(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
26 January 2017 (26.01.2017)

WIPO | PCT

(10) International Publication Number
WO 2017/013217 A2

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

(21) International Application Number:
PCT/EP2016/067437

(22) International Filing Date:
21 July 2016 (21.07.2016)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
PCT/EP2015/066711 22 July 2015 (22.07.2015) EP

(71) Applicant: **CURETIS GMBH** [DE/DE]; Max-Eyth-Str. 42, 71088 Holzgerlingen (DE).

(72) Inventors: **KELLER, Andreas**; Albert Schweitzer Str. 6, 66346 Püttlingen (DE). **SCHMOLKE, Susanne**; Gebbertstraße 37, 91052 Erlangen (DE). **STÄHLER, Cord Friedrich**; Muldweg 8, 69493 Hirschberg an der Bergstraße (DE). **BACKES, Christina**; Am Geisenberg 45, 66125 Saarbrücken (DE). **GALATA, Valentina**; Auf der Schlecht 4, 66123 Saarbrücken (DE).

(74) Agent: **SCHNAPPAUF, Georg**; ZSP Patentanwälte PartG mbB, Radlkoferstrasse 2, 81373 München (DE).

(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, 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, 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.

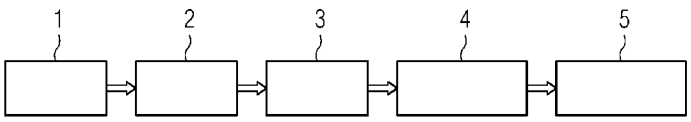
(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, 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: GENETIC TESTING FOR PREDICTING RESISTANCE OF SALMONELLA SPECIES AGAINST ANTIMICROBIAL AGENTS



(57) Abstract: The invention relates to a method of determining an infection of a patient with Salmonella species potentially resistant to antimicrobial drug treatment, a method of selecting a treatment of a patient suffering from an antibiotic resistant Salmonella infection, and a method of determining an antibiotic resistance profile for bacterial microorganisms of Salmonella species, as well as computer program products used in these methods. In an exemplary method, a sample 1 is used for molecular testing 2, and then a molecular fingerprint 3 is taken. The result is then compared to a reference library 4, and the result 5 is reported.

WO 2017/013217 A2

Genetic testing for predicting resistance of Salmonella species against antimicrobial agents

The present invention relates to a method of determining an infection of a patient with Salmonella species potentially resistant to antimicrobial drug treatment, a method of selecting a treatment of a patient suffering from an infection with a potentially resistant Salmonella strain, and a method of determining an antimicrobial drug, e.g. antibiotic, resistance profile for bacterial microorganisms of Salmonella species, as well as computer program products used in these methods.

Antibiotic resistance is a form of drug resistance whereby a sub-population of a microorganism, e.g. a strain of a bacterial species, can survive and multiply despite exposure to an antibiotic drug. It is a serious and health concern for the individual patient as well as a major public health issue. Timely treatment of a bacterial infection requires the analysis of clinical isolates obtained from patients with regard to antibiotic resistance, in order to select an efficacious therapy. Generally, for this purpose an association of the identified resistance with a certain microorganism (i.e. ID) is necessary.

Antibacterial drug resistance (ADR) represents a major health burden. According to the World Health Organization's antimicrobial resistance global report on surveillance, ADR leads to 25,000 deaths per year in Europe and 23,000 deaths per year in the US. In Europe, 2.5 million extra hospital days lead to societal cost of 1.5 billion euro. In the US, the direct cost of 2 million illnesses leads to 20 billion dollar direct cost. The overall cost is estimated to be substantially higher, reducing the gross domestic product (GDP) by up to 1.6%.

Salmonellae are Gram-negative, flagellated, facultatively anaerobic bacilli belonging to the family of Enterobacteriaceae. Determination of antigenic structure permits one to identify the organisms clinically and assign them to one of nine serogroups (A-I), each containing many serovars (= serotypes).

Salmonellae are ubiquitous human and animal pathogens, and salmonellosis is common throughout the world. Salmonellosis in humans ranges clinically from the common Salmonella gastroenteritis (diarrhea, abdominal cramps, and fever) to enteric fevers (including typhoid fever) which are life-threatening febrile systemic illness requiring prompt antibiotic therapy. Particular serovars show a strong propensity to produce a particular syndrome (S typhi, S paratyphi-A, and S schottmuelleri produce enteric fever; S choleraesuis produces septicemia or focal infections; S typhimurium and S enteritidis produce gastroenteritis); however, on occasion, any serotype can produce any of the syndromes.

In general, more serious infections occur in infants, in adults over the age of 50, and in subjects with debilitating illnesses.

In a recent report by CDC (Centers for Disease Control and Prevention), titled *Antibiotic Resistance Threats in the United States*, 2013, drug-resistant, nontyphoidal Salmonella was listed among bacteria that pose a serious threat level. Nontyphoidal Salmonella causes approximately 1.2 million illnesses, 23,000 hospitalizations, and 450 deaths each year in the United States. Direct medical costs are estimated to be $365 million annually. Of concern, surveillance data reveal that an increasing proportion of nontyphoidal Salmonella are resistant to ceftriaxone or ciprofloxacin, drugs representing classes of antibiotics commonly used to treat severe salmonellosis. Taking into account all of the classes of antibiotics for which testing is done at CDC, about 5% of

nontyphoidal Salmonella tested by CDC are resistant to antibiotics in 5 or more classes.

In general the mechanisms for resistance of bacteria against antimicrobial treatments rely to a very substantial part on the organism's genetics. The respective genes or molecular mechanisms are either encoded in the genome of the bacteria or on plasmids that can be interchanged between different bacteria. The most common resistance mechanisms include:

1) Efflux pumps are high-affinity reverse transport systems located in the membrane that transports the antibiotic out of the cell, e.g. resistance to tetracycline.
2) Specific enzymes modify the antibiotic in a way that it loses its activity. In the case of streptomycin, the antibiotic is chemically modified so that it will no longer bind to the ribosome to block protein synthesis.
3) An enzyme is produced that degrades the antibiotic, thereby inactivating it. For example, the penicillinases are a group of beta-lactamase enzymes that cleave the beta lactam ring of the penicillin molecule.

In addition, some pathogens show natural resistance against drugs. For example, an organism can lack a transport system for an antibiotic or the target of the antibiotic molecule is not present in the organism.

Pathogens that are in principle susceptible to drugs can become resistant by modification of existing genetic material (e.g. spontaneous mutations for antibiotic resistance, happening in a frequency of one in about 100 mio bacteria in an infection) or the acquisition of new genetic material from another source. One example is horizontal gene transfer, a process where genetic material contained in small packets of DNA can be transferred between individual bacteria of the same species or even between different species. Horizontal gene transfer may happen by transduction, transformation or conjugation.

Generally, testing for susceptibility/resistance to antimicrobial agents is performed by culturing organisms in different concentration of these agents.

In brief, agar plates are inoculated with patient sample (e.g. urine, sputum, blood, stool) overnight. On the next day individual colonies are used for identification of organisms, either by culturing or using mass spectroscopy. Based on the identity of organisms new plates containing increasing concentration of drugs used for the treatment of these organisms are inoculated and grown for additional 12 - 24 hours. The lowest drug concentration which inhibits growth (minimal inhibitory concentration - MIC) is used to determine susceptibility/resistance for tested drugs. The process takes at least 2 to 3 working days during which the patient is treated empirically. A significant reduction of time-to-result is needed especially in patients with life-threatening disease and to overcome the widespread misuse of antibiotics.

Recent developments include PCR based test kits for fast bacterial identification (e.g. Biomerieux Biofire Tests, Curetis Unyvero Tests). With these test the detection of selected resistance loci is possible for a very limited number of drugs, but no correlation to culture based AST is given. Mass spectroscopy is increasingly used for identification of pathogens in clinical samples (e.g. Bruker Biotyper), and research is ongoing to establish methods for the detection of susceptibility/resistance against antibiotics.

For some drugs such it is known that at least two targets are addressed, e.g. in case of Ciprofloxacin (drug bank ID 00537; http://www.drugbank.ca/drugs/DB00537) targets include DNA Topoisomerase IV, DNA Topoisomerase II and DNA Gyrase. It can be expected that this is also the case for other drugs although the respective secondary targets have not been identified yet. In case of a common regulation, both relevant ge-

netic sites would naturally show a co-correlation or redundancy.

It is known that drug resistance can be associated with genetic polymorphisms. This holds for viruses, where resistance testing is established clinical practice (e.g. HIV genotyping). More recently, it has been shown that resistance has also genetic causes in bacteria and even higher organisms, such as humans where tumors resistance against certain cytostatic agents can be linked to genomic mutations.

Wozniak et al. (BMC Genomics 2012, 13(Suppl 7):S23) disclose genetic determinants of drug resistance in Staphylococcus aureus based on genotype and phenotype data. Stoesser et al. disclose prediction of antimicrobial susceptibilities for Escherichia coli and Klebsiella pneumoniae isolates using whole genomic sequence data (J Antimicrob Chemother 2013; 68: 2234-2244).

Chewapreecha et al (Chewapreecha et al (2014) Comprehensive Identification of single nucleotid polymorphisms associated with beta-lactam resistance within pneumococcal mosaic genes. PLoS Genet 10(8): e1004547) used a comparable approach to identify mutations in gram-positive Streptococcus Pneumonia.

The fast and accurate detection of infections with Salmonella species and the prediction of response to anti-microbial therapy represent a high unmet clinical need.

This need is addressed by the present invention.

Summary of the Invention

The present inventors addressed this need by carrying out whole genome sequencing of a large cohort of Salmonella clinical isolates and comparing the genetic mutation profile to

classical culture based antimicrobial susceptibility testing with the goal to develop a test which can be used to detect bacterial susceptibility/resistance against antimicrobial drugs using molecular testing.

The inventors performed extensive studies on the genome of bacteria of Salmonella species either susceptible or resistant to antimicrobial, e.g. antibiotic, drugs. Based on this information, it is now possible to provide a detailed analysis on the resistance pattern of Salmonella strains based on individual genes or mutations on a nucleotide level. This analysis involves the identification of a resistance against individual antimicrobial, e.g. antibiotic, drugs as well as clusters of them. This allows not only for the determination of a resistance to a single antimicrobial, e.g. antibiotic, drug, but also to groups of antimicrobial drugs, e.g. antibiotics such as lactam or quinolone antibiotics, or even to all relevant antibiotic drugs.

Therefore, the present invention will considerably facilitate the selection of an appropriate antimicrobial, e.g. antibiotic, drug for the treatment of a Salmonella infection in a patient and thus will largely improve the quality of diagnosis and treatment.

According to a first aspect, the present invention discloses a diagnostic method of determining an infection of a patient with Salmonella species potentially resistant to antimicrobial drug treatment, which can be also described as a method of determining an antimicrobial drug, e.g. antibiotic, resistant Salmonella infection of a patient, comprising the steps of:

a) obtaining or providing a sample containing or suspected of containing at least one Salmonella species from the patient;

b) determining the presence of at least one mutation in at least two genes from the group of genes listed in Table 1 or

Table 2 below, wherein the presence of said at least two mutations is indicative of an infection with an antimicrobial drug resistant, e.g. antibiotic resistant, Salmonella strain in said patient.

An infection of a patient with Salmonella species potentially resistant to antimicrobial drug treatment herein means an infection of a patient with Salmonella species wherein it is unclear if the Salmonella species is susceptible to treatment with a specific antimicrobial drug or if it is resistant to the antimicrobial drug.

In step b) above, as well as corresponding steps, at least one mutation in at least two genes is determined, so that in total at least two mutations are determined, wherein the two mutations are in different genes.

Table 1: List of genes

recN	hemH	UMN798_3428	metE	yijD
UMN798_4831	UMN798_1939	copS	UMN798_0628	UMN798_4878
leuB	recF	emrA	glyQ	dcp
thiH	UMN798_1612	UMN798_2909	UMN798_1680	UMN798_4073
yjbC	nadB	UMN798_3160	hutU	envC
UMN798_3889	UMN798_1629	bcfB	degQ	UMN798_1331
trg	uvrC	polB	hpcD	UMN798_1628
UMN798_1701	glgS	plsB	yjcC	feoB
misL	dxr	hemF	rnfG	yhjB
UMN798_1163	UMN798_0394	alkA	nhaA	lspA

Table 2: List of genes

recN	hemH	UMN798_3428	metE	yijD
UMN798_4831	UMN798_1939	copS	UMN798_0628	UMN798_4878
leuB	recF	emrA	glyQ	dcp
thiH	UMN798_1612	UMN798_2909	UMN798_1680	UMN798_4073
yjbC	nadB	UMN798_3160	hutU	envC
UMN798_3889	UMN798_1629	bcfB	degQ	UMN798_1331
trg	uvrC	polB	hpcD	UMN798_1628
UMN798_1701	glgS	plsB	yjcC	feoB
misL	dxr	hemF	rnfG	yhjB
UMN798_1163	UMN798_0394	alkA	nhaA	lspA

According to a second aspect, the present invention relates to a method of selecting a treatment of a patient suffering from an infection with a potentially resistant Salmonella strain, e.g. from an antimicrobial drug, e.g. antibiotic, resistant Salmonella infection, comprising the steps of:

a) obtaining or providing a sample containing or suspected of containing at least one Salmonella species from the patient;

b) determining the presence of at least one mutation in at least two genes from the group of genes listed in Table 1 or Table 2 above, wherein the presence of said at least two mutations is indicative of a resistance to one or more antimicrobial, e.g. antibiotic, drugs;

c) identifying said at least one or more antimicrobial, e.g. antibiotic, drugs; and

d) selecting one or more antimicrobial, e.g. antibiotic, drugs different from the ones identified in step c) and being suitable for the treatment of a Salmonella infection.

A third aspect of the present invention relates to a method of determining an antimicrobial drug, e.g. antibiotic, resistance profile for bacterial microorganisms of Salmonella species, comprising:

obtaining or providing a first data set of gene sequences of a plurality of clinical isolates of Salmonella species;
providing a second data set of antimicrobial drug, e.g. antibiotic, resistance of the plurality of clinical isolates of Salmonella species;
aligning the gene sequences of the first data set to at least one, preferably one, reference genome of Salmonella, and/or assembling the gene sequence of the first data set, at least in part;
analyzing the gene sequences of the first data set for genetic variants to obtain a third data set of genetic variants;
correlating the third data set with the second data set and statistically analyzing the correlation; and
determining the genetic sites in the genome of Salmonella associated with antimicrobial drug, e.g. antibiotic, resistance.

In addition, the present invention relates in a fourth aspect to a method of determining an antimicrobial drug, e.g. antibiotic, resistance profile for a bacterial microorganism belonging to the species Salmonella comprising the steps of
a) obtaining or providing a sample containing or suspected of containing the bacterial microorganism;
b) determining the presence of a mutation in at least one gene of the bacterial microorganism as determined by the method according to the third aspect of the present invention;
wherein the presence of a mutation is indicative of a resistance to an antimicrobial, e.g. antibiotic, drug.

Furthermore, the present invention discloses in a fifth aspect a diagnostic method of determining an infection of a patient with Salmonella species potentially resistant to anti-

microbial drug treatment, which can, like in the first aspect, also be described as method of determining an antimicrobial drug, e.g. antibiotic, resistant Salmonella infection of a patient, comprising the steps of:

a) obtaining or providing a sample containing or suspected of containing a bacterial microorganism belonging to the species Salmonella from the patient;

b) determining the presence of at least one mutation in at least one gene of the bacterial microorganism belonging to the species Salmonella as determined by the method according to the third aspect of the present invention, wherein the presence of said at least one mutation is indicative of an antimicrobial drug, e.g. antibiotic, resistant Salmonella infection in said patient.

Also disclosed is in a sixth aspect a method of selecting a treatment of a patient suffering from an infection with a potentially resistant Salmonella strain, e.g. from an antimicrobial drug, e.g. antibiotic, resistant Salmonella infection, comprising the steps of:

a) obtaining or providing a sample containing or suspected of containing a bacterial microorganism belonging to the species Salmonella from the patient;

b) determining the presence of at least one mutation in at least one gene of the bacterial microorganism belonging to the species Salmonella as determined by the method according to the third aspect of the present invention, wherein the presence of said at least one mutation is indicative of a resistance to one or more antimicrobial, e.g. antibiotic, drugs;

c) identifying said at least one or more antimicrobial, e.g. antibiotic, drugs; and

d) selecting one or more antimicrobial, e.g. antibiotic, drugs different from the ones identified in step c) and being suitable for the treatment of a Salmonella infection.

A seventh aspect of the present invention relates to a method of acquiring, respectively determining, an antimicrobial drug, e.g. antibiotic, resistance profile for a bacterial microorganism of Salmonella species, comprising:
obtaining or providing a first data set of gene sequences of a clinical isolate of Salmonella species;
providing a second data set of antimicrobial drug, e.g. antibiotic, resistance of a plurality of clinical isolates of Salmonella species;
aligning the gene sequences of the first data set to at least one, preferably one, reference genome of Salmonella, and/or
assembling the gene sequence of the first data set, at least in part;
analyzing the gene sequences of the first data set for genetic variants to obtain a third data set of genetic variants of the first data set;
correlating the third data set with the second data set and statistically analyzing the correlation; and
determining the genetic sites in the genome of Salmonella of the first data set associated with antimicrobial drug, e.g. antibiotic, resistance.

According to an eighth aspect, the present invention discloses a computer program product comprising executable instructions which, when executed, perform a method according to the third, fourth, fifth, sixth or seventh aspect of the present invention.
Further aspects and embodiments of the invention are disclosed in the dependent claims and can be taken from the fol-

lowing description, figures and examples, without being limited thereto.

Figures

The enclosed drawings should illustrate embodiments of the present invention and convey a further understanding thereof. In connection with the description they serve as explanation of concepts and principles of the invention. Other embodiments and many of the stated advantages can be derived in relation to the drawings. The elements of the drawings are not necessarily to scale towards each other. Identical, functionally equivalent and acting equal features and components are denoted in the figures of the drawings with the same reference numbers, unless noted otherwise.

Fig. 1 shows schematically a read-out concept for a diagnostic test according to a method of the present invention.

Detailed description of the present invention

Definitions

Unless defined otherwise, technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs.

An "antimicrobial drug" in the present invention refers to a group of drugs that includes antibiotics, antifungals, antiprotozoals, and antivirals. According to certain embodiments, the antimicrobial drug is an antibiotic.

The term "nucleic acid molecule" refers to a polynucleotide molecule having a defined sequence. It comprises DNA molecules, RNA molecules, nucleotide analog molecules and combinations and derivatives thereof, such as DNA molecules or RNA molecules with incorporated nucleotide analogs or cDNA.

The term "nucleic acid sequence information" relates to information which can be derived from the sequence of a nucleic acid molecule, such as the sequence itself or a variation in the sequence as compared to a reference sequence.

The term "mutation" relates to a variation in the sequence as compared to a reference sequence. Such a reference sequence can be a sequence determined in a predominant wild type organism or a reference organism, e.g. a defined and known bacterial strain or substrain. A mutation is for example a deletion of one or multiple nucleotides, an insertion of one or multiple nucleotides, or substitution of one or multiple nucleotides, duplication of one or a sequence of multiple nucleotides, translocation of one or a sequence of multiple nucleotides, and, in particular, a single nucleotide polymorphism (SNP).

In the context of the present invention a "sample" is a sample which comprises at least one nucleic acid molecule from a bacterial microorganism. Examples for samples are: cells, tissue, body fluids, biopsy specimens, blood, urine, saliva, sputum, plasma, serum, cell culture supernatant, swab sample and others. According to certain embodiments, the sample is a patient sample (clinical isolate).

New and highly efficient methods of sequencing nucleic acids referred to as next generation sequencing have opened the possibility of large scale genomic analysis. The term "next generation sequencing" or "high throughput sequencing" refers to high-throughput sequencing technologies that parallelize the sequencing process, producing thousands or millions of sequences at once. Examples include Massively Parallel Signature Sequencing (MPSS), Polony sequencing, 454 pyrosequencing, Illumina (Solexa) sequencing, SOLiD sequencing, Ion semiconductor sequencing, DNA nanoball sequencing, Helioscope(TM) single molecule sequencing, Single Molecule SMRT(TM) sequencing, Single Molecule real time (RNAP) sequencing, Nanopore DNA sequencing, Sequencing By Hybridization, Amplicon Sequencing, GnuBio.

Within the present description the term "microorganism" comprises the term microbe. The type of microorganism is not particularly restricted, unless noted otherwise or obvious, and, for example, comprises bacteria, viruses, fungi, microscopic algae und protozoa, as well as combinations thereof. According to certain aspects, it refers to one or more Salmonella species, particularly Salmonella_choleraesuis, Salmonella_dublin, Salmonella_enterica_ssp_arizonae, Salmonella_enterica_ssp_diarizoniae, Salmonella_enteritidis, Salmonella_gallinarum, Salmonella_Group_A, Salmonella_Group_B, Salmonella_Group_C, Salmonella_Group_D, Salmonella_heidelberg, Salmonella_miami, Salmonella_newport, Salmonella_panama, Salmonella_parahaemolyticus_A, Salmonella_paratyphi_A, Salmonella_paratyphi_B, Salmonella_pullorum, Salmonella_senfienberg, Salmonella_species, Salmonella_species_Lac_--,_ONPG_+, Salmonella_species_Lac_+,_ONPG_+, Salmonella_subgenus_I, Salmonella_subgenus_II, Salmonel-

la_subgenus_IV, Salmonella_subgroup_I_Suc+, Salmonella_tennessee, and/or Salmonella_typhi.

