genomeA
Journals.ASM.org

Draft Genome Sequence of the Animal and Human Pathogen *Malassezia pachydermatis* Strain CBS 1879

Sergio Triana,[a,b] Andrés González,[b] Robin A. Ohm,[c] Han A. B. Wösten,[c] Hans de Cock,[c] Silvia Restrepo,[a] Adriana Celis[a,c]

Department of Biological Sciences, Universidad de los Andes, Bogotá, Colombia[a]; Grupo de Diseño de Productos y Procesos (GDPP), Chemical Engineering Department, Universidad de los Andes, Bogotá, Colombia[b]; Microbiology, Department of Biology, Utrecht University, Utrecht, The Netherlands[c]

***Malassezia pachydermatis* is a basidiomycetous yeast that causes infections in humans and animals. Here, we report the genome sequence of *Malassezia pachydermatis* strain CBS 1879, which will facilitate the study of mechanisms underlying pathogenicity of the only non-lipid-dependent *Malasezzia* species.**

Received 31 August 2015 **Accepted** 3 September 2015 **Published** 15 October 2015

Citation Triana S, González A, Ohm RA, Wösten HAB, de Cock H, Restrepo S, Celis A. 2015. Draft genome sequence of the animal and human pathogen *Malassezia pachydermatis* strain CBS 1879. Genome Announc 3(5):e01197-15. doi:10.1128/genomeA.01197-15.

Address correspondence to Adriana Celis, acelis@uniandes.edu.co.

M. pachydermatis is the only non-lipid-dependent species of the genus *Malasezzia.* All other 13 species (1) of this genus are obligate lipophilic and require fatty acids for growth. This is due to the lack of a fungal type fatty acid synthase (2). *M. pachydermatis* is able to assimilate fatty acids from the growth medium and can thus be considered a facultative lipophilic species (3). The molecular mechanisms underpinning this behavior are not yet clear.

M. pachydermatis is a member of the microbiota of animals. It is an opportunistic pathogen of dogs causing dermatitis and otitis externa. *M. pachydermatis* has also been implicated in human bloodstream infections (4, 5). Three genomes of the obligate lipophilic species *M. globosa*, *M. restricta*, and *M. sympodiales* have been reported (2, 6). Our goal is to understand the facultative lipophilic nature of *M. pachydermatis*.

M. pachydermatis genomic DNA was extracted as previously described (7). The DNA was sequenced with the Illumina HiSeq 2000 platform at ServiceXS (Leiden, the Netherlands). Two runs with 120-bp paired-end reads on 250-bp fragments were performed following standard Illumina protocols with a 280-fold genome coverage. Reads were quality controlled with FastQC (8) and trimmed using Flexbar (9). *De novo* assembly was performed using CLC Assembly Cell (CLC bio, Denmark). The resulting contigs were scaffolded using SSPACE_Basic (10), and gaps were filled with GapFiller (11). The final assembly consisted of 148 contigs that were linked by pair-end reads into 91 scaffolds, 28 of which were longer than 1 kb. The maximum contig and scaffold length were 1,466,538 and 1,489,072 bp, respectively, and the N_{50} was 0.64 Mbp and 1.3 Mbp, respectively. The genome size was 8.15 Mbp with a G+C content of 55.17%.

The genome was annotated using Maker2 (12) and we made use of a set of 109,264 previously reported *Ustilaginomycotina* proteins, 1,413 ESTs from *Malassezia* spp., and CEGMA (13). The homology-based predictor GeneMark and the ab-initio predictors SNAP (14) and Augustus (15) were used to predict genes. In order to train Augustus and SNAP we ran MAKER two consecutive times; the initial annotation output from MAKER was converted into a model for SNAP and a training set for Augustus, which was used in the subsequent run. Functional annotation of the predicted genes was performed by Blast2GO (16), which involved Blast and InterProScan annotation (17, 18).

A total of 4,202 protein-coding genes were predicted with an average size of 1,581 bp. The coding regions corresponded to 81% of the genome. In addition, CEGMA showed that 97.18% of the eukaryotic core genome was present in the genome (13).

Lipid degrading enzymes play an important role in the host invasion process of *M. pachydermatis* (19, 20). A total of 50 lipid degrading enzymes were identified in the genome, including 35 lipases and 15 esterases. Most interestingly, a typical fungal fatty acid synthase was not detected in the genome. Instead a polyketide synthase, homologous to fatty acid synthases (21), was detected that showed 75% identity with its bidirectional homologue of *M. sympodialis*. How *M. pachydermatis* is able to grow in the absence of fatty acids is a subject for future research.

Nucleotide sequence accession numbers. This whole-genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession number LGAV00000000. The version described in this paper is LGAV01000000.

ACKNOWLEDGMENTS

This work was supported by the Netherlands fellowship program NFP-phd.14/99 and the Colciencias grant No. 120465741393.

