

Short Communication

Draft genome sequencing of the enigmatic yeast *Saitoella complicata*

Hiromi Nishida,^{1,*} Makiko Hamamoto,² and Junta Sugiyama³

¹ Agricultural Bioinformatics Research Unit, Graduate School of Agricultural and Life Sciences, University of Tokyo, Bunkyo-ku, Tokyo 113-8657, Japan

² Department of Life Sciences, School of Agriculture, Meiji University, Kawasaki, Kanagawa 214-8571, Japan

³ Chiba Branch Office & Lab, TechnoSuruga Laboratory Co., Ltd., Yotsukaido, Chiba 284-0001, Japan

(Received March 12, 2011; Accepted May 3, 2011)

Key Words—ascomycetous yeast; genome sequence; *Saitoella complicata*; Taphrinomycotina

Recently fungal genome sequences using massively parallel DNA sequencers have been published (Di-Guistini et al., 2009; Martinez et al., 2009; Nishida et al., 2011; Nowrousian et al., 2010). The aim of this article is to describe the draft genome sequence of the anamorphic yeast *Saitoella complicata* in the subphylum Taphrinomycotina and briefly discuss the phylogenetic and evolutionary implications.

The phylum Ascomycota consists of three subphyla Taphrinomycotina, Saccharomycotina, and Pezizomycotina (Hibbett et al., 2007). The subphylum Taphrinomycotina consists of the following four classes: Neolectomycetes, Pneumocystidomycetes, Schizosaccharomycetes, and Taphrinomycetes (Hibbett et al., 2007). Although the genus *Saitoella* belongs to the Taphrinomycotina, it does not belong to any class (Hibbett et al., 2007). The genus *Saitoella* consists of a single species, *Saitoella complicata*. We believe that this fungus is a key fungus to elucidate the phylogenetic relationships among the Taphrinomycotina.

Saitoella was described as a new yeast genus with the type species *S. complicata* in 1987 (Goto et al.,

1987). This anamorphic, saprobic yeast shares some characteristics with the yeast taxa of both phyla Ascomycota and Basidiomycota. For example, the negative diazonium blue B (DBB) reaction and negative extracellular DNase activity resemble characteristics of ascomycetous yeasts, whereas the positive urease activity, major ubiquinone system Q-10, and enteroblastic budding resemble those of basidiomycetous yeasts (Goto et al., 1987; Sugiyama et al., 1985). The 18S rDNA sequence comparison-based analysis showed that *Saitoella* and *Taphrina* form a monophyletic lineage that diverged prior to the separation of other ascomycetes (Nishida and Sugiyama, 1993).

S. complicata Goto, Sugiy., Hamam. & Komag. NBRC 10748 (= JCM 7358, = IAM 12963; type strain) was used in this study. After the strain had been cultivated in YM broth (yeast extract, 3 g; malt extract, 3 g; peptone, 5 g; dextrose, 10 g; water, 1 L) at 25°C for 3 days, the cells were washed 3 times using TE buffer (pH 8.0). The genome DNA was extracted from 7×10^9 cells and purified using a QIAamp DNA Micro Kit (QIAGEN).

The genome sequencing was performed using the massively parallel DNA sequencer Roche GS FLX Titanium. The result of sequencing is summarized in Table 1. The assemblies of the reads were performed using the assembler software Newbler version 2.3 (454 sequencing, Roche). Assembly of the sequence data resulted in 7,981 contigs (12,981,880 bases). Those

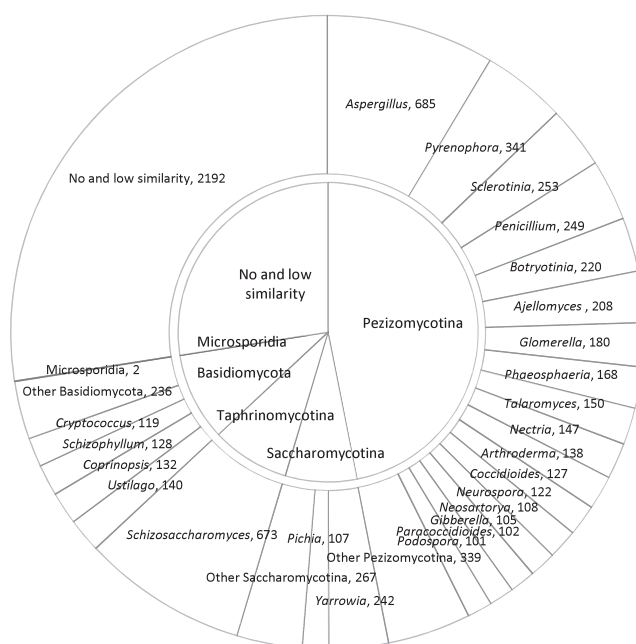
* Address reprint requests to: Dr. Hiromi Nishida, Agricultural Bioinformatics Research Unit, Graduate School of Agriculture and Life Sciences, The University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113-8657, Japan.

