



(51) International Patent Classification:

C12Q 1/689 (2018.01) G16B 20/20 (2019.01)

C12Q 1/04 (2006.01) C12R 1/145 (2006.01)

(21) International Application Number:

PCT/CN2019/092585

(22) International Filing Date:

24 June 2019 (24.06.2019)

(25) Filing Language:

English

(26) Publication Language:

English

(71) Applicant: BGI SHENZHEN [CN/CN]; Main Building, Beishan Industrial Zone, Yantian, District, Shenzhen, Guangdong 518083 (CN).

(72) Inventors: LI, Junhua; Main Building, Beishan Industrial Zone, Yantian, District, Shenzhen, Guangdong 518083 (CN). ZHAO, Hailong; Main Building, Beishan Industrial Zone, Yantian, District, Shenzhen, Guangdong 518083

(CN). PENG, Ye; Main Building, Beishan Industrial Zone, Yantian, District, Shenzhen, Guangdong 518083 (CN). SHAW, Peter M.; Main Building, Beishan Industrial Zone, Yantian, District, Shenzhen, Guangdong 518083 (CN). NICKLE, David; Main Building, Beishan Industrial Zone, Yantian, District, Shenzhen, Guangdong 518083 (CN). SHEN, Judong; Main Building, Beishan Industrial Zone, Yantian, District, Shenzhen, Guangdong 518083 (CN).

(74) Agent: DHC IP ATTORNEYS; Suite 1810, Modern International Commercial, Building, Cross Of Fuhua Road And Jintian Road, Futian District, Shenzhen, Guangdong 518048 (CN).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP,

(54) Title: SNP MARKERS OF DRUG REDUCED SUSCEPTIBILITY RELATED EVOLUTIONARY BRANCHES OF CLOSTRIDIUM DIFFICILE, METHOD FOR IDENTIFYING STRAIN CATEGORY, AND USE THEREOF

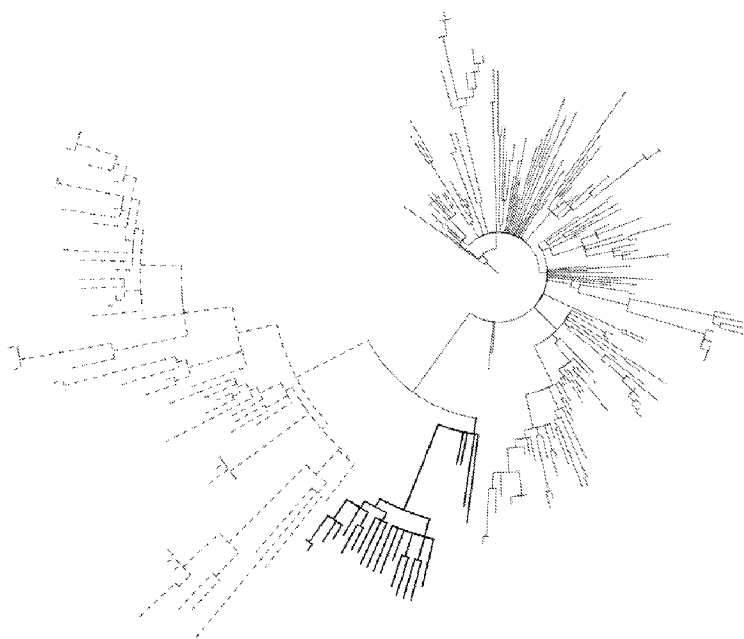


Fig. 1

(57) Abstract: Provided are SNP markers of drug reduced susceptibility related evolutionary branches of *Clostridium difficile*, a method for identifying the category of a *Clostridium difficile* strain, and use thereof. The SNP markers are specific markers of three categories of the *Clostridium difficile* clade2 (mainly hypervirulent ribotype 027), allowing for rapid and accurate identification of the evolutionary branches of *Clostridium difficile* strains that are resistant to a variety of therapeutic drugs and related drugs. Accurate categorization of the drug reduced susceptibility related evolutionary branches not only provides evidence for the evolutionary traceability of drug-resistant pathogens, but also offers effective and actionable guidance on clinical drug usage.



KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

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

Published:

- with international search report (Art. 21(3))
- with sequence listing part of description (Rule 5.2(a))

**SNP MARKERS OF DRUG REDUCED SUSCEPTIBILITY
RELATED EVOLUTIONARY BRANCHES OF *CLOSTRIDIUM
DIFFICILE*, METHOD FOR IDENTIFYING STRAIN
CATEGORY, AND USE THEREOF**

Technical Field

The present invention relates to the technical field of microbial drug resistance, in particular to SNP markers of drug reduced susceptibility related evolutionary branches of *Clostridium difficile*, a method for identifying the category of a *Clostridium difficile* strain, and use thereof.

Background of the Invention

According to data released by the Centers for Disease Control and Prevention (CDC), USA in 2015, nearly 500,000 people in the United States are infected with *Clostridium difficile* each year, and the annual mortality is as high as 30,000 deaths. The bacterium has become the most common hospital-acquired pathogen instead of MRSA, the cost of treatment amounting to \$1.5 billion per year. The total number of *Clostridium difficile* infections in Asians has also increased year by year. The bacteria has high resistance to common antibiotics such as erythromycin, clindamycin and fluoroquinolone, and can only be treated by metronidazole and vancomycin, the sensitivity to which the bacteria shows, however, has been ever-decreasing in recent years, particularly in the case of *Clostridium difficile* clade2 strains. Clade2 is defined according to the *Clostridium difficile* multilocus sequence typing (MLST) database (<http://pubmlst.org/cdifficile>). The database determines the sequence type (ST) of each strain based on the polymorphism of seven *Clostridium difficile* housekeeping genes, and assigns the ST to one of five clades (clade1 to clade5) based on the evolutionary relationship within the whole species. As one of the clades, clade2 mainly comprises hypervirulent ribotype 027

(Ribotype027). As one of the 600 ribotypes of *Clostridium difficile*, Ribotype027 was experimentally obtained by taxonomic categorization by means of a polymerase chain reaction method using the polymorphism of the intergenic regions of 16S-23S ribosomal RNA genes, and has received much attention due to its high degree of drug resistance and clinical severity.

The continued use of antibiotics against the drug-resistant bacteria not only fails to achieve therapeutic effects, but also wastes resources and delays the clinical cycle, and may even cause more serious drug-resistant mutations. Existing techniques for detecting the drug resistance of the pathogen mainly include the following two means: (1) identification of the drug resistance by drug sensitivity test; and (2) identification of the drug resistance using existing detection kits of drug resistant determinants. The existing techniques for detecting the drug resistance of the pathogen suffer from the following problems: (1) detection of *Clostridium difficile* by drug sensitivity test involves stringent culture conditions and prolonged experimental period, which fails to meet the requirements of clinically rapid detection; and (2) the existing drug resistance detection kits do not allow for detection of therapeutic drugs and are thus of little significance in terms of clinical guidance.