A reference to a microorganism or microorganisms in the present description comprises a reference to one microorganism as well a plurality of microorganisms, e.g. two, three, four, five, six or more microorganisms.

A vertebrate within the present invention refers to animals having a vertebrae, which includes mammals - including humans, birds, reptiles, amphibians and fishes. The present invention thus is not only suitable for human medicine, but also for veterinary medicine.

According to certain embodiments, the patient in the present methods is a vertebrate, more preferably a mammal and most preferred a human patient.

Before the invention is described in exemplary detail, it is to be understood that this invention is not limited to the particular component parts of the process steps of the methods described herein as such methods may vary. It is also to be understood that the terminology used herein is for purposes of describing particular embodiments only, and is not intended to be limiting. It must be noted that, as used in the specification and the appended claims, the singular forms "a," "an" and "the" include singular and/or plural referents unless the context clearly dictates otherwise. For example, the term "a" as used herein can be understood as one single entity or in the meaning of "one or more" entities. It is also to be understood that plural forms include singular and/or plural referents unless the context clearly dictates otherwise. It is moreover to be understood that, in case parameter

ranges are given which are delimited by numeric values, the ranges are deemed to include these limitation values.

Regarding the dosage of the antimicrobial, e.g. antibiotic, drugs, it is referred to the established principles of pharmacology in human and veterinary medicine. For example, Forth, Henschler, Rummel "Allgemeine und spezielle Pharmakologie und Toxikologie", 9th edition, 2005, pp. 781 – 919, might be used as a guideline. Regarding the formulation of a ready-to-use medicament, reference is made to "Remington, The Science and Practice of Pharmacy", 22nd edition, 2013, pp. 777 - 1070.

Assembling of a gene sequence can be carried out by any known method and is not particularly limited.

According to certain embodiments, mutations that were found using alignments can also be compared or matched with alignment-free methods, e.g. for detecting single base exchanges, for example based on contigs that were found by assemblies. For example, reads obtained from sequencing can be assembled to contigs and the contigs can be compared to each other.

According to a first aspect, the present invention relates to a diagnostic method of determining an infection of a patient with Salmonella species potentially resistant to antimicrobial drug treatment, which can also be described as method of determining an antimicrobial drug, e.g. antibiotic, resistant Salmonella infection of a patient, comprising the steps of:

a) obtaining or providing a sample containing or suspected of containing at least one Salmonella species from the patient;

b) determining the presence of at least one mutation in at least two genes from the group of genes consisting of recN, hemH, UMN798_3428, metE, yijD, UMN798_4831, UMN798_1939, copS, UMN798_0628, UMN798_4878, leuB, recF, emrA, glyQ, dcp, thiH, UMN798_1612, UMN798_2909, UMN798_1680, UMN798_4073, yjbC, nadB, UMN798_3160, hutU, envC, UMN798_3889, UMN798_1629, bcfB, degQ, UMN798_1331, trg, uvrC, polB, hpcD, UMN798_1628, UMN798_1701, glgS, plsB, yjcC, feoB, misL, dxr, hemF, rnfG, yhjB, UMN798_1163, UMN798_0394, alkA, nhaA, and lspA, wherein the presence of said at least two mutations is indicative of an infection with an antimicrobial, e.g. antibiotic, resistant Salmonella strain in said patient.

In this method, as well as the other methods of the invention, the sample can be provided or obtained in any way, preferably non-invasive, and can be e.g. provided as an *in vitro* sample or prepared as *in vitro* sample.

According to certain aspects, mutations in at least two, three, four, five, six, seven, eight, nine or ten genes are determined in any of the methods of the present invention, e.g. in at least two genes or in at least three genes. Instead of testing only single genes or mutants, a combination of several variant positions can improve the prediction accuracy and further reduce false positive findings that are influenced by other factors. Therefore, it is in particular preferred to determine the presence of a mutation in 2, 3, 4, 5, 6, 7, 8 or 9 (or more) genes selected from Table 1 or 2.

For the above genes, i.e. the genes also denoted in Tables 1 and 2, the highest probability of a resistance to at least one antimicrobial drug, e.g. antibiotic, could be observed, with p-values smaller than 10^{-30}, particularly smaller than

10^{-40}, indicating the high significance of the values (n= 636; α = 0.05). Details regarding Tables 1 and 2 can be taken from Tables 3 and 4 (4a, 4b, 4c) disclosed in the Examples. Having at least two genes with mutations determined, a high probability of an antimicrobial drug, e.g. antibiotic, resistance could be determined. The genes in Table 1 thereby represent the 50 best genes for which a mutation was observed in the genomes of Salmonella species, whereas the genes in Table 2 represent the 50 best genes for which a cross-correlation could be observed for the antimicrobial drug, e.g. antibiotic, susceptibility testing for Salmonella species as described below.

According to certain embodiments, the obtaining or providing a sample containing or suspected of containing at least one Salmonella species from the patient in this method - as well as the other methods of the invention – can comprise the following:

A sample of a vertebrate, e.g. a human, e.g. is provided or obtained and nucleic acid sequences, e.g. DNA or RNA sequences, are recorded by a known method for recording nucleic acid, which is not particularly limited. For example, nucleic acid can be recorded by a sequencing method, wherein any sequencing method is appropriate, particularly sequencing methods wherein a multitude of sample components, as e.g. in a blood sample, can be analyzed for nucleic acids and/or nucleic acid fragments and/or parts thereof contained therein in a short period of time, including the nucleic acids and/or nucleic acid fragments and/or parts thereof of at least one microorganism of interest, particularly of at least one Salmonella species. For example, sequencing can be carried out using polymerase chain reaction (PCR), particularly multiplex PCR, or high throughput sequencing or next generation se-

quencing, preferably using high-throughput sequencing. For sequencing, preferably an *in vitro* sample is used.

The data obtained by the sequencing can be in any format, and can then be used to identify the nucleic acids, and thus genes, of the microorganism, e.g. of Salmonella species, to be identified, by known methods, e.g. fingerprinting methods, comparing genomes and/or aligning to at least one, or more, genomes of one or more species of the microorganism of interest, i.e. a reference genome, etc., forming a third data set of aligned genes for a Salmonella species – discarding additional data from other sources, e.g. the vertebrate. Reference genomes are not particularly limited and can be taken from several databases. Depending on the microorganism, different reference genomes or more than one reference genomes can be used for aligning. Using the reference genome - as well as also the data from the genomes of the other species, e.g. Salmonella species - mutations in the genes for each species and for the whole multitude of samples of different species, e.g. Salmonella species, can be obtained.

For example, it is useful in genome-wide association studies to reference the points of interest, e.g. mutations, to one constant reference for enhanced standardization. In case of the human with a high consistency of the genome and 99% identical sequences among individuals this is easy and represents the standard, as corresponding reference genomes are available in databases. In case of organisms that trigger infectious diseases (e.g. bacteria and viruses) this is much more difficult, though. One possibility is to fall back on a virtual pan genome which contains all sequences of a certain genus. A further possibility is the analysis of all available references, which is much more complex. Therein all *n* refer-

ences from a database (e.g. RefSeq) are extracted and compared with the newly sequenced bacterial genomes *k*. After this, matrices (% of mapped reads, % of covered genome) are applied to estimate which reference is best suited to all new bacteria. However, *n* x *k* complete alignments are carried out. Having a big number of references, though, stable results can be obtained, as is the case for Salmonella.

According to certain embodiments, the genomes of Salmonella species are referenced to one reference genome. However, it is not excluded that for other microorganisms more than one reference genome is used. In the present methods, the reference genome of Salmonella is NC_017046 as annotated at the NCBI according to certain embodiments. The reference genome is attached to this application as sequence listing with SEQ ID NO 1.

The reference sequence was obtained from Salmonella strain NC_017046 (http://www.genome.jp/dbget-bin/www_bget?refseq+NC_017046)

LOCUS NC_017046 4876219 bp DNA circular CON 01-MAR-2015
DEFINITION Salmonella enterica subsp. enterica serovar Typhimurium str. 798, complete genome.
ACCESSION NC_017046
VERSION NC_017046.1 GI:383494824
DBLINK BioProject: PRJNA224116
BioSample: SAMN02604223
Assembly: GCF_000252875.1
KEYWORDS RefSeq.
SOURCE Salmonella enterica subsp. enterica serovar Typhimurium str. 798
ORGANISM Salmonella enterica subsp. enterica serovar Typhimurium str. 798

Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Salmonella.
REFERENCE 1 (bases 1 to 4876219)
AUTHORS Patterson,S.K., Borewicz,K., Johnson,T., Xu,W. and Isaacson,R.E.
TITLE Characterization and differential gene expression between two phenotypic phase variants in Salmonella enterica serovar Typhimurium
JOURNAL PLoS ONE 7 (8), E43592 (2012)
PUBMED 22937065
REFERENCE 2 (bases 1 to 4876219)
AUTHORS Borewicz,K., Johnson,T.J. and Isaacson,R.E.
TITLE Direct Submission
JOURNAL Submitted (28-JAN-2012) Veterinary and Biomedical Sciences, University of Minnesota, 1971 Commonwealth Avenue, 205 Vet Science, Saint Paul, MN 55108, USA

Alternatively or in addition, the gene sequence of the first data set can be assembled, at least in part, with known methods, e.g. by de-novo assembly or mapping assembly. The sequence assembly is not particularly limited, and any known genome assembler can be used, e.g. based on Sanger, 454, Solexa, Illumina, SOLid technologies, etc., as well as hybrids/mixtures thereof.

According to certain embodiments, the data of nucleic acids of different origin than the microorganism of interest, e.g. Salmonella species, can be removed after the nucleic acids of interest are identified, e.g. by filtering the data out. Such data can e.g. include nucleic acids of the patient, e.g. the vertebrate, e.g. human, and/or other microorganisms, etc. This can be done by e.g. computational subtraction, as developed by Meyerson et al. 2002. For this, also aligning to the

genome of the vertebrate, etc., is possible. For aligning, several alignment-tools are available. This way the original data amount from the sample can be drastically reduced.

Also after such removal of "excess" data, fingerprinting and/or aligning, and/or assembly, etc. can be carried out, as described above, forming a third data set of aligned and/or assembled genes for a Salmonella species.

Using these techniques, genes with mutations of the microorganism of interest, e.g. Salmonella species, can be obtained for various species.

When testing these same species for antimicrobial drug, e.g. antibiotic, susceptibility of a number of antimicrobial drugs, e.g. antibiotics, e.g. using standard culturing methods on dishes with antimicrobial drug, e.g. antibiotic, intake, as e.g. described below, the results of these antimicrobial drug, e.g. antibiotic, susceptibility tests can then be cross-referenced/correlated with the mutations in the genome of the respective microorganism, e.g. Salmonella. Using several, e.g. 50 or more than 50, 100 or more than 100, 200 or more than 200, 300 or more than 300, 400 or more than 400, 500 or more than 500, or 600 or more than 600 different species of a microorganism, e.g. different Salmonella species, statistical analysis can be carried out on the obtained cross-referenced data between mutations and antimicrobial drug, e.g. antibiotic, susceptibility for these number of species, using known methods.

Regarding culturing methods, samples can be e.g. cultured overnight. On the next day individual colonies can be used for identification of organisms, either by culturing or using mass spectroscopy. Based on the identity of organisms new plates containing increasing concentration of antibiotics used for the treatment of these organisms are inoculated and

grown for additional 12 - 24 hours. The lowest drug concentration which inhibits growth (minimal inhibitory concentration - MIC) can be used to determine susceptibility/resistance for tested antibiotics.

Correlation of the nucleic acid / gene mutations with antimicrobial drug, e.g. antibiotic, resistance can be carried out in a usual way and is not particularly limited. For example, resistances can be correlated to certain genes or certain mutations, e.g. SNPs, in genes. After correlation, statistical analysis can be carried out.

In addition, statistical analysis of the correlation of the gene mutations with antimicrobial drug, e.g. antibiotic, resistance is not particularly limited and can be carried out, depending on e.g. the amount of data, in different ways, for example using analysis of variance (ANOVA) or Student's t-test, for example with a sample size n of 50 or more, 100 or more, 200 or more, 300 or more, 400 or more, 500 or more, or 600 or more, and a level of significance (α-error-level) of e.g. 0.05 or smaller, e.g. 0.05, preferably 0.01 or smaller. A statistical value can be obtained for each gene and/or each position in the genome as well as for all antibiotics tested, a group of antibiotics or a single antibiotic. The obtained p-values can also be adapted for statistical errors, if needed.

For statistically sound results a multitude of individuals should be sampled, with n = 50, 100, 200, 300, 400, 500 or 600, and a level of significance (α-error-level) of e.g. 0.05 or smaller, e.g. 0.05, preferably 0.01 or smaller. According to certain embodiments, particularly significant results can be obtained for n = 200, 300, 400, 500 or 600.

For statistically sound results a multitude of individuals should be sampled, with *n* = 50 or more, 100 or more, 200 or more, 300 or more, 400 or more, 500 or more, or 600 or more, and a level of significance (α-error-level) of e.g. 0.05 or smaller, e.g. 0.05, preferably 0.01 or smaller. According to certain embodiments, particularly significant results can be obtained for n = 200 or more, 300 or more, 400 or more, 500 or more, or 600 or more.

After the above procedure has been carried out for more than 600, e.g. 636, individual species of Salmonella, the data disclosed in Tables 1 and 2 were obtained for the statistically best correlations between gene mutations and antimicrobial drug, e.g. antibiotic, resistances. Thus, mutations in these genes were proven as valid markers for antimicrobial drug, e.g. antibiotic, resistance.

According to a further aspect, the present invention relates in a second aspect to a method of selecting a treatment of a patient suffering from an infection with a potentially resistant Salmonella strain, e.g. from an antimicrobial drug, e.g. antibiotic, resistant Salmonella infection, comprising the steps of:

a) obtaining or providing a sample containing or suspected of containing at least one Salmonella species from the patient;

b) determining the presence of at least one mutation in at least two genes from the group of genes consisting of recN, hemH, UMN798_3428, metE, yijD, UMN798_4831, UMN798_1939, copS, UMN798_0628, UMN798_4878, leuB, recF, emrA, glyQ, dcp, thiH, UMN798_1612, UMN798_2909, UMN798_1680, UMN798_4073, yjbC, nadB, UMN798_3160, hutU, envC, UMN798_3889, UMN798_1629, bcfB, degQ, UMN798_1331, trg, uvrC, polB, hpcD,

UMN798_1628, UMN798_1701, glgS, plsB, yjcC, feoB, misL, dxr, hemF, rnfG, yhjB, UMN798_1163, UMN798_0394, alkA, nhaA, and lspA ,wherein the presence of said at least two mutations is indicative of a resistance to one or more antimicrobial, e.g. antibiotic, drugs;

c) identifying said at least one or more antimicrobial, e.g. antibiotic, drugs; and

d) selecting one or more antimicrobial, e.g. antibiotic, drugs different from the ones identified in step c) and being suitable for the treatment of a Salmonella infection.

In this method, the steps a) of obtaining or providing a sample and b) of determining the presence of at least one mutation are as in the method of the first aspect.

The identification of the at least one or more antimicrobial, e.g. antibiotic, drug in step c) is then based on the results obtained in step b) and corresponds to the antimicrobial, e.g. antibiotic, drug(s) that correlate(s) with the mutations. Once these antimicrobial drugs, e.g. antibiotics, are ruled out, the remaining antimicrobial drugs, e.g. antibiotic drugs/antibiotics, can be selected in step d) as being suitable for treatment.

In the description, references to the first and second aspect also apply to the 14th, 15th, 16th and 17th aspect, referring to the same genes, unless clear from the context that they don't apply.

According to certain embodiments, the antimicrobial drug, e.g. antibiotic, in the method of the first or second aspect, as well as in the other methods of the invention, is at least one selected from the group of β-lactams, β-lactam in-

hibitors, quinolines and derivatives thereof, aminoglycosides, polyketides, respectively tetracyclines, and folate synthesis inhibitors, particularly at least one selected from the group of β-lactams, β-lactam inhibitors, quinolines and derivatives thereof, aminoglycosides, and polyketides, respectively tetracyclines.

In the methods of the invention the resistance of Salmonella to one or more antimicrobial, e.g. antibiotic, drugs can be determined according to certain embodiments.

According to certain embodiments of the first and/or second aspect of the invention the antimicrobial, e.g. antibiotic, drug is selected from lactam antibiotics, and the presence of a mutation in the following genes is determined: recN, hemH, UMN798_3428, metE, yijD, UMN798_4831, UMN798_1939, copS, UMN798_0628, UMN798_4878, leuB, recF, emrA, glyQ, dcp, thiH, UMN798_1612, UMN798_2909, UMN798_1680, UMN798_4073, yjbC, nadB, UMN798_3160, hutU, envC, UMN798_3889, UMN798_1629, bcfB, degQ, UMN798_1331, trg, uvrC, polB, hpcD, UMN798_1628, UMN798_1701, glgS, plsB, yjcC, feoB, misL, dxr, hemF, rnfG, yhjB, UMN798_1163, UMN798_0394, alkA, nhaA, lspA.

According to certain embodiments of the first and/or second aspect of the invention the antimicrobial, e.g. antibiotic, drug is selected from aminoglycoside antibiotics, and the presence of a mutation in the following genes is determined: recN, hemH, UMN798_3428, metE, yijD, UMN798_4831, UMN798_1939, copS, UMN798_0628, UMN798_4878, leuB, recF, emrA, glyQ, dcp, thiH, UMN798_1612, UMN798_2909, UMN798_1680, UMN798_4073, yjbC, nadB, UMN798_3160, hutU, envC, UMN798_3889, UMN798_1629, bcfB, degQ, UMN798_1331, trg, uvrC, polB, hpcD, UMN798_1628, UMN798_1701, glgS, plsB, yjcC, feoB,

misL, dxr, hemF, rnfG, yhjB, UMN798_1163, UMN798_0394, alkA, nhaA, lspA.

According to certain embodiments of the first and/or second aspect of the invention the antimicrobial, e.g. antibiotic, drug is selected from polyketide antibiotics, preferably tetracycline antibiotics, and the presence of a mutation in the following genes is determined: recN, hemH, UMN798_3428, metE, yijD, UMN798_4831, UMN798_1939, copS, UMN798_0628, UMN798_4878, leuB, recF, emrA, glyQ, dcp, thiH, UMN798_1612, UMN798_2909, UMN798_1680, UMN798_4073, yjbC, nadB, UMN798_3160, hutU, envC, UMN798_3889, UMN798_1629, bcfB, UMN798_1331, trg, uvrC, polB, hpcD, UMN798_1628, UMN798_1701, glgS, plsB, yjcC, feoB, misL, dxr, hemF, rnfG, yhjB, UMN798_1163, UMN798_0394, alkA, nhaA, lspA.

According to certain embodiments, the antimicrobial drug is an antibiotic/antibiotic drug.

According to certain embodiments of the first and/or second aspect of the invention, determining the nucleic acid sequence information or the presence of a mutation comprises determining the presence of a single nucleotide at a single position in a gene. Thus the invention comprises methods wherein the presence of a single nucleotide polymorphism or mutation at a single nucleotide position is detected.

According to certain embodiments, the antibiotic drug in the methods of the present invention is selected from the group consisting of Amoxicillin/K Clavulanate (AUG), Ampicillin (AM), Aztreonam (AZT), Cefazolin (CFZ), Cefepime (CPE), Cefotaxime (CFT), Ceftazidime (CAZ), Ceftriaxone (CAX), Cefuroxime (CRM), Cephalotin (CF), Ciprofloxacin (CP),

Ertapenem (ETP), Gentamicin (GM), Imipenem (IMP), Levofloxacin (LVX), Meropenem (MER), Piperacillin/Tazobactam (P/T), Ampicillin/Sulbactam (A/S), Tetracycline (TE), Tobramycin (TO), and Trimethoprim/Sulfamethoxazole (T/S).

The inventors have surprisingly found that mutations in certain genes are indicative not only for a resistance to one single antimicrobial, e.g. antibiotic, drug, but to groups containing several drugs.

According to certain embodiments of the first and/or second aspect of the invention, the gene is from Table 1 or Table 2, the antibiotic drug is selected from lactam antibiotics and a mutation in at least one of the following genes is detected with regard to reference genome NC_017046: recN, hemH, UMN798_3428, metE, yijD, UMN798_4831, UMN798_1939, copS, UMN798_0628, UMN798_4878, leuB, recF, emrA, glyQ, dcp, thiH, UMN798_1612, UMN798_2909, UMN798_1680, UMN798_4073, yjbC, nadB, UMN798_3160, hutU, envC, UMN798_3889, UMN798_1629, bcfB, degQ, UMN798_1331, trg, uvrC, polB, hpcD, UMN798_1628, UMN798_1701, glgS, plsB, yjcC, feoB, misL, dxr, hemF, rnfG, yhjB, UMN798_1163, UMN798_0394, alkA, nhaA, lspA.

