Multilocus Sequence Typing for Analyses of Clonality of Candida albicans Strains in Taiwan

Kuo-Wei Chen,1 Yee-Chun Chen,2 Hsiu-Jung Lo,3 Frank C. Odds,4 Tzu-Hui Wang,1 Chi-Yang Lin,1 and Shu-Ying Li 1*

Laboratory for Mycology, Division of Research and Diagnostics, Center for Disease Control, Taipei, Taiwan; Department of Internal Medicine, National Taiwan University Hospital, Taipei, Taiwan; Division of Clinical Research, National Health Research Institutes, Taipei, Taiwan; and Aberdeen Fungal Group, Institute of Medical Sciences, University of Aberdeen, Aberdeen AB25 2ZD, United Kingdom

Received 14 February 2006/Returned for modification 22 March 2006/Accepted 1 April 2006

Multilocus sequence typing (MLST) was used to characterize the genetic profiles of 51 Candida albicans isolates collected from 12 hospitals in Taiwan. Among the 51 isolates, 16 were epidemiologically unrelated, 28 were isolates from 11 critically ill, human immunodeficiency virus (HIV)-negative patients, and 7 were long-term serial isolates from 3 HIV-positive patients. Internal regions of seven housekeeping genes were sequenced. A total of 83 polymorphic nucleotide sites were identified. Ten to 20 different genotypes were observed at the different loci, resulting, when combined, in 45 unique genotype combinations or diploid sequence types (DSTs). Thirty (36.1%) of the 83 individual changes were synonymous and 53 (63.9%) were nonsynonymous. Due to the diploid nature of C. albicans, MLST was more discriminatory than the pulsed-field gel electrophoresis–BssHII-restricted fragment method in discriminating epidemiologically related strains. MLST is able to trace microevolution over time of C. albicans isolates in the same patient. All but one of the DSTs of our Taiwanese strain collections were novel to the internet C. albicans DST database (http://test1.mlst.net/). The DSTs of C. albicans in Taiwan were analyzed together with those of the reference strains and of the strains from the United Kingdom and United States by unweighted-pair group method using average linkages and minimum spanning tree. Our result showed that the DNA type of each isolate was patient specific and associated with ABC type and decade of isolation but not associated with mating type, anatomical source of isolation, hospital origin, or fluconazole resistance patterns.

Invasive Candida infections continue to cause major problems of morbidity and mortality in a diverse range of debilitated and immunocompromised hosts and constitute an important public health problem (12, 13, 20, 41, 42). Candida species were the leading pathogens of nosocomial bloodstream infection in a large teaching hospital in Taiwan, with Candida albicans being the leading cause of Candida infections (5, 6). The infections caused by C. albicans result in increased lengths of hospital stays and medical costs (31, 34). Furthermore, an outbreak of C. albicans fungemia in a neonatal intensive care unit (16) and yeast carriage on hands of healthcare workers in that intensive care unit were identified (15). The increasing frequency of invasive candidal infections in Taiwan and their severe outcome has underscored the importance of the understanding of the molecular epidemiology of fungal infections.

Molecular typing methods used to assess clonality of C. albicans include pulsed-field gel electrophoresis (PFGE)-based typing methods (4), restriction fragment length polymorphism (10, 29), restriction fragment length polymorphism followed by Ca3 probe hybridization (19, 40), and randomly amplified polymorphic DNA analysis (33). More recently developed typing methods include amplified fragment length polymorphism (1), multilocus microsatellite gene analysis (11, 25), and analysis of variable numbers of tandem repeats (24). The choice of appropriate typing methods depends on the purpose of the study. Molecular typing methods should be reproducible, discriminatory, digitally portable, and amenable to standardization and have high throughput (35).

Recently, multilocus sequence typing (MLST) has been developed to meet the increasing need for global surveillance and comparison of genotypes in a central database via the internet. MLST is based on the sequencing of 6 to 7 selected housekeeping genes and identification of polymorphic nucleotide sites. Combination of the alleles at the different loci results in unique diploid sequence types (DSTs) that can be used to discriminate C. albicans strains (2, 3, 37). MLST provides a robust and unambiguous characterization system to evaluate the worldwide diversity and epidemiology of pathogens and is truly portable between laboratories. MLST for typing of pathogenic fungal species has been developed, such as Candida glabrata (9), Candida tropicalis (38), Histoplasma capsulatum, Aspergillus flavus, Coccidioides immitis (21), and the Fusarium oxysporum complex (28). A consensus set of 7 genes encoding housekeeping functions comprising the fragments AAT1α, ACC1, ADP1, MPB, SYA1, VPS13, and ZWF1b is recommended for MLST with C. albicans (http://test1.mlst.net) (3).