Summary of the Invention

The present invention provides SNP markers of drug reduced susceptibility related evolutionary branches of *Clostridium difficile*, a method for identifying the category of a *Clostridium difficile* strain, and use thereof. The present invention allows for rapid and accurate identification of the evolutionary branches of *Clostridium difficile* strains that are resistant to a variety of therapeutic drugs and related drugs, providing meaningful clinical therapeutic guidance.

According to a first aspect, an embodiment provides SNP markers of drug reduced susceptibility related evolutionary branches of *Clostridium*

difficile, wherein the *Clostridium difficile* is the *Clostridium difficile* clade 2, and the SNP markers are selected from the group comprising any one of the SNP markers in the following three categories, or any combination thereof:

(a) Category 1: an A base at genomic base position 1029237, a T base at genomic base position 1205938, an A base at genomic base position 2487991, an A base at genomic base position 2861888, a T base at genomic base position 882348, a G base at genomic base position 1798870, and an A base at genomic base position 3083454;

(b) Category 2: a C base at genomic base position 6310, and an A base at genomic base position 1550363;

(c) Category 3: a G base at genomic base position 118669, a C base at genomic base position 1205250, a C base at genomic base position 1235096, an A base at genomic base position 1462869, a T base at genomic base position 1549858, an A base at genomic base position 2367860, an A base at genomic base position 2851331, a C base at genomic base position 3031309, a G base at genomic base position 3419928, a G base at genomic base position 1602810, and a T base at genomic base position 2585036.

In a preferred embodiment, the *Clostridium difficile* clade 2 is hypervirulent ribotype 027 (Ribotype027).

According to a second aspect, an embodiment provides a method for identifying the category of a *Clostridium difficile* strain, comprising obtaining base information at the site of at least one SNP marker of the SNP markers according to the first aspect in a *Clostridium difficile* strain to be identified; and determining the category of the *Clostridium difficile* strain according to the base information.

In a preferred embodiment, the method comprises obtaining base information at the site of at least one SNP marker in Category 1, Category 2

or Category 3 according to the first aspect in a *Clostridium difficile* strain to be identified; and determining whether the *Clostridium difficile* strain belongs to said Category 1, Category 2 or Category 3 according to the base information.

In a preferred embodiment, the method comprises obtaining base information at the site of at least one SNP marker in each category of at least two categories of Category 1, Category 2 or Category 3 according to the first aspect in a *Clostridium difficile* strain to be identified; and determining whether the *Clostridium difficile* strain belongs to said Category 1, Category 2 or Category 3 according to the base information.

In a preferred embodiment, the method obtains the base information of the site of said SNP marker by amplifying a genomic region in which the site of said SNP marker is located by using primers adapted to specifically amplify the region, followed by sequencing the region.

In a third aspect, an embodiment provides a method for diagnosing the category of a *Clostridium difficile* strain in a subject infected with the strain, comprising: obtaining base information at the site of at least one SNP marker of the SNP markers according to the first aspect in the *Clostridium difficile* strain from the subject; and determining the category of the *Clostridium difficile* strain according to the base information.

In a preferred embodiment, the method comprises obtaining base information at the site of at least one SNP marker in Category 1, Category 2 or Category 3 according to the first aspect in the *Clostridium difficile* strain from the subject; and determining whether the *Clostridium difficile* strain belongs to said Category 1, Category 2 or Category 3 according to the base information.

In a preferred embodiment, the method comprises obtaining base information at the site of at least one SNP marker in each category of at least

two categories of Category 1, Category 2 and Category 3 according to the first aspect in the *Clostridium difficile* strain from the subject; and determining whether the *Clostridium difficile* strain belongs to said Category 1, Category 2 or Category 3 according to the base information.

In a preferred embodiment, the method obtains the base information of the site of said SNP marker by amplifying a genomic region in which the site of said SNP marker is located by using primers adapted to specifically amplify the region, followed by sequencing the region.

In a fourth aspect, an embodiment provides a method for treating a subject infected with a *Clostridium difficile* strain, comprising obtaining base information at the site of at least one SNP marker of the SNP markers according to the first aspect in the *Clostridium difficile* strain from the subject; determining the category of the *Clostridium difficile* strain according to the base information; and administering moxifloxacin and/or metronidazole to the subject when the category of *Clostridium difficile* strain comprises Category 2 described above; and optionally, administering vancomycin to the subject when the category of the *Clostridium difficile* strain comprises Category 2 and/or Category 3 described above.

In a preferred embodiment, the method obtains the base information of the site of said SNP marker by amplifying a genomic region in which the site of said SNP marker is located by using primers adapted to specifically amplify the region, followed by sequencing the region.

In a fifth aspect, an embodiment provides primers adapted to specifically amplify a genomic region in which the site of a SNP marker according to the first aspect is located, wherein the primers comprise a forward primer and a reverse primer, the forward primer and the reverse primer respectively specifically binding to a genomic sense strand and a genomic antisense strand flanking said SNP marker.

In a preferred embodiment, the primers comprise a fluorescent label and are adapted for fluorescence quantitative PCR.

In a preferred embodiment, the primers serve as hybridization probes to be immobilized on a chip to capture a sequence of the genomic region in which the site of said SNP marker is located.

In a sixth aspect, an embodiment provides a use of the primers according to the fifth aspect in the identification of the category of a *Clostridium difficile* strain, or in the diagnosis of the category of a *Clostridium difficile* strain from a subject infected with the strain.

In a seventh aspect, an embodiment provides a use of the SNP markers according to the first aspect in the identification of the category of a *Clostridium difficile* strain, or in the diagnosis of the category of a *Clostridium difficile* strain from a subject infected with the strain.

In an eight aspect, an embodiment provides a kit comprising primers, the kit comprises the primers comprising a forward primer and a reverse primer, the forward primer and the reverse primer respectively specifically binding to a genomic sense strand and a genomic antisense strand flanking a SNP marker according to the first aspect, and adapted to specifically amplify a genomic region in which the site of said SNP marker is located; optionally, the kit further comprises PCR amplification components besides the primers described above, such as a Taq DNA polymerase, dNTPs, and a reaction buffer, among others.

The SNP markers of drug reduced susceptibility related evolutionary branches of *Clostridium difficile* provided in the present invention are not only drug resistance markers, but also markers for identifying the drug reduced susceptibility related evolutionary branches of the *Clostridium difficile* clade 2 (mainly hypervirulent ribotype 027 (Ribotype027)). Strains in the same evolutionary branch have a closer phylogenetic relationship

genome-wide, that is, share more identical genomic characteristics. The present invention allows for rapid and accurate identification of the evolutionary branches of *Clostridium difficile* resistant to a variety of therapeutic drugs and related drugs, providing meaningful clinical therapeutic guidance.

Brief Description of the Drawings

Fig. 1 is a diagram showing the evolutionary relationship and categorization of 269 strains of *Clostridium difficile* hypervirulent ribotype 027 (RT027) according to an example of the present invention.