According to certain embodiments of the first and/or second aspect of the invention, the gene is from Table 1 or Table 2, the antibiotic drug is selected from aminoglycoside antibiotics and a mutation in at least one of the following genes is detected with regard to reference genome NC_017046: recN, hemH, UMN798_3428, metE, yijD, UMN798_4831, UMN798_1939, copS, UMN798_0628, UMN798_4878, leuB, recF, emrA, glyQ, dcp, thiH, UMN798_1612, UMN798_2909, UMN798_1680, UMN798_4073, yjbC, nadB, UMN798_3160, hutU, envC, UMN798_3889, UMN798_1629, bcfB, degQ, UMN798_1331, trg, uvrC, polB, hpcD,

UMN798_1628, UMN798_1701, glgS, plsB, yjcC, feoB, misL, dxr, hemF, rnfG, yhjB, UMN798_1163, UMN798_0394, alkA, nhaA, lspA.

According to certain embodiments of the first and/or second aspect of the invention, the gene is from Table 1 or Table 2, the antibiotic drug is selected from polyketide, preferably tetracycline antibiotics and a mutation in at least one of the following genes is detected with regard to reference genome NC_017046: recN, hemH, UMN798_3428, metE, yijD, UMN798_4831, UMN798_1939, copS, UMN798_0628, UMN798_4878, leuB, recF, emrA, glyQ, dcp, thiH, UMN798_1612, UMN798_2909, UMN798_1680, UMN798_4073, yjbC, nadB, UMN798_3160, hutU, envC, UMN798_3889, UMN798_1629, bcfB, UMN798_1331, trg, uvrC, polB, hpcD, UMN798_1628, UMN798_1701, glgS, plsB, yjcC, feoB, misL, dxr, hemF, rnfG, yhjB, UMN798_1163, UMN798_0394, alkA, nhaA, lspA.

For specific antimicrobial drugs, e.g. antibiotics, specific positions in the above genes can be determined where a high statistical significance is observed. The inventors found that, apart from the above genes indicative of a resistance against antibiotics, also single nucleotide polymorphisms (= SNP's) may have a high significance for the presence of a resistance against defined antibiotic drugs. The analysis of these polymorphisms on a nucleotide level may further improve and accelerate the determination of a drug resistance to antimicrobial drugs, e.g. antibiotics, in Salmonella.

According to certain embodiments of the first and/or second aspect of the invention, the gene is from Table 1 or Table 2, the antibiotic drug is selected from lactam antibiotics and a mutation in at least one of the following nucleotide positions is detected with regard to reference genome NC_017046:

2833888, 546961, 3334479, 4191057, 4366486, 4724403, 1895588, 1139812, 637461, 4779417, 131219, 4062015, 2983118, 3861998, 1548464, 4397111, 1574737, 2840330, 1650934, 3966175, 4436188, 2780306, 3075942, 855087, 3582301, 3772654, 1590194, 25574, 3536122, 1313897, 1673475, 1994028, 115342, 1148509, 1589819, 1672517, 3379140, 4478134, 4523874, 3686566, 3976726, 258966, 2558379, 1487086, 3799879, 1163470, 409259, 2208848, 46695, 56998, 3582354, 3335426, 1673444.

According to certain embodiments of the first and/or second aspect of the invention, the gene is from Table 1 or Table 2, the antibiotic drug is selected from aminoglycoside antibiotics and a mutation in at least one of the following nucleotide positions is detected with regard to reference genome NC_017046: 2833888, 546961, 3334479, 4191057, 4366486, 4724403, 1895588, 1139812, 637461, 4779417, 131219, 4062015, 2983118, 3861998, 1548464, 4397111, 1574737, 2840330, 1650934, 3966175, 4436188, 2780306, 3075942, 855087, 3582301, 3772654, 1590194, 25574, 3536122, 1313897, 1673475, 1994028, 115342, 1148509, 1589819, 1672517, 3379140, 4478134, 4523874, 3686566, 3976726, 258966, 2558379, 1487086, 3799879, 1163470, 409259, 2208848, 46695, 56998, 3582354, 3335426, 1673444.

According to certain embodiments of the first and/or second aspect of the invention, the gene is from Table 1 or Table 2, the antibiotic drug is selected from polyketide, preferably tetracycline antibiotics and a mutation in at least one of the following nucleotide positions is detected with regard to reference genome NC_017046: 2833888, 546961, 3334479, 4191057, 4366486, 4724403, 1895588, 1139812, 637461, 4779417, 131219, 4062015, 2983118, 3861998, 1548464, 4397111, 1574737, 2840330, 1650934, 3966175, 4436188, 2780306, 3075942, 855087, 3582301, 3772654, 1590194, 25574, 1313897, 1673475, 1994028,

115342, 1148509, 1589819, 1672517, 3379140, 4478134, 4523874, 3686566, 3976726, 258966, 2558379, 1487086, 3799879, 1163470, 409259, 2208848, 46695, 56998, 3582354, 3335426, 1673444.

According to certain embodiments of the first and/or second aspect of the invention, the antibiotic drug is CFZ and a mutation in at least one of the following nucleotide positions is detected with regard to reference genome NC_017046: 4779417, 2983118, 1548464, 1574737, 3772654, 1313897.

According to certain embodiments of the first and/or second aspect of the invention, the antibiotic drug is GM and a mutation in at least one of the following nucleotide positions is detected with regard to reference genome NC_017046: 2840330, 25574, 3536122, 115342, 1148509, 3379140, 4478134, 3686566, 1487086, 3799879, 1163470, 2208848, 46695, 56998, 3335426.

According to certain embodiments of the first and/or second aspect of the invention, the antibiotic drug is CF and a mutation in at least one of the following nucleotide positions is detected with regard to reference genome NC_017046: 2833888, 546961, 3334479, 4191057, 4366486, 4724403, 1895588, 1139812, 637461, 4779417, 131219, 4062015, 2983118, 3861998, 1548464, 4397111, 1574737, 1650934, 4436188, 3075942, 855087, 3582301, 3772654, 1590194, 1313897, 1673475, 1994028, 1589819, 1672517, 3976726, 3582354, 1673444.

According to certain embodiments of the first and/or second aspect of the invention, the antibiotic drug is TE and a mutation in at least one of the following nucleotide positions is detected with regard to reference genome NC_017046: 2833888, 546961, 3334479, 4191057, 4366486, 4724403, 1895588,

1139812, 637461, 4779417, 131219, 4062015, 2983118, 3861998, 1548464, 4397111, 1574737, 2840330, 1650934, 3966175, 4436188, 2780306, 3075942, 855087, 3582301, 3772654, 1590194, 25574, 1313897, 1673475, 1994028, 115342, 1148509, 1589819, 1672517, 3379140, 4478134, 4523874, 3686566, 3976726, 258966, 2558379, 1487086, 3799879, 1163470, 409259, 2208848, 46695, 56998, 3582354, 3335426, 1673444.

According to certain embodiments of the first and/or second aspect of the invention, the antibiotic drug is A/S and a mutation in at least one of the following nucleotide positions is detected with regard to reference genome NC_017046: 2833888, 546961, 3334479, 4191057, 4366486, 4724403, 1895588, 1139812, 637461, 4779417, 131219, 4062015, 2983118, 3861998, 1548464, 4397111, 1574737, 2840330, 1650934, 3966175, 4436188, 2780306, 3075942, 855087, 3582301, 3772654, 1590194, 25574, 3536122, 1313897, 1673475, 1994028, 115342, 1148509, 1589819, 1672517, 3379140, 4478134, 4523874, 3686566, 3976726, 258966, 2558379, 1487086, 3799879, 1163470, 409259, 2208848, 46695, 56998, 3582354, 3335426, 1673444.

According to certain embodiments of the first and/or second aspect of the invention, the antibiotic drug is CRM and a mutation in at least one of the following nucleotide positions is detected with regard to reference genome NC_017046: 2833888, 546961, 3334479, 4191057, 4366486, 4724403, 1895588, 1139812, 637461, 4779417, 131219, 4062015, 2983118, 3861998, 1548464, 4397111, 1574737, 2840330, 1650934, 3966175, 4436188, 2780306, 3075942, 855087, 3582301, 3772654, 1590194, 25574, 3536122, 1313897, 1673475, 1994028, 115342, 1148509, 1589819, 1672517, 3379140, 4478134, 4523874, 3686566, 3976726, 258966, 2558379, 1487086, 3799879, 1163470, 409259, 2208848, 46695, 56998, 3582354, 3335426, 1673444.

According to certain embodiments of the first and/or second aspect of the invention, the antibiotic drug is P/T and a mutation in at least one of the following nucleotide positions is detected with regard to reference genome NC_017046: 2833888, 546961, 3334479, 4191057, 4366486, 4724403, 1895588, 1139812, 637461, 4779417, 131219, 4062015, 2983118, 3861998, 1548464, 4397111, 1574737, 2840330, 1650934, 3966175, 4436188, 2780306, 3075942, 855087, 3582301, 3772654, 1590194, 25574, 3536122, 1313897, 1673475, 1994028, 115342, 1148509, 1589819, 1672517, 3379140, 4478134, 4523874, 3686566, 3976726, 258966, 2558379, 1487086, 3799879, 1163470, 409259, 2208848, 46695, 56998, 3582354, 3335426, 1673444.

According to certain embodiments of the first and/or second aspect of the invention, the antibiotic drug is TO and a mutation in at least one of the following nucleotide positions is detected with regard to reference genome NC_017046: 2833888, 546961, 3334479, 4191057, 4366486, 4724403, 1895588, 1139812, 637461, 4779417, 131219, 4062015, 2983118, 3861998, 1548464, 4397111, 1574737, 2840330, 1650934, 3966175, 4436188, 2780306, 3075942, 855087, 3582301, 3772654, 1590194, 25574, 3536122, 1313897, 1673475, 1994028, 115342, 1148509, 1589819, 1672517, 3379140, 4478134, 4523874, 3686566, 3976726, 258966, 2558379, 1487086, 3799879, 1163470, 409259, 2208848, 46695, 56998, 3582354, 3335426, 1673444.

According to certain embodiments of the first and/or second aspect of the invention, the antibiotic drug is AM and a mutation in at least one of the following nucleotide positions is detected with regard to reference genome NC_017046: 2833888, 546961, 3334479, 4191057, 4366486, 4724403, 1895588, 1139812, 637461, 4779417, 131219, 4062015, 2983118, 3861998, 1548464, 4397111, 1574737, 2840330, 1650934, 3966175,

4436188, 2780306, 3075942, 855087, 3772654, 25574, 3536122, 1313897, 1994028, 115342, 1148509, 3379140, 4478134, 4523874, 3686566, 3976726, 258966, 2558379, 1487086, 3799879, 1163470, 409259, 2208848, 46695, 56998, 3335426.

According to certain embodiments of the first and/or second aspect of the invention, the antibiotic drug is AUG and a mutation in at least one of the following nucleotide positions is detected with regard to reference genome NC_017046: 2833888, 546961, 3334479, 4191057, 4366486, 4724403, 1895588, 1139812, 637461, 4779417, 131219, 4062015, 2983118, 3861998, 1548464, 4397111, 1574737, 2840330, 1650934, 3966175, 4436188, 2780306, 3075942, 855087, 3582301, 3772654, 1590194, 25574, 3536122, 1313897, 1673475, 1994028, 115342, 1148509, 1589819, 1672517, 3379140, 4478134, 4523874, 3686566, 3976726, 258966, 2558379, 1487086, 3799879, 1163470, 409259, 2208848, 46695, 56998, 3582354, 3335426, 1673444.

According to certain embodiments of the first and/or second aspect of the invention, the resistance of a bacterial microorganism belonging to the species Salmonella against 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15 or 16, 17, 18, 19, 20 or 21 antibiotic drugs is determined.

According to certain embodiments of the first and/or second aspect of the invention, a detected mutation is a mutation leading to an altered amino acid sequence in a polypeptide derived from a respective gene in which the detected mutation is located. According to this aspect, the detected mutation thus leads to a truncated version of the polypeptide (wherein a new stop codon is created by the mutation) or a mutated version of the polypeptide having an amino acid exchange at the respective position.

According to certain embodiments of the first and/or second aspect of the invention, determining the nucleic acid sequence information or the presence of a mutation comprises determining a partial sequence or an entire sequence of the at least two genes.

According to certain embodiments of the first and/or second aspect of the invention, determining the nucleic acid sequence information or the presence of a mutation comprises determining a partial or entire sequence of the genome of the Salmonella species, wherein said partial or entire sequence of the genome comprises at least a partial sequence of said at least two genes.

According to certain embodiments of the first and/or second aspect of the invention, determining the nucleic acid sequence information or the presence of a mutation comprises using a next generation sequencing or high throughput sequencing method. According to preferred embodiments of the first and/or second aspect of the invention, a partial or entire genome sequence of the bacterial organism of Salmonella species is determined by using a next generation sequencing or high throughput sequencing method.

In a further, third aspect, the present invention relates to a method of determining an antimicrobial drug, e.g. antibiotic, resistance profile for bacterial microorganisms of Salmonella species, comprising:

obtaining or providing a first data set of gene sequences of a plurality of clinical isolates of Salmonella species;

providing a second data set of antimicrobial drug, e.g. antibiotic, resistance of the plurality of clinical isolates of Salmonella species;

aligning the gene sequences of the first data set to at least one, preferably one, reference genome of Salmonella, and/or assembling the gene sequence of the first data set, at least in part;

analyzing the gene sequences of the first data set for genetic variants to obtain a third data set of genetic variants;

correlating the third data set with the second data set and statistically analyzing the correlation; and

determining the genetic sites in the genome of Salmonella associated with antimicrobial drug, e.g. antibiotic, resistance.

The different steps can be carried out as described with regard to the method of the first aspect of the present invention.

When referring to the second data set, wherein the second data set e.g. comprises, respectively is, a set of antimicrobial drug, e.g. antibiotic, resistances of a plurality of clinical isolates, this can, within the scope of the invention, also refer to a self-learning data base that, whenever a new sample is analyzed, can take this sample into the second data set and thus expand its data base. The second data set thus does not have to be static and can be expanded, either by external input or by incorporating new data due to self-learning. This is, however, not restricted to the third aspect of the invention, but applies to other aspects of the invention that refer to a second data set, which does not necessarily have to refer to antimicrobial drug resistance.

The same applies, where applicable, to the first data set, e.g. in the third aspect.

According to certain embodiments, statistical analysis in the present methods is carried out using Fisher's test with $p < 10^{-6}$, preferably $p < 10^{-9}$, particularly $p < 10^{-10}$.

The method of the third aspect of the present invention, as well as related methods, e.g. according to the 7th and 10th aspect, can, according to certain embodiments, comprise correlating different genetic sites to each other, e.g. in at least two, three, four, five, six, seven, eight, nine or ten genes. This way even higher statistical significance can be achieved.

According to certain embodiments of the method of the third aspect and related methods – as above, the second data set is provided by culturing the clinical isolates of Salmonella species on agar plates provided with antimicrobial drugs, e.g. antibiotics, at different concentrations and the second data is obtained by taking the minimal concentration of the plates that inhibits growth of the respective Salmonella species.

According to certain embodiments of the method of the third aspect and related methods, the antibiotic is at least one selected from the group of β-lactams, β-lactam inhibitors, quinolines and derivatives thereof, aminoglycosides, tetracyclines, and folate synthesis inhibitors, preferably Amoxicillin/K Clavulanate, Ampicillin, Aztreonam, Cefazolin, Cefepime, Cefotaxime, Ceftazidime, Ceftriaxone, Cefuroxime, Cephalothin, Ciprofloxacin, Ertapenem, Gentamicin, Imipenem, Levofloxacin, Meropenem, Piperacillin/Tazobactam, Ampicil-

lin/Sulbactam, Tetracycline, Tobramycin, and Trimethoprim/Sulfamethoxazole.

According to certain embodiments of the method of the third aspect and related methods, the gene sequences in the third data set are comprised in at least one gene from the group of genes consisting of recN, hemH, UMN798_3428, metE, yijD, UMN798_4831, UMN798_1939, copS, UMN798_0628, UMN798_4878, leuB, recF, emrA, glyQ, dcp, thiH, UMN798_1612, UMN798_2909, UMN798_1680, UMN798_4073, yjbC, nadB, UMN798_3160, hutU, envC, UMN798_3889, UMN798_1629, bcfB, degQ, UMN798_1331, trg, uvrC, polB, hpcD, UMN798_1628, UMN798_1701, glgS, plsB, yjcC, feoB, misL, dxr, hemF, rnfG, yhjB, UMN798_1163, UMN798_0394, alkA, nhaA, and lspA, or from the genes listed in Table 5, preferably Table 5a.

According to certain embodiments of the method of the third aspect and related methods, the genetic sites in the genome of Salmonella associated with antimicrobial drug, e.g. antibiotic, resistance are at least comprised in one gene from the group of genes consisting of recN, hemH, UMN798_3428, metE, yijD, UMN798_4831, UMN798_1939, copS, UMN798_0628, UMN798_4878, leuB, recF, emrA, glyQ, dcp, thiH, UMN798_1612, UMN798_2909, UMN798_1680, UMN798_4073, yjbC, nadB, UMN798_3160, hutU, envC, UMN798_3889, UMN798_1629, bcfB, degQ, UMN798_1331, trg, uvrC, polB, hpcD, UMN798_1628, UMN798_1701, glgS, plsB, yjcC, feoB, misL, dxr, hemF, rnfG, yhjB, UMN798_1163, UMN798_0394, alkA, nhaA, and lspA.

According to certain embodiments of the method of the third aspect and related methods, the genetic variant has a point mutation, an insertion and or deletion of up to four bases, and/or a frameshift mutation.

A fourth aspect of the present invention relates to a method of determining an antimicrobial drug, e.g. antibiotic, resistance profile for a bacterial microorganism belonging to the species Salmonella comprising the steps of

a) obtaining or providing a sample containing or suspected of containing the bacterial microorganism;

b) determining the presence of a mutation in at least one gene of the bacterial microorganism as determined by the method of the third aspect of the invention;

wherein the presence of a mutation is indicative of a resistance to an antimicrobial drug, e.g. antibiotic, drug.

Steps a) and b) can herein be carried out as described with regard to the first aspect, as well as for the following aspects of the invention.

With this method, any mutations in the genome of Salmonella species correlated with antimicrobial drug, e.g. antibiotic, resistance can be determined and a thorough antimicrobial drug, e.g. antibiotic, resistance profile can be established.

A simple read out concept for a diagnostic test as described in this aspect is shown schematically in Fig. 1.

According to Fig. 1, a sample 1, e.g. blood from a patient, is used for molecular testing 2, e.g. using next generation sequencing (NGS), and then a molecular fingerprint 3 is taken, e.g. in case of NGS a sequence of selected genomic/plasmid regions or the whole genome is assembled. This is then compared to a reference library 4, i.e. selected sequences or the whole sequence are/is compared to one or more reference sequences, and mutations (SNPs, sequence- gene ad-

ditions/deletions, etc.) are correlated with susceptibility/ reference profile of reference strains in the reference library. The reference library 4 herein contains many genomes and is different from a reference genome. Then the result 5 is reported comprising ID (pathogen identification), i.e. a list of all (pathogenic) species identified in the sample, and AST (antimicrobial susceptibility testing), i.e. a list including a susceptibility /resistance profile for all species listed

A fifth aspect of the present invention relates to a diagnostic method of determining an infection of a patient with Salmonella species potentially resistant to antimicrobial drug treatment, which also can be described as method of determining an antimicrobial drug, e.g. antibiotic, resistant Salmonella infection in a patient, comprising the steps of:

a) obtaining or providing a sample containing or suspected of containing a bacterial microorganism belonging to the species Salmonella from the patient;

b) determining the presence of at least one mutation in at least one gene of the bacterial microorganism belonging to the species Salmonella as determined by the method of the third aspect of the present invention, wherein the presence of said at least one mutation is indicative of an antimicrobial drug, e.g. antibiotic, resistant Salmonella infection in said patient.

Again, steps a) and b) can herein be carried out as described with regard to the first aspect of the present invention.

According to this aspect, a Salmonella infection in a patient can be determined using sequencing methods as well as a resistance to antimicrobial drugs, e.g. antibiotics, of the

Salmonella species be determined in a short amount of time compared to the conventional methods.