In this study, we use MLST to assess the clonality of C. albicans and to ascertain whether different characteristics (e.g., fluconazole resistance, patient or hospital origin, source, or decade of isolation) can be attributed to certain specific MLST DSTs in Taiwan. The purpose of this study was to evaluate the usefulness of MLST relative to PFGE-BssHII fingerprinting tools for discriminating among strains of C. albicans. The data

* Corresponding author. Mailing address: Laboratory for Mycology, Division of Research and Diagnostics, Center for Disease Control, No. 161, Kun-Yang Street, Taipei, Taiwan. Phone: 886 2 26531388. Fax: 886 2 26513572. E-mail: syl@cdc.gov.tw.
obtained in this study will also contribute to the global database and serve as a platform for comparison of domestic as well as international fungal genotypes.

MATERIALS AND METHODS

Fungal isolates. A total of 51 Candida albicans clinical isolates and 2 reference strains were used in this study. Information on each isolate was collected, including MICs of fluconazole, hospital origin, and body site origin. The 51 C. albicans clinical isolates can be divided into 3 groups. Group 1 comprised 16 isolates (P1 to P16) as part of the collections of the Taiwan Surveillance of Antimicrobial Resistance of Yeasts Project, which collected clinical isolates from 22 hospitals located in different geographic areas in Taiwan from 15 April to 15 June 1999 (44). In Group 1, only one isolate was accepted during each episode of infection, therefore they were epidemiologically unrelated. Group 2 consisted of 28 isolates (P17 to P27) from 11 human immunodeficiency virus (HIV)-negative patients during a 6-month surveillance study in adult intensive care units of a large teaching hospital (7). Group 3 contained seven oral isolates collected from three HIV-positive patients (P28 to P30) between 1999 and 2002 (17, 22). The identification of all isolates was done by the germ tube test followed by VITEK Yeast Biochemical Card and API-32C systems (43). The MICs of fluconazole for the C. albicans isolates were determined by the microdilution broth method, according to the guidelines of the Clinical and Laboratory Standards Institute (CLSI, formerly NCCLS) document M27-A as described previously (43).

PFGE of BssHII-restricted fragments. Preparation of plugs and restriction enzyme digestion were conducted as described previously (4). PFGE was performed with a Biometra Rotophor at a pulse time of 14 s to 60 s, an angle of 120°, and 180 V in 0.8% agarose gel with 0.5× TBE (50 mM Tris, 45 mM boric acid, 0.5 mM EDTA) for 36 h. After electrophoresis, the gel was stained in ethidium bromide solution for 15 min and destained in distilled water. DNA fragments were imaged with the IS-1000 digital imaging system (Alpha Innotech Corporation, San Leandro, Calif.). Dendrograms were analyzed with BioNumerics software, version 4.0 (Applied Maths, Kortrijk, Belgium) as described previously. Isolates were assigned different PFGE genotypes when the band similarity value was less than 95% (4).

DNA extraction. The total genomic DNA of the strain was extracted by means of the PUREGENE DNA purification kit (Genta, Minneapolis, Minn.) and was described previously (14). The concentration of DNA extracted from C. albicans isolates was measured with a spectrophotometer (A260). DNA was stored at −80°C until used.

MLST. MLST was based on seven housekeeping genes, including loci AAT1a, ACC1, ADP1, MPIb, SYA1, VPS13, and ZWF1b (3). PCRs were carried out with mixtures containing 2 μl of extracted DNA (10 ng/μl), 4 μl of each primer (5 μM), 10 μl of distilled water, and the TEMPLY PCR kit (LTK BioLaboratories, Taipei, Taiwan). PCRs were performed with an initial 2-min denaturation step at 94°C, followed by 25 cycles of 94°C for 1 min, 52°C for 1 min, and 72°C for 1 min, with a final extension step of 10 min at 72°C. PCRs were performed in a PTC-200 96-well thermal cycler (MJ Research). DNA sequencing was performed by using the unweighted-pair group method using average linkages and minimal spanning tree algorithm of the BioNumerics software applied to modified sequence data. To cope with heterozygous code data, each base in the concatenated sequences from the seven gene fragments sequenced were concatenated into a single sequence. To cope with heterozygous code data, each base in the concatenated sequences of the polymorphic sites was transformed into two bases: the same if the base is homoyzogous code, so, e.g., the sequence ACGT would emerge as AAACCGTGT, and as the component bases for heterozygous codes, so, e.g., AWT would come out as AAATCGTT.