Detailed Description

The present invention will be further described in detail below with reference to the accompanying drawings. In the following embodiments, many details are described so that the present invention will be better understood. However, those skilled in the art can readily recognize that some of the features may be omitted, or replaced by other materials or methods, depending on different situations.

Additionally, the characteristics, operations or features described in the specification can be combined in any suitable manner to form various embodiments. Moreover, the steps or actions in the description of the method may also be switched or adjusted in sequence in a manner that is obvious to those skilled in the art. Therefore, the various sequences in the description and the drawings are merely for the purpose of clearly describing a particular embodiment and are not intended to be required, unless it is otherwise specified that a specific sequence must be followed.

The present invention provides SNP markers of drug reduced susceptibility related evolutionary branches of *Clostridium difficile*, wherein the *Clostridium difficile* is the *Clostridium difficile* clade 2 (mainly hypervirulent ribotype 027 (Ribotype027)), and the SNP markers are selected from the group comprising any one of the SNP markers in the

following three categories, or any combination thereof:

(a) Category 1: an A base at genomic base position 1029237, a T base at genomic base position 1205938, an A base at genomic base position 2487991, an A base at genomic base position 2861888, a T base at genomic base position 882348, a G base at genomic base position 1798870, and an A base at genomic base position 3083454;

(b) Category 2: a C base at genomic base position 6310, and an A base at genomic base position 1550363;

(c) Category 3: a G base at genomic base position 118669, a C base at genomic base position 1205250, a C base at genomic base position 1235096, an A base at genomic base position 1462869, a T base at genomic base position 1549858, an A base at genomic base position 2367860, an A base at genomic base position 2851331, a C base at genomic base position 3031309, a G base at genomic base position 3419928, a G base at genomic base position 1602810, and a T base at genomic base position 2585036.

It should be noted that all of the above base position numbers are based on *Clostridium difficile* CD196 (NC_013315.1) as a reference genome.

Of the above-said 20 SNP markers in the present invention, 15 SNP markers have 100% category specificity. That is, the first 4 SNP markers in Category 1 only exist in the *Clostridium difficile* high-toxic ribotype 027 belonging to Category 1, both of the SNP markers in Category 2 only exist in the *Clostridium difficile* high-toxic ribotype 027 belonging to Category 2, and the first 9 SNP markers in Category 3 only exist in the *Clostridium difficile* high-toxic ribotype 027 belonging to Category 3. Therefore, by identifying the above-said SNP markers, it can be determined to which specific category of Category 1, Category 2 or Category 3 the *Clostridium difficile* hypervirulent ribotype 027 to be identified belongs.

Of the above-said 20 SNP markers in the present invention, the remaining 5 SNP markers besides the 15 SNP markers immediately described above have greater than 90% category specificity. That is, (1) for Category 1, when a T base is present at genomic base position 882348, the strain has a 93.3% possibility of belonging to Category 1, and when a C base is present at genomic base position 882348, the strain has a 100% possibility of belonging to Category 2 or Category 3; when a G base is present at genomic base position 1798870, the strain has a 91.8% possibility of belonging to Category 1, and when an A base is present at genomic base position 1798870, the strain has a 100% possibility of belonging to Category 2 or Category 3; and when an A base is present at genomic base position 3083454, the strain has a 96.6% possibility of belonging to Category 1, and when a G base is present at genomic base position 3083454, the strain has a 100% possibility of belonging to Category 2 or Category 3; and (2) for Category 3, when a G or T base is present at genomic base position 1602810, the strain has a 100% possibility of belonging to Category 3, and when a C base is present at genomic base position 1602810, the strain has a 100% possibility of belonging to Category 1 or Category 2; and when a T base is present at genomic base position 2585036, the strain has a 100% possibility of belonging to Category 3, and when a G base is present at genomic base position 2585036, the strain has a 95.2% possibility of belonging to Category 1 or Category 2.

It should be noted that by "any combination of SNP markers" is meant a combination of the SNP markers selected from any 1, 2 or 3 categories of the above-said three categories. For example, provided is a combination of any 1, 2, 3, 4, 5, 6 or 7 SNP markers selected from Category 1, such as, from Category 1, a combination of the SNP markers at genomic base positions 1029237 and 1205938, a combination of the SNP markers at genomic base positions 1029237 and 2487991, a combination of the SNP markers at

genomic base positions 1029237 and 2861888, a combination of the SNP markers at genomic base positions 1029237 and 882348, a combination of the SNP markers at genomic base positions 1029237 and 1798870, a combination of the SNP markers at genomic base positions 1029237 and 3083454, a combination of the SNP markers at genomic base positions 1205938 and 2487991, a combination of the SNP markers at genomic base positions 1205938 and 2861888, a combination of the SNP markers at genomic base positions 1205938 and 882348, a combination of the SNP markers at genomic base positions 1205938 and 1798870, a combination of the SNP markers at genomic base positions 1205938 and 3083454, a combination of the SNP markers at genomic base positions 2487991 and 2861888, a combination of the SNP markers at genomic base positions 2487991 and 882348, a combination of the SNP markers at genomic base positions 2487991 and 1798870, a combination of the SNP markers at genomic base positions 2487991 and 3083454, a combination of the SNP markers at genomic base positions 2861888 and 882348, a combination of the SNP markers at genomic base positions 2861888 and 1798870, a combination of the SNP markers at genomic base positions 2861888 and 3083454, a combination of the SNP markers at genomic base positions 882348 and 1798870, a combination of the SNP markers at genomic base positions 882348 and 3083454, a combination of the SNP markers at genomic base positions 1798870 and 3083454; or a combination having 3, 4, 5, 6 or 7 SNP markers, formed on the basis of the above-said combinations by optionally incorporating a 3rd, 4th, 5th, 6th or 7th SNP marker.

And for example, from Category 2, provided is a combination of the SNP markers at genomic base positions 6310 and 1550363.

