Draft Genome Sequence of the Deep-Sea Basidiomycetous Yeast Cryptococcus sp. Strain Mo29 Reveals Its Biotechnological Potential.

Rédou V, Kumar A, Hainaut M, Bernard Henrissat, Record E, Barbier G, Burgaud G.

AMERICAN SOCIETY FOR MICROBIOLOGY | genomeAnnouncements

Draft Genome Sequence of the Deep-Sea Basidiomycetous Yeast *Cryptococcus* sp. Strain Mo29 Reveals Its Biotechnological Potential

Vanessa Rédou,[a] Abhishek Kumar,[b] Matthieu Hainaut,[c,d] Bernard Henrissat,[c,d,e] Eric Record,[f,g] Georges Barbier,[a] Gaëtan Burgaud[a]

Université de Brest, Laboratoire Universitaire de Biodiversité et Ecologie Microbienne, Technopôle de Brest Iroise, Plouzané, France[a]; Division of Molecular Genetic Epidemiology, German Cancer Research Center (DKFZ), Heidelberg, Germany[b]; Architecture et Fonction des Macromolécules Biologiques, Centre national de la recherche scientifique, Aix-Marseille Université, Marseille, France[c]; Institut national de recherche agronomique, Marseille, France[d]; Department of Biological Sciences, King Abdulaziz University, Jeddah, Saudi Arabia[e]; Institut national de recherche agronomique, Biodiversité et Biotechnologie Fongiques, Aix-Marseille Université, Polytech Marseille, Marseille, France[f]; Aix-Marseille Université, Biodiversité et Biotechnologie Fongiques, Polytech Marseille, Marseille, France[g]

V.R. and A.K. contributed equally to this work.

***Cryptococcus* sp. strain Mo29 was isolated from the Rainbow hydrothermal site on the Mid-Atlantic Ridge. Here, we present the draft genome sequence of this basidiomycetous yeast strain, which has highlighted its biotechnological potential as revealed by the presence of genes involved in the synthesis of secondary metabolites and biotechnologically important enzymes.**

Received 14 April 2016 **Accepted** 23 May 2016 **Published** 7 July 2016

Citation Rédou V, Kumar A, Hainaut M, Henrissat B, Record E, Barbier G, Burgaud G. 2016. Draft genome sequence of the deep-sea basidiomycetous yeast *Cryptococcus* sp. strain Mo29 reveals its biotechnological potential. Genome Announc 4(4):e00461-16 doi:10.1128/genomeA.00461-16.

Address correspondence to Gaëtan Burgaud, gaetan.burgaud@univ-brest.fr.

Yeasts of the *Cryptococcus* genus are widely distributed in the biosphere, ranging from terrestrial to marine habitats (1). Sequence tag-based analysis of deep-sea microeukaryotes occurring in deep-sea habitats revealed *Cryptococcus* as highly abundant taxa in cold methane seep (2), the deep-sea floor (3, 4), and the deep subseafloor (5, 6). A recent RNA-based approach suggested that this phylotype was also metabolically active in deep sedimentary ecosystems (5).

Here, we describe the draft genome sequence of *Cryptococcus* sp. strain Mo29 (= UBOCC 208024), a putative producer of original bioactive compounds. High-quality genomic DNA was extracted using the hexadecyltrimethylammonium bromide (CTAB) and Genomic-tip (Qiagen) methods. Genomic DNA of *Cryptococcus* sp. Mo29 was used to generate shotgun and mate-pair libraries with insert sizes of approximately 350 bp and 8 kb, respectively. A shotgun library was made using the TruSeq DNA PCR-free sample preparation kit and a gel-plus mate-paired-end library was generated with the Nextera mate-pair sample preparation kit. Genome sequencing was performed using Illumina HiSeq 2500 sequencing technologies.

The shotgun library produced 41,481,610 reads. The mate-pair library produced 15,799,276 reads. After quality filtering of reads with more than 90% of bases with base quality greater than or equal to Q20, a total of 28,246,332 shotgun reads (2,85248 Mb) and 7,870,684 mate-pair reads (732 Mb) were retained. The ALLPATHS-LG whole-genome shotgun assembler (7) was used for the creation of the *de novo* genome assembly from these short reads. The assembly contained a total of 174 scaffolds with an average read length of 144,951 bp. The N_{50} was 820 kb, and the maximum contig length was 1,690 kb. The total sequence length of the resulting draft genome was 27,528,793 bp with an overall GC content of 50.09%. A total of 174,473 bp were repeats, which constituted 0.63% of the assembled genome size as predicted by repeat masker tool (http://www.repeatmasker.org/). There are 10,009 genes in this fungal genome, as predicted by Augustus 3.0 (8). Our BLAST2GO-based annotation analyses (9) have revealed 57.5% annotated genes, while 42.5% remained unannotated.

