

TITLE

METHOD AND KIT FOR IDENTIFYING DRUG-RESISTANT FUNGAL PATHOGEN

CROSS REFERENCE TO RELATED APPLICATION

[0001] This application claims the benefit of U.S. Provisional Application No. 63/550,611, filed on February 7th, 2024. The content of the application is incorporated herein by reference.

BACKGROUND OF THE INVENTION

1. FIELD OF THE INVENTION

[0002] The present application hereby incorporates by reference the entire contents of the text file named “NHR-P0029-PUS-Sequencing Listing.xml” in XML format. The text file containing the Sequencing Listing of the present application was created on Jan. 22, 2025 and is 45,519 bytes in size.

[0003] The present relates to a method and a kit for identifying drug-resistant fungal pathogen, and particularly to a method and a kit for identifying drug-resistant *Candida tropicalis* with clade 4 genotype.

2. DESCRIPTION OF THE PRIOR ART

[0004] Approximately 2.5 million deaths are related to fungal infections annually, posing serious public health problems, and drug-resistant pathogenic fungi are increasing due to the extensive use of drugs. The number of fungal infection cases is underestimated since the difficulty of diagnosis. The estimated cost of fungal infections in the United States was approximately USD 11.5 billion in 2019, of which medical costs accounted for USD 7.5 billion.

[0005] *Candida* species (spp.) are typically harmless eukaryotic commensal yeasts found in various environmental, human, and other mammalian sources. Within humans, *Candida* spp. play a vital role as a part of the microflora present on the skin, oral, and vaginal mucosa, as well as within the gastrointestinal tract. However, in cases of severely compromised immunity, *Candida* spp. may transform into opportunistic pathogens, leading to illness and even lethality. The drug-resistant clone of *Candida* spp. is spreading in many countries. Currently, whole genome sequencing or conventional biochemical identification is commonly used to identify the predominant drug-resistant genotype of *Candida* spp. However, these methods are time and cost consuming.

[0006] In view of the foregoing, there is an unmet need in the art to develop a rapid detection method to identify the predominant drug-resistant genotype of *Candida* spp. Therefore, this information may help healthcare professionals to determine appropriate antifungal drugs for the treatment of fungal infections.

SUMMARY OF THE INVENTION

[0007] In view of the foregoing, the present disclosure provides a method for identifying a drug-resistant fungal pathogen. The method includes: detecting a single nucleotide polymorphism genotype on *XYR1* gene and *SNQ2* gene of a fungal pathogen.

[0008] The present disclosure further provides a kit for identifying a drug-resistant fungal pathogen. The kit includes: a lysis buffer for breaking down a fungal pathogen; and a primer for detecting *XYR1* gene and/or *SNQ2* gene of the fungal pathogen.

[0009] The objectives of the present disclosure will become readily understandable to those of ordinary skill in the art after reading the following detailed description of the embodiments that are illustrated in the various figures and drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

[0010] The patent or application file contains at least one drawing executed in color. Copies of this patent or patent application publication with color drawing(s) will be provided by the Office upon request and payment of the necessary fee.

[0011] FIG. 1 is a schematic representation of the two *XYR1* genes according to an embodiment of the present disclosure.

[0012] FIG. 2 is a box plot illustrating the distance of intragenotype and intergenotype for two *XYR1* genes according to an embodiment of the present disclosure (the box represents the interquartile range, with the lines extending the range of the data; the points outside the range are mild outliers, with values greater than 1.5× the upper bound of the interquartile range; and “**” indicates p value < 0.01).

[0013] FIG. 3A is an image illustrating amplification of target genes by multiplex polymerase chain reaction (PCR) according to an embodiment of the present disclosure (the template for the multiplex PCR is ATCC750; and the nine amplicons are divided into three groups based on size

of PCR product, including Group 1 (blue, ~750 bp): *ICLI*, *MDRI*, *SAPT2*, *SAPT4*, *ZWF1a*, and *SNQ2*; Group 2 (green, ~1 kb): CTRG_05978 and GTRG_05993; and Group 3 (orange, ~1.5 kb): *ERG11*).

[0014] FIG. 3B is an image illustrating results of multiplex PCR at two annealing temperatures, 60°C (lanes 1 and 2) and 61°C (lanes 3 and 4) according to an embodiment of the present disclosure (negative controls are shown in lanes 2 and 4).

[0015] FIG. 4 is a bar chart showing the percentage of reads for nine amplicons across seven isolates according to an embodiment of the present disclosure (percentage are calculated using Samtools to determine the identity of individual reads).

[0016] FIG. 5A is a box plot illustrating distribution of read percentages for six MLST target genes (*SAPT2*, *SAPT4*, *XYRI*, *ZWF1*, *ERG11*, *SNQ2*) across 96 samples according to an embodiment of the present disclosure (the boxplot represents the interquartile range for read percentages of each gene, with the median marked by a line inside the box; and whiskers extend to the minimum and maximum values of the data, and individual outliers are plotted).

[0017] FIG. 5B is bar chart illustrating sequencing success rates for six MLST genes under four single nucleotide polymorphism (SNP) filter thresholds using multiplex PCR-MinION sequencing across 96 *C. tropicalis* isolates according to an embodiment of the present disclosure (bars represent percentage of accuracy compared to available sequencing results in the art using beltools; and the four SNP filter thresholds include: Filter 1: minimum 40% reads for all MLST genes, except *XYRI* (20%). Filter 2: minimum 30% reads for all MLST except *XYRI*(15%). Filter 3: minimum 25% reads for all MLST gens, except *XYRI* (10%). Filter 4: minimum 10% reads for all MLST genes).

[0018] FIG. 5C is bar chart illustrating optimization of read counts for accurate assembly in multiplex PCR-MinION sequencing according to an embodiment of the present disclosure (accuracy rates for identifying six MLST gene fragments in 96 *C. tropicalis* isolates under varying read count thresholds (5,000, 7,000, 10,000, 12,000) with a consistent 10% SNP filter threshold is applied).

[0019] Left panel of FIG. 6 is an allele discrimination plot for CTRG_05978 displaying alleles as GA/GA homozygous (lower right red cluster), GA/AA heterozygous (middle green cluster),

and AA/AA homozygous (upper left blue cluster) according to an embodiment of the present disclosure (the middle green cluster strains correspond to the CTRG_05978 type IV genotype (GA/AA), and 6 strains are located in the middle green cluster, including YM140156, YM180233, YM180351, YM180473, YM180681, and YM180969).

[0020] Right panel of FIG. 6 is an allele discrimination plot for *SNQ2* A2977G, presenting alleles as G/G homozygous (lower right red cluster), G/A heterozygous (middle green cluster), and A/A homozygous (upper left blue cluster) according to an embodiment of the present disclosure (eight strains exhibit mutations in *SNQ2* mutation; and by combining the results of CTRG_05978 and *SNQ2* genotypes using the TaqMan Genotyping Assay, the identification of CTRG_05978 type IV and *SNQ2* A2977G becomes straightforward).

[0021] FIG. 7 is a schematic workflow illustrating diagnosis and management of the clade 4 *C. tropicalis* by detecting CTRG_05978 SNP and *SNQ2* A2977G mutation according to an embodiment of the present disclosure (the SNPs are marked by red arrows; and yeasts in red refers predominant drug-resistant clade 4 genotype and in other colors represents different genotypes).

DETAILED DESCRIPTION

[0022] The following embodiments are provided to illustrate the present disclosure in detail. A person having ordinary skills in the art can easily understand the advantages and effects of the present disclosure after reading the disclosure of this specification, and also can implement or apply in other different embodiments. Therefore, it is possible to modify and/or alter the following embodiments for carrying out this disclosure without contravening its scope for different aspects and applications, and any element or method within the scope of the present disclosure disclosed herein can combine with any other element or method disclosed in any embodiments of the present disclosure.

[0023] In order that the present disclosure may be more readily understood, certain terms are first defined. In addition, it should be noted that whenever a value or range of values of a parameter are recited, it is intended that values and ranges intermediate to the recited values are also intended to be part of the present disclosure.

[0024] As used herein, the singular forms “a,” “an,” and “the” include plural referents, unless expressly and unequivocally limited to one referent. The term “or” is used interchangeably with

the term “and/or” unless the context clearly indicates otherwise.

[0025] As used herein, the term “comprising,” “comprises” “include,” “including,” “have,” “having,” “contain,” “containing,” and any other variations thereof are intended to cover a non-exclusive inclusion. For example, when describing an object “comprises” a limitation, unless otherwise specified, it may additionally include other ingredients, elements, components, structures, regions, parts, allele, locus, steps, or connections, etc., and should not exclude other limitations.

[0026] In at least one embodiment of the present disclosure, the single nucleotide polymorphism genotype on the *XYRI* gene may include a sequence of ATRA, and R is A or G. In some embodiments, the sequence of ATRA may be a first 4-mer of the *XYRI* gene.

[0027] In at least one embodiment of the present disclosure, the single nucleotide polymorphism genotype on the *SNQ2* gene may include a nucleotide A at position 2977 substituted with a nucleotide G.

[0028] In at least one embodiment of the present disclosure, the fungal pathogen may be a non-*albicans Candida* species. In some embodiments, the non-*albicans Candida* species is at least one selected from the group consisting of *Candida tropicalis*, *Candida parapsilosis*, *Candida glabrata*, *Candida auris*, *Candida guilliermondii*, and *Candida krusei*. In some embodiments, the non-*albicans Candida* species may be *Candida tropicalis*. In some embodiments, the *Candida tropicalis* may include a clade 4 genotype.

[0029] In at least one embodiment of the present disclosure, the drug-resistant fungal pathogen may have a drug resistance against a drug comprising azole.