Further for example, provided is a combination of any 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or 11 SNP markers from Category 3, such as a combination of the SNP markers at genomic base positions 118669 and 1205250, a combination

of the SNP markers at genomic base positions 118669 and 1235096, a combination of the SNP markers at genomic base positions 118669 and 1462869, a combination of the SNP markers at genomic base positions 118669 and 1549858, a combination of the SNP markers at genomic base positions 118669 and 2367860, a combination of the SNP markers at genomic base positions 118669 and 2851331, a combination of the SNP markers at genomic base positions 118669 and 3031309, a combination of the SNP markers at genomic base positions 118669 and 3419928, a combination of the SNP markers at genomic base positions 118669 and 1602810, a combination of the SNP markers at genomic base positions 118669 and 2585036, a combination of the SNP markers at genomic base positions 1205250 and 1235096, a combination of the SNP markers at genomic base positions 1205250 and 1462869, a combination of the SNP markers at genomic base positions 1205250 and 1549858, a combination of the SNP markers at genomic base positions 1205250 and 2367860, a combination of the SNP markers at genomic base positions 1205250 and 2851331, a combination of the SNP markers at genomic base positions 1205250 and 3031309, a combination of the SNP markers at genomic base positions 1205250 and 3419928, a combination of the SNP markers at genomic base positions 1205250 and 1602810, a combination of the SNP markers at genomic base positions 1205250 and 2585036, a combination of the SNP markers at genomic base positions 1235096 and 1462869, a combination of the SNP markers at genomic base positions 1235096 and 1549858, a combination of the SNP markers at genomic base positions 1235096 and 2367860, a combination of the SNP markers at genomic base positions 1235096 and 2851331, a combination of the SNP markers at genomic base positions 1235096 and 3031309, a combination of the SNP markers at genomic base positions 1235096 and 3419928, a combination of the SNP markers at genomic base positions 1235096 and 1602810, a combination of the SNP markers at genomic base positions 1235096 and

2585036, a combination of the SNP markers at genomic base positions 1462869 and 1549858, a combination of the SNP markers at genomic base positions 1462869 and 2367860, a combination of the SNP markers at genomic base positions 1462869 and 2851331, a combination of the SNP markers at genomic base positions 1462869 and 3031309, a combination of the SNP markers at genomic base positions 1462869 and 3419928, a combination of the SNP markers at genomic base positions 1462869 and 1602810, a combination of the SNP markers at genomic base positions 1462869 and 2585036, a combination of the SNP markers at genomic base positions 1549858 and 2367860, a combination of the SNP markers at genomic base positions 1549858 and 2851331, a combination of the SNP markers at genomic base positions 1549858 and 3031309, a combination of the SNP markers at genomic base positions 1549858 and 3419928, a combination of the SNP markers at genomic base positions 1549858 and 1602810, a combination of the SNP markers at genomic base positions 1549858 and 2585036, a combination of the SNP markers at genomic base positions 2367860 and 2851331, a combination of the SNP markers at genomic base positions 2367860 and 3031309, a combination of the SNP markers at genomic base positions 2367860 and 3419928, a combination of the SNP markers at genomic base positions 2367860 and 1602810, a combination of the SNP markers at genomic base positions 2367860 and 2585036, a combination of the SNP markers at genomic base positions 2851331 and 3031309, a combination of the SNP markers at genomic base positions 2851331 and 3419928, a combination of the SNP markers at genomic base positions 2851331 and 1602810, a combination of the SNP markers at genomic base positions 2851331 and 2585036, a combination of the SNP markers at genomic base positions 3031309 and 3419928, a combination of the SNP markers at genomic base positions 3031309 and 1602810, a combination of the SNP markers at genomic base positions 3031309 and 2585036, a combination of the SNP markers at genomic base

positions 3419928 and 1602810, a combination of the SNP markers at genomic base positions 3419928 and 2585036, a combination of the SNP markers at genomic base positions 1602810 and 2585036; or a combination having 3, 4, 5, 6, 7, 8, 9, 10, or 11 SNP markers, formed on the basis of the above-said combinations by optionally incorporating a 3rd, 4th, 5th, 6th, 7th, 8th, 9th, 10th, or 11th SNP marker.

Typical but not limited examples of cross-category combinations of SNP markers include: a combination of any of 1, 2, 3, 4, 5, 6, or 7 SNP markers from Category 1 with any of 1 or 2 SNP markers from Category 2; or a combination of any of 1, 2, 3, 4, 5, 6, or 7 SNP markers from Category 1 with any of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or 11 SNP markers from Category 3; or a combination of any of 1 or 2 SNP markers from Category 2 with any of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or 11 SNP markers from Category 3; or a combination of any of 1, 2, 3, 4, 5, 6, or 7 SNP markers from Category 1 with any of 1 or 2 SNP markers from Category 2 and with any of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or 11 SNP markers from Category 3.

In the present invention, based on the specimens of *Clostridium difficile* hypervirulent ribotype 027 collected from 16 countries around the world (a total of 269 strains), the evolutionary relationship of the strains of the *Clostridium difficile* hypervirulent ribotype 027 is constructed by means of whole genome sequencing, SNP calling, among other approaches, and evolutionary branches with significantly different proportions of drug-resistance and their corresponding markers are identified. RT027 is the principal ribotype in clade2. Cross distribution of RT027 along with other ribotypes such as RT198 and RT176 occurs in the whole-genome phylogenetic tree. Further, using the markers obtained, cluster analysis of clade2 and statistics of proportions of drug-resistance are conducted, revealing high consistency with RT027 grouping, because the markers can distinguish the reduced susceptibility related evolutionary branches in the

whole clade2. Specifically, the work involves the following procedures:

(1) All of the strains are subjected to drug sensitivity test using three types of drugs (moxifloxacin MOX, metronidazole MET, vancomycin VAN, of which moxifloxacin belongs to fluoroquinolone antibiotics, closely related to the outbreak of *Clostridium difficile*, the latter two are the currently available therapeutic drugs against *Clostridium difficile*) to obtain resistant or sensitive phenotypes.

(2) The hypervirulent RT027 strains have an average resistance rate of higher than 20% to any of moxifloxacin, metronidazole or vancomycin (see Table 3). However, by defining resistant or sensitive branches based on the evolutionary relationship and the drug sensitivity test results, the RT027 strains can be further subdivided into three categories, each category having a distinctly different resistance rate. Figure 1 shows the three drug-resistant branches of *Clostridium difficile* hypervirulent ribotype 027 based on the resistance to the MOX drug, wherein the strains represented by the grey dotted lines belong to Category 1 and comprise a total of 56 strains; the strains represented by the black solid lines belong to Category 2 (sensitive branch) and comprise a total of 23 strains; and the strains represented by the grey solid lines belong to Category 3 and comprise a total of 190 strains.

(3) The SNP markers in each branch are identified by:

- a) screening SNP sites genome-wide;
- b) skipping the site at which the bases occurring at the highest frequency for the strains of the three categories are the same; and
- c) skipping the site at which there is unknown information (e.g., due to sequencing errors or insufficient sequencing coverage, the base for a strain at the site is unknown).

(4) The base type and base frequency information of the sites remaining

after the screening in step (3) is analyzed, and the SNP sites at which the main bases in the same strain category have a frequency of greater than a certain threshold value (e.g., 90% or 95% or 100% etc.) are selected as the markers for distinguishing different categories of the strains, as shown in Table 1.

(5) Strains of different categories are distinguished and identified by using the SNP markers obtained in step (4) individually or in combination, depending on the application scenario.

(6) Based on the result of identification of the strain category, specific phenotypic information of the category is obtained.

(7) Clustering analysis is conducted on the 383 strains in clade2 by using the marker sites shown in Table 1, and statistics of proportions of drug-resistance is conducted on each sub-category, revealing high consistency with the categorization of RT027, as shown in Table 4.