Mating type-like locus status and ABC typing. PCR for determination of mating type-like locus status, heterozygous (α/α) or homoyzogous (α/a or α/a), was conducted as previously described (39). PCRs were carried out with 50-μl PCR volumes containing 100 ng of genomic DNA, 2.5 U of DYNAzyme II DNA polymerase (Finnzymes), 5 μl of 10× reaction buffer (supplied with the enzyme), 200 μM deoxynucleoside triphosphate mix, and 5 μM concentrations of each of the forward and reverse primers. The reactions were performed with an automated thermal cycler (Biometra T3000) with a first cycle of denaturation for 3 min at 94°C, followed by 30 cycles of denaturation at 94°C for 1 min, annealing at 52°C for 1 min, elongation at 72°C for 1 min, and a final extension step of 10 min at 72°C. For ABC typing, PCRs for the 25S rRNA gene transcribed spacer region were done as previously described with modification (27). The volume and composition of the PCR mixture and PCR machine were the same as described above, only 1 μM concentrations of each of the forward and reverse primers were added. DNA samples were denatured by incubation for 3 min at 94°C, followed by 30 cycles of 94°C for 1 min, 65°C for 1 min, and 72°C for 2.5 min, and a final extension step of 10 min at 72°C. PCR with the pair of primers, CA-INT-L/CA-INT-R, resulted in a single product for C. albicans genotypes A (~450 bp) and B (~480 bp), but C. albicans genotype C isolates had two PCR products (~450 and ~840 bp) that were identical in size to the respective products of C. albicans genotypes A and B.

Stability of the MLST method. To evaluate the stability of the MLST method, 6 isolates from patients and the two reference strains were subcultured three times at 7- to 10-day intervals. One subclone was chosen from the last subculture and subjected to MLST.

RESULTS

Sequence variability. A total of 2,883 bp from the 7 MLST loci (AAT1a, ACC1, ADP1, MPIb, SYA1, VPS13, and ZWF1b) were sequenced in each of the 51 isolates and the two reference strains. Eighty-three (2.9%) nucleotide sites were found to be polymorphic; all were found to be heterozygous in at least one isolate. The number of polymorphisms per locus was 6 in the ACC1 locus, followed by 10 in the AAT1a and ADP1 loci, 13 in the SYA1 and ZWF1b loci, 15 in the MPIb locus, and 16 in the VPS13 locus. Polymorphic amino acids per locus were 1 in the AAT1a locus, followed by 2 in the ACC1 locus, 3 in the ZWF1b locus, 4 in the ADP1 and SYA1 loci, and 8 in the MPIb and VPS13 loci. The percentage of polymorphic sites per gene was as follows, in increasing order: 1.5% (ACC1), 2.3% (ADP1), 2.7% (AAT1a and ZWF1b), 3.3% (SYA1), and 4% (MPIb and VPS13). The polymorphisms defined were 10 (ADP1), 11 (ACC1), 14 (MPIb), 17 (ZWF1b), 18 (AAT1a), and 20 (SYA1 and VPS13) genotypes per locus. Among the seven fragments sequenced, ACC1 gave the highest discriminatory ratio, yielding 11 different genotypes from just 6 polymorphic sites, followed by ATT1a (18 genotypes from 10 variable sites), SYA1 (20 genotypes from 13 variable sites), ZWF1b (17 genotypes from 13 variable sites), VPS13 (20 genotypes from 16 variable sites), ADP1 (10 genotypes from 10 variable sites), and MPIb (14 genotypes from 15 variable sites). The ACC1 (2), ADP1 (1), MPIb (2), SYA1 (8), VPS13 (6), and ZWF1b (7) fragments generated more new genotypes, which have now been added to the database (37), than genotypes on the internet. Thirty (36.1%) of the 83 individual changes were synonymous and 53 (63.9%) were nonsynonymous. Of the 30 amino acid changes, 20 were substantive changes.