[0030] In at least one embodiment of the present disclosure, the single nucleotide polymorphism genotype on *XYRI* gene and *SNQ2* gene of the fungal pathogen may be detected by high-throughput identification. In some embodiments, the high-throughput identification may include combining a multiplex polymerase chain reaction approach and a MinION nanopore sequencing approach.

[0031] In at least one embodiment of the present disclosure, the single nucleotide polymorphism genotype on *XYRI* gene and *SNQ2* gene of the fungal pathogen may be detected

by a TaqMan genotyping assay.

[0032] In at least one embodiment of the present disclosure, the kit may further include a DNA polymerase; and a PCR buffer comprising Mg^{2+} .

[0033] In some embodiments, among the *Candida* spp. found in humans, *Candida albicans* is the most frequently isolated in hospital settings resulting in nosocomial infections. However, there has been a shift toward the more-treatment-resistant non-*albicans Candida* spp., such as *Candida tropicalis*, *Candida parapsilosis*, *Candida glabrata*, and *Candida krusei*. The prevalence of these species may vary based on the geographical location. *Candida tropicalis* (*C. tropicalis*) stands out among the non-*albicans Candida* species as the second most or even the most prevalent pathogenic yeast in some regions, such as Asia and Latin America.

[0034] In some embodiments, molecular epidemiological investigation of *C. tropicalis* may be instrumental in identifying epidemic genotype, potential modes of transmission, genetic composition of drug-resistant strains, biological habitats, and population structure thereof. The multilocus sequence typing (MLST) has been developed to characterize population structures within a species and distinguish geographical origins. The combinations of alleles at different loci yield unique diploid sequence types (DSTs), enabling differentiation and classification of genetic relatedness among strains. The present MLST approach for *C. tropicalis* utilizes six genes, including *ICL1*, *MDR1*, *SAPT2*, *SAPT4*, *XYR1*, and *ZWF1a*. Moreover, MLST has been extensively employed to assess the clonality of *C. tropicalis* and ascertain various traits, such as fluconazole susceptibility, hospital-based origins, and anatomical sources. Finally, the allelic profiles of isolates may be compared by utilizing the online MLST database to explore diverse geographical origins, drug resistance transmission, and genetic variation patterns.

[0035] In some embodiments, MLST has been also exerted to uncover that fluconazole-resistant *C. tropicalis* strains (isolates) share the same genotype, clade 4, underscoring the utility of MLST typing as a method for distinguishing drug-resistant *C. tropicalis* strains. Although MLST offers the easy application in determining the genetic type of *C. tropicalis*, the need for rapid identification and more comprehensive characterization of drug-resistant strains (isolates) remains vital in the realms of diagnosis and antifungal stewardship. Therefore, the present disclosure may elucidate the genotypes of one of *XYR1* genes, CTRG_05978, and *SNQ2* which is situated 25 kb upstream from CTRG_05978 and has been linked to drug susceptibility in *Nakaseomyces glabratus* (*Candida glabrata*) and *Candida auris*

are the potential markers enabling the rapid identification of azole-resistant *C. tropicalis*, contributing to the prevention of further dissemination of the predominant clade 4 strains (isolates).

EXAMPLES

[0036] Exemplary embodiments of the present disclosure are further described in the following examples, which should not be construed to limit the scope of the present disclosure.

Materials and Methods

***C. tropicalis* strains isolation**

[0037] The methods for isolating *C. tropicalis* from the environment are described in Taiwan surveillance of Antimicrobial Resistance of Yeasts (TSARY) studies (Yang, Y.L., Lin, C.C., Chang, T.P., Lauderdale, T.L., Chen, H.T., Lee, C.F., Hsieh, C.W., Chen, P.C. and Lo, H.J. (2012) Comparison of human and soil *Candida tropicalis* isolates with reduced susceptibility to fluconazole. PloS one, 7, e34609; and Zhou, Z.L., Lin, C.C., Chu, W.L., Yang, Y.L. and Lo, H.J. (2016) The distribution and drug susceptibilities of clinical *Candida* species in TSARY 2014. Diagnostic microbiology and infectious disease, 86, 399-404). All yeast strains are identified by the sequences of the rDNA (Leaw, S.N., Chang, H.C., Barton, R., Bouchara, J.P. and Chang, T.C. (2007) Identification of medically important *Candida* and non-*Candida* yeast species by an oligonucleotide array. Journal of clinical microbiology, 45, 2220-2229). In order to make sure the representative population of *C. tropicalis*, at least 10 strains, each of 11 most common clades including 37 strains having whole-genome data, are analyzed. A total of 120 *C. tropicalis* strains, including ATCC750 and 119 from TSARY (Table 5) are analyzed initially to identify potential markers. To verify, 161 strains including all clade 4 whole-genome data that may be found from the NCBI database along with 10 other clades containing the most DST members are further analyzed.

DNA extraction and MLST analysis of *C. tropicalis* strains

[0038] Genomic DNA of each *C. tropicalis* strain (isolate) is extracted using a Yeast Genomic DNA Kit (Geneaid) according to the manufacturer's instructions, and DNA is quantified using the Qubit fluorometer (ThermoFisher Scientific). The DNA fragments of six genes, *ICL1*, *MDR1*,

SAPT2, *SAPT4*, *XYRI*, and *ZWF1a*, are amplified and sequenced as described in previous studies (Yang, Y.L., Lin, C.C., Chang, T.P., Lauderdale, T.L., Chen, H.T., Lee, C.F., Hsieh, C.W., Chen, P.C. and Lo, H.J. (2012) Comparison of human and soil *Candida tropicalis* isolates with reduced susceptibility to fluconazole. PloS one, 7, e34609; and Zhou, Z.L., Tseng, K.Y., Chen, Y.Z., Tsai, D.J., Wu, C.J., Chen, Y.C., Peng, H.L., Yang, Y.L., Hsieh, L.Y., Chen, C.H. et al. (2022) Genetic relatedness among azole-resistant *Candida tropicalis* clinical strains in Taiwan from 2014 to 2018. International journal of antimicrobial agents, 59, 106592). The primers used are listed below (Table 1). The resultant sequences are aligned with BioNumerics 3.0 (Applied Maths, Kortrijk, Belgium) and compared with those in the database of *C. tropicalis* (<http://pubmlst.org/website>) to determine the level of sequence identities (DST).

Table 1. Primers used in the present disclosure (“*” indicates position of translation initiation site of ATG as +1; “F” indicates forward primer; and “R” indicates reverse primer).

Name	Gene	Sequences	Position*	SEQ ID NO.
MLST				
HJL2796	<i>ICLI-F</i>	CAACAGATTGGTTG CCATCAGAGC	+675 – +698	1
HJL2797	<i>ICLI-R</i>	CGAAGTCATCAACA GCCAAAGCAG	+1412 – +1435	2
HJL2799	<i>MDRI-F</i>	GGCTGGAGATGGA CTTTTTGG	+772 – +792	3
HJL2800	<i>MDRI-R</i>	TGGAGCACCAAAC AATGGG	+1491 – +1509	4
HJL2801	<i>SAPT2-F</i>	CACAGAAACGCCTC CGTTTTG	-46 – -26	5
HJL2803	<i>SAPT2-R</i>	CACTGGTAGCTGAA GGAG	+728 – +745	6
HJL2804	<i>SAPT4-F</i>	TACCCATTGGAATG TATTGTTGA	-103 – -81	7
HJL2805	<i>SAPT4-R</i>	CCAGTTGTTGAATC TGATGAATT	+637 – +659	8
HJL2807	<i>XYRI-F</i>	CCAACAACATAAAA GAGATCGACA	-84 – -61	9
HJL2808	<i>XYRI-R</i>	GCAAGTATGGGTGA	+580 – +601	10

		TGTTCAAT		
HJL2811	<i>ZWF1a-F</i>	GTCWTATGATTCMT TYGGWGA	+3 – +23	11
HJL2812	<i>ZWF1a-R</i>	GARACTGGTCTTTC CATRGT	+751 – +770	12
HJL3097	<i>CTRG_059 78-F</i>	GACGATTACGGTCA CAGAATC	-172 – -152	13
HJL3098	<i>CTRG_059 93-F</i>	CTAATAGCAAGTTA TTGCCCC	-89 – -69	14
HJL3174	<i>SNQ2-F</i>	TGCTCTTGGTGGAT CTACTG	+2523 – +2542	15
HJL3175	<i>SNQ2-R</i>	CTTGATCTATCACC AATATCACC	+3253 – +3275	16

Whole genome sequencing

[0039] The high molecular weight DNAs are used to construct a multiplexing Nanopore sequencing library using the Oxford Nanopore Technologies (ONT) Ligation Sequencing Kit (SQK-LSK109) and Native Barcoding Expansion Kit (EXP-NBD104), following ONT instructions. A standard 72-hour sequencing script is executed on MinKNOW; raw reads are collected then base called and demultiplexed using the standalone application, Guppy.

Investigation of *XYR1* by Nanopore

[0040] The raw reads of 17 *C. tropicalis* strains (isolates) are aligned to the *XYR1* gene (XM_002546469.1) using minimap2 (Li 2018) to produce reads containing *XYR1*. The *XYR1*-containing reads are then de novo assembled using SMARTdenovo (version 1.0) (Liu et al. 2021) to form contigs.

[0041] The assembled *XYR1*-containing contigs are further polished with Medaka (version 1.5.0) using the raw reads and rearranged to maintain the same orientation as the *C. tropicalis* MYA-3404 genomic scaffold supercont3.11 (NW_003020040.1). The sequence analyses for the 17 *XYR1*-containing contigs reveal a similar schematic representation of the two *XYR1* genes compared to the MYA-3404 strain.