Table 1. 20 strain-specific SNP markers

SNP site (genomic position)	Base (percentage of the bases in the category)			Category specificity
	Category 1 (56 strains)	Category 2 (23 strains)	Category 3 (190 strains)	
1029237	A (100.0%)	G (100.0%)	G (100.0%)	Category 1
118669	A (100.0%)	A (100.0%)	G (100.0%)	Category 3
1205250	G (100.0%)	G (100.0%)	C (100.0%)	Category 3
1205938	T (100.0%)	C (100.0%)	C (100.0%)	Category 1
1235096	A (100.0%)	A (100.0%)	C (100.0%)	Category 3
1462869	C (100.0%)	C (100.0%)	A (100.0%)	Category 3
1549858	A (100.0%)	A (100.0%)	T (100.0%)	Category 3
1550363	C (100.0%)	A (100.0%)	C (100.0%)	Category 2
2367860	G (100.0%)	G (100.0%)	A (100.0%)	Category 3
2487991	A (100.0%)	G (100.0%)	G (100.0%)	Category 1
2851331	G (100.0%)	G (100.0%)	A (100.0%)	Category 3
2861888	A (100.0%)	G (100.0%)	G (100.0%)	Category 1
3031309	A (100.0%)	A (100.0%)	C (100.0%)	Category 3
3419928	A (100.0%)	A (100.0%)	G (100.0%)	Category 3
6310	T (100.0%)	C (100.0%)	T (100.0%)	Category 2
882348	T (100.0%)	C (100.0%)	C (97.9%), T (2.1%)	Category 1

1602810	C (100.0%)	C (100.0%)	G (98.9%), T (1.1%)	Category 3
1798870	G (100.0%)	A (100.0%)	A (97.4%), G (2.6%)	Category 1
2585036	G (100.0%)	G (100.0%)	T (97.9%), G (2.1%)	Category 3
3083454	A (100.0%)	G (100.0%)	G (98.9%), A (1.1%)	Category 1

In Table 1, the location of the SNP site on the genome is the location on the whole genome sequence of *Clostridium difficile* CD196 strain as a reference genome.

Correspondingly, an example of the present invention provides a method for identifying the category of a *Clostridium difficile* strain, comprising obtaining base information at the site of at least one SNP marker of the SNP markers shown in Table 1 in a *Clostridium difficile* strain to be identified; and determining the category of the *Clostridium difficile* strain according to the base information.

Strains of different categories are distinguished and identified by using the SNP markers shown in Table 1 individually or in combination, depending on the application scenario.

In one case, base information is obtained at the site of at least one SNP marker of the SNP markers specific to Category 1, Category 2 or Category 3 shown in Table 1 in a *Clostridium difficile* strain to be identified; and it is determined whether the *Clostridium difficile* strain belongs to Category 1, Category 2 or Category 3 according to the base information. For example, in one example, if it is only necessary to identify whether a strain belonging to clade 2 (RT027) is a strain of Category 1, the base type of the strain at site 1029237 (or site 1205938, or site 2487991, or site 2861888) can be identified to determine whether the strain belongs to Category 1. And so on and so forth.

In another case, base information is obtained at the site of at least one SNP marker in each category of at least two categories of the SNP markers

specific to Category 1, Category 2 or Category 3 shown in Table 1 in a *Clostridium difficile* strain to be identified; and it is determined whether the *Clostridium difficile* strain belongs to Category 1, Category 2 or Category 3 according to the base information. For example, in one example, if it is only necessary to identify the exact category of a strain belonging to clade 2 (RT027), the base type of the strain at two category-specific sites, such as Category 1-specific site 1205938 and Category 2-specific site 6310 (or Category 1-specific site 2487991 and Category 3-specific site 3419928, or any other combinations) can be identified, and the strain can be determined to belong to Category 3 if the result of identification of site 1205938 shows that the strain does not belong to Category 1 and the result of identification of site 6310 shows that the strain does not belong to Category 2. And so on and so forth.

In one example, the base information of the site of the SNP marker is obtained by amplifying a genomic region in which the site of the SNP marker is located by using primers adapted to specifically amplify the region, followed by sequencing the region. In the present invention, the primers are designed according to conventional practice in the art, and are not particularly limited. Such primers comprise a forward primer and a reverse primer, the forward primer and the reverse primer respectively specifically binding to a genomic sense strand and a genomic antisense strand flanking the SNP marker.

Table 2 below shows the forward and reverse primers for detecting some of the SNP marker sites, as well as the category specificity of the primers.

Table 2

SNP site	Forward primer	Reverse primer	Category specificity
1029237	TGGATTATGCACTTGAATATGTGAG (SEQ ID NO: 1)	TGGTGTGTCATTTC CACTG (SEQ ID NO: 3)	Category 2, 3
	TGGATTATGCACTTGAATATGTGAA (SEQ ID NO: 2)		Category 1

1205938	ATATCAAGATATGTAGCTGCT <u>ATAGC</u> (SEQ ID NO: 4)	TTTCGACTTTCTTCT GGTCA (SEQ ID NO: 6)	Category 2, 3
	ATATCAAGATATGTAGCTGCT <u>ATAGT</u> (SEQ ID NO: 5)		Category 1
2861888	AGGTAATCATTTACGAATC <u>G</u> (SEQ ID NO: 7)	TCTGACATAAGAAA ATCCTCCTTGT (SEQ ID NO: 9)	Category 2, 3
	AGGTAATCATTTACGAATC <u>A</u> (SEQ ID NO: 8)		Category 1
1205250	AGCCAGGTTCTTCATTTAAGA <u>TGG</u> (SEQ ID NO: 10)	GGTTTCCTCTACCAT GGTTCCA (SEQ ID NO: 12)	Category 1, 2
	AGCCAGGTTCTTCATTTAAGA <u>TGC</u> (SEQ ID NO: 11)		Category 3
1235096	GTCAATTTTACAATTGTTGAC <u>AA</u> (SEQ ID NO: 13)	AGGATCAAATTCCT CAAACCTCA (SEQ ID NO: 15)	Category 1, 2
	GTCAATTTTACAATTGTTGAC <u>AC</u> (SEQ ID NO: 14)		Category 3
1462869	TCTTCTCTTTTGGGTATATTAG <u>TTGC</u> (SEQ ID NO: 16)	AGCAGCTGGATTTT CTGGGT (SEQ ID NO: 18)	Category 1, 2
	TCTTCTCTTTTGGGTATATTAG <u>TTGA</u> (SEQ ID NO: 17)		Category 3
2367860	TCGCATAAAGGTATTGCAAAA <u>CTTG</u> (SEQ ID NO: 19)	AGGTCAAAAACAGG AAGTGGA (SEQ ID NO: 21)	Category 1, 2
	TCGCATAAAGGTATTGCAAAA <u>CTTA</u> (SEQ ID NO: 20)		Category 3
2851331	ACCTCCTAGTATATTATTGAG (SEQ ID NO: 22)	AAAGGTGGAGGTTT TAAGGA (SEQ ID NO: 24)	Category 1, 2
	ACCTCCTAGTATATTATTGAA (SEQ ID NO: 23)		Category 3
882348	AGAGTCTCCAAGCATTTCAGT (SEQ ID NO: 25)	ACTACTGCTGGTCA TGGTTCT (SEQ ID NO: 27)	Category 2, 3
	AGAGTCTCCAAGCATTTCAGC (SEQ ID NO: 26)		Category 1
1602810	TCCATCAGTTCTTGCAAAGTA <u>C</u> (SEQ ID NO: 28)	GCTTGGTGCAACAG ATAAGT (SEQ ID NO: 31)	Category 1, 2
	TCCATCAGTTCTTGCAAAGTA <u>G</u> (SEQ ID NO: 29)		Category 3
	TCCATCAGTTCTTGCAAAGTA <u>T</u> (SEQ ID NO: 30)		Category 3
2585036	ACTTTTATTGCTTATTACTTC <u>ATTG</u> (SEQ ID NO: 32)	AGGTGAGTGTGAAT GTTCAATTGA (SEQ ID NO: 34)	Category 1, 2
	ACTTTTATTGCTTATTACTTC <u>ATT</u> (SEQ ID NO: 33)		Category 3

