brevundimonas* sp. FXJ8.080, *Bacteroides plebeius*, *Varibaculum cambriense*, *Blautia wexlerae*, *Staphylococcus* sp. WB18-16, *Streptococcus* sp. oral taxon G63, *Propionibacterium acnes*, *Anaerococcus* sp. 9401487, *Staphylococcus epidermidis*, *Campylobacter ureolyticus*, *Janibacter* sp. M3-5, *Peptoniphilus* sp. DNFO0840, *Finegoldia* sp. S8 F7, *Prevotella disiens*, *Fusobacterium periodonticum*, *Corynebacterium freiburgense*, *Eremococcus coleocola*, *Streptococcus* sp. BS35a, *Finegoldia magna*, *Staphylococcus aureus*, *Haemophilus influenzae*, *Corynebacterium* sp. NML 97-0186, *Streptococcus* sp. oral taxon G59, *Roseburia* sp. 11SE39, *Catenibacterium mitsuokai*, *Collinsella aerofaciens*, *Peptoniphilus* sp. 2002-2300004, *Corynebacterium canis*, *Prevotella buccalis*, *Dialister invisus*, and *Neisseria mucosa*.