In a sixth aspect the present invention relates to a method of selecting a treatment of a patient suffering from an infection with a potentially resistant Salmonella strain, e.g. an antimicrobial drug, e.g. antibiotic, resistant Salmonella infection, comprising the steps of:

a) obtaining or providing a sample containing or suspected of containing a bacterial microorganism belonging to the species Salmonella from the patient;

b) determining the presence of at least one mutation in at least one gene of the bacterial microorganism belonging to the species Salmonella as determined by the method of the third aspect of the invention, wherein the presence of said at least one mutation is indicative of a resistance to one or more antimicrobial, e.g. antibiotic, drugs;

c) identifying said at least one or more antimicrobial, e.g. antibiotic, drugs; and

d) selecting one or more antimicrobial, e.g. antibiotic, drugs different from the ones identified in step c) and being suitable for the treatment of a Salmonella infection.

This method can be carried out similarly to the second aspect of the invention and enables a fast was to select a suitable treatment with antibiotics for any infection with an unknown Salmonella species.

A seventh aspect of the present invention relates to a method of acquiring, respectively determining, an antimicrobial drug, e.g. antibiotic, resistance profile for a bacterial microorganism of Salmonella species, comprising:

obtaining or providing a first data set of gene sequences of a clinical isolate of Salmonella species;
providing a second data set of antimicrobial drug, e.g. antibiotic, resistance of a plurality of clinical isolates of Salmonella species;
aligning the gene sequences of the first data set to at least one, preferably one, reference genome of Salmonella, and/or assembling the gene sequence of the first data set, at least in part;
analyzing the gene sequences of the first data set for genetic variants to obtain a third data set of genetic variants of the first data set;
correlating the third data set with the second data set and statistically analyzing the correlation; and
determining the genetic sites in the genome of Salmonella of the first data set associated with antimicrobial drug, e.g. antibiotic, resistance.

With this method, antimicrobial drug, e.g. antibiotic, resistances in an unknown isolate of Salmonella can be determined.

According to certain embodiments, the reference genome of Salmonella is NC_017046 as annotated at the NCBI. According to certain embodiments, statistical analysis in the present methods is carried out using Fisher's test with $p < 10^{-6}$, preferably $p < 10^{-9}$, particularly $p < 10^{-10}$. Also, according to certain embodiments, the method further comprises correlating different genetic sites to each other, e.g. in at least two, three, four, five, six, seven, eight, nine or ten genes.

An eighth aspect of the present invention relates to a computer program product comprising computer executable instructions which, when executed, perform a method according to the third, fourth, fifth, sixth or seventh aspect of the present invention.

In certain embodiments the computer program product is one on which program commands or program codes of a computer program for executing said method are stored. According to certain embodiments the computer program product is a storage medium. The same applies to the computer program products of the aspects mentioned afterwards, i.e. the eleventh aspect of the present invention. As noted above, the computer program products of the present invention can be self-learning, e.g. with respect to the first and second data sets.

In order to obtain the best possible information from the highly complex genetic data and develop an optimum model for diagnostic and therapeutical uses as well as the methods of the present invention - which can be applied stably in clinical routine – a thorough in silico analysis can be necessary. The proposed principle is based on a combination of different approaches, e.g. alignment with at least one, preferably more reference genomes and/or assembly of the genome and correlation of mutations found in every sample, e.g. from each patient, with all references and drugs, e.g. antibiotics, and search for mutations which occur in several drug and several strains.

Using the above steps a list of mutations as well of genes is generated. These can be stored in databases and statistical models can be derived from the databases. The statistical models can be based on at least one or more mutations at

least one or more genes. Statistical models that can be trained can be combined from mutations and genes. Examples of algorithms that can produce such models are association Rules, Support Vector Machines, Decision Trees, Decision Forests, Discriminant-Analysis, Cluster-Methods, and many more.

The goal of the training is to allow a reproducible, standardized application during routine procedures.

For this, for example, a genome or parts of the genome of a microorganism can be sequenced from a patient to be diagnosed. Afterwards, core characteristics can be derived from the sequence data which can be used to predict resistance. These are the points in the database used for the final model, i.e. at least one mutation or at least one gene, but also combinations of mutations, etc.

The corresponding characteristics can be used as input for the statistical model and thus enable a prognosis for new patients. Not only the information regarding all resistances of all microorganisms, e.g. of Salmonella species, against all drugs, e.g. antibiotics, can be integrated in a computer decision support tool, but also corresponding directives (e.g. EUCAST) so that only treatment proposals are made that are in line with the directives.

A ninth aspect of the present invention relates to the use of the computer program product according to the eighth aspect for acquiring an antimicrobial drug, e.g. antibiotic, resistance profile for bacterial microorganisms of Salmonella species or in a method of the third aspect of the invention.

In a tenth aspect a method of selecting a treatment of a patient having an infection with a bacterial microorganism of Salmonella species, comprising:
obtaining or providing a first data set comprising a gene sequence of at least one clinical isolate of the bacterial microorganism from the patient;
providing a second data set of antimicrobial drug, e.g. antibiotic, resistance of a plurality of clinical isolates of the bacterial microorganism;
aligning the gene sequences of the first data set to at least one, preferably one, reference genome of the bacterial microorganism, and/or assembling the gene sequence of the first data set, at least in part;
analyzing the gene sequences of the first data set for genetic variants to obtain a third data set of genetic variants of the first data set;
correlating the third data set with the second data set of antimicrobial drug, e.g. antibiotic, resistance of a plurality of clinical isolates of the bacterial microorganism and statistically analyzing the correlation;
determining the genetic sites in the genome of the clinical isolate of the bacterial microorganism of the first data set associated with antimicrobial drug, e.g. antibiotic, resistance; and
selecting a treatment of the patient with one or more antimicrobial, e.g. antibiotic, drugs different from the ones identified in the determination of the genetic sites associated with antimicrobial drug, e.g. antibiotic, resistance is disclosed.

Again, the steps can be carried out as similar steps before. In this method, as well as similar ones, no aligning is necessary, as the unknown sample can be directly correlated, af-

ter the genome or genome sequences are produced, with the second data set and thus mutations and antimicrobial drug, e.g. antibiotic, resistances can be determined. The first data set can be assembled, for example, using known techniques.

According to certain embodiments, statistical analysis in the present method is carried out using Fisher's test with $p < 10^{-6}$, preferably $p < 10^{-9}$, particularly $p < 10^{-10}$. Also, according to certain embodiments, the method further comprises correlating different genetic sites to each other.

An eleventh aspect of the present invention is directed to a computer program product comprising computer executable instructions which, when executed, perform a method according to the tenth aspect.

According to a twelfth aspect of the present invention, a diagnostic method of determining an infection of a patient with Salmonella species potentially resistant to antimicrobial drug treatment, which can also be described as a method of determining an antimicrobial drug, e.g. antibiotic, resistant Salmonella infection of a patient is disclosed, comprising the steps of:

a) obtaining or providing a sample containing or suspected of containing at least one Salmonella species from the patient;

b) determining the presence of at least one mutation in at least two genes from the group of genes listed in Table 5, preferably Table 5a, wherein the presence of said at least two mutations is indicative of an antimicrobial drug, e.g. antibiotic, resistant Salmonella infection in said patient.

A thirteenth aspect of the invention discloses a method of selecting a treatment of a patient suffering from an antimicrobial drug, e.g. antibiotic, resistant Salmonella infection, comprising the steps of:

a) obtaining or providing a sample containing or suspected of containing at least one Salmonella species from the patient;

b) determining the presence of at least one mutation in at least two genes from the group of genes listed in Table 5, preferably at least two genes from the group of genes listed in Table 5a, wherein the presence of said at least two mutations is indicative of a resistance to one or more antimicrobial, e.g. antibiotic, drugs;

c) identifying said at least one or more antimicrobial, e.g. antibiotic, drugs; and

d) selecting one or more antimicrobial, e.g. antibiotic, drugs different from the ones identified in step c) and being suitable for the treatment of a Salmonella infection.

Again, the steps can be carried out as in similar methods before, e.g. as in the first and second aspect of the invention. In the twelfth and thirteenth aspect of the invention, all classes of antibiotics considered in the present method are covered.

Table 5: List of genes

recN	hemH	UMN798_3428	metE	yijD
UMN798_4831	UMN798_1939	copS	UMN798_0628	UMN798_4878
leuB	recF	emrA	glyQ	dcp
thiH	UMN798_1612	UMN798_2909	UMN798_1680	UMN798_4073
yjbC	nadB	UMN798_3160	hutU	envC
UMN798_3889	UMN798_1629	bcfB	degQ	UMN798_1331
trg	uvrC	polB	hpcD	UMN798_1628

UMN798_1701	glgS	plsB	yjcC	feoB
misL	dxr	hemF	rnfG	yhjB
UMN798_1163	UMN798_0394	alkA	nhaA	lspA
UMN798_3890	ycbB	malS	dmsC	UMN798_0020
UMN798_1618	UMN798_1617	UMN798_4717	UMN798_0389	kdpD
hisB	UMN798_2727	UMN798_3918	UMN798_4753	UMN798_4936
yhdM	UMN798_0631	UMN798_1337	UMN798_1550	iroE
pheT	hofC	gyrA	torS	UMN798_1114
UMN798_2482	rseB	hycC	ttk	cpdB
UMN798_4882	rob	creA	UMN798_2202	dppA
adiY	UMN798_0653	gmm	UMN798_0179	UMN798_3553
UMN798_4061	hrpB	UMN798_0975	gcvP	UMN798_0654
pnp	ytfF	UMN798_1632	fhuD	

Herein, the genes in Table 5 are the following: recN, hemH, UMN798_3428, metE, yijD, UMN798_4831, UMN798_1939, copS, UMN798_0628, UMN798_4878, leuB, recF, emrA, glyQ, dcp, thiH, UMN798_1612, UMN798_2909, UMN798_1680, UMN798_4073, yjbC, nadB, UMN798_3160, hutU, envC, UMN798_3889, UMN798_1629, bcfB, degQ, UMN798_1331, trg, uvrC, polB, hpcD, UMN798_1628, UMN798_1701, glgS, plsB, yjcC, feoB, misL, dxr, hemF, rnfG, yhjB, UMN798_1163, UMN798_0394, alkA, nhaA, lspA, UMN798_3890, ycbB, malS, dmsC, UMN798_0020, UMN798_1618, UMN798_1617, UMN798_4717, UMN798_0389, kdpD, hisB, UMN798_2727, UMN798_3918, UMN798_4753, UMN798_4936, yhdM, UMN798_0631, UMN798_1337, UMN798_1550, iroE, pheT, hofC, gyrA, torS, UMN798_1114, UMN798_2482, rseB, hycC, ttk, cpdB, UMN798_4882, rob, creA, UMN798_2202, dppA, adiY, UMN798_0653, gmm, UMN798_0179, UMN798_3553, UMN798_4061, hrpB, UMN798_0975, gcvP, UMN798_0654, pnp, ytfF, UMN798_1632, and fhuD.

Table 5a : List of genes

recN	hemH	UMN798_3428	metE	yijD

UMN798_4831	UMN798_1939	copS	UMN798_0628	UMN798_4878
leuB	recF	emrA	glyQ	dcp
thiH	UMN798_1612	UMN798_2909	UMN798_1680	UMN798_4073
yjbC	nadB	UMN798_3160	hutU	envC
UMN798_3889	UMN798_1629	bcfB	degQ	UMN798_1331
trg	uvrC	polB	hpcD	UMN798_1628
UMN798_1701	glgS	plsB	yjcC	feoB
misL	dxr	hemF	rnfG	yhjB
UMN798_1163	UMN798_0394	alkA	nhaA	lspA
UMN798_3890	ycbB	malS	dmsC	UMN798_0020
UMN798_1618	UMN798_1617	UMN798_4717	UMN798_0389	kdpD
hisB	UMN798_2727	UMN798_3918	UMN798_4753	UMN798_4936
yhdM	UMN798_0631	UMN798_1337	UMN798_1550	iroE
pheT	hofC	fhuD	torS	UMN798_1114
UMN798_2482	rseB	hycC	ttk	cpdB
UMN798_4882	rob	creA	UMN798_2202	dppA
adiY	UMN798_0653	gmm	UMN798_0179	UMN798_3553
UMN798_4061	hrpB	UMN798_0975	gcvP	UMN798_0654
pnp	ytfF	UMN798_1632		

Herein, the genes in Table 5a are the following: recN, hemH, UMN798_3428, metE, yijD, UMN798_4831, UMN798_1939, copS, UMN798_0628, UMN798_4878, leuB, recF, emrA, glyQ, dcp, thiH, UMN798_1612, UMN798_2909, UMN798_1680, UMN798_4073, yjbC, nadB, UMN798_3160, hutU, envC, UMN798_3889, UMN798_1629, bcfB, degQ, UMN798_1331, trg, uvrC, polB, hpcD, UMN798_1628, UMN798_1701, glgS, plsB, yjcC, feoB, misL, dxr, hemF, rnfG, yhjB, UMN798_1163, UMN798_0394, alkA, nhaA, lspA, UMN798_3890, ycbB, malS, dmsC, UMN798_0020, UMN798_1618, UMN798_1617, UMN798_4717, UMN798_0389, kdpD, hisB, UMN798_2727, UMN798_3918, UMN798_4753, UMN798_4936, yhdM, UMN798_0631, UMN798_1337, UMN798_1550, iroE, pheT, hofC, torS, UMN798_1114, UMN798_2482, rseB, hycC, ttk, cpdB, UMN798_4882, rob, creA, UMN798_2202, dppA, adiY, UMN798_0653,

gmm, UMN798_0179, UMN798_3553, UMN798_4061, hrpB, UMN798_0975, gcvP, UMN798_0654, pnp, ytfF, UMN798_1632, and fhuD.

According to certain embodiments, mutations in at least two, three, four, five, six, seven, eight, nine or ten genes are determined in any of the methods of the present invention, e.g. in at least two genes or in at least three genes. Instead of testing only single genes or mutants, a combination of several variant positions can improve the prediction accuracy and further reduce false positive findings that are influenced by other factors. Therefore, it is in particular preferred to determine the presence of a mutation in 2, 3, 4, 5, 6, 7, 8 or 9 (or more) genes selected from Table 5, preferably Table 5a.

Further, according to certain embodiments, the reference genome of Salmonella is again NC_017046 as annotated at the NCBI. According to certain embodiments, statistical analysis in the present methods is carried out using Fisher's test with $p < 10^{-6}$, preferably $p < 10^{-9}$, particularly $p < 10^{-10}$. Also, according to certain embodiments, the method further comprises correlating different genetic sites to each other. Also the other aspects of the embodiments of the first and second aspect of the invention apply.

According to certain embodiments of the method of the twelfth and/or thirteenth aspect of the present invention, as well as also of the eighteenth aspect of the present invention, the antimicrobial drug is an antibiotic. According to certain embodiments, the antibiotic is a lactam antibiotic and a mutation in at least one of the genes listed in Table 6 is de-

tected, or a mutation in at least one of the positions (denoted POS in the table) listed in Table 6.

Table 6: List for lactam antibiotics (7 antibiotics)

gene name	POS	antibiotic	p-value (FDR)	genbank protein accession number
UMN798_4878	4779417	CF;TE;CFZ;CRM;P/T;TO;AM;A/S;AUG	3,07597E-46	YP_005399754.1
emrA	2983118	CF;TE;CFZ;CRM;P/T;TO;AM;A/S;AUG	7,30419E-45	YP_005398183.1
dcp	1548464	CF;TE;CFZ;CRM;P/T;TO;AM;A/S;AUG	2,06883E-43	YP_005396900.1
UMN798_1612	1574737	CF;TE;CFZ;CRM;P/T;TO;AM;A/S;AUG	2,20543E-42	YP_005396930.1
UMN798_3889	3772654	CF;TE;CFZ;CRM;P/T;TO;AM;A/S;AUG	1,08166E-40	YP_005398910.1
UMN798_1331	1313897	CF;TE;CFZ;CRM;P/T;TO;AM;A/S;AUG	2,91148E-40	YP_005396683.1
UMN798_3890	3774253	CF;TE;CFZ;CRM;P/T;TO;AM;A/S;AUG	3,75244E-39	YP_005398911.1
ycbB	1042088	CF;TE;CFZ;CRM;P/T;TO;AM;A/S;AUG	2,20447E-38	YP_005396423.1
malS	3872045	CF;TE;CFZ;CRM;P/T;TO;AM;A/S;AUG	2,52527E-37	YP_005398986.1
dmsC	1530443	CF;TE;CFZ;CRM;P/T;AM; A/S;AUG	1,22527E-35	YP_005396885.1
UMN798_0020	20803	CF;TE;CFZ;CRM;P/T;TO;AM;A/S;AUG	1,36107E-35	YP_005395531.1
UMN798_1618	1581225	CF;CFZ;CRM;P/T;TO;AM; A/S;AUG	2,96558E-35	YP_005396935.1
UMN798_1617	1580429	CF;CFZ;CRM;P/T;TO;AM; A/S;AUG	3,49041E-35	YP_005396934.1
UMN798_4717	4617911	CF;CFZ;CRM;P/T;TO;AM; A/S;AUG	1,10141E-34	YP_005399615.1
UMN798_0389	401518	CF;CFZ;CRM;P/T;TO;AM; A/S;AUG	1,58664E-34	YP_005395855.1

FDR: determined according to FDR (Benjamini Hochberg) method (Benjamini Hochberg, 1995)

According to certain embodiments of the method of the twelfth and/or thirteenth aspect of the present invention, as well as also of the eighteenth aspect of the present invention, the antibiotic is at least one of CF, CFZ, CRM, P/T, AM, A/S, and AUG and a mutation in at least one of the genes of UMN798_4878, emrA, dcp, UMN798_1612, UMN798_3889, UMN798_1331, UMN798_3890, ycbB, malS, dmsC, UMN798_0020, UMN798_1618, UMN798_1617, UMN798_4717, UMN798_0389 is detected, or a mutation in at least one of the positions of 4779417, 2983118, 1548464, 1574737, 3772654, 1313897, 3774253, 1042088, 3872045, 1530443, 20803, 1581225, 1580429, 4617911, 401518.

According to certain embodiments of the method of the twelfth and/or thirteenth aspect of the present invention, as well as also of the eighteenth aspect of the present invention, the antibiotic is a quinolone antibiotic and a mutation in at least one of the genes listed in Table 7 is detected, e.g. in torS, or a mutation in at least one of the positions (denoted POS in the table) listed in Table 7, e.g. position 4048606.

Table 7: List for quinolone antibiotics (all 2)

gene name	**POS**	**antibiotic**	**p-value (FDR)**	**genbank protein accession number**
gyrA	2373180	CP;LVX	3,47827E-22	YP_005397662.1
gyrA	2373169	CP	2,22208E-21	YP_005397662.1
torS	4048606	TE;CP	1,09183E-10	YP_005399143.1

According to certain embodiments of the method of the twelfth and/or thirteenth aspect of the present invention, as well as also of the eighteenth aspect of the present invention, the antibiotic is at least one of CP and LVX and a mutation in gyrA is detected, or a mutation in position 2373180.

According to certain embodiments of the method of the twelfth and/or thirteenth aspect of the present invention, as well as also of the eighteenth aspect of the present invention, the antibiotic is CP and a mutation in at least one of the genes of gyrA, torS, e.g. torS, is detected, or a mutation in at least one of the positions of 2373169, 4048606, e.g. 4048606.

According to certain embodiments of the method of the twelfth and/or thirteenth aspect of the present invention, as well as also of the eighteenth aspect of the present invention, the antibiotic is an aminoglycoside antibiotic and a mutation in at least one of the genes listed in Table 8 is detected, or a mutation in at least one of the positions (denoted POS in the table) listed in Table 8.

Table 8: List of aminoglycoside antibiotics

gene name	**POS**	**antibiotic**	**p-value (FDR)**	**genbank protein accession number**
UMN798_2909	2840330	TE;GM;A/S;CRM; P/T;TO;AM;AUG	3,28817E-42	YP_005398061.1
bcfB	25574	TE;GM;A/S;CRM; P/T;TO;AM;AUG	2,23532E-40	YP_005395534.1
degQ	3536122	GM;A/S;CRM;P/T; TO;AM;AUG	2,63263E-40	YP_005398691.1
polB	115342	TE;GM;A/S;CRM; P/T;TO;AM;AUG	7,82734E-40	YP_005395610.1
hpcD	1148509	TE;GM;A/S;CRM; P/T;TO;AM;AUG	8,03033E-40	YP_005396519.1
glgS	3379140	TE;GM;A/S;CRM; P/T;TO;AM;AUG	9,6031E-40	YP_005398544.1
plsB	4478134	TE;GM;A/S;CRM; P/T;TO;AM;AUG	9,6031E-40	YP_005399505.1
feoB	3686566	TE;GM;A/S;CRM; P/T;TO;AM;AUG	1,3875E-39	YP_005398837.1
rnfG	1487086	TE;GM;A/S;CRM; P/T;TO;AM;AUG	1,71803E-39	YP_005396848.1

UMN798_3428	3335426	TE;GM;A/S;CRM; P/T;TO;AM;AUG	1,71803E-39	YP_005398501.1
yhjB	3799879	TE;GM;A/S;CRM; P/T;TO;AM;AUG	1,71803E-39	YP_005398931.1
UMN798_1163	1163470	TE;GM;A/S;CRM; P/T;TO;AM;AUG	2,82119E-39	YP_005396535.1
alkA	2208848	TE;GM;A/S;CRM; P/T;TO;AM;AUG	3,17482E-39	YP_005397522.1
nhaA	46695	TE;GM;A/S;CRM; P/T;TO;AM;AUG	3,75244E-39	YP_005395552.1
lspA	56998	TE;GM;A/S;CRM; P/T;TO;AM;AUG	3,75244E-39	YP_005395561.1

According to certain embodiments of the method of the twelfth and/or thirteenth aspect of the present invention, as well as also of the eighteenth aspect of the present invention, the antibiotic is at least one of TO and GM and a mutation in at least one of the genes of UMN798_2909, bcfB, degQ, polB, hpcD, glgS, plsB, feoB, rnfG, UMN798_3428, yhjB, UMN798_1163, alkA, nhaA, lspA is detected, or a mutation in at least one of the positions of 2840330, 25574, 3536122, 115342, 1148509, 3379140, 4478134, 3686566, 1487086, 3335426, 3799879, 1163470, 2208848, 46695, 56998.