Multiple sequence alignments and evolutionary distances

[0042] After designing the primer sequences for sequencing the partial sequences of CTRG_05978 and CTRG_05993, 120 strains of *C. tropicalis*, including ATCC750, and 119 from clinical sources, are analyzed. The *XYRI* DNA sequence is aligned using ClustalW multiple alignment in MEGA 10. Pairwise genetic distances are separated into intraspecific and interspecific parameters and calculated to characterize both intra-and interspecific variation within and between CTRG_05978 coding DNA sequence (CDS) and CTRG_05993 CDS by using MEGA10. The statistical significance (p-value), Wilcoxon Rank Sum, and Signed Rank Tests are determined by R software version 4.1.2 (Wilcox.test package).

TaqMan SNP genotyping

[0043] The TaqMan SNP genotyping assay is performed using TaqMan Genotyping Master Mix (Applied Biosystems) and the ABI QuantStudio™ 6 Flex System (Applied Biosystems), following the manufacturer's instructions. Thermal cycle conditions in CTRG_05978 are as follows: 1 cycle of 95 °C for 10 min, 40 cycles of 95 °C for 15 sec, 54 °C for 20 sec and 60 °C for 1 min, and 1 cycle of 60 °C for 1 min; and in *SNQ2* are as follows: 1 cycle of 95 °C for 10 min, 40 cycles of 95 °C for 15 sec, 60 °C for 1 min, and 1 cycle of 60 °C for 1 min.

Statistical analysis

[0044] SPSS software for Windows (version 12.0) is used to analyze the data. The null hypothesis that the frequencies of the MLST genome type and CTRG_05978 start codon region type are equal is investigated by the chi-square (χ^2) test. The chi-squared or Fisher's exact test with 1-tailed correction is applied for categorical variables. Logistic regression is applied to assess the independent effects of factors with values less than 0.05 in univariate analysis. A p-value less than 0.05 is considered significant.

Multiplex polymerase chain reaction

[0045] Nine target gene fragments (i.e., *ICL1*, *MDR1*, *SAPT2*, *SAPT4*, *XYRI*, *ZWF1a*, *ERG11*, and *SNQ2*) are amplified using multiplex PCR with the KOD -Multi & Epi™ DNA polymerase (Toyobo, Osaka, Japan) in a 50- μ l reaction. The reaction mixture comprises 25 μ l of 2 $\times$ PCR

buffer, including dNTPs (dATP, dGTP, dCTP, and dTTP) and Mg^{2+} at a final concentration of 2.0 mM. Additionally, the mixture also contained 10 μ l of primer mix, 1 μ l of KOD -Multi & Epi-enzyme, anti-KOD DNA polymerase antibodies (to neutralize the polymerase and inhibit its 3'→5' exonuclease activities), 5 μ l of yeast lysate (in 0.02 N NaOH), and 14 μ l of nuclease-free water. The primer mix comprises primers for *ICL1*, *MDR1*, *SAPT2*, *SAPT4*, *XYR1*, *ZWF1a*, *ERG11*, and *SNQ2* at concentrations of 0.2 μ M, 0.1 μ M, 0.2 μ M, 0.15 μ M, 0.15 μ M, 0.2 μ M, 0.15 μ M, and 0.1 μ M, respectively. Primer sequences are detailed in Table 2. The PCR protocol is as follows: initial denaturation at 94°C for 2 min, followed by 35 cycles of denaturation at 98°C for 10 sec, annealing at 60°C for 30 sec, extension at 68°C for 1 min and 45 sec, and a final extension at 68°C for 10 min.

Table 2. The oligo and probes used in the present disclosure (“*” indicates position numbers indicate location of gene fragments according to the location of the ATG translation initiation site for each gene, designated as +1; “F” indicates forward primer; and “R” indicates reverse primer).

Name	Genes	Sequences	Position*	SEQ ID NO.
Multiplex PCR				
HJL3199	<i>ICL1-F</i>	AACAGATTGGTTGCTATCAGAGC	+678 – +698	17
HJL3200	<i>ICL1-R</i>	AGTCATCAACAGCCAAAGCAG	+1412 – +1432	18
HJL2799	<i>MDR1-F</i>	GGCTGGAGATGGACTTTTTGG	+772 – +792	3
HJL3202	<i>MDR1-R</i>	ACAGGATAATTTGGCGTTGCT	+1524 – +1545	19
HJL3203	<i>SAPT2-F</i>	ATTCACAGAAACGCCTCCGT	-50 – -29	20
HJL3204	<i>SAPT2-R</i>	GTACCACTGGTAGCTGAAGGA	+728 – +749	21
HJL3205	<i>SAPT4-F</i>	CCCATTGGAATGTATTGTTGACC	-101 – -83	22
HJL3206	<i>SAPT4-R</i>	TTAGCATTGTCAACCCACCA	+605 – +626	23
HJL3207	<i>CTRG_05978-F</i>	GGTCACAGAATCATTCACTTTCC	-114 – -92	24
HJL3208	<i>XYR1-R</i>	ACCAGTAATGGCAATACCAGC	+788 – +808	25
HJL3209	<i>CTRG_05993-F</i>	ACTTTCCCCCAGAGATTTTCAG	-183 – -163	26
HJL3210	<i>ZWF1a-F</i>	ATGATTCCTTTGGTGACCGTG	+7 – +27	27
HJL3211	<i>ZWF1a-R</i>	AGAAACTGGTCTTCCATGGTC	+749 – +770	28
HJL3212	<i>ERG11-F</i>	CAGTTGATTGTCATTACTAAACCGT	-593 – -568	29
HJL3213	<i>ERG11-R</i>	TTGACCACCCATCAAAACACC	+926 – +947	30

HJL3214	<i>SNQ2-F</i>	ATGCTCTTGGTGGATCTACTG	+2521 – +2542	31
HJL3215	<i>SNQ2-R</i>	CACTTGATCTATCACCAATATCACC	+3251 – +3275	32

Analysis of multiplex polymerase chain reaction

[0046] The Rapid Barcoding Kit 96 (SQK-RBK114.96) is used for rapid barcoding of the 96 samples. Each sample is mixed with 5 µl of multiplex PCR product, 2.5 µl of nuclease-free water, and 2.5 µl of a rapid barcode. The mixture is incubated at 30°C for 2 min, followed by 80°C for 2 min. All 96 barcoded DNA samples are pooled, and 120 µl of the pooled DNA is mixed with an equal volume of solid phase reversible immobilization (SPRI) beads. After 5 min of incubation with a hula mixer, the barcoded DNA is cleaned twice with 240 µl of freshly prepared 80% ethanol and eluted with 15 µl of elution buffer (EB). An aliquot of 800 ng of barcoded DNA is adjusted to a total volume of 11 µl with EB. One microliter of diluted rapid adaptor F is added to the barcoded DNA, and the mixture is incubated at room temperature for 5 min. BSA priming mix is added for optimal sequencing performance.

[0047] A pre-sequencing mix (PSM) is prepared by adding 37.5 µl of Sequencing Buffer and 25.5 µl of Library Beads (LIB) mixed immediately before use to a 12 µl DNA library. The PSM is loaded via the SpotON port into a primed Flow Cell (FLO-MIN114) for sequencing. Basecalling and demultiplexing are performed in real-time using MinKNOW GUI implemented with GPU Gupp on a desktop PC with an NVIDIA RTX 3090 graphics card with 24 GB RAM to produce high-accuracy reads in FASTQ format. The default output set for MinKNOW is a FASTQ file containing 10,000 reads. To assemble MLST genes, *ERG11*, and *SNQ2*, the FASTQ file containing reads is polished with Medaka, and heterozygous sites are identified from the presence of variant sites between 10% to 90% in those 9 alleles.

Assembly MLST gene

[0048] To evaluate the accuracy of consensus sequences generated from Multiplex PCR-MinION sequencing data, compare the consensus sequences with available sequencing results in the art. A custom Python script is used to process Nanopore FASTQ files and analyze the impact of different single nucleotide polymorphism (SNP) filtering criteria based on the alternative allele ratio. The scripts perform sequence alignment using Minimap2, consensus calling with Medaka, and variant calling using Bcftools to generate VCF files containing SNPs

and their alt/ref ratios.

[0049] The script is executed with appropriate command-line arguments to specify the input FASTQ file, reference FASTA file, and output directory. The outputs, including consensus sequences, filtered VCF files, and SNP ratio summaries, are compared to assess the effects of different SNP filtering and sequence modification criteria. The accuracy of the consensus sequences generated by the script is evaluated by aligning them to reference sequences and comparing SNP positions, base substitutions, and overall sequence concordance.

Results

***XYRI* has the highest typing efficiency among six genes (i.e., *ICLI*, *MDRI*, *SAPT2*, *SAPT4*, *XYRI*, and *ZWF1a*) included for MLST analysis**

[0050] As the MLST method is capable of differentiating the predominant drug-resistant group, clade 4, the member with the highest typing efficiency emerges as a particularly valuable candidate for further analysis. Among the six gene fragments used for MLST, a total of 49 polymorphic sites are identified in *ICLI*, 73 in *MDRI*, 93 in *SAPT2*, 106 in *SAPT4*, 63 in *XYRI*, and 49 in *ZWF1a*, according to the *C. tropicalis* MLST database. *XYRI* exhibits the highest typing efficiency, with the ability to distinguish 230 genotypes per polymorphism across 63 polymorphic sites (Table 3). The comprehensive analysis of *XYRI* sequence, which exhibited the highest typing efficiency within the MLST analysis, is preformed subsequently.

Table 3. Six *Candida tropicalis* fragments used for MLST.