In Table 2, the SNP site refers to the position on the genome of the *Clostridium difficile* CD196 strain; the SNP marker site in the primer sequences is underlined; and the two pairs of primers used for identifying the same SNP marker site share the same reverse primer, and are only different in the last base at the 3' end of the forward primers.

For the primers in Table 2, in actual applications, fluorescent labels of different colors can be respectively added to the two pairs of primers that recognize the same SNP marker site, and the category of the strain can be determined according to the difference in the fluorescence of two colors in the process of qPCR; alternatively, the sequence can be amplified using conventional PCR, and the SNP marker site can be determined by means of mass spectrometry or sequencing; and further alternatively, the primers can be reconstructed as hybridization probes for use in DNA chips.

The present invention finds use in clinical diagnosis and treatment. In one example, a method is provided for diagnosing the category of a *Clostridium difficile* strain in a subject infected with the strain, comprising: obtaining base information at the site of at least one SNP marker of the SNP markers according to the present invention in the *Clostridium difficile* strain from the subject; and determining the category of the *Clostridium difficile* strain according to the base information.

In one case, base information is obtained at the site of at least one SNP marker of the SNP markers specific to Category 1, Category 2 or Category 3 of the present invention in a *Clostridium difficile* strain from the subject infected with the strain; and it is determined whether the *Clostridium difficile* strain belongs to Category 1, Category 2 or Category 3 according to the base information. For example, in one example, if it is only necessary to identify whether a *Clostridium difficile* strain belonging to clade 2 (RT027) from the subject infected with the strain is a strain of Category 1, the base type of the strain at site 1029237 (or site 1205938, or site 2487991, or site 2861888) can be identified to determine whether the strain belongs to Category 1. And so on and so forth.

In another case, base information is obtained at the site of at least one SNP marker in each category of at least two categories of the SNP markers specific to Category 1, Category 2 or Category 3 of the present invention in

a *Clostridium difficile* strain from the subject infected with the strain; and it is determined whether the *Clostridium difficile* strain belongs to Category 1, Category 2 or Category 3 according to the base information. For example, in one example, if it is only necessary to identify the exact category of a *Clostridium difficile* strain belonging to clade 2 (RT027) from the subject infected with the strain, the base type of the strain at two category-specific sites, such as Category 1-specific site 1205938 and Category 2-specific site 6310 (or Category 1-specific site 2487991 and Category 3-specific site 3419928, or any other combinations) can be identified, and the strain can be determined to belong to Category 3 if the result of identification of site 1205938 shows that the strain does not belong to Category 1 and the result of identification of site 6310 shows that the strain does not belong to Category 2. And so on and so forth.

In the present invention, based on the result of identification of the strain category, specific phenotypic information of the category can be obtained.

The clade 2 (hypervirulent RT027) strains have an average resistance rate of higher than 30% to any of moxifloxacin, metronidazole or vancomycin (see Table 3). However, by defining resistant or sensitive branches based on the evolutionary relationship and the drug sensitivity test results, the RT027 strains can be further subdivided into three categories, each category having a distinctly different resistance rate. The hypervirulent RT027 strains of Category 2 have a low resistance rate to moxifloxacin and metronidazole, which is only 4.35% and 0.0%, respectively, while the hypervirulent RT027 strains of Category 2 and Category 3 have a resistance rate of slightly higher than 20% to vancomycin, which is 21.74% and 21.05%, respectively. Similarly, the clade2 strains of Category 2 have a low resistance rate to moxifloxacin and metronidazole, which is only 2.08% and 3.03%, respectively, while the clade2 strains of Category 2 and Category 3 have a resistance rate of higher than 20% to vancomycin, which is 30.21%

and 22.58%, respectively. Clinical subdivision of the strains to obtain phenotypic information of the corresponding category will help to improve the correct use of antibiotics, reduce the wastage of medical resources, and reduce the suffering of patients. Therefore, the category-specific SNP sites according to the present invention have a significant clinical application value.

Table 3. Antibiotics resistance rate of 269 RT027 strains

Antibiotics	Resistance rate to Category 1 (56 strains)	Resistance rate to Category 2 (23 strains)	Resistance rate to Category 3 (190 strains)	Average resistance rate
Moxifloxacin	96.43%	4.35%	98.42%	89.96%
Metronidazole	51.79%	0.0%	47.37%	44.24%
Vancomycin	94.64%	21.74%	21.05%	36.43%

Table 4. Antibiotics resistance rate of 383 clade2 strains

Antibiotics	Resistance rate to Category 1 (70 strains)	Resistance rate to Category 2 (96 strains)	Resistance rate to Category 3 (217 strains)	Average resistance rate
Moxifloxacin	94.29%	2.08%	98.62%	73.63%
Metronidazole	52.86%	3.13%	48.39%	37.86%
Vancomycin	92.86%	30.21%	22.58%	37.34%

Based on the phenotypic information from the above resistance rates, an example of the present invention provides a method for treating a subject infected with a *Clostridium difficile* strain, comprising obtaining base information at the site of at least one SNP marker of the SNP markers according to the present invention in the *Clostridium difficile* strain from the subject; determining the category of the *Clostridium difficile* strain according to the base information; and administering moxifloxacin and/or metronidazole to the subject when the category of *Clostridium difficile* strain comprises Category 2; and optionally, administering vancomycin to the subject when the category of the *Clostridium difficile* strain comprises Category 2 and/or Category 3.

The SNP markers and primers according to the present invention are all

useful in the identification of the category of a *Clostridium difficile* strain, or in the diagnosis of the category of a *Clostridium difficile* strain from a subject infected with the strain. Therefore, an example of the present invention provides a use of the SNP markers and primers according to the present invention in the identification of the category of a *Clostridium difficile* strain, or in the diagnosis of the category of a *Clostridium difficile* strain from a subject infected with the strain.