According to certain embodiments of the method of the twelfth and/or thirteenth aspect of the present invention, as well as also of the eighteenth aspect of the present invention, the antibiotic is a polyketide antibiotic and a mutation in at least one of the genes listed in Table 9 is detected, or a mutation in at least one of the positions (denoted POS in the table) listed in Table 9.

According to certain embodiments of the method of the twelfth and/or thirteenth aspect of the present invention, as well as also of the eighteenth aspect of the present invention, the

antibiotic is TE and a mutation in at least one of the genes of recN, hemH, UMN798_3428, metE, yijD, UMN798_4831, UMN798_1939, copS, UMN798_0628, UMN798_4878, leuB, recF, emrA, glyQ, dcp, recN, hemH is detected, or a mutation in at least one of the positions of 2833888, 546961, 3334479, 4191057, 4366486, 4724403, 1895588, 1139812, 637461, 4779417, 131219, 4062015, 2983118, 3861998, 1548464, 2833888, 546961.

Table 9: List of polyketides, preferably tetracycline

gene name	**POS**	**antibiotic**	**p-value (FDR)**	**genbank protein accession number**
recN	2833888	CF;TE;A/S;CRM;P/T;TO;AM;AUG	1,03434E-49	YP_005398056.1
hemH	546961	CF;TE;A/S;CRM;P/T;TO;AM;AUG	2,16472E-49	YP_005395984.1
UMN798_3428	3334479	CF;TE;A/S;CRM;P/T;TO;AM;AUG	2,16472E-49	YP_005398501.1
metE	4191057	CF;TE;A/S;CRM;P/T;TO;AM;AUG	2,16472E-49	YP_005399262.1
yijD	4366486	CF;TE;A/S;CRM;P/T;TO;AM;AUG	2,16472E-49	YP_005399415.1
UMN798_4831	4724403	CF;TE;A/S;CRM;P/T;TO;AM;AUG	2,16472E-49	YP_005399714.1
UMN798_1939	1895588	CF;TE;A/S;CRM;P/T;TO;AM;AUG	3,03206E-49	YP_005397219.1
copS	1139812	CF;TE;A/S;CRM;P/T;TO;AM;AUG	1,30591E-48	YP_005396510.1
UMN798_0628	637461	CF;TE;A/S;CRM;P/T;TO;AM;AUG	6,68901E-47	YP_005396068.1
UMN798_4878	4779417	CF;TE;CFZ;CRM;P/T;TO;AM;A/S;AUG	3,07597E-46	YP_005399754.1
leuB	131219	CF;TE;A/S;CRM;P/T;TO;AM;AUG	3,79108E-45	YP_005395624.1
recF	4062015	CF;TE;A/S;CRM;P/T;	5,78617E-45	YP_005399154.1

		TO;AM;AUG		
emrA	2983118	CF;TE;CFZ;CRM;P/T; TO;AM;A/S;AUG	7,30419E-45	YP_005398183.1
glyQ	3861998	CF;TE;A/S;CRM;P/T; TO;AM;AUG	1,47298E-44	YP_005398978.1
dcp	1548464	CF;TE;CFZ;CRM;P/T; TO;AM;A/S;AUG	2,06883E-43	YP_005396900.1

A fourteenth aspect of the present invention is directed to a diagnostic method of determining an infection of a patient with Salmonella species potentially resistant to antimicrobial drug treatment, which can also be described as method of determining an antimicrobial drug, e.g. antibiotic, resistant Salmonella infection of a patient, comprising the steps of:

a) obtaining or providing a sample containing or suspected of containing at least one Salmonella species from the patient;

b) determining the presence of at least one mutation in at least one gene from the group of genes consisting of recN, hemH, UMN798_3428, metE, yijD, UMN798_4831, UMN798_1939, copS, UMN798_0628, UMN798_4878, leuB, recF, glyQ, dcp, thiH, UMN798_1612, UMN798_2909, UMN798_1680, UMN798_4073, yjbC, UMN798_3160, hutU, envC, UMN798_3889, UMN798_1629, bcfB, degQ, UMN798_1331, trg, uvrC, polB, hpcD, UMN798_1628, UMN798_1701, glgS, plsB, yjcC, feoB, misL, dxr, hemF, rnfG, yhjB, UMN798_1163, UMN798_0394, alkA, nhaA, and lspA, preferably from the group of genes consisting of hemH, UMN798_3428, UMN798_4831, UMN798_1939, UMN798_0628, UMN798_4878, leuB, glyQ, dcp, thiH, UMN798_1612, UMN798_2909, UMN798_1680, UMN798_4073, yjbC, UMN798_3160, UMN798_3889, UMN798_1629, bcfB, degQ, UMN798_1331, trg, hpcD, UMN798_1628, UMN798_1701, glgS, yjcC, misL, hemF, rnfG, UMN798_1163, UMN798_0394, and nhaA, wherein the presence of said at least one mutation is

indicative of an antimicrobial drug, e.g. antibiotic, resistant Salmonella infection in said patient.

A fifteenth aspect of the present invention is directed to a method of selecting a treatment of a patient suffering from an antimicrobial drug, e.g. antibiotic, resistant Salmonella infection, comprising the steps of:

a) obtaining or providing a sample containing or suspected of containing at least one Salmonella species from the patient;

b) determining the presence of at least one mutation in at least one gene from the group of genes consisting of recN, hemH, UMN798_3428, metE, yijD, UMN798_4831, UMN798_1939, copS, UMN798_0628, UMN798_4878, leuB, recF, glyQ, dcp, thiH, UMN798_1612, UMN798_2909, UMN798_1680, UMN798_4073, yjbC, UMN798_3160, hutU, envC, UMN798_3889, UMN798_1629, bcfB, degQ, UMN798_1331, trg, uvrC, polB, hpcD, UMN798_1628, UMN798_1701, glgS, plsB, yjcC, feoB, misL, dxr, hemF, rnfG, yhjB, UMN798_1163, UMN798_0394, alkA, nhaA, and lspA, preferably from the group of genes consisting of hemH, UMN798_3428, UMN798_4831, UMN798_1939, UMN798_0628, UMN798_4878, leuB, glyQ, dcp, thiH, UMN798_1612, UMN798_2909, UMN798_1680, UMN798_4073, yjbC, UMN798_3160, UMN798_3889, UMN798_1629, bcfB, degQ, UMN798_1331, trg, hpcD, UMN798_1628, UMN798_1701, glgS, yjcC, misL, hemF, rnfG, UMN798_1163, UMN798_0394, and nhaA, wherein the presence of said at least one mutation is indicative of a resistance to one or more antimicrobial, e.g. antibiotic, drugs;

c) identifying said at least one or more antimicrobial, e.g. antibiotic, drugs; and

d) selecting one or more antimicrobial, e.g. antibiotic, drugs different from the ones identified in step c) and being suitable for the treatment of a Salmonella infection.

Again, in the fourteenth and the fifteenth aspect the steps correspond to those in the first or second aspect, although only a mutation in at least one gene is determined.

A sixteenth aspect of the present invention is directed to a method of treating a patient suffering from an antimicrobial drug, e.g. antibiotic, resistant Salmonella infection, comprising the steps of:

a) obtaining or providing a sample containing or suspected of containing at least one Salmonella species from the patient;

b) determining the presence of at least one mutation in at least one gene from the group of genes consisting of recN, hemH, UMN798_3428, metE, yijD, UMN798_4831, UMN798_1939, copS, UMN798_0628, UMN798_4878, leuB, recF, glyQ, dcp, thiH, UMN798_1612, UMN798_2909, UMN798_1680, UMN798_4073, yjbC, UMN798_3160, hutU, envC, UMN798_3889, UMN798_1629, bcfB, degQ, UMN798_1331, trg, uvrC, polB, hpcD, UMN798_1628, UMN798_1701, glgS, plsB, yjcC, feoB, misL, dxr, hemF, rnfG, yhjB, UMN798_1163, UMN798_0394, alkA, nhaA, and lspA, preferably from the group of genes consisting of hemH, UMN798_3428, UMN798_4831, UMN798_1939, UMN798_0628, UMN798_4878, leuB, glyQ, dcp, thiH, UMN798_1612, UMN798_2909, UMN798_1680, UMN798_4073, yjbC, UMN798_3160, UMN798_3889, UMN798_1629, bcfB, degQ, UMN798_1331, trg, hpcD, UMN798_1628, UMN798_1701, glgS, yjcC, misL, hemF, rnfG, UMN798_1163, UMN798_0394, and nhaA, wherein the presence of said at least one mutation is indicative of a resistance to one or more antimicrobial, e.g. antibiotic, drugs;

c) identifying said at least one or more antimicrobial, e.g. antibiotic, drugs;

d) selecting one or more antimicrobial, e.g. antibiotic, drugs different from the ones identified in step c) and being suitable for the treatment of a Salmonella infection; and

e) treating the patient with said one or more antimicrobial, e.g. antibiotic, drugs.

A seventeenth aspect of the present invention is directed to a method of treating a patient suffering from an antimicrobial drug, e.g. antibiotic, resistant Salmonella infection, comprising the steps of:

a) obtaining or providing a sample containing or suspected of containing at least one Salmonella species from the patient;

b) determining the presence of at least one mutation in at least two genes from the group of genes consisting of recN, hemH, UMN798_3428, metE, yijD, UMN798_4831, UMN798_1939, copS, UMN798_0628, UMN798_4878, leuB, recF, emrA, glyQ, dcp, thiH, UMN798_1612, UMN798_2909, UMN798_1680, UMN798_4073, yjbC, nadB, UMN798_3160, hutU, envC, UMN798_3889, UMN798_1629, bcfB, degQ, UMN798_1331, trg, uvrC, polB, hpcD, UMN798_1628, UMN798_1701, glgS, plsB, yjcC, feoB, misL, dxr, hemF, rnfG, yhjB, UMN798_1163, UMN798_0394, alkA, nhaA, and lspA, wherein the presence of said at least two mutations is indicative of a resistance to one or more antimicrobial, e.g. antibiotic, drugs;

c) identifying said at least one or more antimicrobial, e.g. antibiotic, drugs;

d) selecting one or more antimicrobial, e.g. antibiotic, drugs different from the ones identified in step c) and being suitable for the treatment of a Salmonella infection; and

e) treating the patient with said one or more antimicrobial, e.g. antibiotic, drugs.

An eighteenth aspect of the present invention is directed to a method of treating a patient suffering from an antimicrobial drug, e.g. antibiotic, resistant Salmonella infection, comprising the steps of:

a) obtaining or providing a sample containing or suspected of containing at least one Salmonella species from the patient;

b) determining the presence of at least one mutation in at least two genes from the group of genes listed in Table 5, preferably Table 5a, wherein the presence of said at least two mutations is indicative of a resistance to one or more antimicrobial, e.g. antibiotic, drugs;

c) identifying said at least one or more antimicrobial, e.g. antibiotic, drugs;

d) selecting one or more antimicrobial, e.g. antibiotic, drugs different from the ones identified in step c) and being suitable for the treatment of a Salmonella infection; and

e) treating the patient with said one or more antimicrobial, e.g. antibiotic, drugs.

A nineteenth aspect of the present invention is directed to a method of treating a patient suffering from an antimicrobial drug, e.g. antibiotic, resistant Salmonella infection, comprising the steps of:

a) obtaining or providing a sample containing or suspected of containing at least one Salmonella species from the patient;

b) determining the presence of at least one mutation in at least one gene from the group of genes listed in Table 10, preferably from the group of genes listed in Table 11, wherein the presence of said at least one mutation is indicative of a resistance to one or more antimicrobial, e.g. antibiotic, drugs;

c) identifying said at least one or more antimicrobial, e.g. antibiotic, drugs;

d) selecting one or more antimicrobial, e.g. antibiotic, drugs different from the ones identified in step c) and being suitable for the treatment of a Salmonella infection; and

e) treating the patient with said one or more antimicrobial, e.g. antibiotic, drugs.

Also in the sixteenth to nineteenth aspect of the invention, steps a) to d) are analogous to the steps in the method of the second aspect of the present invention. Step e) can be sufficiently carried out without being restricted and can be done e.g. non-invasively.

A twentieth aspect of the present invention is directed to a diagnostic method of determining an infection of a patient with Salmonella species potentially resistant to antimicrobial drug treatment, which can also be described as method of determining an antimicrobial drug, e.g. antibiotic, resistant Salmonella infection of a patient, comprising the steps of:

a) obtaining or providing a sample containing or suspected of containing at least one Salmonella species from the patient;

b) determining the presence of at least one mutation in at least one gene from the group of genes listed in Table 10, preferably from the group of genes listed in Table 11, wherein the presence of said at least one mutation is indicative of an antimicrobial drug, e.g. antibiotic, resistant Salmonella infection in said patient.

Table 10: List of genes

recN	hemH	UMN798_3428	metE	yijD
UMN798_4831	UMN798_1939	copS	UMN798_0628	UMN798_4878
leuB	recF	pnp	glyQ	dcp
thiH	UMN798_1612	UMN798_2909	UMN798_1680	UMN798_4073
yjbC	ytfF	UMN798_3160	hutU	envC
UMN798_3889	UMN798_1629	bcfB	degQ	UMN798_1331
trg	uvrC	polB	hpcD	UMN798_1628
UMN798_1701	glgS	plsB	yjcC	feoB
misL	dxr	hemF	rnfG	yhjB
UMN798_1163	UMN798_0394	alkA	nhaA	lspA
UMN798_3890	UMN798_1632	malS	dmsC	UMN798_0020
UMN798_1618	UMN798_1617	UMN798_4717	UMN798_0389	kdpD
hisB	UMN798_2727	UMN798_3918	UMN798_4753	UMN798_4936
yhdM	UMN798_0631	UMN798_1337	UMN798_1550	iroE
pheT	hofC	fhuD	torS	UMN798_1114
UMN798_2482	rseB	hycC	ttk	cpdB
UMN798_4882	UMN798_0654	creA	UMN798_2202	dppA
adiY	UMN798_0653	gcvP	UMN798_0179	UMN798_3553
UMN798_4061	hrpB	UMN798_0975		

Table 11: List of genes

UMN798_1632	hemH	UMN798_3428	ytfF	pnp
UMN798_4831	UMN798_1939	UMN798_0654	UMN798_0628	UMN798_4878
leuB	gcvP	UMN798_0975	glyQ	dcp
thiH	UMN798_1612	UMN798_2909	UMN798_1680	UMN798_4073
yjbC	UMN798_4061	UMN798_3160	UMN798_3553	UMN798_0179
UMN798_3889	UMN798_1629	bcfB	degQ	UMN798_1331
trg	UMN798_0653	adiY	hpcD	UMN798_1628
UMN798_1701	glgS	dppA	yjcC	UMN798_2202
misL	creA	hemF	rnfG	UMN798_4882
UMN798_1163	UMN798_0394	cpdB	nhaA	ttk
UMN798_3890	hycC	malS	dmsC	UMN798_0020
UMN798_1618	UMN798_1617	UMN798_4717	UMN798_0389	kdpD

hisB	UMN798_2727	UMN798_3918	UMN798_4753	UMN798_4936
yhdM	UMN798_0631	UMN798_1337	UMN798_1550	iroE
pheT	rseB	UMN798_2482	torS	UMN798_1114

A twenty-first aspect of the present invention is directed to a method of selecting a treatment of a patient suffering from an antimicrobial drug, e.g. antibiotic, resistant Salmonella infection, comprising the steps of:

a) obtaining or providing a sample containing or suspected of containing at least one Salmonella species from the patient;

b) determining the presence of at least one mutation in at least one gene from the group of genes listed in Table 10, preferably from the group of genes listed in Table 11, wherein the presence of said at least one mutation is indicative of a resistance to one or more antimicrobial, e.g. antibiotic, drugs;

c) identifying said at least one or more antimicrobial, e.g. antibiotic, drugs; and

d) selecting one or more antimicrobial, e.g. antibiotic, drugs different from the ones identified in step c) and being suitable for the treatment of a Salmonella infection.

Again, in the twentieth and the twenty-first aspect the steps correspond to those in the first or second aspect, although only a mutation in at least one gene is determined.

Examples

The present invention will now be described in detail with reference to several examples thereof. However, these examples are illustrative and do not limit the scope of the invention.

Example 1

Whole genome sequencing was carried out in addition to classical antimicrobial susceptibility testing of the same isolates for a cohort of 636 specimens. This allowed performing genome wide correlation studies to find genetic variants (e.g. point mutations, small insertions and deletion, larger structural variants, plasmid copy number gains, gene dosage effects) in the genome and plasmids that are significantly correlated to the resistance against one or several drugs. The approach also allows for comparing the relevant sites in the genome to each other.

In the approach the different sources of genetic resistance as well as the different ways of how bacteria can become resistant were covered. By measuring clinical isolates collected in a broad geographical area and across a broad time span of three decades a complete picture going far beyond the rather artificial step of laboratory generated resistance mechanisms was tried to be generated.

To this end, a set of 21 clinically relevant antimicrobial agents with 5 different modes of action was put together, and the minimally inhibitory concentration (MIC) of the 21 drugs for the Salmonella isolates was measured.

The detailed procedure is given in the following:

Bacterial Strains

The inventors selected 636 Salmonella strains, particularly from Salmonella_choleraesuis, Salmonella_dublin, Salmonella_enterica_ssp_arizonae, Salmonella_enterica_ssp_diarizoniae, Salmonella_enteritidis, Salmonella_gallinarum, Salmonella_Group_A, Salmonella_Group_B,

Salmonella_Group_C, Salmonella_Group_D, Salmonella_heidelberg, Salmonella_miami, Salmonella_newport, Salmonella_panama, Salmonella_parahaemolyticus_A, Salmonella_paratyphi_A, Salmonella_paratyphi_B, Salmonella_pullorum, Salmonella_senfienberg, Salmonella_species, Salmonella_species_Lac_--,_ONPG_+, Salmonella_species_Lac_+,_ONPG_+, Salmonella_subgenus_I, Salmonella_subgenus_II, Salmonella_subgenus_IV, Salmonella_subgroup_I_Suc+, Salmonella_tennessee, and Salmonella_typhi, from the microbiology strain collection at Siemens Healthcare Diagnostics (West Sacramento, CA) for susceptibility testing and whole genome sequencing.

Antimicrobial Susceptibility Testing (AST) Panels

Frozen reference AST panels were prepared following Clinical Laboratory Standards Institute (CLSI) recommendations. The following antimicrobial agents (with µg/ml concentrations shown in parentheses) were included in the panels: Amoxicillin/K Clavulanate (0.5/0.25-64/32), Ampicillin (0.25-128), Ampicillin/Sulbactam (0.5/0.25-64/32), Aztreonam (0.25-64), Cefazolin (0.5-32), Cefepime (0.25-64), Cefotaxime (0.25-128), Ceftazidime (0.25-64), Ceftriaxone (0.25-128), Cefuroxime (1-64), Cephalothin (1-64), Ciprofloxacin (0.015-8), Ertepenem (0.12-32), Gentamicin (0.12-32), Imipenem (0.25-32), Levofloxacin (0.25-16), Meropenem (0.12-32), Piperacillin/Tazobactam (0.25/4-256/4), Tetracycline (0.5-64), Tobramycin (0.12-32), and Trimethoprim/Sulfamethoxazole (0.25/4.7-32/608). Prior to use with clinical isolates, AST panels were tested with QC strains. AST panels were considered acceptable for testing with clinical isolates when the QC results met QC ranges described by CLSI16.

Inoculum Preparation

Isolates were cultured on trypticase soy agar with 5% sheep blood (BBL, Cockeysville, Md.) and incubated in ambient air at 35±1°C for 18-24 h. Isolated colonies (4-5 large colonies or 5-10 small colonies) were transferred to a 3 ml Sterile Inoculum Water (Siemens) and emulsified to a final turbidity of a 0.5 McFarland standard. 2 ml of this suspension was added to 25 ml Inoculum Water with Pluronic-F (Siemens). Using the Inoculator (Siemens) specific for frozen AST panels, 5 µl of the cell suspension was transferred to each well of the AST panel. The inoculated AST panels were incubated in ambient air at 35±1°C for 16-20 h. Panel results were read visually, and minimal inhibitory concentrations (MIC) were determined.