Tel: +81-3-5841-1297 Fax: +81-3-5841-1136

E-mail: hnishida@iu.a.u-tokyo.ac.jp

Table 1. Summary of the *Saitoella complicata* genome sequencing.

Total number of reads = 409,958
Total number of bases = 95,080,586
Number of aligned reads = 391,807 (95.57%)
Number of aligned bases = 91,856,020 (96.61%)
Number of assembled reads = 380,278
Number of partial reads = 11,310
Number of singleton reads = 5,568
Number of repeat reads = 323
Number of outlier reads = 5,221
Number of too short reads = 7,258
Total contigs in the assembly
Number of contigs = 7,981
Number of bases = 12,981,880

Fig. 1. Pie chart of the genera with which each contig of *Saitoella complicata* had the highest similarity.

The BLASTX program on the NCBI (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) was used in the DNA sequence similarity search. A total of 5,789 DNA fragments encoding proteins (E -value $\leq 10^{-5}$) were obtained.

7,981 DNA sequences have been deposited in the DDBJ database, under the accession numbers BACD01000001–BACD01007981.

The BLASTX program in the Fungal Genomes Central on the NCBI (<http://www.ncbi.nlm.nih.gov/projects/genome/guide/fungi/>) was used in the DNA sequence similarity search for all proteins of the 67 organisms (Table 2).

Table 2. Fungal species names used in the comparative study.

33 species from Pezizomycotina:	<i>Candida dubliniensis</i>
<i>Ajellomyces capsulatus</i>	<i>Candida glabrata</i>
<i>Ajellomyces dermatitidis</i>	<i>Candida tropicalis</i>
<i>Arthroderma benhamiae</i>	<i>Clavispora lusitanae</i>
<i>Arthroderma gypseum</i>	<i>Debaryomyces hansenii</i>
<i>Arthroderma otae</i>	<i>Kluyveromyces lactis</i>
<i>Aspergillus clavatus</i>	<i>Lachancea thermotolerans</i>
<i>Aspergillus flavus</i>	<i>Lodderomyces elongisporus</i>
<i>Aspergillus nidulans</i>	<i>Meyerozyma guilliermondii</i>
<i>Aspergillus niger</i>	<i>Pichia angusta</i>
<i>Aspergillus terreus</i>	<i>Pichia pastoris</i>
<i>Botryotinia fuckeliana</i>	<i>Saccharomyces cerevisiae</i>
<i>Chaetomium globosum</i>	<i>Scheffersomyces stipitis</i>
<i>Coccidioides immitis</i>	<i>Vanderwaltozyma polyspora</i>
<i>Coccidioides posadasii</i>	<i>Yarrowia lipolytica</i>
<i>Gibberella zeae</i>	<i>Zygosaccharomyces rouxii</i>
<i>Glomerella graminicola</i>	
<i>Magnaporthe oryzae</i>	2 from Taphrinomycotina:
<i>Nectria haematococca</i>	<i>Schizosaccharomyces japonicus</i>
<i>Neosartorya fischeri</i>	<i>Schizosaccharomyces pombe</i>
<i>Neosartorya fumigata</i>	
<i>Neurospora crassa</i>	10 from Basidiomycota:
<i>Paracoccidioides brasiliensis</i>	<i>Coprinopsis cinerea</i>
<i>Penicillium chrysogenum</i>	<i>Cryptococcus gattii</i>
<i>Penicillium marneffei</i>	<i>Cryptococcus neoformans</i>
<i>Phaeosphaeria nodorum</i>	<i>Laccaria bicolor</i>
<i>Podospira anserina</i>	<i>Malassezia globosa</i>
<i>Pyrenophora teres</i>	<i>Moniliophthora perniciosa</i>
<i>Pyrenophora tritici-repentis</i>	<i>Postia placenta</i>
<i>Sclerotinia sclerotiorum</i>	<i>Puccinia graminis</i>
<i>Talaromyces stipitatus</i>	<i>Schizophyllum commune</i>
<i>Trichophyton verrucosum</i>	<i>Ustilago maydis</i>
<i>Uncinocarpus reesii</i>	
<i>Verticillium albo-atrum</i>	4 from Microsporidia:
	<i>Encephalitozoon cuniculi</i>
18 from Saccharomycotina:	<i>Encephalitozoon intestinalis</i>
<i>Ashbya gossypii</i>	<i>Enterocytozoon bieneusi</i>
<i>Candida albicans</i>	<i>Nosema ceranae</i>