Gene fragment	Length	No. of Genotypes	G+C	No. of Polymorphic sites	No. of Genotypes Polymorphic sites
<i>ICLI</i>	447	63	40.99	49	1.29
<i>MDRI</i>	425	217	34.95	73	2.97
<i>SAPT2</i>	525	87	37.25	93	0.94
<i>SAPT4</i>	390	146	32.50	106	1.38
<i>XYRI</i>	370	230	32.36	63	3.65
<i>ZWF1a</i>	520	71	35.11	49	1.45

Two *XYR1* genes are detected on one chromosome

[0051] Referring to FIG. 1, two *XYR1* genes, CTRG_05978 and CTRG_05993 (GenBank accession number: NW_003020040.1), on chromosome VI with approximately 30 kb apart from each other in opposite orientations is found. The direction of the arrows indicates the direction of two *XYR1* genes. Since the two *XYR1* genes are high consensus, in order to differentiate the sequence of MLST fragments of two *XYR1* genes, primer set HJL2808/HJL3097 and HJL2808/HJL3098 for CTRG_05978 and CTRG_05993, respectively, is employed. Seventeen different DST strains previously collected, as well as ATCC750, are analyzed, and all of them (such as MYA-3404) possess two *XYR1* genes.

Two *XYR1* genes are likely to be ancestral duplications

[0052] Two *XYR1* genes, CTRG_05978 and CTRG_05993 (GenBank accession number: NW_003020040.1), on chromosome VI with approximately 30 kb apart from each other in opposite orientations are detected (FIG. 1). A distance matrix analysis, including intra and inter-genotype distances for CTRG_05978 and CTRG_05993, is performed using MEGA X software (Kumar, S., Stecher, greater than ergenotype cohorts (median = 0.003, mean $\pm$ SD = 0.004 $\pm$ 0.005), the CTRG_05993 intergenotype cohorts (median = 0, mean $\pm$ SD = 0.002 $\pm$ 0.001), and the intragenotype cohorts (median = 0.005, mean $\pm$ SD = 0.009 $\pm$ 0.007) is generated. Notably, the largest distance is obtained from the intragenotype comparisons (0–0.022), while the distances within the CTRG_05993 intergenotype cohorts show relatively low values (0–0.005) (FIG. 2).

[0053] The fact that the intragenotype distance is greater than the individual intergenotype distances of CTRG_05978 and CTRG_05993, suggesting that these two alleles may be ancestral duplications.

The SNP of the translation initiation codon of CTRG_05978 associat with the predominant clade 4 azole-resistant *C. tropicalis* genotype

[0054] Due to the higher diversity of CTRG_05978, the sequence of CTRG_05978 is further analyzed. Several SNPs are in the translation initiation codon of CTRG_05978 (Table 4 and Table 5). One of these SNPs with G to A substitution in the start codon result in a loss-of-function mutation. A more comprehensive analysis of the first 60 nucleotides in the

codon sequences of CTRG_05978 across all 120 strains (isolates) is then conducted. Notably, the present disclosure reveals that approximately 82.5% (99 out of 120) of the strains (isolates) possess at least one allele of CTRG_05978 that harbors a loss-of-function mutation in Xyr1.

Table 4. The sequence of the first 4-mer of CTRG_05978 and CTRG_05993.

CTRG_05978 Type	Clade	DST	Strain	The first 4-mer of CTRG_05978	The first 4-mer of CTRG_05993
I	5	140	YM180156	ATAA	ATGT
II	3	872	YM180490	ATGA	ATGT
III	2	153	YM141055	ATGT	ATGT
IV	4	225	YFA120301	ATRA	ATGT
V	6	374	NC1S04	ATGW	ATGT
VI	5	910	YM140912	ATRW	ATGT
VII	3	370	YM140797	ATWA	ATGT

Table 5. SNPs of CTRG_05978 and *SNQ2* (DST stands for diploid sequence type; WT stands for wild type; Hom stands for homozygous; and Het stands for heterozygous).

CTRG_05978 Type	<i>SNQ2</i> (A2977G)	Clade	DST	Strain	The first of 60-mer of CTRG_05978	SEQ ID NO.
I (ATAA)	A/A	1	190	YM060500	ATAATTAAATTT TTCACCTCTCCA ACAACATAAAAG AGATCGACATTT CTCTCAATTAT	33
I (ATAA)	A/G	1	196	YM060342	ATAATKWWWW WWTTYWYYWCT YCWMCAACWAY WAWRARRWSRW CRWYWYTYTTC WWYWAT	34
I (ATAA)	A/G	1	582	YM140516	ATAATTAAATTT TTCACCTCTCCA ACAACATAAAAG AGATCGACATTT CTCTCAATTAT	33
I (ATAA)	A/G	1	582	YM140682	ATAATTAAATTT	33

					TTCACCTTCTCCA ACAACCTATAAAAG AGATCGACATTT CTCTTCAATTAT	
I (ATAA)	A/G	3	585	YM140298	ATAATTAAATTT TTCACCTTCTCCA ACAACCTATAAAAG AGATCGACATTT CTCTTCAATTAT	33
I (ATAA)	A/A	3	965	YM140465	ATAATTAAATTT TTCACCTTCTCCA ACAACCTATAAAAG AGATCGACATTT CTCTTCAATTAT	33
I (ATAA)	A/A	3	965	YM180451	ATAATTAAATTT TTCACCTTCTCCA ACAACCTATAAAAG AGATCGACATTT CTCTTCAATTAT	33
I (ATAA)	A/A	3	965	YM180453	ATAATTAAATTT TTCACCTTCTCCA ACAACCTATAAAAG AGATCGACATTT CTCTTCAATTAT	33
I (ATAA)	A/G	3	971	YM180043	ATAATKWWWW WWTTYWYYWCT YCWMCACWAY WAWRARRWSRW CRWYWYTYTC WWYWAT	34
I (ATAA)	G/G	4	546	YM180542	ATAATTAAATTT TTCACCTTCTCCA ACAACCTATAAAAG AGATCGACATTT CTCTTCAATTAT	33
I (ATAA)	A/A	4	508	YM180921	ATAATTAAATTT TTCACCTTCTCCA ACAACCTATAAAAG AGATCGACATTT CTCTTCAATTAT	33
I (ATAA)	A/A	5	98	YM180064	ATAATTAAATTT	33

					TTCAC TTCTCCA ACA ACTATAAAG AGATCGACATTT CTCTTCAATTAT	
I (ATAA)	A/A	5	98	YM181190	ATAATTAAATTT TTCAC TTCTCCA ACA ACTATAAAG AGATCGACATTT CTCTTCAATTAT	33
I (ATAA)	A/A	5	140	YM060098	ATAATTAAATTT TTCAC TTCTCCA ACA ACTATAAAG AGATCGACATTT CTCTTCAATTAT	33
I (ATAA)	A/A	5	140	YM060828	ATAATTAAATTT TTCAC TTCTCCA ACA ACTATAAAG AGATCGACATTT CTCTTCAATTAT	33
I (ATAA)	A/A	5	140	YM140225	ATAATTAAATTT TTCAC TTCTCCA ACA ACTATAAAG AGATCGACATTT CTCTTCAATTAT	33
I (ATAA)	A/A	5	140	YM140977	ATAATTAAATTT TTCAC TTCTCCA ACA ACTATAAAG AGATCGACATTT CTCTTCAATTAT	33
I (ATAA)	A/A	5	140	YM141125	ATAATTAAATTT TTCAC TTCTCCA ACA ACTATAAAG AGATCGACATTT CTCTTCAATTAT	33
I (ATAA)	A/A	5	140	YM180054	ATAATTAAATTT TTCAC TTCTCCA ACA ACTATAAAG AGATCGACATTT CTCTTCAATTAT	33
I (ATAA)	A/A	5	140	YM180156	ATAATTAAATTT TTCAC TTCTCCA	33

					ACAACTATAAAG AGATCGACATTT CTCTTCAATTAT	
I (ATAA)	A/A	5	140	YM180483	ATAATTAAATTT TTCACCTCTCCA ACAACTATAAAG AGATCGACATTT CTCTTCAATTAT	33
I (ATAA)	A/A	5	140	YM180486	ATAATTAAATTT TTCACCTCTCCA ACAACTATAAAG AGATCGACATTT CTCTTCAATTAT	33
I (ATAA)	A/A	5	140	YM180942	ATAATTAAATTT TTCACCTCTCCA ACAACTATAAAG AGATCGACATTT CTCTTCAATTAT	33
I (ATAA)	A/A	5	140	YM990275	ATAATTAAATTT TTCACCTCTCCA ACAACTATAAAG AGATCGACATTT CTCTTCAATTAT	33
I (ATAA)	A/A	5	140	YM990592	ATAATTAAATTT TTCACCTCTCCA ACAACTATAAAG AGATCGACATTT CTCTTCAATTAT	33
I (ATAA)	A/A	5	1142	YM990659	ATAATTAAATTT TTCACCTCTCCA ACAACTATAAAG AGATCGACATTT CTCTTCAATTAT	33
I (ATAA)	A/A	5	911	YM140479	ATAATTAAATTT TTCACCTCTCCA ACAACTATAAAG AGATCGACATTT CTCTTCAATTAT	33
I (ATAA)	A/A	5	911	YM140896	ATAATTAAATTT TTCACCTCTCCA ACAACTATAAAG	33