In Table 1, details are given of the ABC type and DST for each isolate, together with the hospital origin and anatomical source. For the 51 clinical isolates, 45 DSTs were defined by the genotypes identified from the seven loci. With the sole exception of type 468, all DSTs of our Taiwanese strain col-

Downloaded from http://jcm.asm.org/ on October 30, 2017 by guest
lections were novel to the internet DST database. A dendrogram generated to include the data from Table 1 and data available from the internet MLST database showed that only 9 of the isolates in the present study (DSTs 676, 680, 661, 688, 694 to 696, and 704) coclustered with isolates presently assigned to clade 1, 3, or 4 among the four major C. albicans clades (37). No isolates coclustered with strains in clade 2.

In this study, genetic profiles of 51 C. albicans clinical isolates and 2 reference strains were obtained by PFGE of BssHII-restricted fragments and MLST analysis (Table 1). All isolates were typeable by these two methods. For the 51 clinical isolates, MLST generated 45 DSTs and PFGE-BssHII generated 40 DNA patterns in distinguishing isolates. For 16 unrelated isolates, both typing methods generated 16 genotypes.
For 35 epidemiologically related isolates, MLST generated 29 genotypes and PFGE-BssHII generated 25 genotypes.

Relationships between the isolates. The dendrogram in Fig. 1 indicates the similarities of 51 C. albicans isolates determined by MLST with 7 gene fragments. The dendrogram shows all isolates can be divided into 2 clusters (I and II). All isolates from HIV patients, except P28-1, belonged to cluster I. Isolates from the same patient clustered together relatively closely. For long-term isolates from the same HIV patients, isolates from different years belonged to different genotypes. Neither PFGE type nor MLST DST correlated with fluconazole resistance or hospital origin.

Epidemiologically unrelated group 1 isolates (P1 to P16) displayed different PFGE-BssHII genotypes and MLST DSTs. No genotypes were correlated with specimen types, hospital origins, and fluconazole resistance. Group 2 strains were collected from intensive care unit patients within a short period of time. In this group, most strains from the same patient demonstrated the same PFGE-BssHII genotypes. However, few strains with the same PFGE-BssHII genotypes also exhibited the same DSTs; they differed in from 0 to 2 alleles. This may be due to the microevolution of the persistent strain within the same patient. Isolates P28 to P30 were collected from the oral cavity of HIV-positive patients over a period spanning many years. Serial isolates of the same patient exhibited different PFGE-BssHII genotypes and MLST DSTs. For MLST, within

FIG. 1. Dendrogram indicating the similarities of 63 C. albicans isolates determined by MLST with 7 gene fragments.
the same patient, the numbers of different alleles accumulated with time. This was especially apparent in strain P29-1 to -3, P29-3 differs from P29-2 by 1 allele in 2 years and differed from P29-3 by 2 alleles 3 years later. The long time period and perhaps the antifungal regimens have resulted in the accumulation of microevolutionary events. Three isolates, 18-1, 18-2, and 18-3 from the sputum, urine, and rectal swab, respectively, from one intensive care unit patient, P18, possessed different PFGE and MLST types. This is further confirmed by the finding that these three isolates belonged to different 25S rRNA types (A, B, and C types, respectively). These data demonstrate that different clones can be isolated from different body sites, especially nonsterile sites, in one patient.

**ABC typing.** Isolates from the United Kingdom and the United States belonged to cluster I, except UK2 and US2. All United Kingdom and U.S. isolates were type A, except for UK2 and US2, which were type B. Compared with isolates from other countries, such as the United Kingdom and United States, where isolates are predominantly type A, with type C being very rare (37), the Taiwanese isolates had similar proportion of types A (22/51, 43.1%) and B (23/51, 45.1%). There were only 6 type C strains (11.8%). All isolates from HIV patients were type A, except for 28-1 and 28-2, which were type B. All isolates were heterozygous at the mating type locus, except for P29-3, which has an MLT type a/a.

The MST (Fig. 2) showed the phylogenetic relationships between isolates. Based on MST, all isolates can be grouped into 8 groups (groups 1 to 8). The MST grouping correlates with ABC typing. Groups 2, 6, and 7 are type A. Groups 5 and 8 are type B. Groups 3 and 4 are predominantly type B. Group 1 is predominantly a composite of types A and C. Strains from the United Kingdom and United States fall predominantly into cluster 2. The MST grouping correlates very well with the unweighted-pair group method using average linkage clustering, with MST groups 1, 2, 6, 7, and 8 in cluster I and MST groups 3, 4, and 5 in cluster II.

**Stability of the MLST method.** The in vitro stability of the MLST method was demonstrated by the finding that all of the consecutive subcultured clones of each 6 isolates from patients and the two reference strains showed same MLST DST.