DNA extraction

Four streaks of each Gram-negative bacterial isolate cultured on trypticase soy agar containing 5% sheep blood and cell suspensions were made in sterile 1.5 ml collection tubes containing 50 µl Nuclease-Free Water (AM9930, Life Technologies). Bacterial isolate samples were stored at -20 °C until nucleic acid extraction. The Tissue Preparation System (TPS) (096D0382-02_01_B, Siemens) and the VERSANT® Tissue Preparation Reagents (TPR) kit (10632404B, Siemens) were used to extract DNA from these bacterial isolates. Prior to extraction, the bacterial isolates were thawed at room temperature and were pelleted at 2000 G for 5 seconds. The DNA extraction protocol DNAext was used for complete total nucleic acid extraction of 48 isolate samples and eluates, 50 µl each, in 4 hours. The total nucleic acid eluates were then transferred into 96-Well qPCR Detection Plates (401341, Agilent Technologies) for RNase A digestion, DNA quantitation, and plate DNA concentration standardization processes. RNase A (AM2271, Life Technologies) which was diluted in nuclease-free water

following manufacturer's instructions was added to 50 µl of the total nucleic acid eluate for a final working concentration of 20 µg/ml. Digestion enzyme and eluate mixture were incubated at 37°C for 30 minutes using Siemens VERSANT® Amplification and Detection instrument. DNA from the RNase digested eluate was quantitated using the Quant-iT™ PicoGreen dsDNA Assay (P11496, Life Technologies) following the assay kit instruction, and fluorescence was determined on the Siemens VERSANT® Amplification and Detection instrument. Data analysis was performed using Microsoft® Excel 2007. 25 µl of the quantitated DNA eluates were transferred into a new 96-Well PCR plate for plate DNA concentration standardization prior to library preparation. Elution buffer from the TPR kit was used to adjust DNA concentration. The standardized DNA eluate plate was then stored at -80°C until library preparation.

Next Generation Sequencing

Prior to library preparation, quality control of isolated bacterial DNA was conducted using a Qubit 2.0 Fluorometer (Qubit dsDNA BR Assay Kit, Life Technologies) and an Agilent 2200 TapeStation (Genomic DNA ScreenTape, Agilent Technologies). NGS libraries were prepared in 96 well format using NexteraXT DNA Sample Preparation Kit and NexteraXT Index Kit for 96 Indexes (Illumina) according to the manufacturer's protocol. The resulting sequencing libraries were quantified in a qPCR-based approach using the KAPA SYBR FAST qPCR MasterMix Kit (Peqlab) on a ViiA 7 real time PCR system (Life Technologies). 96 samples were pooled per lane for paired-end sequencing (2x 100bp) on Illumina Hiseq2000 or Hiseq2500 sequencers using TruSeq PE Cluster v3 and TruSeq SBS v3 sequncing chemistry (Illumina). Basic sequencing quality parameters were determined using the FastQC quality control

tool for high throughput sequence data (Babraham Bioinformatics Institute).

Data analysis

Raw paired-end sequencing data for the 636 Salmonella samples were mapped against the Salmonella reference (NC_017046) with BWA 0.6.1.20. The resulting SAM files were sorted, converted to BAM files, and PCR duplicates were marked using the Picard tools package 1.104 (http://picard.sourceforge.net/). The Genome Analysis Toolkit 3.1.1 (GATK)21 was used to call SNPs and indels for blocks of 200 Salmonella samples (parameters: -ploidy 1 -glm BOTH -stand_call_conf 30 -stand_emit_conf 10). VCF files were combined into a single file and quality filtering for SNPs was carried out (QD < 2.0 || FS > 60.0 || MQ < 40.0) and indels (QD < 2.0 || FS > 200.0). Detected variants were annotated with SnpEff22 to predict coding effects. For each annotated position, genotypes of all Salmonella samples were considered. Salmonella samples were split into two groups, low resistance group (having lower MIC concentration for the considered drug), and high resistance group (having higher MIC concentrations) with respect to a certain MIC concentration (breakpoint). To find the best breakpoint all thresholds were evaluated and p-values were computed with Fisher's exact test relying on a 2x2 contingency table (number of Salmonella samples having the reference or variant genotype vs. number of samples belonging to the low and high resistance group). The best computed breakpoint was the threshold yielding the lowest p-value for a certain genomic position and drug. For further analyses positions with non-synonymous alterations and p-value < 10^{-9} were considered.

Since a potential reason for drug resistance is gene duplication, gene dose dependency was evaluated. For each sample the

genomic coverage for each position was determined using BED Tools. Gene ranges were extracted from the reference assembly NC_017046.gff and the normalized median coverage per gene was calculated. To compare low- and high-resistance isolates the best area under the curve (AUC) value was computed. Groups of at least 20% of all samples having a median coverage larger than zero for that gene and containing more than 15 samples per group were considered in order to exclude artifacts and cases with AUC > 0.75 were further evaluated.

To include data on the different ways how resistance mechanisms are acquired Salmonella isolates collected over more than three decades were analyzed such that also horizontal gene transfer could potentially be discovered.

In detail, the following steps were carried out:
Salmonella strains to be tested were seeded on agar plates and incubated under growth conditions for 24 hours. Then, colonies were picked and incubated in growth medium in the presence of a given antibiotic drug in dilution series under growth conditions for 16-20 hours. Bacterial growth was determined by observing turbidity.

Next mutations were searched that are highly correlated with the results of the phenotypic resistance test.

For sequencing, samples were prepared using a Nextera library preparation, followed by multiplexed sequencing using the Illuminat HiSeq 2500 system, paired end sequencing. Data were mapped with BWA (Li H. and Durbin R. (2010) Fast and accurate long-read alignment with Burrows-Wheeler Transform. Bioinformatics, Epub. [PMID: 20080505])and SNP were called using samtools (Li H.*, Handsaker B.*, Wysoker A., Fennell T., Ruan

J., Homer N., Marth G., Abecasis G., Durbin R. and 1000 Genome Project Data Processing Subgroup (2009) The Sequence alignment/map (SAM) format and SAMtools. Bioinformatics, 25, 2078-9. [PMID: 19505943]).

As reference genome, NC_017046 as annotated at the NCBI was determined as best suited.

The mutations were matched to the genes and the amino acid changes were calculated. Using different algorithms (SVM, homology modeling) mutations leading to amino acid changes with likely pathogenicity / resistance were calculated.

In total, whole genomes and plasmids of 636 different clinical isolates of Salmonella species were sequenced, and classical antimicrobial susceptibility testing (AST) against 21 therapy forms as described above was performed for all organisms. From the classical AST a table with 636 rows (isolates) and 21 columns (MIC values for 21 drugs) was obtained. Each table entry contained the MIC for the respective isolate and the respective drug. The genetic data were mapped to different reference genomes of Salmonella that have been annotated at the NCBI (http://www.ncbi.nlm.nih.gov/), and the best reference was chosen as template for the alignment - NC_017046 as annotated at the NCBI. Additionally, assemblies were carried out and it was verified that the sequenced genomes fulfil all quality criteria to become reference genomes.

Next, genetic variants were evaluated. This approach resulted in a table with the genetic sites in columns and the same isolates in 636 rows. Each table entry contained the genetic determinant at the respective site (A, C, T, G, small insertions and deletions, ...) for the respective isolate.

In a next step different statistical tests were carried out

1) For comparing resistance / susceptibility to genetic sites we calculated contingency tables and determined the significance using Fishers test
2) For comparing different sites to each other we calculated the correlation between different genetic sites
3) For detecting gene dosage effects, e.g. loss or gain of genes (in the genome or on plasmids) we calculated the coverage (i.e. how many read map to the current position) at each site for resistant and not resistant isolates.

From the data, first the 50 genes with the best p-value were chosen for the list of mutations as well as the list of correlated antibiotic resistance, representing Tables 1 and 2.

A full list of all genetic sites, drugs, drug classes, affected genes etc. is provided in Tables 3 and 4a, 4b and 4c, wherein Table 3 corresponds to Table 1 and represents the genes having the lowest p-values after determining mutations in the genes, and Table 4, respectively Tables 4a, 4b and 4c correspond to Table 2 and represent the genes having the lowest p-values after correlating the mutations with antibiotic resistance for the respective antibiotics.

In addition, the data with the best p-values for each antibiotic class with the most antibiotic drugs as well as each antibiotic, respectively, were evaluated, being disclosed in Tables 5 – 9.

In Tables 3 - 9 the columns are designated as follows:
Gene name: affected gene;

POS: genomic position of the SNP / variant in the Salmonella reference genome (see above);

p-value: significance value calculated using Fishers exact test (determined according to FDR (Benjamini Hochberg) method (Benjamini Hochberg, 1995));

genbank protein accession number: (NCBI) Accession number of the corresponding protein of the genes

Table 3: Detailed results for the genes in Example 1 (corresponding to Table 1)

POS	drug class	#drug classes	p-value	gene name	genbank protein accession number
2833888	polyketide (tetracycline); Lactams;aminoglycoside	3	1,03434E-49	recN	YP_005398056.1
546961	polyketide (tetracycline); Lactams;aminoglycoside	3	2,16472E-49	hemH	YP_005395984.1
3334479	polyketide (tetracycline); Lactams;aminoglycoside	3	2,16472E-49	UMN798_3428	YP_005398501.1
4191057	polyketide (tetracycline); Lactams;aminoglycoside	3	2,16472E-49	metE	YP_005399262.1
4366486	polyketide (tetracycline); Lactams;aminoglycoside	3	2,16472E-49	yijD	YP_005399415.1
4724403	polyketide (tetracycline); Lactams;aminoglycoside	3	2,16472E-49	UMN798_4831	YP_005399714.1
1895588	polyketide (tetracycline); Lactams;aminoglycoside	3	3,03206E-49	UMN798_1939	YP_005397219.1
1139812	polyketide (tetracycline); Lactams;aminoglycoside	3	1,30591E-48	copS	YP_005396510.1
637461	polyketide (tetracycline); Lactams;aminoglycoside	3	6,68901E-47	UMN798_0628	YP_005396068.1
4779417	polyketide (tetracycline); Lactams;aminoglycoside	3	3,07597E-46	UMN798_4878	YP_005399754.1

131219	polyketide (tetracycline); Lac-tams;aminoglycoside	3	3,79108E-45	leuB	YP_005395624.1
4062015	polyketide (tetracycline); Lac-tams;aminoglycoside	3	5,78617E-45	recF	YP_005399154.1
2983118	polyketide (tetracycline); Lac-tams;aminoglycoside	3	7,30419E-45	emrA	YP_005398183.1
3861998	polyketide (tetracycline); Lac-tams;aminoglycoside	3	1,47298E-44	glyQ	YP_005398978.1
1548464	polyketide (tetracycline); Lac-tams;aminoglycoside	3	2,06883E-43	dcp	YP_005396900.1
4397111	polyketide (tetracycline); Lac-tams;aminoglycoside	3	3,53351E-43	thiH	YP_005399435.1
1574737	polyketide (tetracycline); Lac-tams;aminoglycoside	3	2,20543E-42	UMN798_1612	YP_005396930.1
2840330	polyketide (tetracycline); Lac-tams;aminoglycoside	3	3,28817E-42	UMN798_2909	YP_005398061.1
1650934	polyketide (tetracycline); Lac-tams;aminoglycoside	3	4,50718E-42	UMN798_1680	YP_005396992.1
3966175	polyketide (tetracycline); Lac-tams;aminoglycoside	3	5,93558E-42	UMN798_4073	YP_005399073.1
4436188	polyketide (tetracycline); Lac-tams;aminoglycoside	3	5,9971E-42	yjbC	YP_005399464.1
2780306	polyketide (tetracycline); Lac-	3	1,5881E-41	nadB	YP_005398008.1

	tams;aminoglycoside				
3075942	polyketide (tetracycline); Lac-tams;aminoglycoside	3	1,90155E-41	UMN798_3160	YP_005398272.1
855087	polyketide (tetracycline); Lac-tams;aminoglycoside	3	7,25896E-41	hutU	YP_005396263.1
3582301	polyketide (tetracycline); Lac-tams;aminoglycoside	3	7,67646E-41	envC	YP_005398732.1
3772654	polyketide (tetracycline); Lac-tams;aminoglycoside	3	1,08166E-40	UMN798_3889	YP_005398910.1
1590194	polyketide (tetracycline); Lac-tams;aminoglycoside	3	1,52394E-40	UMN798_1629	YP_005396944.1
25574	polyketide (tetracycline); Lac-tams;aminoglycoside	3	2,23532E-40	bcfB	YP_005395534.1
3536122	aminoglycoside;Lactams	2	2,63263E-40	degQ	YP_005398691.1
1313897	polyketide (tetracycline); Lac-tams;aminoglycoside	3	2,91148E-40	UMN798_1331	YP_005396683.1
1673475	polyketide (tetracycline); Lac-tams;aminoglycoside	3	3,8896E-40	trg	YP_005397014.1
1994028	polyketide (tetracycline); Lac-tams;aminoglycoside	3	3,8896E-40	uvrC	YP_005397309.1
115342	polyketide (tetracycline); Lac-tams;aminoglycoside	3	7,82734E-40	polB	YP_005395610.1
1148509	polyketide (tetracycline); Lac-	3	8,03033E-40	hpcD	YP_005396519.1

	tams;aminoglycoside				
1589819	polyketide (tetracycline); Lac-tams;aminoglycoside	3	8,03033E-40	UMN798_1628	YP_005396943.1
1672517	polyketide (tetracycline); Lac-tams;aminoglycoside	3	8,09948E-40	UMN798_1701	YP_005397013.1
3379140	polyketide (tetracycline); Lac-tams;aminoglycoside	3	9,6031E-40	glgS	YP_005398544.1
4478134	polyketide (tetracycline); Lac-tams;aminoglycoside	3	9,6031E-40	plsB	YP_005399505.1
4523874	polyketide (tetracycline); Lac-tams;aminoglycoside	3	1,24792E-39	yjcC	YP_005399531.1
3686566	polyketide (tetracycline); Lac-tams;aminoglycoside	3	1,3875E-39	feoB	YP_005398837.1
3976726	polyketide (tetracycline); Lac-tams;aminoglycoside	3	1,3875E-39	misL	YP_005399078.1
258966	polyketide (tetracycline); Lac-tams;aminoglycoside	3	1,4329E-39	dxr	YP_005395732.1
2558379	polyketide (tetracycline); Lac-tams;aminoglycoside	3	1,45191E-39	hemF	YP_005397834.1
1487086	polyketide (tetracycline); Lac-tams;aminoglycoside	3	1,71803E-39	rnfG	YP_005396848.1
3799879	polyketide (tetracycline); Lac-tams;aminoglycoside	3	1,71803E-39	yhjB	YP_005398931.1

1163470	polyketide (tetracycline); Lac-tams;aminoglycoside	3	2,82119E-39	UMN798_1163	YP_005396535.1
409259	polyketide (tetracycline); Lac-tams;aminoglycoside	3	3,02872E-39	UMN798_0394	YP_005395859.1
2208848	polyketide (tetracycline); Lac-tams;aminoglycoside	3	3,17482E-39	alkA	YP_005397522.1
46695	polyketide (tetracycline); Lac-tams;aminoglycoside	3	3,75244E-39	nhaA	YP_005395552.1
56998	polyketide (tetracycline); Lac-tams;aminoglycoside	3	3,75244E-39	lspA	YP_005395561.1
3582354	polyketide (tetracycline); Lac-tams;aminoglycoside	3	4,47627E-40	envC	YP_005398732.1
3335426	polyketide (tetracycline); Lac-tams;aminoglycoside	3	1,71803E-39	UMN798_3428	YP_005398501.1
1673444	polyketide (tetracycline); Lac-tams;aminoglycoside	3	2,82119E-39	trg	YP_005397014.1

Table 4a: Detailed results for the genes in Example 1 (corresponding to Table 2)

POS	drug	#drugs	drug class	#drug classes
2833888	CF;TE;A/S;CRM;P/T;TO;AM;AUG	8	Polyketide*;Lactams;aminoglycoside	3
546961	CF;TE;A/S;CRM;P/T;TO;AM;AUG	8	Polyketide*;Lactams;aminoglycoside	3
3334479	CF;TE;A/S;CRM;P/T;TO;AM;AUG	8	Polyketide*;Lactams;aminoglycoside	3
4191057	CF;TE;A/S;CRM;P/T;TO;AM;AUG	8	Polyketide*;Lactams;aminoglycoside	3
4366486	CF;TE;A/S;CRM;P/T;TO;AM;AUG	8	Polyketide*;Lactams;aminoglycoside	3
4724403	CF;TE;A/S;CRM;P/T;TO;AM;AUG	8	Polyketide*;Lactams;aminoglycoside	3
1895588	CF;TE;A/S;CRM;P/T;TO;AM;AUG	8	Polyketide*;Lactams;aminoglycoside	3
1139812	CF;TE;A/S;CRM;P/T;TO;AM;AUG	8	Polyketide*;Lactams;aminoglycoside	3
637461	CF;TE;A/S;CRM;P/T;TO;AM;AUG	8	Polyketide*;Lactams;aminoglycoside	3
4779417	CF;TE;CFZ;CRM;P/T;TO;AM;A/S;AUG	9	Polyketide*;Lactams;aminoglycoside	3
131219	CF;TE;A/S;CRM;P/T;TO;AM;AUG	8	Polyketide*;Lactams;aminoglycoside	3
4062015	CF;TE;A/S;CRM;P/T;TO;AM;AUG	8	Polyketide*;Lactams;aminoglycoside	3
2983118	CF;TE;CFZ;CRM;P/T;TO;AM;A/S;AUG	9	Polyketide*;Lactams;aminoglycoside	3
3861998	CF;TE;A/S;CRM;P/T;TO;AM;AUG	8	Polyketide*;Lactams;aminoglycoside	3
1548464	CF;TE;CFZ;CRM;P/T;TO;AM;A/S;AUG	9	Polyketide*;Lactams;aminoglycoside	3
4397111	CF;TE;A/S;CRM;P/T;TO;AM;AUG	8	Polyketide*;Lactams;aminoglycoside	3
1574737	CF;TE;CFZ;CRM;P/T;TO;AM;A/S;AUG	9	Polyketide*;Lactams;aminoglycoside	3
2840330	TE;GM;A/S;CRM;P/T;TO;AM;AUG	8	Polyketide*;Lactams;aminoglycoside	3
1650934	CF;TE;A/S;CRM;P/T;TO;AM;AUG	8	Polyketide*;Lactams;aminoglycoside	3
3966175	TE;A/S;CRM;P/T;TO;AM;AUG	7	Polyketide*;Lactams;aminoglycoside	3
4436188	CF;TE;A/S;CRM;P/T;TO;AM;AUG	8	Polyketide*;Lactams;aminoglycoside	3

2780306	TE;A/S;CRM;P/T;TO;AM;AUG	7	Polyketide*;Lactams;aminoglycoside	3
3075942	CF;TE;A/S;CRM;P/T;TO;AM;AUG	8	Polyketide*;Lactams;aminoglycoside	3
855087	CF;TE;A/S;CRM;P/T;TO;AM;AUG	8	Polyketide*;Lactams;aminoglycoside	3
3582301	CF;TE;A/S;CRM;P/T;TO;AUG	7	Polyketide*;Lactams;aminoglycoside	3
3772654	CF;TE;CFZ;CRM;P/T;TO;AM;A/S;AUG	9	Polyketide*;Lactams;aminoglycoside	3
1590194	CF;TE;A/S;CRM;P/T;TO;AUG	7	Polyketide*;Lactams;aminoglycoside	3
25574	TE;GM;A/S;CRM;P/T;TO;AM;AUG	8	Polyketide*;Lactams;aminoglycoside	3
3536122	GM;A/S;CRM;P/T;TO;AM;AUG	7	aminoglycoside;Lactams	2
1313897	CF;TE;CFZ;CRM;P/T;TO;AM;A/S;AUG	9	Polyketide*;Lactams;aminoglycoside	3
1673475	CF;TE;A/S;CRM;P/T;TO;AUG	7	Polyketide*;Lactams;aminoglycoside	3
1994028	CF;TE;A/S;CRM;P/T;TO;AM;AUG	8	Polyketide*;Lactams;aminoglycoside	3
115342	TE;GM;A/S;CRM;P/T;TO;AM;AUG	8	Polyketide*;Lactams;aminoglycoside	3
1148509	TE;GM;A/S;CRM;P/T;TO;AM;AUG	8	Polyketide*;Lactams;aminoglycoside	3
1589819	CF;TE;A/S;CRM;P/T;TO;AUG	7	Polyketide*;Lactams;aminoglycoside	3
1672517	CF;TE;A/S;CRM;P/T;TO;AUG	7	Polyketide*;Lactams;aminoglycoside	3
3379140	TE;GM;A/S;CRM;P/T;TO;AM;AUG	8	Polyketide*;Lactams;aminoglycoside	3
4478134	TE;GM;A/S;CRM;P/T;TO;AM;AUG	8	Polyketide*;Lactams;aminoglycoside	3
4523874	TE;A/S;CRM;P/T;TO;AM;AUG	7	Polyketide*;Lactams;aminoglycoside	3
3686566	TE;GM;A/S;CRM;P/T;TO;AM;AUG	8	Polyketide*;Lactams;aminoglycoside	3
3976726	CF;TE;A/S;CRM;P/T;TO;AM;AUG	8	Polyketide*;Lactams;aminoglycoside	3
258966	TE;A/S;CRM;P/T;TO;AM;AUG	7	Polyketide*;Lactams;aminoglycoside	3
2558379	TE;A/S;CRM;P/T;TO;AM;AUG	7	Polyketide*;Lactams;aminoglycoside	3
1487086	TE;GM;A/S;CRM;P/T;TO;AM;AUG	8	Polyketide*;Lactams;aminoglycoside	3