					AGATCGACATTT CTCTTCAATTAT	
I (ATAA)	A/A	5	953	YM180956	ATAATTAAATTT TTCACCTCTCCA ACAACTATAAAG AGATCGACATTT CTCTTCAATTAT	33
I (ATAA)	A/A	6	197	YM060109	ATAATKWWWW WWTTYWYYWCT YCWMCAACWAY WAWRARRWSRW CRWYWYTYTC WWYWAT	34
I (ATAA)	A/A	7	138	YM990236	ATAATTAAATTT TTCACCTCTCCA ACAACTATAAAG AGATCGACATTT CTCTTCAATTAT	33
I (ATAA)	A/A	7	666	YM140303	ATAATTAAATTT TTCACCTCTCCA ACAACTATAAAG AGATCGACATTT CTCTTCAATTAT	33
I (ATAA)	A/G	8	142	YM990329	ATAATKWWWW WWTTYWYYWCT YCWMCAACWAY WAWRARRWSRW CRWYWYTYTC WWYWAT	34
I (ATAA)	A/G	8	195	YM060051	ATAATKWWWW WWTTYWYYWCT YCWMCAACWAY WAWRARRWSRW CRWYWYTYTC WWYWAT	34
II (ATGA)	A/A	3	872	YM180456	ATGATTAAATTT TTCACCTCTCCA ACAACTATAAAG AGATCGACATTT CTCTTCAATTAT	35
II (ATGA)	A/A	3	872	YM180460	ATGATTAAATTT	35

					TTCACCTTCTCCA ACAACTATAAAAG AGATCGACATTT CTCTTCAATTAT	
II (ATGA)	A/A	3	872	YM180472	ATGATTAAATTT TTCACCTTCTCCA ACAACTATAAAAG AGATCGACATTT CTCTTCAATTAT	35
II (ATGA)	A/A	3	872	YM180490	ATGATTAAATTT TTCACCTTCTCCA ACAACTATAAAAG AGATCGACATTT CTCTTCAATTAT	35
II (ATGA)	A/A	7	171	YM060172	ATGATTAAATTT TTCACCTTCTCCA ACAACTATAAAAG AGATCGACATTT CTCTTCAATTAT	35
III (ATGT)	A/A	2	27	YM060481	ATGTTTAAATTTT TCACTTCTCCAA CAACTATAAAGA GATCGACATTTT TCTTCAATTAT	36
III (ATGT)	A/A	2	90	YM140365	ATGTTTAAATTTT TCACTTCTCCAA CAACTATAAAGA GATCGACATTTT TCTTCAATTAT	36
III (ATGT)	A/A	2	134	YM060299	ATGTTTAAATTTT TCACTTCTCCAA CAACTATAAAGA GATCGACATTTT TCTTCAATTAT	36
III (ATGT)	A/A	2	134	YM060507	ATGTTTAAATTTT TCACTTCTCCAA CAACTATAAAGA GATCGACATTTT TCTTCAATTAT	36
III (ATGT)	A/A	2	134	YM060508	ATGTTTAAATTTT TCACTTCTCCAA	36

					CAACTATAAAGA GATCGACATTTC TCTTCAATTAT	
III (ATGT)	A/A	2	134	YM060512	ATGTTTAAATTTT TCACTTCTCCAA CAACTATAAAGA GATCGACATTTC TCTTCAATTAT	36
III (ATGT)	A/A	2	134	YM140518	ATGTTTAAATTTT TCACTTCTCCAA CAACTATAAAGA GATCGACATTTC TCTTCAATTAT	36
III (ATGT)	A/A	2	134	YM140717	ATGTTTAAATTTT TCACTTCTCCAA CAACTATAAAGA GATCGACATTTC TCTTCAATTAT	36
III (ATGT)	A/A	2	153	YM141055	ATGTTTAAATTTT TCACTTCTCCAA CAACTATAAAGA GATCGACATTTC TCTTCAATTAT	36
III (ATGT)	A/A	2	200	YM060379	ATGTTTAAATTTT TCACTTCTCCAA CAACTATAAAGA GATCGACATTTC TCTTCAATTAT	36
III (ATGT)	A/A	2	200	YM060800	ATGTTTAAATTTT TCACTTCTCCAA CAACTATAAAGA GATCGACATTTC TCTTCAATTAT	36
III (ATGT)	A/A	2	202	YM060565	ATGTTTAAATTTT TCACTTCTCCAA CAACTATAAAGA GATCGACATTTC TCTTCAATTAT	36
III (ATGT)	A/A	6	149	YM140470	ATGTTTAAATTTT TCACTTCTCCAA CAACTATAAAGA	36

					GATCGACATTTCTCTCAATTAT	
IV (ATRA)	A/A	3	189	YM060383	ATRATWAWWW WWTTYWYYWCT YCWMCAACWAY WAWRARRWSRW CRWYWYTYTC WWYWAT	37
IV (ATRA)	A/A	3	845	YM141035	ATRATTAAATTT TTCACTTCTCCA ACAACTATAAAG AGATCGACATTT CTCTTCAATTAT	38
IV (ATRA)	A/A	3	845	YM180351	ATRATTAAATTT TTCACTTCTCCA ACAACTATAAAG AGATCGACATTT CTCTTCAATTAT	38
IV (ATRA)	A/G	4	225	YM140093	ATRATTAAATTT TTCACTTCTCCA ACAACTATAAAG AGATCGACATTT CTCTTCAATTAT	38
IV (ATRA)	A/G	4	225	YM140168	ATRATTAAATTT TTCACTTCTCCA ACAACTATAAAG AGATCGACATTT CTCTTCAATTAT	38
IV (ATRA)	A/G	4	225	YM140219	ATRATTAAATTT TTCACTTCTCCA ACAACTATAAAG AGATCGACATTT CTCTTCAATTAT	38
IV (ATRA)	A/G	4	225	YM140372	ATRATTAAATTT TTCACTTCTCCA ACAACTATAAAG AGATCGACATTT CTCTTCAATTAT	38
IV (ATRA)	A/G	4	225	YM140443	ATRATTAAATTT TTCACTTCTCCA ACAACTATAAAG	38

					AGATCGACATTT CTCTTCAATTAT	
IV (ATRA)	A/G	4	225	YM140444	ATRATTAAATTT TTCACCTTCTCCA ACAACTATAAAG AGATCGACATTT CTCTTCAATTAT	38
IV (ATRA)	A/G	4	225	YM140449	ATRATTAAATTT TTCACCTTCTCCA ACAACTATAAAG AGATCGACATTT CTCTTCAATTAT	38
IV (ATRA)	A/G	4	225	YM140453	ATRATTAAATTT TTCACCTTCTCCA ACAACTATAAAG AGATCGACATTT CTCTTCAATTAT	38
IV (ATRA)	A/G	4	225	YM140454	ATRATTAAATTT TTCACCTTCTCCA ACAACTATAAAG AGATCGACATTT CTCTTCAATTAT	38
IV (ATRA)	A/G	4	225	YM140586	ATRATTAAATTT TTCACCTTCTCCA ACAACTATAAAG AGATCGACATTT CTCTTCAATTAT	38
IV (ATRA)	A/G	4	225	YM140779	ATRATTAAATTT TTCACCTTCTCCA ACAACTATAAAG AGATCGACATTT CTCTTCAATTAT	38
IV (ATRA)	A/G	4	225	YM140935	ATRATTAAATTT TTCACCTTCTCCA ACAACTATAAAG AGATCGACATTT CTCTTCAATTAT	38
IV (ATRA)	A/G	4	225	YM140950	ATRATTAAATTT TTCACCTTCTCCA ACAACTATAAAG AGATCGACATTT	38

					CTCTTCAATTAT	
IV (ATRA)	A/G	4	225	YM140954	ATRATTAAATTT TTCACTTCTCCA ACAACTATAAAG AGATCGACATTT CTCTTCAATTAT	38
IV (ATRA)	A/G	4	225	YM140969	ATRATTAAATTT TTCACTTCTCCA ACAACTATAAAG AGATCGACATTT CTCTTCAATTAT	38
IV (ATRA)	A/G	4	225	YM141153	ATRATTAAATTT TTCACTTCTCCA ACAACTATAAAG AGATCGACATTT CTCTTCAATTAT	38
IV (ATRA)	A/G	4	225	YM180235	ATRATTAAATTT TTCACTTCTCCA ACAACTATAAAG AGATCGACATTT CTCTTCAATTAT	38
IV (ATRA)	A/G	4	225	YM180946	ATRATTAAATTT TTCACTTCTCCA ACAACTATAAAG AGATCGACATTT CTCTTCAATTAT	38
IV (ATRA)	A/G	4	225	YM180969	ATRATTAAATTT TTCACTTCTCCA ACAACTATAAAG AGATCGACATTT CTCTTCAATTAT	38
IV (ATRA)	A/G	4	506	YM140066	ATRATTAAATTT TTCACTTCTCCA ACAACTATAAAG AGATCGACATTT CTCTTCAATTAT	38
IV (ATRA)	A/G	4	506	YM140132	ATRATTAAATTT TTCACTTCTCCA ACAACTATAAAG AGATCGACATTT CTCTTCAATTAT	38

IV (ATRA)	A/G	4	506	YM140258	ATRATTAAATTT TTCACTTCTCCA ACAACTATAAAG AGATCGACATTT CTCTTCAATTAT	38
IV (ATRA)	A/G	4	506	YM140441	ATRATTAAATTT TTCACTTCTCCA ACAACTATAAAG AGATCGACATTT CTCTTCAATTAT	38
IV (ATRA)	A/G	4	506	YM180681	ATRATTAAATTT TTCACTTCTCCA ACAACTATAAAG AGATCGACATTT CTCTTCAATTAT	38
IV (ATRA)	A/G	4	506	YM180837	ATRATTAAATTT TTCACTTCTCCA ACAACTATAAAG AGATCGACATTT CTCTTCAATTAT	38
IV (ATRA)	A/G	4	506	YM180950	ATRATTAAATTT TTCACTTCTCCA ACAACTATAAAG AGATCGACATTT CTCTTCAATTAT	38
IV (ATRA)	A/G	4	506	YM180978	ATRATTAAATTT TTCACTTCTCCA ACAACTATAAAG AGATCGACATTT CTCTTCAATTAT	38
IV (ATRA)	A/G	4	506	YM181042	ATRATTAAATTT TTCACTTCTCCA ACAACTATAAAG AGATCGACATTT CTCTTCAATTAT	38
IV (ATRA)	A/G	4	546	YM180538	ATRATTAAATTT TTCACTTCTCCA ACAACTATAAAG AGATCGACATTT CTCTTCAATTAT	38
IV (ATRA)	A/G	4	592	YM180233	ATRATTAAATTT	38