The genome analysis of secondary metabolite biosynthesis gene clusters using antiSMASH 3.0 software (10) highlighted the presence of four biosynthetic gene clusters, which included one type III polyketide synthase and one terpene synthase and two clusters that were unknown (also called as others in AntiSMASH annotation). Analysis of the sequence with the CAZy database (11) identified 860 genes with activity involving carbohydrates, including 396 glycoside hydrolases, 130 glycosyltransferases, 31 polysaccharide lyases, 63 carbohydrate esterases, 99 carbohydrate-binding modules, and 141 auxiliary activities. Particularly noteworthy is the finding that *Crypotococcus* sp. Mo29 is equipped with all necessary enzymes for the deconstruction of terrestrial plant cell walls, including cellulases, xylanases, pectinases and lignases. The finding of numerous genes associated with synthesis of secondary metabolites and CAZyme in the genome sequence of *Cryptococcus* sp. Mo29 suggests that further investigation will result in the discovery of useful gene products that may be exploited for biotechnological application.

Nucleotide sequence accession numbers. The nucleotide sequence of the *Cryptococcus* sp. UBOCC 208024 (Mo29) genome is deposited in DDBJ/EMBL/GenBank under accession numbers FKKD01000001 to FKKD01000687. This paper describes the first version of the genome.

ACKNOWLEDGMENTS

We thank Macrogen Inc. (South Korea) for genome sequencing and assembly.

This study was supported by a granted from the UE project MaCuMBA (Marine microorganisms: Cultivation Methods for Improv-

ing their Biotechnological Applications, FP7, grant agreement 311975, Brussels, Belgium).

FUNDING INFORMATION

This work, including the efforts of Vanessa Rédou, was funded by EC | Seventh Framework Programme (FP7) (311975).

REFERENCES

1. **Richards TA, Jones MD, Leonard G, Bass D**. 2012. Marine fungi: their ecology and molecular diversity. Annu Rev Mar Sci **4:**495–522. http://dx.doi.org/10.1146/annurev-marine-120710-100802.
2. **Takishita K, Tsuchiya M, Reimer JD, Maruyama T**. 2006. Molecular evidence demonstrating the basidiomycetous fungus *Cryptococcus curvatus* is the dominant microbial eukaryote in sediment at the Kuroshima knoll methane seep. Extremophiles **10:**165–169. http://dx.doi.org/10.1007/s00792-005-0495-7.
3. **Nagahama T, Hamamoto M, Nakase T, Takami H, Horikoshi K**. 2001. Distribution and identification of red yeasts in deep-sea environments around the northwest Pacific Ocean. Antonie Van Leeuwenhoek **80:**101–110. http://dx.doi.org/10.1023/A:1012270503751.
4. **Nagahama T, Hamamoto M, Nakase T, Takaki Y, Horikoshi K**. 2003. *Cryptococcus surugaensis* sp. nov., a novel yeast species from sediment collected on the deep-sea floor of Suruga Bay. Int J Syst Evol Microbiol **53:**2095–2098. http://dx.doi.org/10.1099/ijs.0.02712-0.
5. **Edgcomb VP, Beaudoin D, Gast R, Biddle JF, Teske A**. 2011. Marine subsurface eukaryotes: the fungal majority. Environ Microbiol **13:**172–183. http://dx.doi.org/10.1111/j.1462-2920.2010.02318.x.
6. **Rédou V, Ciobanu MC, Pachiadaki MG, Edgcomb V, Alain K, Barbier G, Burgaud G**. 2014. In-depth analyses of deep subsurface sediments using 454-pyrosequencing reveals a reservoir of buried fungal communities at record-breaking depths. FEMS Microbiol Ecol **90:**908–921. http://dx.doi.org/10.1111/1574-6941.12447.
7. **Gnerre S, MacCallum I, Przybylski D, Ribeiro FJ, Burton JN, Walker BJ, Sharpe T, Hall G, Shea TP, Sykes S, Berlin AM, Aird D, Costello M, Daza R, Williams L, Nicol R, Gnirke A, Nusbaum C, Lander ES, Jaffe DB**. 2011. High-quality draft assemblies of mammalian genomes from massively parallel sequence data. Proc Natl Acad Sci USA **108:**1513–1518. http://dx.doi.org/10.1073/pnas.1017351108.
8. **Stanke M, Schöffmann O, Morgenstern B, Waack S**. 2006. Gene prediction in eukaryotes with a generalized hidden Markov model that uses hints from external sources. BMC Bioinformatics **7:**62. http://dx.doi.org/10.1186/1471-2105-7-62.
9. **Götz S, García-Gómez JM, Terol J, Williams TD, Nagaraj SH, Nueda MJ, Robles M, Talón M, Dopazo J, Conesa A**. 2008. High-throughput functional annotation and data mining with the Blast2GO suite. Nucleic Acids Res **36:**3420–3435. http://dx.doi.org/10.1093/nar/gkn176.
10. **Weber T, Blin K, Duddela S, Krug D, Kim HU, Bruccoleri R, Lee SY, Fischbach MA, Müller R, Wohlleben W, Breitling R, Takano E, Medema MH**. 2015. antiSMASH 3.0—a comprehensive resource for the genome mining of biosynthetic gene clusters. Nucleic Acids Res **43:**W237–W243. http://dx.doi.org/10.1093/nar/gkv437.
11. **Lombard V, Golaconda Ramulu HG, Drula E, Coutinho PM, Henrissat B**. 2014. The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Res **42:**D490–D495. http://dx.doi.org/10.1093/nar/gkt1178.