3799879	TE;GM;A/S;CRM;P/T;TO;AM;AUG	8	Polyketide*;Lactams;aminoglycoside	3
1163470	TE;GM;A/S;CRM;P/T;TO;AM;AUG	8	Polyketide*;Lactams;aminoglycoside	3
409259	TE;A/S;CRM;P/T;TO;AM;AUG	7	Polyketide*;Lactams;aminoglycoside	3
2208848	TE;GM;A/S;CRM;P/T;TO;AM;AUG	8	Polyketide*;Lactams;aminoglycoside	3
46695	TE;GM;A/S;CRM;P/T;TO;AM;AUG	8	Polyketide*;Lactams;aminoglycoside	3
56998	TE;GM;A/S;CRM;P/T;TO;AM;AUG	8	Polyketide*;Lactams;aminoglycoside	3
3582354	CF;TE;A/S;CRM;P/T;TO;AUG	7	Polyketide*;Lactams;aminoglycoside	3
3335426	TE;GM;A/S;CRM;P/T;TO;AM;AUG	8	Polyketide*;Lactams;aminoglycoside	3
1673444	CF;TE;A/S;CRM;P/T;TO;AUG	7	Polyketide*;Lactams;aminoglycoside	3

*: (tetracycline)

Table 4b: Detailed results for the genes in Example 1 (corresponding to Table 2, continued)

POS	best drug	#significant Lactams	#significant fluoroquinolones	#significant aminoglycosides	#significant polyketide (tetracycline)	#significant other (benzene derived)/ sulfonamide
2833888	AUG	6	0	1	1	0
546961	AUG	6	0	1	1	0
3334479	AUG	6	0	1	1	0
4191057	AUG	6	0	1	1	0
4366486	AUG	6	0	1	1	0
4724403	AUG	6	0	1	1	0
1895588	AUG	6	0	1	1	0
1139812	AUG	6	0	1	1	0
637461	AUG	6	0	1	1	0
4779417	AUG	7	0	1	1	0
131219	AUG	6	0	1	1	0
4062015	AUG	6	0	1	1	0
2983118	AUG	7	0	1	1	0
3861998	AUG	6	0	1	1	0
1548464	AUG	7	0	1	1	0
4397111	AUG	6	0	1	1	0
1574737	AUG	7	0	1	1	0
2840330	AUG	5	0	2	1	0
1650934	AUG	6	0	1	1	0

3966175	AUG	5	0	1	1	0	
4436188	AUG	6	0	1	1	0	
2780306	AUG	5	0	1	1	0	
3075942	AUG	6	0	1	1	0	
855087	AUG	6	0	1	1	0	
3582301	P/T	5	0	1	1	0	
3772654	AUG	7	0	1	1	0	
1590194	P/T	5	0	1	1	0	
25574	AUG	5	0	2	1	0	
3536122	AUG	5	0	2	0	0	
1313897	AUG	7	0	1	1	0	
1673475	P/T	5	0	1	1	0	
1994028	AUG	6	0	1	1	0	
115342	TO	5	0	2	1	0	
1148509	TO	5	0	2	1	0	
1589819	P/T	5	0	1	1	0	
1672517	P/T	5	0	1	1	0	
3379140	TO	5	0	2	1	0	
4478134	TO	5	0	2	1	0	
4523874	AUG	5	0	1	1	0	
3686566	AUG	5	0	2	1	0	
3976726	AUG	6	0	1	1	0	
258966	AUG	5	0	1	1	0	

2558379	AUG	5	0	1	1	0
1487086	TO	5	0	2	1	0
3799879	TO	5	0	2	1	0
1163470	TO	5	0	2	1	0
409259	AUG	5	0	1	1	0
2208848	TO	5	0	2	1	0
46695	TO	5	0	2	1	0
56998	TO	5	0	2	1	0
3582354	P/T	5	0	1	1	0
3335426	TO	5	0	2	1	0
1673444	P/T	5	0	1	1	0

Table 4c: Detailed results for the genes in Example 1 (corresponding to Table 2, continued)

POS	p-value	gene name	genbank protein accession number
2833888	1,03434E-49	recN	YP_005398056.1
546961	2,16472E-49	hemH	YP_005395984.1
3334479	2,16472E-49	UMN798_3428	YP_005398501.1
4191057	2,16472E-49	metE	YP_005399262.1
4366486	2,16472E-49	yijD	YP_005399415.1
4724403	2,16472E-49	UMN798_4831	YP_005399714.1
1895588	3,03206E-49	UMN798_1939	YP_005397219.1
1139812	1,30591E-48	copS	YP_005396510.1
637461	6,68901E-47	UMN798_0628	YP_005396068.1
4779417	3,07597E-46	UMN798_4878	YP_005399754.1

131219	3,79108E-45	leuB	YP_005395624.1
4062015	5,78617E-45	recF	YP_005399154.1
2983118	7,30419E-45	emrA	YP_005398183.1
3861998	1,47298E-44	glyQ	YP_005398978.1
1548464	2,06883E-43	dcp	YP_005396900.1
4397111	3,53351E-43	thiH	YP_005399435.1
1574737	2,20543E-42	UMN798_1612	YP_005396930.1
2840330	3,28817E-42	UMN798_2909	YP_005398061.1
1650934	4,50718E-42	UMN798_1680	YP_005396992.1
3966175	5,93558E-42	UMN798_4073	YP_005399073.1
4436188	5,9971E-42	yjbC	YP_005399464.1
2780306	1,5881E-41	nadB	YP_005398008.1
3075942	1,90155E-41	UMN798_3160	YP_005398272.1
855087	7,25896E-41	hutU	YP_005396263.1
3582301	7,67646E-41	envC	YP_005398732.1
3772654	1,08166E-40	UMN798_3889	YP_005398910.1
1590194	1,52394E-40	UMN798_1629	YP_005396944.1
25574	2,23532E-40	bcfB	YP_005395534.1
3536122	2,63263E-40	degQ	YP_005398691.1
1313897	2,91148E-40	UMN798_1331	YP_005396683.1
1673475	3,8896E-40	trg	YP_005397014.1
1994028	3,8896E-40	uvrC	YP_005397309.1
115342	7,82734E-40	polB	YP_005395610.1

1148509	8,03033E-40	hpcD	YP_005396519.1
1589819	8,03033E-40	UMN798_1628	YP_005396943.1
1672517	8,09948E-40	UMN798_1701	YP_005397013.1
3379140	9,6031E-40	glgS	YP_005398544.1
4478134	9,6031E-40	plsB	YP_005399505.1
4523874	1,24792E-39	yjcC	YP_005399531.1
3686566	1,3875E-39	feoB	YP_005398837.1
3976726	1,3875E-39	misL	YP_005399078.1
258966	1,4329E-39	dxr	YP_005395732.1
2558379	1,45191E-39	hemF	YP_005397834.1
1487086	1,71803E-39	rnfG	YP_005396848.1
3799879	1,71803E-39	yhjB	YP_005398931.1
1163470	2,82119E-39	UMN798_1163	YP_005396535.1
409259	3,02872E-39	UMN798_0394	YP_005395859.1
2208848	3,17482E-39	alkA	YP_005397522.1
46695	3,75244E-39	nhaA	YP_005395552.1
56998	3,75244E-39	lspA	YP_005395561.1
3582354	4,47627E-40	envC	YP_005398732.1
3335426	1,71803E-39	UMN798_3428	YP_005398501.1
1673444	2,82119E-39	trg	YP_005397014.1

Also the antibiotic/drug classes, the number of significant antibiotics correlated to the mutations (over all antibiotics or over certain classes), as well as the correlated antibiotics are denoted in the Tables.

The p-value was calculated using the Fisher exact test based on contingency table with 4 fields: #samples Resistant / wild type; #samples Resistant / mutant; #samples not Resistant / wild type; #samples not Resistant / mutant

The test is based on the distribution of the samples in the 4 fields. Even distribution indicates no significance, while clustering into two fields indicates significance.

The following results were obtained

- A total of approx. 55.800 different correlations between genetic sites and anti-microbial agents were detected (p-value < 10^{-10}).
- The biggest part of these were point mutations (i.e. single base exchanges)
- The highest significance that was reached was 10^{-49}
- Besides these, insertions or deletions of up to four bases were discovered
- Further, potential genetic tests for four different drug classes relating to resistances were discovered
    - β-lactams (includes Penicillins, Cephalosporins, Carbapenems, Monobactams)
    - Quinolones, particularly Fluoroquinolones
    - Aminoglycosides
    - Polyketides, particularly Tetracyclines
- Mutations were observed in 3,874 different genes

While in the tables only the best mutations in each gene are represented, a manifold of different SNPs has been found for each gene. Examples for multiple SNPs for two of the genes given in Table 3 are shown in the following Tables 12 and 13.

Table 12: Statistically significant SNPs in gene bcfB (genbank protein accession number YP_005395534.1) (headers as in Tables 3 and 4, respectively)

POS	drug	#drugs	drug class	best drug	p-value
25565	CRM;P/T;AUG	3	Lactams	CRM	1.3351E-016
25114	TE;TO;CRM	3	Polyketide*;Lactams; aminoglycoside	TE	2.7943E-011
25182	TO;GM	2	aminoglycoside	GM	1.3816E-013
25616	CF;TE;A/S;CRM; P/T;TO;AUG	7	Polyketide*;Lactams; aminoglycoside	P/T	1.9318E-035
25602	CRM;P/T;AUG	3	Lactams	CRM	1.3351E-016
25617	CF;TE;A/S;CRM; P/T;TO;AUG	7	Polyketide*;Lactams; aminoglycoside	P/T	8.6136E-035
25575	TO;GM	2	aminoglycoside	TO	5.0004E-015
25722	AUG	1	Lactams	AUG	8.5373E-010
25559	CRM;P/T;AUG	3	Lactams	CRM	2.9472E-016
25120	TE;P/T;TO;CRM; AUG	5	Polyketide*;Lactams; aminoglycoside	P/T	1.1032E-025
25455	A/S;AM;AUG	3	Lactams	AM	7.4430E-022
25052	A/S;P/T;AUG	3	Lactams	P/T	3.1479E-016
25598	A/S;P/T;AUG	3	Lactams	P/T	9.4153E-016
25272	CF;TE;A/S;CRM; P/T;TO;AUG	7	Polyketide*;Lactams; aminoglycoside	P/T	1.9318E-035
25109	TE;P/T;TO;CRM; AUG	5	Polyketide*;Lactams; aminoglycoside	P/T	2.6436E-024
25643	TE;P/T;TO;CRM; AUG	5	Polyketide*;Lactams; aminoglycoside	P/T	3.1008811907 5937E-026
25043	CRM;P/T;AUG	3	Lactams	CRM	1.3860E-016
25787	TE;P/T;TO;CRM; AUG	5	Polyketide*;Lactams; aminoglycoside	P/T	3.1009E-026
25703	CRM;P/T;AUG	3	Lactams	CRM	1.3860E-016
25712	TO;GM	2	aminoglycoside	TO	8.3819E-022

25122	TE;TO;CRM	3	Polyketide*;Lactams; aminoglycoside	TE	1.1259E-012
25071	CRM;P/T;AUG	3	Lactams	P/T	4.4353E-018
25089	A/S;P/T;AUG	3	Lactams	P/T	7.961E-015
25551	CRM;P/T;AUG	3	Lactams	CRM	1.3351E-016
25557	CRM;P/T;AUG	3	Lactams	P/T	4.3112E-017
25655	A/S;P/T;AUG	3	Lactams	P/T	1.4988E-015
25123	TE;A/S;CRM;P/T ;TO;AUG	6	Polyketide*;Lactams; aminoglycoside	P/T	4.8019E-015
25599	CF;TE;A/S;CRM; P/T;TO;AUG	7	Polyketide*;Lactams; aminoglycoside	P/T	1.9318E-035
25053	CRM;P/T;AUG	3	Lactams	P/T	2.9857E-016
25074	TE;P/T;TO;CRM; AUG	5	Polyketide*;Lactams; aminoglycoside	P/T	6.3651E-026
25045	CRM;P/T;AUG	3	Lactams	CRM	1.3860E-016
25094	CRM;P/T;AUG	3	Lactams	P/T	1.4259E-016
25583	A/S;P/T;AUG	3	Lactams	P/T	9.4153E-016
25562	TE;P/T;TO;CRM; AUG	5	Polyketide*;Lactams; aminoglycoside	P/T	7.0754E-025
25484	TE;P/T;CRM;AUG	4	Polyketide*;Lactams	P/T	2.3718E-018
25741	A/S;P/T;AUG	3	Lactams	P/T	9.4153E-016
25596	CRM;P/T;AUG	3	Lactams	CRM	1.3351E-016
25340	TE;P/T;TO;CRM; AUG	5	Polyketide*;Lactams; aminoglycoside	P/T	2.4030E-025
25574	TE;GM;A/S;CRM; P/T;TO;AM;AUG	8	Polyketide*;Lactams; aminoglycoside	AUG	2.2353E-040
25493	A/S;P/T;AUG	3	Lactams	P/T	2.9312E-015
25790	A/S;P/T;AUG	3	Lactams	P/T	9.4153E-016
25256	TE;TO;CRM	3	Polyketide*;Lactams; aminoglycoside	TE	4.2851E-012

*: (tetracycline)

Table 13: Statistically significant SNPs in gene nhaA (genbank protein accession number YP_005395552.1) (headers as in Tables 3 and 4, respectively)

POS	drug	#drugs	drug class	best drug	p-value
46308	TE;P/T;TO;CRM;	5	Polyketide*;Lactams;	P/T	6.7496E-025

	AUG		aminoglycoside		
46413	CRM;P/T;AUG	3	Lactams	CRM	1.1274E-016
46765	A/S;P/T;AUG	3	Lactams	P/T	8.5117E-015
46724	CRM;P/T;AUG	3	Lactams	CRM	1.1274E-016
46205	A/S;P/T;AUG	3	Lactams	P/T	2.4320E-015
47033	A/S;TO;AM;GM	4	aminoglycoside; Lactams	TO	7.3361E-017
46695	TE;GM;A/S;CRM; P/T;TO;AM;AUG	8	Polyketide*;Lactams; aminoglycoside	TO	3.7524E-039
47257	A/S;P/T;AUG	3	Lactams	P/T	7.8918E-016
47247	A/S;TO;AM;AUG	4	aminoglycoside; Lactams	AM	6.5822E-016
46950	A/S;P/T;AUG	3	Lactams	P/T	7.8918E-016
46440	A/S;P/T;AUG	3	Lactams	P/T	7.8918E-016
46474	A/S;P/T;AUG	3	Lactams	P/T	7.8918E-016
46830	A/S;P/T;AUG	3	Lactams	P/T	1.2790E-016

*: (tetracycline)

Similar results were obtained for other genes but are omitted for the sake of brevity.

Further, a synergistic effect of individual SNPs was demonstrated by exhaustively comparing significance levels for association of single SNPs with antibiotic susceptibility/resistance and significance levels for association of combinations of SNPs with antibiotic susceptibility/resistance. For a representative example of 2 SNPs the significance level for synergistic association of two SNPs was improved with the values given in Table 14 compared to the association of either SNP alone, given for exemplary different antibiotics.

Table 14: Synergistic increase for association of two SNPs

drug	POS 1	Ref	Alt	POS 2	Ref	Alt	Improv [%]
CRM	2833888	G	A	3379140	C	T	17180.9
CRM	2833888	G	A	3536122	G	A	296.8

CRM	2833888	G	A	1487086	C	T	10989.5
CRM	2833888	G	A	2208848	C	A,T	13443.8
CRM	2833888	G	A	3799879	C	T	10989.5
CRM	2833888	G	A	1994028	C	T	114.9
CRM	2833888	G	A	3686566	G	A	634.2
CRM	2833888	G	A	1163470	G	T	3919.4
CRM	2833888	G	A	1148509	G	C	3616.5
CRM	2833888	G	A	4478134	A	C	17180.9
CRM	2833888	G	A	56998	A	C	13170.2
CRM	2833888	G	A	46695	G	A	13170.2
CRM	2833888	G	A	25574	T	A,G	284.0
CRM	1313897	C	T	3379140	C	T	1944.3
CRM	1313897	C	T	1487086	C	T	1285.6
CRM	1313897	C	T	2208848	C	A,T	1620.7
CRM	1313897	C	T	3799879	C	T	1285.6
CRM	1313897	C	T	3686566	G	A	105.9
CRM	1313897	C	T	1163470	G	T	490.1
CRM	1313897	C	T	1148509	G	C	475.0
CRM	1313897	C	T	4478134	A	C	1944.3
CRM	1313897	C	T	56998	A	C	1634.2
CRM	1313897	C	T	46695	G	A	1634.2
CRM	3379140	C	T	855087	C	G	1311.3
CRM	3379140	C	T	4191057	G	A,C	623.4
CRM	3379140	C	T	4724403	C	T	1484.1
CRM	3379140	C	T	4436188	A	G	1606.9
CRM	3379140	C	T	4779417	T	C	101.0
CRM	3379140	C	T	546961	G	T	20434.1
CRM	3379140	C	T	3582301	A	G	159.4
CRM	3536122	G	A	1650934	G	A	459.3
CRM	3536122	G	A	1994028	C	T	736.2
CRM	3536122	G	A	4523874	C	T	550.8
CRM	3536122	G	A	3966175	T	G	351.5
CRM	3536122	G	A	258966	G	C,T	380.4
CRM	1487086	C	T	855087	C	G	843.8
CRM	1487086	C	T	4191057	G	A,C	349.9

CRM	1487086	C	T	4724403	C	T	2306.4
CRM	1487086	C	T	4436188	A	G	1449.8
CRM	1487086	C	T	546961	G	T	27275.0
AM	1487086	C	T	3582301	A	G	214.6
A/S	1487086	C	T	3582301	A	G	118.9
CRM	2208848	C	A,T	855087	C	G	1131.7
CRM	2208848	C	A,T	4191057	G	A,C	428.3
CRM	2208848	C	A,T	4724403	C	T	2980.5
CRM	2208848	C	A,T	4436188	A	G	1767.9
CRM	2208848	C	A,T	4779417	T	C	106.5
CRM	2208848	C	A,T	546961	G	T	37055.6
AM	2208848	C	A,T	3582301	A	G	161.4
AM	1589819	A	G	1163470	G	T	120.6
AM	1650934	G	A	1672517	T	C	112.8
AM	1650934	G	A	1673475	T	C	116.0
CRM	1650934	G	A	3686566	G	A	1163.3
CRM	1650934	G	A	546961	G	T	190.1
CRM	1650934	G	A	25574	T	A,G	1058.0
AM	1672517	T	C	3966175	T	G	352.6
A/S	1672517	T	C	3966175	T	G	245.6
AM	1672517	T	C	1163470	G	T	135.3
AM	1672517	T	C	258966	G	C,T	215.6
A/S	1672517	T	C	258966	G	C,T	131.5
AM	1673475	T	C	3966175	T	G	438.8
A/S	1673475	T	C	3966175	T	G	278.9
AM	1673475	T	C	1163470	G	T	148.6
AM	1673475	T	C	258966	G	C,T	223.5
A/S	1673475	T	C	258966	G	C,T	154.9
CRM	3799879	C	T	855087	C	G	843.8
CRM	3799879	C	T	4191057	G	A,C	349.9
CRM	3799879	C	T	4724403	C	T	2306.4
CRM	3799879	C	T	4436188	A	G	1449.8
CRM	3799879	C	T	546961	G	T	27275.0
AM	3799879	C	T	3582301	A	G	214.6
A/S	3799879	C	T	3582301	A	G	118.9

CRM	1994028	G	T	3686566	G	A	1277.0
CRM	1994028	G	T	25574	T	A,G	777.8
CRM	4523874	C	T	3686566	G	A	1353.7
CRM	4523874	C	T	546961	G	T	248.4
CRM	4523874	C	T	25574	T	A,G	1324.2
CRM	855087	C	G	1163470	G	T	383.0
CRM	855087	C	G	1148509	G	C	320.2
CRM	855087	C	G	4478134	A	C	1311.3
CRM	855087	C	G	56998	A	C	1047.0
CRM	855087	C	G	46695	G	A	1047.0
CRM	3966175	T	G	3686566	G	A	570.0
AM	3966175	T	G	3582301	A	G	299.2
A/S	3966175	T	G	3582301	A	G	153.4
CRM	3966175	T	G	25574	T	A,G	243.7
CRM	3686566	G	A	4779417	T	C	239.2
CRM	3686566	G	A	258966	G	C,T	937.7
CRM	4191057	G	A,C	1163470	G	T	137.3
CRM	4191057	G	A,C	1148509	G	C	3689.8
CRM	4191057	G	A,C	4478134	A	C	623.4
CRM	4191057	G	A,C	56998	A	C	440.7
CRM	4191057	G	A,C	46695	G	A	440.7
CRM	4191057	G	A,C	25574	T	A,G	530.9
CRM	4724403	C	T	1163470	G	T	400.4
CRM	4724403	C	T	1148509	G	C	829.0
CRM	4724403	C	T	4478134	A	C	1484.1
CRM	4724403	C	T	56998	A	C	2939.8
CRM	4724403	C	T	46695	G	A	2939.8
CRM	1163470	G	T	4436188	A	G	420.5
CRM	1163470	G	T	546961	G	T	3816.3
AM	1163470	G	T	3582301	A	G	408.7
A/S	1163470	G	T	3582301	A	G	221.9
CRM	1148509	G	C	4436188	A	G	459.2
CRM	1148509	G	C	546961	G	T	8786.6
AM	1148509	G	C	3582301	A	G	177.4
A/S	1148509	G	C	3582301	A	G	107.5

CRM	4478134	A	C	4436188	A	G	1606.9
CRM	4478134	A	C	4779417	T	C	101.0
CRM	4478134	A	C	546961	G	T	20434.1
AM	4478134	A	C	3582301	A	G	159.4
CRM	4436188	A	G	56998	A	C	1772.4
CRM	4436188	A	G	46695	G	A	1772.4
CRM	4779417	T	C	56998	A	C	101.0
CRM	4779417	T	C	46695	G	A	101.0
CRM	258966	G	C,T	546961	G	T	147.6
AM	258966	G	C,T	3582301	A	G	159.0
CRM	258966	G	C,T	25574	T	A,G	777.8
CRM	546961	G	T	56998	A	C	38427.4
CRM	546961	G	T	46695	G	A	38427.4
AM	3582301	A	G	56998	A	C	153.8
AM	3582301	A	G	46695	G	A	153.8

POS 1, 2 = position 1, 2 used for combination; Ref = reference base; Alt = alternated base in samples; improv = improvement compared to minimum p-value of single SNP

For example, the improvement of 153.8 % in the last example with positions 3582301 and 46695 for AM results from a p-value change from 2.27248e-06 to 1.4778e-06.