					TTCACCTTCTCCA ACAACTATAAAAG AGATCGACATTT CTCTTCAATTAT	
IV (ATRA)	A/G	4	595	YM140907	ATRATTAAATTT TTCACCTTCTCCA ACAACTATAAAAG AGATCGACATTT CTCTTCAATTAT	38
IV (ATRA)	A/G	4	600	YM140789	ATRATTAAATTT TTCACCTTCTCCA ACAACTATAAAAG AGATCGACATTT CTCTTCAATTAT	38
IV (ATRA)	A/A	4	667	YM141031	ATRATTAAATTT TTCACCTTCTCCA ACAACTATAAAAG AGATCGACATTT CTCTTCAATTAT	38
IV (ATRA)	A/G	4	879	YM180123	ATRATTAAATTT TTCACCTTCTCCA ACAACTATAAAAG AGATCGACATTT CTCTTCAATTAT	38
IV (ATRA)	A/G	4	924	YM140982	ATRATTAAATTT TTCACCTTCTCCA ACAACTATAAAAG AGATCGACATTT CTCTTCAATTAT	38
IV (ATRA)	A/G	4	1096	YM181122	ATRATTAAATTT TTCACCTTCTCCA ACAACTATAAAAG AGATCGACATTT CTCTTCAATTAT	38
IV (ATRA)	A/A	6	921	YM140073	ATRATKWWWW WWTTYWYYWCT YCWMCAACWAY WAWRARRWSRW CRWYWYTYTC WWYWAT	39
IV (ATRA)	A/A	6	1095	YM181080	ATRATWAWWW	37

					WWTTYWYYWCT YCWMCAACWAY WAWRARRWSRW CRWYWYTYTTC WWYWAT	
IV (ATRA)	A/A	7	139	YM060369	ATRATTAAATTT TTCACCTCTCCA ACAACCTATAAAG AGATCGACATTT CTCTTCAATTAT	38
IV (ATRA)	A/A	7	139	YM140156	ATRATTAAATTT TTCACCTCTCCA ACAACCTATAAAG AGATCGACATTT CTCTTCAATTAT	38
IV (ATRA)	A/A	7	139	YM990268	ATRATTAAATTT TTCACCTCTCCA ACAACCTATAAAG AGATCGACATTT CTCTTCAATTAT	38
IV (ATRA)	A/A	7	152	YM990652	ATRATTAAATTT TTCACCTCTCCA ACAACCTATAAAG AGATCGACATTT CTCTTCAATTAT	38
IV (ATRA)	A/A	7	184	YM060210	ATRATTAAATTT TTCACCTCTCCA ACAACCTATAAAG AGATCGACATTT CTCTTCAATTAT	38
IV (ATRA)	A/A	7	833	YM180473	ATRATTAAATTT TTCACCTCTCCA ACAACCTATAAAG AGATCGACATTT CTCTTCAATTAT	38
IV (ATRA)	A/A	7	956	YM180522	ATRATTAAATTT TTCACCTCTCCA ACAACCTATAAAG AGATCGACATTT CTCTTCAATTAT	38
V (ATGW)	A/A	6	374	YM180220	ATGWTAAATTT	40

					TTCAC TTCTCCA ACA ACTATAAAG AGATCGACATTT CTCTTCAATTAT	
V (ATGW)	A/A	ND	914	ATCC750	ATGWTTAAATTT TTCAC TTCTCCA ACA ACTATAAAG AGATCGACATTT CTCTTCAATTAT	40
VI (ATRW)	A/A	3	148	YM990574	ATRWTTAAATTT TTCAC TTCTCCA ACA ACTATAAAG AGATCGACATTT CTCTTCAATTAT	41
VI (ATRW)	A/A	3	967	YM180402	ATRWTTAAATTT TTCAC TTCTCCA ACA ACTATAAAG AGATCGACATTT CTCTTCAATTAT	41
VI (ATRW)	A/A	5	910	YM140912	ATRWTTAAATTT TTCAC TTCTCCA ACA ACTATAAAG AGATCGACATTT CTCTTCAATTAT	41
VI (ATRW)	A/A	6	149	YM060097	ATRWTTAAATTT TTCAC TTCTCCA ACA ACTATAAAG AGATCGACATTT CTCTTCAATTAT	41
VI (ATRW)	A/A	6	149	YM060177	ATRWTTAAATTT TTCAC TTCTCCA ACA ACTATAAAG AGATCGACATTT CTCTTCAATTAT	41
VI (ATRW)	A/A	6	149	YM140092	ATRWTTAAATTT TTCAC TTCTCCA ACA ACTATAAAG AGATCGACATTT CTCTTCAATTAT	41
VI (ATRW)	A/A	6	149	YM180172	ATRWTTAAATTT TTCAC TTCTCCA	41

					ACAACTATAAAG AGATCGACATTT CTCTTCAATTAT	
VI (ATRW)	A/A	6	421	YM180461	ATRWTTAAATTT TTCACCTCTCCA ACAACTATAAAG AGATCGACATTT CTCTTCAATTAT	41
VI (ATRW)	A/A	6	960	YM180459	ATRWTTAAATTT TTCACCTCTCCA ACAACTATAAAG AGATCGACATTT CTCTTCAATTAT	41
VI (ATRW)	A/A	6	1194	YM990579	ATRWTTAAATTT TTCACCTCTCCA ACAACTATAAAG AGATCGACATTT CTCTTCAATTAT	41
VI (ATRW)	A/A	8	150	YM990593	ATRWTTAAATTT TTCACCTCTCCA ACAACTATAAAG AGATCGACATTT CTCTTCAATTAT	41
VI (ATRW)	A/A	9	133	YM990136	ATRWTTAAATTT TTCACCTCTCCA ACAACTATAAAG AGATCGACATTT CTCTTCAATTAT	41
VI (ATRW)	A/A	9	359	YM140463	ATRWTTAAATTT TTCACCTCTCCA ACAACTATAAAG AGATCGACATTT CTCTTCAATTAT	41
VI (ATRW)	A/A	9	923	YM140438	ATRWTTAAATTT TTCACCTCTCCA ACAACTATAAAG AGATCGACATTT CTCTTCAATTAT	41
VI (ATRW)	A/A	9	923	YM140458	ATRWTTAAATTT TTCACCTCTCCA ACAACTATAAAG	41

					AGATCGACATTT CTCTTCAATTAT	
VI (ATRW)	A/A	9	923	YM140474	ATRWTTAAATTT TTCACCTCTCCA ACAACTATAAAG AGATCGACATTT CTCTTCAATTAT	41
VI (ATRW)	A/A	10	601	YM140019	ATRWTTAAATTT TTCACCTCTCCA ACAACTATAAAG AGATCGACATTT CTCTTCAATTAT	41
VII (ATWA)	A/A	3	370	YM140797	ATWATWAWWW WWTTYWYYWCT YCWMCAACWAY WAWRARRWSRW CRWYWYTYTC WWYWAT	42

[0055] Furthermore, the first 4-mer of CTRG_05978 open reading frame may be classified into 7 different sequence types: ATAA, ATGA, ATGT, ATGW, ATRA, ATRW, and ATRW (Table 4). Interestingly, an association between the sequence of CTRG_05978 and DST by MLST is observed (Table 6 and Table 5). The majority of clade 4 strains tested, with the exception of YM180542 (DST506) and YM180921 (DST508), exhibit the genotype IV (ATRA) sequence (36/ 48 vs. 2/72, $p < 0.001$). This significant association between the sequence of CTRG_05978 and the clade 4 genotype was further confirmed by logistic regression analysis ($p < 0.001$; OR = 96.92) (Table 7).

Table 6. The association of genotype and sequence types of CTRG_05978 among 120 *C. tropicalis* isolates.

Clade	I (ATAA)	II (ATGA)	III (ATGT)	IV (ATRA)	V (ATGW)	VI (ATRW)	VII (ATWA)	Total
1	4	0	0	0	0	0	0	4
2	0	0	12	0	0	0	0	12
3	5	4	0	3	0	2	1	15
4	2	0	0	36	0	0	0	38
5	18	0	0	0	0	1	0	19
6	1	0	1	2	1	7	0	12
7	2	1	0	7	0	0	0	10

WO 2025/171193	0	0	0	0	1	PCT/US2025/014895	
Others	0	0	0	0	1	6	7
Total	34	5	13	48	2	17	120

Table 7. Logistic regression for type IV and clade 4 (OR stands for Odds ratios; and CI stands for confidence interval).

Variable	Clade 4 (n = 38)	Non-Clade 4 (n = 82)	P value
Type IV	36 (94.74)	12 (14.63)	< 0.001
Variable	OR	95% CI of OR	
Clade 4	105	22.28 – 494.9	

The SNP of *SNQ2* is associated with the SNP of the sequence type CTRG_05978, and their combination effectively distinguish the Clade 4 genotype

[0056] Since the clade 4 being identified as the primary azole-resistant genotype, an additional investigation to determine the potential involvement of other gene(s) associated with drug resistance is embarked. The present disclosure uncovers a missense mutation I993V due to the SNP (A2977G) in *SNQ2*, which is situated 25 kb upstream from CTRG_05978, which has been linked to drug susceptibility in *N. glabratus* (*C. glabrata*) and *Candida auris*. A possible connection between the mutation in *SNQ2* and the sequence profile of CTRG_05978 is suggested (Table 8 and Table 5).

Table 8. The association of genotype and SNP of *SNQ2* at 2977 among 120 *C. tropicalis* isolates.

