EUKARYOTIC CELL, Dec. 2011, p. 1740–1741 Vol. 10, No. 12
1535-9778/11/$12.00 doi:10.1128/EC.05255-11

Draft Genome Sequence of *Penicillium marneffei* Strain PM1

Patrick C. Y. Woo,[1,2]† Susanna K. P. Lau,[1,2]† Bin Liu,[3,4,5]† James J. Cai,[6]† Ken T. K. Chong,[1] Herman Tse,[1,2] Richard Y. T. Kao,[1,2] Che-Man Chan,[1] Wang-Ngai Chow,[1] and Kwok-Yung Yuen[1,2]*

Department of Microbiology[1] and State Key Laboratory of Emerging Infectious Diseases,[2] The University of Hong Kong, Hong Kong; Center of Systematic Genomics, Xinjiang Institute of Ecology and Geography, Urumqi, China[3]; Beijing Institute of Genomics[4] and Chinese Academy of Sciences,[5] Beijing, China; and Department of Veterinary Integrative Biosciences, Texas A&M University, College Station, Texas[6]

Received 10 October 2011/Accepted 17 October 2011

***Penicillium marneffei* is the most important thermal dimorphic, pathogenic fungus endemic in China and Southeast Asia and is particularly important in HIV-positive patients. We report the 28,887,485-bp draft genome sequence of *P. marneffei*, which contains its complete mitochondrial genome, sexual cycle genes, a high diversity of Mp1p homologues, and polyketide synthase genes.**

Penicillium marneffei is the most important thermal dimorphic, pathogenic fungus causing respiratory, skin, and systemic mycosis in China and Southeast Asia. The appearance of the HIV pandemic saw the emergence of opportunistic infections caused by *P. marneffei* in HIV-positive patients. About 8% of AIDS patients in Hong Kong are infected with *P. marneffei* (6). In Thailand, penicilliosis is the third most common indicator disease of AIDS, following tuberculosis and cryptococcosis (7). Aiming at improving our understanding of its mechanisms of pathogenesis and thermal dimorphism, the genome sequencing project of *P. marneffei*, in collaboration with the Beijing Genome Institute, was commenced in 2002.

The genome sequence of *P. marneffei* PM1 was determined with the whole-genome shotgun method, similarly to our previous sequencing of the *Laribacter hongkongensis* genome (10). PM1 was isolated from an HIV-negative patient suffering from culture-documented penicilliosis in Hong Kong. A genomic DNA library was made in pUC18 carrying inserts of 3.0 to 5.0 kb prepared by physical shearing using the sonication method. Single ends of 315,580 clones, representing a 6-fold coverage of the *P. marneffei* genome, were sequenced. The Phred/Phrap/Consed software package was used for base calling and sequence assembly (1–5). Protein-coding genes and introns were predicted using FGENESH (Softberry).

The draft genome of PM1 consists of 2,780 sequence contigs with a total length of 28,887,485 bp and a G+C content of 47%. Contigs were ordered into 273 supercontigs with a total length of 28.42 Mb. The number of protein-coding genes predicted was 10,060, with 9,257 (92%) longer than 100 amino acids. The average gene density is one gene per 2.8 kb. The protein-coding sequence occupies 62.1% (51.2% excluding introns) of the sequenced portion of the genome. An estimated total of 28,180 introns, varying from 15 to 1,617 nucleotides long with a mean length of 111 nucleotides, are distributed among 91% of *P. marneffei* genes. The telomere tandem repeat identified is TTAGGG.