The present invention has been described above with reference to specific examples, which are merely intended to aid the understanding of the present invention and are not intended to limit the present invention thereto. Several simple derivations, variations or substitutions can be made by a person skilled in the art to which the present invention pertains in light of the concept of the present invention.

CLAIMS

1. SNP markers of drug reduced susceptibility related evolutionary branches of *Clostridium difficile*, wherein the *Clostridium difficile* is the *Clostridium difficile* clade2, and the SNP markers are selected from the group comprising any one of the SNP markers in the following three categories, or any combination thereof:

(a) Category 1: an A base at genomic base position 1029237, a T base at genomic base position 1205938, an A base at genomic base position 2487991, an A base at genomic base position 2861888, a T base at genomic base position 882348, a G base at genomic base position 1798870, and an A base at genomic base position 3083454;

(b) Category 2: a C base at genomic base position 6310, and an A base at genomic base position 1550363;

(c) Category 3: a G base at genomic base position 118669, a C base at genomic base position 1205250, a C base at genomic base position 1235096, an A base at genomic base position 1462869, a T base at genomic base position 1549858, an A base at genomic base position 2367860, an A base at genomic base position 2851331, a C base at genomic base position 3031309, a G base at genomic base position 3419928, a G base at genomic base position 1602810, and a T base at genomic base position 2585036.

2. A method for identifying the category of a *Clostridium difficile* strain, wherein the method comprises obtaining base information at the site of at least one SNP marker of the SNP markers according to claim 1 in a *Clostridium difficile* strain to be identified; and determining the category of the *Clostridium difficile* strain according to the base information.

3. The method according to claim 2, wherein the method comprises obtaining base information at the site of at least one SNP marker in Category

1, Category 2 or Category 3 according to claim 1 in a *Clostridium difficile* strain to be identified; and determining whether the *Clostridium difficile* strain belongs to said Category 1, Category 2 or Category 3 according to the base information.

4. The method according to claim 2, wherein the method comprises obtaining base information at the site of at least one SNP marker in each category of at least two categories of Category 1, Category 2 or Category 3 according to claim 1 in a *Clostridium difficile* strain to be identified; and determining whether the *Clostridium difficile* strain belongs to said Category 1, Category 2 or Category 3 according to the base information.

5. The method according to claim 2, wherein the method obtains the base information of the site of said SNP marker by amplifying a genomic region in which the site of said SNP marker is located by using primers adapted to specifically amplify the region, followed by sequencing the region.

6. A method for diagnosing the category of a *Clostridium difficile* strain in a subject infected with the strain, wherein the method comprises obtaining base information at the site of at least one SNP marker of the SNP markers according to claim 1 in the *Clostridium difficile* strain from the subject; and determining the category of the *Clostridium difficile* strain according to the base information.

7. The method according to claim 6, wherein the method comprises obtaining base information at the site of at least one SNP marker in Category 1, Category 2 or Category 3 according to claim 1 in the *Clostridium difficile* strain from the subject; and determining whether the *Clostridium difficile* strain belongs to said Category 1, Category 2 or Category 3 according to the base information.

8. The method according to claim 6, wherein the method comprises

obtaining base information at the site of at least one SNP marker in each category of at least two categories of Category 1, Category 2 and Category 3 according to claim 1 in the *Clostridium difficile* strain from the subject; and determining whether the *Clostridium difficile* strain belongs to said Category 1, Category 2 or Category 3 according to the base information.

9. The method according to claim 6, wherein the method obtains the base information of the site of said SNP marker by amplifying a genomic region in which the site of said SNP marker is located by using primers adapted to specifically amplify said region, followed by sequencing said region.

10. A method for treating a subject infected with a *Clostridium difficile* strain, wherein the method comprises obtaining base information at the site of at least one SNP marker of the SNP markers according to claim 1 in the *Clostridium difficile* strain from the subject; determining the category of the *Clostridium difficile* strain according to the base information; and administering moxifloxacin and/or metronidazole to the subject when the category of *Clostridium difficile* strain comprises said Category 2; and optionally, administering vancomycin to the subject when the category of the *Clostridium difficile* strain comprises said Category 2 and/or Category 3.

11. The method according to claim 10, wherein the method obtains the base information of the site of said SNP marker by amplifying a genomic region in which the site of said SNP marker is located by using primers adapted to specifically amplify the region, followed by sequencing the region.

12. Primers adapted to specifically amplify a genomic region in which the site of a SNP marker according to claim 1 is located, wherein the primers comprise a forward primer and a reverse primer, the forward primer and the reverse primer respectively specifically binding to a genomic sense strand

and a genomic antisense strand flanking said SNP marker.

13. The primers according to claim 12, wherein the primers comprise a fluorescent label and are adapted for fluorescence quantitative PCR.

14. The primers according to claim 12, wherein the primers serve as hybridization probes to be immobilized on a chip to capture a sequence of the genomic region in which the site of said SNP marker is located.

15. Use of the primers according to any one of claims 12 to 14 in the identification of the category of a *Clostridium difficile* strain, or in the diagnosis of the category of a *Clostridium difficile* strain from a subject infected with the strain.

16. Use of the SNP markers according to claim 1 in the identification of the category of a *Clostridium difficile* strain, or in the diagnosis of the category of a *Clostridium difficile* strain from a subject infected with the strain.

17. A kit, wherein the kit comprises primers comprising a forward primer and a reverse primer, the forward primer and the reverse primer respectively specifically binding to a genomic sense strand and a genomic antisense strand flanking a SNP marker according to claim 1, and adapted to specifically amplify a genomic region in which the site of said SNP marker is located; optionally, the kit further comprises PCR amplification components besides said primers.