Clade	A/A	A/G	G/G	Total
1	1	3	0	4
2	12	0	0	12
3	13	2	0	15
4	2	35	1	38
5	19	0	0	19
6	12	0	0	12
7	10	0	0	10
8	1	2	0	3
Others	7	0	0	7
Total	77	42	1	120

[0057] The *SNQ2* A2977G having high association with the genotype IV of CTRG_05978 (35/43 vs. 13/77, $p < 0.001$, at $\phi = 0.631$) and the clade 4 genotype (36/43 vs. 2/77, $p < 0.001$, at $\phi = 0.820$) is observed. The combination of the *SNQ2* A2977G and the genotype IV of CTRG_05978 appears to be effective in distinguishing the clade 4 *C. tropicalis* strains (35/35 vs. 3/85, $p < 0.001$), and the relationship exhibits a notably high association ($\phi = 0.943$).

[0058] To obtain a more comprehensive understanding of whether the SNPs of *SNQ2* and CTRG_5978 may efficiently identify the clade 4 in *C. tropicalis*, a thorough analysis of whole-genome data from 161 *C. tropicalis* strains available in the NCBI database (<https://www.ncbi.nlm.nih.gov/sra>) is extended. These strains encompass geographic variations, with the majority originating from Asia (76%), followed by Oceania (15%), Europe (5%), and North America (3%).

[0059] Overall, the present disclosure reveals a significant prevalence of the *SNQ2* A2977G and CTRG_05978 genotype IV among the clade 4 strains (Table 9). *SNQ2* and CTRG_05978, indeed, may serve as markers for identifying clade 4 *C. tropicalis* (95/96 vs. 22/185, $p < 0.001$, at $\phi = 0.838$).

Table 9. The association of genotype and SNPs of *SNQ2* and CTRG_05978 among 281 *C. tropicalis* isolates

Clade		CTRG_05978		<i>SNQ2</i> at 2977 position		Type IV of CTRG_05978 & <i>SNQ2</i> G2977	
		Type IV (ATRA)	Others	G	A	Y	N
1	Taiwan	0	4	3	1	0	4
	Worldwide	0	1	0	1	0	1
	Subtotal	0	5	3	2	0	5
2	Taiwan	0	12	0	12	0	12
	Worldwide	0	16	0	16	0	16
	Subtotal	0	28	0	28	0	28
3	Taiwan	3	12	2	13	0	15
	Worldwide	0	13	0	13	0	13

	Subtotal	3	25	2	26	0	28
4	Taiwan	36	2	36	2	35	3
	Worldwide	60	19	69	10	60	19
	Subtotal	96	21	105	12	95	22
5	Taiwan	0	19	0	19	0	19
	Worldwide	0	15	0	15	0	15
	Subtotal	0	34	0	34	0	34
6	Taiwan	2	10	0	12	0	12
	Worldwide	0	16	0	16	0	16
	Subtotal	2	26	0	28	0	28
7	Taiwan	7	3	0	10	0	10
	Worldwide	0	4	1	3	0	4
	Subtotal	7	7	1	13	0	14
8	Taiwan	0	3	2	1	0	3
	Worldwide	0	2	2	0	0	2
	Subtotal	0	5	4	1	0	5
Others	Taiwan	0	7	0	7	0	7
	Worldwide	11	4	3	12	1	14
	Subtotal	11	11	3	19	1	21
Total		119	162	118	163	96	185

[0060] Combination of SNP of CTRG_05978 and *SNQ2* may effectively distinguish the clade 4 *C. tropicalis* strains from others, which the specificity of identification increases from 0.877 to 0.991, allowing for a better discrimination of clade 4 *C. tropicalis* strains (Table 10).

Table 10. The sensitivity and specificity of identifying clade 4 *C. tropicalis* strains with SNP of CTRG_05978 and/or *SNQ2*.

	CTRG_05978 only	<i>SNQ2</i> only	Combine CTRG_05978 and <i>SNQ2</i>
Sensitivity	0.962	0.962	0.962
Specificity	0.877	0.772	0.991

[0061] The sensitivity (true positive rate or recall) is calculated as: sensitivity = True Positives (TP)/ (True Positives (TP) + False Negatives (FN)). Specificity (true negative rate) measures the

proportion of actual negative cases correctly classified, and is determined as: specificity = True Negatives (TN)/ (True Negatives (TN) + False Positives (FP)).

Multiplex PCR of nine genes is established

[0062] To better understand and manage *C. tropicalis* infection, rapid diagnosis alone is insufficient; efficient, cost-effective, and high-throughput methods are also required for comprehensive surveillance. A multiplex PCR approach targeting nine genes, including six genes of MLST, *SNQ2*, *CTRG_05978*, and *ERG11* using 17 specific primers is employed (Table 2). These nine genes are amplified either independently or simultaneously, as demonstrated in FIG. 3A and FIG. 3B, respectively. When the nine genes are amplified independently, these genes were divided into three groups based on DNA fragment lengths (FIG. 3A). Despite the predicted annealing temperatures of all primers exceed 61°C, multiplex PCR reveals only two major groups on the DNA electrophoresis gel instead of the expected three groups. (FIG. 3B, lane 3), with the ~1.5 Kb *ERG11* fragment missing. However, when the annealing temperature is reduced to 60°C, all three groups of fragments are successfully amplified (FIG. 3B, lane 1). Importantly, whether all the nine target genes are amplified successfully and the quantity of each relative to one another requires validation using MinION sequencing.

Capacity of MinION for molecular typing

[0063] MinION nanopore sequencing is used to assess the quality of the multiplex PCR and to evaluate optimal primer concentrations for typing these nine genes (Table 11). The read counts of all the nine genes were analyzed across seven different strains, and the proportion of reads for eight genes (*ERG11* excluded, due to copy number variation (CNV) effects) relative to the total reads for each of the seven strains is compared (FIG. 4). The results show that the optimal primer concentrations for getting the most balanced read counts are listed as follows: 0.2 µM for *ICL1*, *SAPT2*, *ZWF1α*; 0.15 µM for *MDR1*, *SAPT4*, *CTRG_05978*, *CTRG_05993*, and *ERG11*; 0.1 µM for *SNQ2*.

Table 11. Optimization of multiplex PCR using MinION sequencing data (six experimental conditions tested to optimize multiplex PCR reaction, including variations in primer concentrations, numbers of cycles, and annealing temperature).

Condition	C1	C2	C3	C4	C5	C6
Template	Colony	Colony	Colony	Colony	Colony	Colony

Primers	0.15 μM/each	0.15 μM/each	0.2 μM <i>ICL1</i>	0.4 μM <i>ICL1</i>	0.4 μM <i>ICL1</i>	0.4 μM <i>ICL1</i>
			0.15 μM <i>MDR1</i>	0.1 μM <i>MDR1</i>	0.05 μM <i>MDR1</i>	0.05 μM <i>MDR1</i>
			0.2 μM <i>SAPT2</i>	0.3 μM <i>SAPT2</i>	0.3 μM <i>SAPT2</i>	0.3 μM <i>SAPT2</i>
			0.15 μM <i>SAPT4</i>	0.15 μM <i>SAPT4</i>	0.15 μM <i>SAPT4</i>	0.15 μM <i>SAPT4</i>
			0.15 μM <i>CTRG_05978</i>	0.15 μM <i>CTRG_05978</i>	0.15 μM <i>CTRG_05978</i>	0.15 μM <i>CTRG_05978</i>
			0.15 μM <i>CTRG_05993</i>	0.15 μM <i>CTRG_05993</i>	0.15 μM <i>CTRG_05993</i>	0.15 μM <i>CTRG_05993</i>
			0.2 μM <i>ZWF1a</i>	0.3 μM <i>ZWF1a</i>	0.3 μM <i>ZWF1a</i>	0.3 μM <i>ZWF1a</i>
			0.1 μM <i>SNQ2</i>	0.1 μM <i>SNQ2</i>	0.1 μM <i>SNQ2</i>	0.1 μM <i>SNQ2</i>
			0.15 μM <i>ERG11</i>	0.15 μM <i>ERG11</i>	0.4 μM <i>ERG11</i>	0.4 μM <i>ERG11</i>
Annealing (°C)	60	60	60	60	60	59
Cycle	35X	40X	35X	35X	35X	35X

[0064] For MLST typing, 96 individual FASTQ output files are generated, with the total number of reads ranged from 11,211 to 169,720 per sample (combine total 8,172,043 reads and an average of 85,126 reads per sample). And the average reads depth of coverage for each target gene is as follows: *ICL1* (6,061.3), *MDR1* (10,594.5), *SAPT2* (10,762.3), *SAPT4* (15,074.4), *XYR1* (12,843.1), *ZWF1a* (4,605.9), *ERG11* (10,667.7), and *SNQ2* (7,709.5) (FIG. 5A). Target genes are assembled using Medaka, and variant calls were processed with Bcftools (<https://samtools.github.io/bcftools/bcftools.html>) (Danecek and McCarthy, 2017). How different SNP filtering criteria and read depths affect concordance between generated consensus sequences and available sequencing results in the art is evaluated. Initially, the accuracy of MLST gene sequences is evaluated using all reads from a single MinION sequencing flow cell under various SNP filtering thresholds, including:

[0065] Filter 1: MLST genes at 40% (except *XYR1* at 20%); Filter 2: MLST genes at 30% (except *XYR1* at 15%); Filter 3: MLST genes at 25% (except *XYR1* at 10%), Filter 4: all MLST genes at 10%).