**DISCUSSION**

The increased incidence of transmission of pathogens through international travel, global food chain supply, or even deliberate terror attack have highlighted the importance of

---

**FIG. 2.** MST of 51 Taiwanese clinical isolates and strains from the United Kingdom and United States as well as reference strains. Each isolate is represented by a circle. Relationships between the strains were depicted as connections between isolates and the lengths of the branches linking them. Angles of line connection are of no relevance.
global collaborative surveillance networking in control of infectious diseases. Choosing appropriate molecular typing methods is indispensable for fulfilling such needs. MLST allows the exchange of molecular typing information via the internet for global epidemiology. We evaluated MLST methodology to ascertain its potential for outbreak investigation and to know whether different characteristics (patient origin, drug resistance, geographic origin, and source of isolation) can be attributed to certain specific molecular types in Taiwan. The data obtained in this study will contribute to our attempt to establish a central genetic database of fungal pathogens in Taiwan.

PFGE of restricted fragments represents a whole-genome scanning method to reflect mutation events such as polymorphism in the recognition site, translocation (18), reorganization of non-rRNA gene-containing chromosomes, or the non-reciprocal reorganization of rRNA gene cistrons in the rRNA gene-containing chromosomes (30). MLST is based on variability within particular housekeeping genes due to mutation or recombination events; thus, it provides many genetic types per locus and these can be utilized to define the allelic profile or sequence type and determine the relatedness of strains (26, 36). In bacteria, PFGE-restricted methods are more discriminatory than MLST methods. However, our data suggested that the discriminatory ability of MLST for the typing of *C. albicans* offers advantages over PFGE typing, as the differential power has been greatly enhanced by MLST. This may due to the large genome size of yeast pathogens (16 Mb for diploid *C. albicans*), which is on average about 4 to 8 times the average size of bacterial genomes (gram-positive bacteria, 4.5 Mb; gram-negative bacteria, 2.2 Mb). The best resolution for PFGE is limited for a certain size range and number of fragments (e.g., 30 to 35 bands). This may limit the resolutive power of PFGE applied to *Candida* pathogens. The diploid nature of *C. albicans* may further increase the discriminatory power of MLST for this species. In our study, isolates from the same patient with the same PFGE-BssHII genotype could be further differentiated into different MLST DSTs, differing by 0 to 2 allele types. Southern blotting with the Ca3 probe or the C fragment derived from this probe has been demonstrated to be useful in tracing the microevolutionary events in *C. albicans* (23). Our data show that MLST is superior to PFGE for tracing the microevolution of *C. albicans* strains within the same patient. A comparative study also showed that MLST is at least comparable to Ca3 Southern hybridization probe techniques in discriminative power (32). Furthermore, MLST offers distinct advantages in standardizability and portability.

Our result also showed that the DNA type of each isolate was patient specific and associated with ABC type and decade of isolation, but not associated with anatomical source of isolation, hospital origin, or fluconazole resistance patterns, which is in accordance with previous reports (8).

Sequence-based typing data like MLST greatly facilitate standardization and international data exchange. The present study showed that only a few of the Taiwanese isolates (9 in 51 isolates) clustered with isolates presently assigned to the four major *C. albicans* clades (37). Further studies on *C. albicans* strains from the Asian region are needed to understand the global epidemiology of *C. albicans*.

**ACKNOWLEDGMENTS**

This work was supported by grant DOH94-DC-2011 from the Center for Disease Control, Department of Health, and grant 94-0324-19-6090-35 from the National Science Council, Taiwan (S.-Y.L.) and by a grant from the Wellcome Trust (F.C.O.).

We are grateful to Chien-Shun Chiu for insightful comments on the manuscript, and we express our sincere appreciation to all 13 participating hospitals for providing the strains and information related to these strains. They are Chang Gung Memorial Hospital at Linkou, Chang Gung Memorial Hospital at Keelung, Lo-Tung Poli Ai Hospital, Lo-Tung St. Mary Hospital, National Taiwan University Hospital, Tri Service General Hospital, Veterans General Hospital-Taiichung, Zen Ai General Hospital, Kaohsiung Medical College Chung-Ho Memorial Hospital, Kaohsiung Military Hospital, Veterans General Hospital-Kaohsiung, Buddhist Tzu-Chi General Hospital in Hua-Lien, and Mackay Memorial Hospital Taichung Branch.

**REFERENCES**