REFERENCES

1. **Gaitanis G, Magiatis P, Hantschke M, Bassukas ID, Velegraki A.** 2012. The *Malassezia* genus in skin and systemic diseases. Clin Microbiol Rev **25:**106–141. http://dx.doi.org/10.1128/CMR.00021-11.
2. **Gioti A, Nystedt B, Li W, Xu J, Andersson A, Averette AF, Münch K, Wang X, Kappauf C, Kingsbury JM, Kraak B, Walker LA, Johansson HJ, Holm T, Lehtiö J, Stajich JE, Mieczkowski P, Kahmann R, Kennell JC, Cardenas ME, Lundeberg J, Saunders CW, Boekhout T, Dawson TL, Munro CA, de Groot PWJ, Butler G, Heitman J, Scheynius A.** 2013. Genomic insights into the atopic eczema-associated skin commensal yeast *Malassezia sympodialis*. mBio **4:**e00572-12. http://dx.doi.org/10.1128/mBio.00572-12.
3. **Huang H-P, Little CJL, Fixter LM.** 1993. Effects of fatty acids on the

growth and composition of *Malassezia pachydermatis* and their relevance to canine otitis externa. Res Vet Sci **55:**119–123. http://dx.doi.org/10.1016/0034-5288(93)90045-H.

4. **Choudhury S, Marte RL**. 2014. *Malassezia pachydermatis* fungaemia in an adult on posaconazole prophylaxis for acute myeloid leukaemia. Pathology (Phila) **46:**466–467. http://dx.doi.org/10.1097/PAT.0000000000000139.
5. **Gueho E, Simmons RB, Pruitt WR, Meyer SA, Ahearn DG**. 1987. Association of *Malassezia pachydermatis* with systemic infections of humans. J Clin Microbiol **25:**1789–1790.
6. **Xu J, Saunders CW, Hu P, Grant RA, Boekhout T, Kuramae EE, Kronstad JW, DeAngelis YM, Reeder NL, Johnstone KR, Leland M, Fieno AM, Begley WM, Sun Y, Lacey MP, Chaudhary T, Keough T, Chu L, Sears R, Yuan B, Dawson TL**. 2007. Dandruff-associated *Malassezia* genomes reveal convergent and divergent virulence traits shared with plant and human fungal pathogens. Proc Natl Acad Sci U S A **104:**18730–18735. http://dx.doi.org/10.1073/pnas.0706756104.
7. **Green MR**. 2012. Molecular cloning: a laboratory manual. In Sambrook J (ed). 4th ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
8. **Andrews S**. 2010. FastQC: a quality control tool for high throughput sequence data. Babraham BioInformatics, Cambridgeshire, United Kingdom.
9. **Dodt M, Roehr JT, Ahmed R, Dieterich C**. 2012. FLEXBAR—flexible barcode and adapter processing for next-generation sequencing platforms. Biol **1:**895–905. http://dx.doi.org/10.3390/biology1030895.
10. **Boetzer M, Henkel CV, Jansen HJ, Butler D, Pirovano W**. 2011. Scaffolding pre-assembled contigs using SSPACE. Bioinformatics Oxf Engl **27:**578–579. http://dx.doi.org/10.1093/bioinformatics/btq683.
11. **Boetzer M, Pirovano W**. 2012. Toward almost closed genomes with GapFiller. Genome Biol **13:**R56. http://dx.doi.org/10.1186/gb-2012-13-6-r56.
12. **Cantarel BL, Korf I, Robb SM, Parra G, Ross E, Moore B, Holt C, Sánchez Alvarado A, Yandell M**. 2008. MAKER: an easy-to-use annotation pipeline designed for emerging model organism genomes. Genome Res **18:**188–196. http://dx.doi.org/10.1101/gr.6743907.
13. **Parra G, Bradnam K, Korf I**. 2007. CEGMA: a pipeline to accurately annotate core genes in eukaryotic genomes. Bioinformatics **23:**1061–1067. http://dx.doi.org/10.1093/bioinformatics/btm071.
14. **Korf I**. 2004. Gene finding in novel genomes. BMC Bioinformatics **5:**59. http://dx.doi.org/10.1186/1471-2105-5-59.
15. **Stanke M, Steinkamp R, Waack S, Morgenstern B**. 2004. Augustus: a web server for gene finding in eukaryotes. Nucleic Acids Res **32:**W309–W312. http://dx.doi.org/10.1093/nar/gkh379.
16. **Conesa A, Götz S**, 2008. Blast2GO: a comprehensive suite for functional analysis in plant genomics. Int J Plant Genomics **2008:**1–12. http://dx.doi.org/10.1155/2008/619832.
17. **Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ**. 1990. Basic local alignment search tool. J Mol Biol **215:**403–410. http://dx.doi.org/10.1016/S0022-2836(05)80360-2.
18. **Jones P, Binns D, Chang H-Y, Fraser M, Li W, McAnulla C, McWilliam H, Maslen J, Mitchell A, Nuka G, Pesseat S, Quinn AF, Sangrador-Vegas A, Scheremetjew M, Yong S-Y, Lopez R, Hunter S**. 2014. InterProScan 5: genome-scale protein function classification. Bioinformatics **30:**1236–1240. http://dx.doi.org/10.1093/bioinformatics/btu031.
19. **Boekhout T, Guého-Kellermann E, Mayser P, Velegraki A**. 2010. *Malassezia* and the skin: science and clinical practice. Springer Verlag, Berlin, Germany.
20. **Velegraki A, Cafarchia C, Gaitanis G, Iatta R, Boekhout T**. 2015. *Malassezia* infections in humans and animals: pathophysiology, detection, and treatment. PLoS Pathog **11:**e1004523. http://dx.doi.org/10.1371/journal.ppat.1004523.
21. **Arakawa K**. 2012. Biosynthesis: diversity between PKS and FAS. Nat Chem Biol **8:**604–605. http://dx.doi.org/10.1038/nchembio.1007.