[0066] The results demonstrate that when the SNP filtering threshold is set at 10%, all MLST loci (*ICL1*, *MDR1*, *SAPT2*, *SAPT4*, *XYR1*, *ZWF1a*) achieve 100% concordance with the available sequencing results in the art. However, higher SNP filtering thresholds (>10%) reduces similarity for certain loci, particularly *XYR1*, which exhibits only 77.08% concordance at 40% filtering threshold (FIG. 5B).

[006W]O 2025/171193 a 10% SNP filtering threshold is applied while **vPCT/US2025/014895** (5000, 7000, 10,000, and 12,000) to determine the minimum number of reads required to achieve 100% accuracy. The results show that at least 10,000 reads is necessary to ensure complete concordance between the generated consensus sequences and available sequencing results in the art (FIG. 5C).

[0068] In summary, setting the SNP filtering threshold at 10% and utilizing a minimum of 10,000 reads ensures the accurate assembly of consensus sequences, achieving full concordance with available sequencing data in the art.

TaqMan SNP genotyping Assay is useful in identifying CTRG_05978 type IV and *SNQ2* A2977G

[0069] To identify predominant azole-resistant clade 4 *C. tropicalis* strains rapidly, a TaqMan SNP Genotyping Assay is established for detecting the genotype of CTRG_05978 type IV and *SNQ2* A2977G. The specificity of the primers and probes is tested using purified genomic DNAs from 18 different *C. tropicalis* strains including ATCC 750 and 17 strains from clinical settings (Table 12). The primers used are listed in Table 13.

Table 12. TaqMan Genotyping Assay enables rapid detection of the genotype IV of CTRG_05978 and the *SNQ2* A2977G (*: Based on the first 4 mers of CTRG_05978, 7 types in those strains are defined: type I is ATAA; type II is ATGA; type III is ATGT; type IV is ATRA; type V is ATGW, type VI is ATRW, and type VII is ATWA; and **: CTRG_5979 detects GA and AA SNP).

Strain	CTRG_05978* Type	<i>SNQ2</i> A2977G	Clade	TaqMan Genotype	
				CTRG_05978**	<i>SNQ2</i> A2977G
ATCC750	V	A/A	ND (non-determined)	GA	A/A
YM140156	IV	A/A	7	GA/AA	A/A
YM140717	III	A/A	2	-	A/A
YM180064	I	A/A	5	AA	A/A
YM180172	VI	A/A	6	AA	A/A
YM180220	V	A/A	6	GA	A/A

YM180233	IV	A/G	4	GA/AA	A/G
YM180351	IV	A/A	3	GA/AA	A/A
YM180402	VI	A/A	3	AA	A/A
YM180456	II	A/A	3	GA	A/A
YM180459	VI	A/A	6	AA	A/A
YM180460	II	A/A	3	GA	A/A
YM180473	IV	A/A	7	GA/AA	A/A
YM180542	I	G/G	4	AA	G/G
YM180681	IV	A/G	4	GA/AA	A/G
YM180921	I	A/A	4	AA	A/A
YM180969	IV	A/G	4	GA/AA	A/G
YM180472	II	A/A	3	GA	A/A

Table 13. Primers used in the TaqMan SNP genotyping assay (*: position of translation initiation site of ATG as +1).

Name		Sequences	Position*	SEQ ID NO.
CTRG_059 78 Type IV	Forward Primer	CGTCATATACTCGGA AGGACTTTCTCTA	-62 – -33	43
	Reverse Primer	CGATCTCTTTATAGT TGTTGGAGAAGTGA	+13 – +42	44
	Probe 1 (ATGA)	CCCCTTAAATAATGA TTAAAT	-11 – +10	45
	Probe 2 (ATAA)	CCCCTTAAATAATAA TTAAAT	-11 – +10	46
SNQ2 A2977G	Forward Primer	TGAAAAGATCATTGA TGTTTTGGATATGAA AGG	+2931 – +2963	47
	Reverse Primer	CGTTGCTCGACATTT AAACCATTACC	+2992 – +3017	48
	Probe 1 (A)	CCGATGCCATTGTTG	+2969 – +2983	49
	Probe 2 (G)	CCGATGCCGTTGTTG	+2969 – +2983	50

[0070] For the CTRG_05978 genotype, all type IV strains (YM140156, YM180233, YM180351, YM180681, and YM180969) are located in the heterozygous area (green spots), while the others are located in homozygous areas (red and blue spots) (left panel of FIG. 6). In the case of the *SNQ2* A2977G assay, it also specifically detected 7 mutation strains (blue and green spots) (right panel of FIG. 6). Overall, the results of TaqMan SNP Genotyping Assay are consistent with Sanger sequencing or Whole Genome Sequencing, suggesting that the aqMan SNP Genotyping Assay may be a rapid and convenient method for identifying the predominant azole-resistant clade 4 *C. tropicalis* genotype through detecting SNPs of CTRG_05978 and *SNQ2*.

Summary of Examples

[0071] A method for detecting the potential marker for rapidly identifying the azole-resistant *C. tropicalis* with the clade 4 genotype, has been developed and tested in the examples of the present disclosure. The potential markers involve only a few nucleotide variations, which may be detected using designed capture probes targeting the SNPs of Snq2 I993V and the first 4-mer of CTRG_05978. Even though techniques such as real-time PCR applications, ELISA-like assays, or microarray analysis may also be suitable for the purpose, an easily accessible and convenient method for detection as the present disclosure are more practical for clinical settings. The TaqMan SNP genotype assay as a rapid method for identifying these SNPs in CTRG_05978 and the *SNQ2* A2977G is provided in the present disclosure. Some *C. tropicalis* isolates of clade 4 do not belong to the CTRG_05987 type IV cluster and/or do not carry the *SNQ2* A2977G mutation and one non-clade 4 strain belongs to the CTRG_05987 type IV cluster and carries the *SNQ2* A2977G mutation. To increase the accuracy of rapidly identifying azole-resistant clade 4 *C. tropicalis*, determining whether the strains carry ERG11 mutations in addition to the genotypes of CTRG_05987 and *SNQ2* identified in the present disclosure is suggested.

[0072] The SNP in CTRG_05978 and the *SNQ2* A2977G are valuable tools for swiftly identifying the clade 4 genotype (FIG. 7). The present disclosure may greatly contribute to the development of targeted diagnostic resulting in a proper prescription for therapy. In conclusion, the present disclosure may present a promising approach for monitoring and managing the spread of clade 4 *C. tropicalis* infections, thereby offering a potential tool in mitigating significant health risks for vulnerable patient populations.

[0073] The effective management of *Candida tropicalis* infections necessitates rapid and

scalable diagnostic solutions. To address this, the present disclosure provides two innovative approaches. First, a rapid detection marker system leverages the three selected SNP to identify dominant drug-resistant genotypes. This streamlined approach enables rapid screening, making it ideal for point-of-care testing facilities such as hospitals and clinics. Furthermore, this system is designed to optimize existing detection kits, enhancing its scalability and ease of implementation. On the other hand, the examples of the present disclosure may provide a high-throughput approach combining multiplex PCR and MinION sequencing. This synergistic combination allows for large-scale general testing, enabling core laboratories and institutional labs to efficiently handle a high volume of samples. Together, these two approaches offer a promising strategy for monitoring and managing the spread of clade 4 *C. tropicalis*, addressing both the need for rapid, targeted testing and the requirement for large-scale epidemiological surveillance.

[0074] Those skilled in the art will readily observe that numerous modifications and alterations of the device and method may be made while retaining the teachings of the present disclosure. Accordingly, the above disclosure should be construed as limited only by the metes and bounds of the appended claims.

CLAIMS

What is claimed is:

1. A method for identifying a drug-resistant fungal pathogen, comprising:
detecting a single nucleotide polymorphism genotype on *XYR1* gene and *SNQ2* gene of a fungal pathogen.
2. The method of claim 1, wherein the single nucleotide polymorphism genotype on the *XYR1* gene comprises a sequence of ATRA, and R is A or G.
3. The method of claim 2, wherein the sequence of ATRA is a first 4-mer of the *XYR1* gene.
4. The method of claim 1, wherein the single nucleotide polymorphism genotype on the *SNQ2* gene comprises a nucleotide A at position 2977 substituted with a nucleotide G (A2977G).
5. The method of claim 4, wherein the single nucleotide polymorphism genotype on the *XYR1* gene comprises a sequence of ATRA, and R is A or G.
6. The method of claim 5, wherein the sequence of ATRA is a first 4-mer of the *XYR1* gene.
7. The method of claim 1, wherein the fungal pathogen is a non-*albicans* *Candida* species.
8. The method of claim 7, wherein the non-*albicans* *Candida* species is at least one selected from the group consisting of *Candida tropicalis*, *Candida parapsilosis*, *Candida glabrata*, *Candida auris*, *Candida guilliermondii*, and *Candida krusei*.
9. The method of claim 8, wherein the non-*albicans* *Candida* species is *Candida tropicalis*.
10. The method of claim 9, wherein the *Candida tropicalis* comprises a clade 4 genotype.
11. The method of claim 1, wherein the drug-resistant fungal pathogen has a drug resistance against a drug comprising azole.
12. The method of claim 1, wherein the single nucleotide polymorphism genotype on *XYR1* gene and *SNQ2* gene of the fungal pathogen is detected by high-throughput identification, and the high-throughput identification comprises combining a multiplex polymerase chain reaction

approach and a MinION nanopore sequencing approach.

13. The method of claim 1, wherein the single nucleotide polymorphism genotype on *XYRI* gene and *SNQ2* gene of the fungal pathogen is detected by a TaqMan genotyping assay.

14. A kit for identifying a drug-resistant fungal pathogen, comprising:

a lysis buffer for breaking down the fungal pathogen; and

a primer for detecting *XYRI* gene and/or *SNQ2* gene of a fungal pathogen.

15. The kit of claim 14, further comprising:

a DNA polymerase; and

a PCR buffer comprising Mg^{2+} .