Again, similar results were obtained for other SNPs in respective genes.

A genetic test for the combined pathogen identification and antimicrobial susceptibility testing direct from the patient sample can reduce the time-to actionable result significantly from several days to hours, thereby enabling targeted treatment. Furthermore, this approach will not be restricted to central labs, but point of care devices can be developed that allow for respective tests. Such technology along with the

present methods and computer program products could revolutionize the care, e.g. in intense care units or for admissions to hospitals in general. Furthermore, even applications like real time outbreak monitoring can be achieved using the present methods.

Instead of using only single variants, a combination of several variant positions can improve the prediction accuracy and further reduce false positive findings that are influenced by other factors.

Compared to approaches using MALDI-TOF MS, the present approach has the advantage that it covers almost the complete genome and thus enables us to identify the potential genomic sites that might be related to resistance. While MALDI-TOF MS can also be used to identify point mutations in bacterial proteins, this technology only detects a subset of proteins and of these not all are equally well covered. In addition, the identification and differentiation of certain related strains is not always feasible.

The present method allows computing a best breakpoint for the separation of isolates into resistant and susceptible groups. The inventors designed a flexible software tool that allows to consider - besides the best breakpoints - also values defined by different guidelines (e.g. European and US guidelines), preparing for an application of the GAST in different countries.

The inventors demonstrate that the present approach is capable of identifying mutations in genes that are already known as drug targets, as well as detecting potential new target sites.

The current approach enables

a. Identification and validation of markers for genetic identification and susceptibility/resistance testing within one diagnostic test
b. validation of known drug targets and modes of action
c. detection of potentially novel resistance mechanisms leading to putative novel target / secondary target genes for new therapies

Claims

1. A diagnostic method of determining an infection of a patient with Salmonella species potentially resistant to antimicrobial drug, e.g. antibiotic, treatment, comprising the steps of:

a) obtaining or providing a sample containing or suspected of containing at least one Salmonella species from the patient;

b) determining the presence of at least one mutation in at least two genes from the group of genes consisting of recN, hemH, UMN798_3428, metE, yijD, UMN798_4831, UMN798_1939, copS, UMN798_0628, UMN798_4878, leuB, recF, emrA, glyQ, dcp, thiH, UMN798_1612, UMN798_2909, UMN798_1680, UMN798_4073, yjbC, nadB, UMN798_3160, hutU, envC, UMN798_3889, UMN798_1629, bcfB, degQ, UMN798_1331, trg, uvrC, polB, hpcD, UMN798_1628, UMN798_1701, glgS, plsB, yjcC, feoB, misL, dxr, hemF, rnfG, yhjB, UMN798_1163, UMN798_0394, alkA, nhaA, and lspA, wherein the presence of said at least two mutations is indicative of an infection with an antimicrobial drug, e.g. antibiotic, resistant Salmonella strain in said patient.

2. A method of selecting a treatment of a patient suffering from an infection with a potentially resistant Salmonella strain, comprising the steps of:

a) obtaining or providing a sample containing or suspected of containing at least one Salmonella species from the patient;

b) determining the presence of at least one mutation in at least two genes from the group of genes consisting of recN, hemH, UMN798_3428, metE, yijD, UMN798_4831, UMN798_1939, copS, UMN798_0628, UMN798_4878, leuB, recF, emrA, glyQ, dcp, thiH, UMN798_1612, UMN798_2909,

UMN798_1680, UMN798_4073, yjbC, nadB, UMN798_3160, hutU, envC, UMN798_3889, UMN798_1629, bcfB, degQ, UMN798_1331, trg, uvrC, polB, hpcD, UMN798_1628, UMN798_1701, glgS, plsB, yjcC, feoB, misL, dxr, hemF, rnfG, yhjB, UMN798_1163, UMN798_0394, alkA, nhaA, and lspA, wherein the presence of said at least two mutations is indicative of a resistance to one or more antimicrobial, e.g. antibiotic, drugs;

c) identifying said at least one or more antimicrobial, e.g. antibiotic, drugs; and

d) selecting one or more antimicrobial, e.g. antibiotic, drugs different from the ones identified in step c) and being suitable for the treatment of a Salmonella infection.

3. The method of one or more of the preceding claims, where the method involves determining the resistance of Salmonella to one or more antimicrobial, e.g. antibiotic, drugs.

4. The method of any one of claims 1 to 3, wherein the antimicrobial, e.g. antibiotic, drug is selected from lactam antibiotics and the presence of a mutation in the following genes is determined: recN, hemH, UMN798_3428, metE, yijD, UMN798_4831, UMN798_1939, copS, UMN798_0628, UMN798_4878, leuB, recF, emrA, glyQ, dcp, thiH, UMN798_1612, UMN798_2909, UMN798_1680, UMN798_4073, yjbC, nadB, UMN798_3160, hutU, envC, UMN798_3889, UMN798_1629, bcfB, degQ, UMN798_1331, trg, uvrC, polB, hpcD, UMN798_1628, UMN798_1701, glgS, plsB, yjcC, feoB, misL, dxr, hemF, rnfG, yhjB, UMN798_1163, UMN798_0394, alkA, nhaA, and/or lspA; and/or
wherein the antimicrobial, e.g. antibiotic, drug is selected from aminoglycoside antibiotics, and the presence

of a mutation in the following genes is determined: recN, hemH, UMN798_3428, metE, yijD, UMN798_4831, UMN798_1939, copS, UMN798_0628, UMN798_4878, leuB, recF, emrA, glyQ, dcp, thiH, UMN798_1612, UMN798_2909, UMN798_1680, UMN798_4073, yjbC, nadB, UMN798_3160, hutU, envC, UMN798_3889, UMN798_1629, bcfB, degQ, UMN798_1331, trg, uvrC, polB, hpcD, UMN798_1628, UMN798_1701, glgS, plsB, yjcC, feoB, misL, dxr, hemF, rnfG, yhjB, UMN798_1163, UMN798_0394, alkA, nhaA, and/or lspA; and/or

wherein the antimicrobial, e.g. antibiotic, drug is selected from polyketide antibiotics, preferably tetracycline antibiotics, and the presence of a mutation in the following genes is determined: recN, hemH, UMN798_3428, metE, yijD, UMN798_4831, UMN798_1939, copS, UMN798_0628, UMN798_4878, leuB, recF, emrA, glyQ, dcp, thiH, UMN798_1612, UMN798_2909, UMN798_1680, UMN798_4073, yjbC, nadB, UMN798_3160, hutU, envC, UMN798_3889, UMN798_1629, bcfB, UMN798_1331, trg, uvrC, polB, hpcD, UMN798_1628, UMN798_1701, glgS, plsB, yjcC, feoB, misL, dxr, hemF, rnfG, yhjB, UMN798_1163, UMN798_0394, alkA, nhaA, and/or lspA.

5. The method of one or more of the preceding claims, wherein the antimicrobial drug, e.g. antibiotic drug, is selected from the group consisting of Amoxicillin/K Clavulanate (AUG), Ampicillin (AM), Aztreonam (AZT), Cefazolin (CFZ), Cefepime (CPM), Cefotaxime (CFT), Ceftazidime (CAZ), Ceftriaxone (CAX), Cefuroxime (CRM), Cephalotin (CF), Ciprofloxacin (CP), Ertapenem (ETP), Gentamicin (GM), Imipenem (IMP), Levofloxacin (LVX), Meropenem (MER), Piperacillin/Tazobactam (P/T), Ampicillin/Sulbactam (A/S), Tetracycline (TE), Tobramycin (TO), and Trimethoprim/Sulfamethoxazole (T/S).

6. The method of any one of claims 1 to 5, wherein the antibiotic drug is CFZ and a mutation in at least one of the following nucleotide positions is detected with regard to reference genome NC_017046 as annotated at the NCBI: 4779417, 2983118, 1548464, 1574737, 3772654, 1313897; and/or
wherein the antibiotic drug is GM and a mutation in at least one of the following nucleotide positions is detected with regard to reference genome NC_017046 as annotated at the NCBI: 2840330, 25574, 3536122, 115342, 1148509, 3379140, 4478134, 3686566, 1487086, 3799879, 1163470, 2208848, 46695, 56998, 3335426; and/or
wherein the antibiotic drug is CF and a mutation in at least one of the following nucleotide positions is detected with regard to reference genome NC_017046 as annotated at the NCBI: 2833888, 546961, 3334479, 4191057, 4366486, 4724403, 1895588, 1139812, 637461, 4779417, 131219, 4062015, 2983118, 3861998, 1548464, 4397111, 1574737, 1650934, 4436188, 3075942, 855087, 3582301, 3772654, 1590194, 1313897, 1673475, 1994028, 1589819, 1672517, 3976726 3582354, 1673444; and/or
wherein the antibiotic drug is TE and a mutation in at least one of the following nucleotide positions is detected with regard to reference genome NC_017046 as annotated at the NCBI: 2833888, 546961, 3334479, 4191057, 4366486, 4724403, 1895588, 1139812, 637461, 4779417, 131219, 4062015, 2983118, 3861998, 1548464, 4397111, 1574737, 2840330, 1650934, 3966175, 4436188, 2780306, 3075942, 855087, 3582301, 3772654, 1590194, 25574, 1313897, 1673475, 1994028, 115342, 1148509, 1589819, 1672517, 3379140, 4478134, 4523874, 3686566, 3976726, 258966, 2558379, 1487086, 3799879, 1163470, 409259, 2208848, 46695, 56998, 3582354, 3335426, 1673444; and/or

wherein the antibiotic drug is A/S and a mutation in at least one of the following nucleotide positions is detected with regard to reference genome NC_017046 as annotated at the NCBI: 2833888, 546961, 3334479, 4191057, 4366486, 4724403, 1895588, 1139812, 637461, 4779417, 131219, 4062015, 2983118, 3861998, 1548464, 4397111, 1574737, 2840330, 1650934, 3966175, 4436188, 2780306, 3075942, 855087, 3582301, 3772654, 1590194, 25574, 3536122, 1313897, 1673475, 1994028, 115342, 1148509, 1589819, 1672517, 3379140, 4478134, 4523874, 3686566, 3976726, 258966, 2558379, 1487086, 3799879, 1163470, 409259, 2208848, 46695, 56998, 3582354, 3335426, 1673444; and/or

wherein the antibiotic drug is CRM and a mutation in at least one of the following nucleotide positions is detected with regard to reference genome NC_017046 as annotated at the NCBI: 2833888, 546961, 3334479, 4191057, 4366486, 4724403, 1895588, 1139812, 637461, 4779417, 131219, 4062015, 2983118, 3861998, 1548464, 4397111, 1574737, 2840330, 1650934, 3966175, 4436188, 2780306, 3075942, 855087, 3582301, 3772654, 1590194, 25574, 3536122, 1313897, 1673475, 1994028, 115342, 1148509, 1589819, 1672517, 3379140, 4478134, 4523874, 3686566, 3976726, 258966, 2558379, 1487086, 3799879, 1163470, 409259, 2208848, 46695, 56998, 3582354, 3335426, 1673444; and/or

wherein the antibiotic drug is P/T and a mutation in at least one of the following nucleotide positions is detected with regard to reference genome NC_017046 as annotated at the NCBI: 2833888, 546961, 3334479, 4191057, 4366486, 4724403, 1895588, 1139812, 637461, 4779417, 131219, 4062015, 2983118, 3861998, 1548464, 4397111, 1574737, 2840330, 1650934, 3966175, 4436188, 2780306, 3075942, 855087, 3582301, 3772654, 1590194, 25574,

3536122, 1313897, 1673475, 1994028, 115342, 1148509, 1589819, 1672517, 3379140, 4478134, 4523874, 3686566, 3976726, 258966, 2558379, 1487086, 3799879, 1163470, 409259, 2208848, 46695, 56998, 3582354, 3335426, 1673444; and/or

wherein the antibiotic drug is TO and a mutation in at least one of the following nucleotide positions is detected with regard to reference genome NC_017046 as annotated at the NCBI: 2833888, 546961, 3334479, 4191057, 4366486, 4724403, 1895588, 1139812, 637461, 4779417, 131219, 4062015, 2983118, 3861998, 1548464, 4397111, 1574737, 2840330, 1650934, 3966175, 4436188, 2780306, 3075942, 855087, 3582301, 3772654, 1590194, 25574, 3536122, 1313897, 1673475, 1994028, 115342, 1148509, 1589819, 1672517, 3379140, 4478134, 4523874, 3686566, 3976726, 258966, 2558379, 1487086, 3799879, 1163470, 409259, 2208848, 46695, 56998, 3582354, 3335426, 1673444; and/or

wherein the antibiotic drug is AM and a mutation in at least one of the following nucleotide positions is detected with regard to reference genome NC_017046 as annotated at the NCBI: 2833888, 546961, 3334479, 4191057, 4366486, 4724403, 1895588, 1139812, 637461, 4779417, 131219, 4062015, 2983118, 3861998, 1548464, 4397111, 1574737, 2840330, 1650934, 3966175, 4436188, 2780306, 3075942, 855087, 3772654, 25574, 3536122, 1313897, 1994028, 115342, 1148509, 3379140, 4478134, 4523874, 3686566, 3976726, 258966, 2558379, 1487086, 3799879, 1163470, 409259, 2208848, 46695, 56998, 3335426; and/or

wherein the antibiotic drug is AUG and a mutation in at least one of the following nucleotide positions is detected with regard to reference genome NC_017046 as annotated at the NCBI: 2833888, 546961, 3334479, 4191057, 4366486, 4724403, 1895588, 1139812, 637461, 4779417,

131219, 4062015, 2983118, 3861998, 1548464, 4397111, 1574737, 2840330, 1650934, 3966175, 4436188, 2780306, 3075942, 855087, 3582301, 3772654, 1590194, 25574, 3536122, 1313897, 1673475, 1994028, 115342, 1148509, 1589819, 1672517, 3379140, 4478134, 4523874, 3686566, 3976726, 258966, 2558379, 1487086, 3799879, 1163470, 409259, 2208848, 46695, 56998, 3582354, 3335426, 1673444.

7. The method of any one of claims 1 to 6, wherein the resistance of a bacterial microorganism belonging to the species Salmonella against 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15 or 16, 17, 18, 19, 20 or 21 antibiotic drugs is determined.

8. The method of one or more of the preceding claims, wherein determining the nucleic acid sequence information or the presence of a mutation comprises determining a partial sequence or an entire sequence of the at least two genes.

9. The method of one or more of the preceding claims, wherein determining the nucleic acid sequence information or the presence of a mutation comprises determining a partial or entire sequence of the genome of the Salmonella species, wherein said partial or entire sequence of the genome comprises at least a partial sequence of said at least two genes.

10. The method of one or more of the preceding claims, wherein determining the nucleic acid sequence information or the presence of a mutation comprises using a next generation sequencing or high throughput sequencing method, preferably wherein a partial or entire genome sequence of the bacterial organism of Salmonella species is deter-

mined by using a next generation sequencing or high throughput sequencing method.

11. A method of determining an antimicrobial drug, e.g. antibiotic, resistance profile for bacterial microorganisms of Salmonella species, comprising:
obtaining or providing a first data set of gene sequences of a plurality of clinical isolates of Salmonella species;
providing a second data set of antimicrobial drug, e.g. antibiotic, resistance of the plurality of clinical isolates of Salmonella species;
aligning the gene sequences of the first data set to at least one, preferably one, reference genome of Salmonella, and/or assembling the gene sequence of the first data set, at least in part;
analyzing the gene sequences of the first data set for genetic variants to obtain a third data set of genetic variants;
correlating the third data set with the second data set and statistically analyzing the correlation; and
determining the genetic sites in the genome of Salmonella associated with antimicrobial drug, e.g. antibiotic, resistance.

12. A diagnostic method of determining an infection of a patient with Salmonella species potentially resistant to antimicrobial drug treatment, comprising the steps of:
a) obtaining or providing a sample containing or suspected of containing a bacterial microorganism belonging to the species Salmonella from the patient;
b) determining the presence of at least one mutation in at least one gene of the bacterial microorganism belonging to the species Salmonella as determined by the

method of claim 11, wherein the presence of said at least one mutation is indicative of an infection with an antimicrobial drug resistant Salmonella strain in said patient.

13. A method of selecting a treatment of a patient suffering from an infection with a potentially resistant Salmonella strain, comprising the steps of:

a) obtaining or providing a sample containing or suspected of containing a bacterial microorganism belonging to the species Salmonella from the patient;

b) determining the presence of at least one mutation in at least one gene of the bacterial microorganism belonging to the species Salmonella as determined by the method of claim 11, wherein the presence of said at least one mutation is indicative of a resistance to one or more antimicrobial drugs;

c) identifying said at least one or more antimicrobial drugs; and

d) selecting one or more antimicrobial drugs different from the ones identified in step c) and being suitable for the treatment of a Salmonella infection.

14. A method of acquiring an antimicrobial drug, e.g. antibiotic, resistance profile for bacterial microorganisms of Salmonella species, comprising:

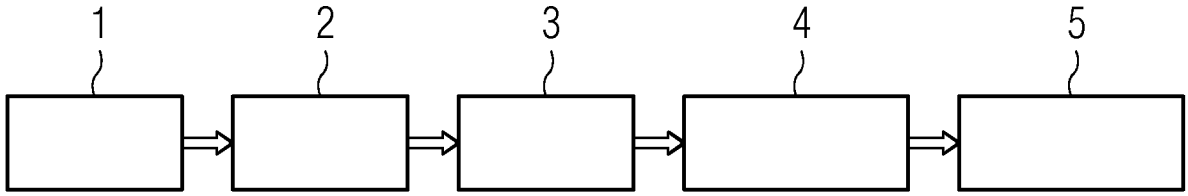
obtaining or providing a first data set of gene sequences of a clinical isolate of Salmonella species;

providing a second data set of antimicrobial drug, e.g. antibiotic, resistance of a plurality of clinical isolates of Salmonella species;

aligning the gene sequences of the first data set to at least one, preferably one, reference genome of Salmonel-

la, and/or assembling the gene sequence of the first data set, at least in part;
analyzing the gene sequences of the first data set for genetic variants to obtain a third data set of genetic variants of the first data set;
correlating the third data set with the second data set and statistically analyzing the correlation; and
determining the genetic sites in the genome of Salmonella of the first data set associated with antimicrobial drug, e.g. antibiotic, resistance.

15. Computer program product comprising computer executable instructions which, when executed, perform a method according to any one of claims 11 to 14.


1
2
3
4
5