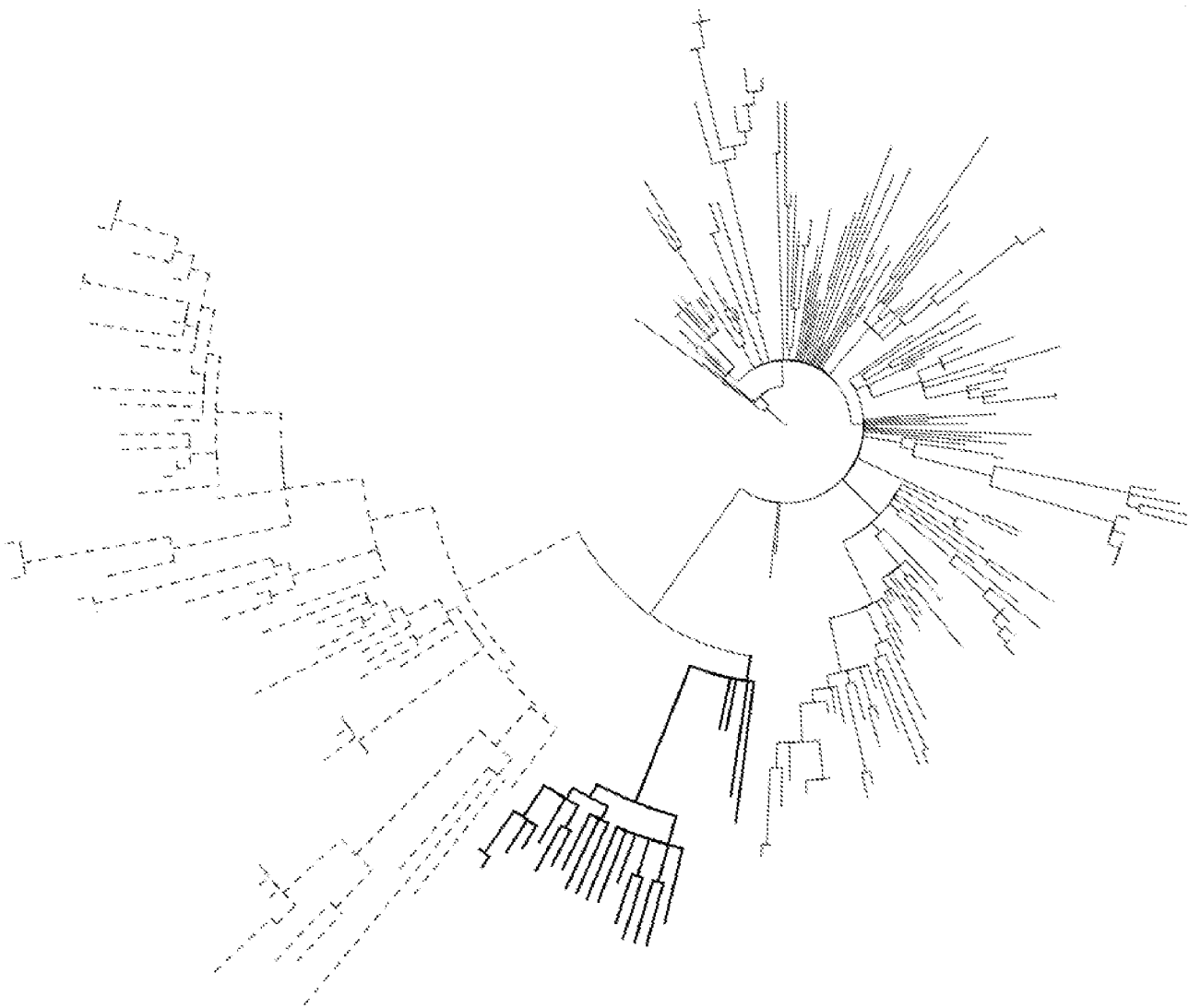


Fig. 1

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2019/092585

A. CLASSIFICATION OF SUBJECT MATTER

C12Q 1/689(2018.01)i; C12Q 1/04(2006.01)i; G16B 20/20(2019.01)i; C12R 1/145(2006.01)n

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

C12Q; G16B; C12R

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

DWPI,CNABS,CPEA,SIPOABS,VEN,CNTXT,EPTXT,USTXT,WOTXT,CNKI,PUBMED,ELSEVIER, SPRINGER,NCBI, EMBL-EBI: Clostridium, difficile, 027, SNP, sensi+, resist+, antibiotic, moxifloxacin, MOX, metronidazole, MET, vancomycin, VAN, clade, evolutionary branches.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CARMAN, R.J.et al. "Diversity of moxifloxacin resistance during a nosocomial outbreak of a predominantly ribotype ARU 027 Clostridium difficile diarrhea" <i>ANAEROBE</i> , Vol. 15, 08 October 2009 (2009-10-08), pages 244-248	1-17
A	WO 03079241 A1 (DIATECH PTY, LTD.) 25 September 2003 (2003-09-25) the whole document	1-17
A	WO 2013163210 A1 (ROLFE, P. A. et al.) 31 October 2013 (2013-10-31) the whole document	1-17
A	WO 2012087135 A1 (ACADEMISCH ZIEKENHUIS LEIDEN H.O.D.N. LUMC) 28 June 2012 (2012-06-28) the whole document	1-17
A	CN 104928376 A (WUHAN UNIVERSITY) 23 September 2015 (2015-09-23) the whole document	1-17



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

31 December 2019

Date of mailing of the international search report

23 March 2020

Name and mailing address of the ISA/CN

National Intellectual Property Administration, PRC
6, Xitucheng Rd., Jimen Bridge, Haidian District, Beijing
100088
China

Facsimile No. (86-10)62019451

Authorized officer

ZHONG,Hui

Telephone No. 86-(10)-53961936

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2019/092585

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JIA, H.B.et al. "Nosocomial transmission of Clostridium difficile ribotype 027 in a Chinese hospital, 2012-2014, traced by whole genome sequencing" <i>BMC GENOMICS</i> , Vol. 17, 26 May 2016 (2016-05-26), Article No. 405	1-17
A	STABLER, R.A.et al. "NC_013315.1:Clostridium difficile CD196 complete genome, strain CD196" <i>NCBI:GENBANK DATABASE</i> , 18 May 2017 (2017-05-18), gene 6066..8492	1-17
A	VALIENTE, E.et al. "The Clostridium difficile PCR ribotype 027 lineage: a pathogen on the move" <i>CLINICAL MICROBIOLOGY AND INFECTION</i> , Vol. 20, No. 5, 13 March 2014 (2014-03-13), pages 396-404	1-17
A	XIAO, K.L.et al. "Detection of genotype and toxin associated gene of hypervirulent Clostridium difficile clinical isolates" <i>INT.J.LAB.MED.</i> , Vol. 36, No. 8, 30 April 2015 (2015-04-30), pages 1021-1025	1-17

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2019/092585

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☒ Claims Nos.: **10-11**
because they relate to subject matter not required to be searched by this Authority, namely:

[1] The subject matter of claims 10-11 relates to methods of treating disease in a subject, and therefore does not warrant an international search according to the criteria set out in Rule 39.1(iv). However, the search has been carried out and based on the manufacturing of a medicament.
2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2019/092585

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
WO	03079241	A1	25 September 2003	NZ	535264	A	31 August 2007
				EP	1490817	A4	01 October 2008
				AU	2003209837	A1	29 September 2003
				AU	2011201392	A1	14 April 2011
				AU	PS115502	D0	18 April 2002
				AU	2003209837	B2	01 October 2009
				CA	2479469	A1	25 September 2003
				AU	PS115502	A0	18 April 2002
				US	2006218182	A1	28 September 2006
				EP	1490817	A1	29 December 2004
WO	2013163210	A1	31 October 2013	US	2015344973	A1	03 December 2015
WO	2012087135	A1	28 June 2012	None			
CN	104928376	A	23 September 2015	CN	104928376	B	14 December 2018