Available online at www.sciencedirect.com

ScienceDirect

International Journal of Medical Microbiology 298 (2008) 663–668

IJMM

www.elsevier.de/ijmm

Comparison of *Candida albicans* strain types among isolates from three countries

Shunji Takakura[a], Satoshi Ichiyama[a], Judith M. Bain[b], Amanda D. Davidson[b], Mette D. Jacobsen[b], Duncan J. Shaw[b], Neil A.R. Gow[b], Frank C. Odds[b,*]

[a]*Department of Clinical Laboratory Medicine, Graduate School of Medicine, Kyoto University, 54 Shogoin-Kawaharacho, Sakyo-ku, Kyoto 6068507, Japan*
[b]*Aberdeen Fungal Group, School of Medical Sciences, Institute of Medical Sciences, University of Aberdeen, Aberdeen AB25 2ZD, UK*

Received 1 June 2007; received in revised form 23 September 2007; accepted 8 November 2007

Abstract

Multi-locus sequence typing data for 217 *Candida albicans* isolates cultured since 1990 from blood and vaginal samples in Japan, England/Wales and the USA were analysed for geographically related variations. While no significant differences were found between distributions of diploid sequence types (DSTs) in blood vs. vaginal isolates, there were highly significant differences in the clade distributions of isolates from the three geographical sources. Clade 2 strains were predominantly isolates from England/Wales, while clade 3 strains came mainly from the USA. The isolates from Japan were highly prevalent among strains in clades 5–17, and provided the first example seen so far in *C. albicans* of an amino acid encoded by three separate codons. Within clade 1, the most commonly encountered clade for isolates from all three regions, 15 Japanese isolates and 1 English isolate formed a separate clonal cluster in eBURST analysis. A similarly well demarcated clonal cluster rich in isolates from Japan was also found among the clade 4 strains. The data suggest *C. albicans* undergoes localized evolution, but human movements and person-to-person spread considerably blur the boundaries of such evolution.

Keywords: *Candida albicans*; Japan; MLST; Strain typing; Molecular epidemiology

Introduction

Candida albicans is the most frequently encountered among *Candida* species associated with man as human commensals and opportunistic pathogens (Calderone, 2002; Odds, 1988). Bloodstream infections caused by *Candida* species remain difficult to diagnose and treat and are associated with high mortalities (Edmond et al., 1999; Hajjeh et al., 2004; Ostrosky-Zeichner and Pappas, 2006). A recent national survey of candidemia in Japan showed a distribution of causative species similar to that in other countries, with *C. albicans* implicated more often than other *Candida* species (Takakura et al., 2004).

Multi-locus sequence typing (MLST) has been used to differentiate strain types among isolates of *C. albicans* (Bougnoux et al., 2002, 2003, 2004). MLST combines a high discriminatory index with exceptional portability since MLST data can be accessed and updated on the

---

*Corresponding author. Tel./fax: +44 1224 555 828.
E-mail address: f.odds@abdn.ac.uk (F.C. Odds).

1438-4221/$ - see front matter © 2007 Elsevier GmbH. All rights reserved.
doi:10.1016/j.ijmm.2007.11.002

internet (linked via http://www.mlst.net/ or http://www.pubmlst.org/). A population genetic analysis of *C. albicans* MLST data, based on 416 isolates from separate sources (Tavanti et al., 2005a), showed that the species could be divided into several clades, and that clades differed in the proportions of isolates they included from different geographical sources. An MLST survey of 51 *C. albicans* isolates from 12 hospitals in Taiwan indicated differences between genotypes of these isolates and those from other countries (Chen et al., 2006). However, although a recent study of MLST data for 1391 isolates from diverse global sources confirmed that different proportions of isolates from different regions are assigned to the various *C. albicans* clades (Odds et al., 2007), the possibility that subgroups within large clades may show stronger geographical differences was not investigated.

One problem in analysing large MLST datasets is that the isolates come in dissimilar proportions from both anatomical and geographical sources. We searched the *C. albicans* MLST database (http://test1.mlst.net/) for subsets of isolates from different geographical sources that had the same anatomical origin. Similarly sized subsets could be found for blood and vaginal isolates from Japan, England/Wales and North America. We therefore compared the MLST results for these three sets of isolates for similarities and differences that reflected either their anatomical or their geographical origins.

Materials and methods

C. albicans isolates

A panel of 217 *C. albicans* isolates was analysed. It comprised 140 blood isolates and 77 vaginal isolates, each obtained from a separate patient. The selection of the panel was restricted to isolates made since 1990, and the isolates came from patients in Japan (27 blood isolates, 31 vaginal isolates), England/Wales (82 blood isolates, 25 vaginal isolates) and North America (31 blood isolates, 21 vaginal isolates). The criteria for date and location of isolation allowed us to assemble panels of at least 20 blood and 20 vaginal isolates from each geographical source. Similarly comparable panels of isolates could not be assembled from the MLST database for other geographical sources or other anatomical sources. MLST data for the isolates were determined by ourselves as described elsewhere (Bougnoux et al., 2003; Chen et al., 2006; Odds et al., 2006, 2007) and are stored on the central *C. albicans* MLST database (http://test1.mlst.net/). Full details of the isolate panel are given in the Supplementary online material (Table S1).

Results of MLST are stored in the form of genotype numbers, which define unique sequences for pairs of alleles, and diploid sequence types (DSTs), which define unique combinations of genotypes.

Analysis of MLST data

The MLST results for single nucleotide polymorphisms (SNPs) in the seven loci sequenced were analysed as previously described (Chen et al., 2006; Odds et al., 2007; Tavanti et al., 2005a, b). Briefly, SNP data were concatenated into a single sequence with each SNP represented as two bases; identical when the sequence chromatogram showed a homozygous result and representing the different bases from each allele for a heterozygous result. These derived sequences were used to generate a single-linkage dendrogram for the 217 isolates by the unweighted pair-group method using arithmetic averages (UPGMA) and determined by *p*-distance as implemented by MEGA3 software (Kumar et al., 2004). The significance of dendrogram nodes was determined by bootstrapping with 1000 randomizations. Assignment of isolates to clades was based on the results of a large-scale analysis of 1391 isolates using a cut-off *p*-distance of 0.04 in a dendrogram generated by UPGMA (Odds et al., 2007). The isolates were thus each designated as a member of 17 MLST clades or as singletons, not assignable to a UPGMA-based clade. Clonal clusters of isolates that had identical sequences in five of the seven loci were delineated by eBURST version 3.0 (Feil et al., 2004; Spratt et al., 2004).

Results

MLST data for isolates from Japan, England and the USA

A total of 187 different DSTs was represented among the 217 isolates in the study panel. DST 69, the most commonly encountered *C. albicans* DST among all isolates typed (Odds et al., 2007; Tavanti et al., 2005a), was represented by 5 isolates in the present study set – 2 each from England and the USA and one from Japan – and thus was also the most common DST in the present study. Japanese isolate JIMS187105, (DST 827), was unique in having an adenine residue in MLST position 317 in one allele of *ACC1*. All other isolates in the *C. albicans* MLST database (http://test1.mlst.net/) have cytosine or thymine in this position. This isolate is the sole example among >1400 isolates tested to date of a third nucleotide at any of the 172 polymorphic sites in the 7 gene fragments sequenced for MLST. The variable base was in the third position of a GCX codon for alanine, so the mutation was synonymous.