With the availability of the draft genome sequence, the complete mitochondrial genome sequence of *P. marneffei*, a circular DNA molecule of 35,438 bp with a G+C content of 25%, was determined; the sequence was more closely related to those of molds than to those of yeasts (12). All meiotic genes (except *HOP1*) and genes encoding putative pheromone processing enzymes, pheromone receptors, and pheromone response pathway proteins in *Aspergillus fumigatus* and *Aspergillus nidulans* and a putative *MAT-1*α box mating-type gene were present, suggesting that *P. marneffei* can potentially be a heterothallic fungus that does not switch mating type (8). Genes encoding Mp1p homologues were identified and used for construction of a highly discriminative multilocus sequence typing scheme for *P. marneffei* (9). Twenty-three putative polyketide synthase (PKS) genes and two putative PKS-nonribosomal peptide synthase hybrid genes were identified, a diversity much higher than those of other pathogenic thermal dimorphic fungi, such as *Histoplasma capsulatum* (one PKS gene) and *Coccidioides immitis* (10 PKS genes) (11). Further in-depth analysis would uncover additional clues that explain the virulent and thermal dimorphic behaviors of this unique *Penicillium* species.

Nucleotide sequence accession number. The draft genome sequence is deposited in GenBank with accession number AGCC00000000.

This work was partly supported by the Research Fund for the Control of Infectious Diseases of the Health, Welfare and Food Bureau of the Hong Kong SAR Government, a Research Grant Council grant, the University Development Fund, a Committee for Conference and Research grant, the Providence Foundation Limited in memory of the late Lui Hac Minh, the HKU Award for CAE Membership, the HKU Medical Faculty Award for CAE Membership, and Hector T. G. Ma.

* Corresponding author. Mailing address: Department of Microbiology, The University of Hong Kong, University Pathology Building, Queen Mary Hospital, Hong Kong. Phone: (852) 22554892. Fax: (852) 28551241. E-mail: kyyuen@hkucc.hku.hk.

† P. C. Y. Woo, S. K. P. Lau, B. Liu, and J. J. Cai made equal contributions to the work.

REFERENCES

1. **Ewing, B., and P. Green.** 1998. Base-calling of automated sequencer traces using phred. II. Error probabilities. Genome Res. **8:**186–194.
2. **Ewing, B., L. Hillier, M. C. Wendl, and P. Green.** 1998. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. Genome Res. **8:**175–185.
3. **Gordon, D.** 2003. Viewing and editing assembled sequences using Consed. Curr. Protoc. Bioinformatics **Chapter 11:**Unit 11.2.

4. **Gordon, D., C. Abajian, and P. Green.** 1998. Consed: a graphical tool for sequence finishing. Genome Res. **8:**195–202.
5. **Gordon, D., C. Desmarais, and P. Green.** 2001. Automated finishing with autofinish. Genome Res. **11:**614–625.
6. **Low, K., and S. S. Lee.** 2002. The pattern of AIDS reporting and the implications on HIV surveillance. Public Health Epidemiol. Bull. **11**(4)**:**41–49.
7. **Supparatpinyo, K., C. Khamwan, V. Baosoung, K. E. Nelson, and T. Sirisanthana.** 1994. Disseminated Penicillium marneffei infection in southeast Asia. Lancet **344:**110–113.
8. **Woo, P. C., et al.** 2006. Genomic and experimental evidence for a potential sexual cycle in the pathogenic thermal dimorphic fungus Penicillium marneffei. FEBS Lett. **580:**3409–3416.
9. **Woo, P. C., et al.** 2007. MP1 homologue-based multilocus sequence system for typing the pathogenic fungus Penicillium marneffei: a novel approach using lineage-specific genes. J. Clin. Microbiol. **45:**3647–3654.
10. **Woo, P. C., et al.** 2009. The complete genome and proteome of Laribacter hongkongensis reveal potential mechanisms for adaptations to different temperatures and habitats. PLoS Genet. **5:**e1000416.
11. **Woo, P. C., et al.** 2010. High diversity of polyketide synthase genes and the melanin biosynthesis gene cluster in Penicillium marneffei. FEBS J. **277:** 3750–3758.
12. **Woo, P. C., et al.** 2003. The mitochondrial genome of the thermal dimorphic fungus Penicillium marneffei is more closely related to those of molds than yeasts. FEBS Lett. **555:**469–477.