The BLASTX search revealed that 5,789 of the 7,981 contigs had protein-coding sequences with E -value $\leq 10^{-5}$. Among the 5,789 sequences, 3,743 (64.7%), 616 (10.6%), 673 (11.6%), and 755 (13.0%) had the highest similarity with proteins of the Pezizomycotina, Saccharomycotina, Taphrinomycotina, and Basidiomycota, respectively (Fig. 1). At the genus level, 685 (11.8%) and 673 (11.6%) had the highest similarity with proteins of *Aspergillus* and *Schizosaccharomyces*, respectively. Those results indicate that the gene content of the *S. complicata* genome has higher affinity to the Pezizomycotina than the Saccharomycotina or Taph-

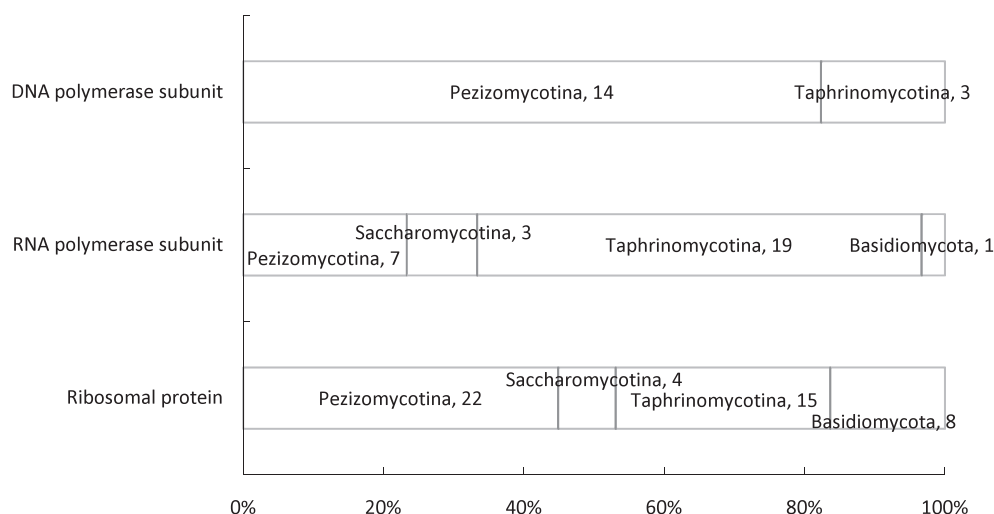


Fig. 2. Band chart of the phyla and subphyla of DNA polymerase subunit, RNA polymerase subunit, and ribosomal protein with which each contig of *Saitoella complicata* had the highest similarity.

A total of 17, 30, and 49 contigs had the highest similarity with DNA polymerase subunit, RNA polymerase subunit, and ribosomal protein, respectively.

rinomycotina. However, in the Taphrinomycotina, only *Schizosaccharomyces* genome information is available. Based on the ribosomal RNA, RNA polymerase, and β -tubulin genes comparison, the genus *Saitoella* is more closely related to the genera *Protomyces* and *Taphrina* lineage than that of the genera *Neolecta*, *Pneumocystis*, or *Schizosaccharomyces* (Sugiyama et al., 2006). If the *Taphrina* and *Protomyces* genome sequences were determined, the result would be different.

It was unexpected that the *S. complicata* genome had significantly higher similarity with the Pezizomycotina genomes than the Saccharomycotina genomes because Taphrinomycotina (including *Saitoella*) had diverged from an ascomycetous common ancestor before Pezizomycotina and Saccharomycotina branched off in the course of evolution (Liu et al., 2009; Nishida and Sugiyama, 1993; Sugiyama et al., 2006). One possibility is that Taphrinomycotina is not monophyletic (Nishida and Sugiyama, 1994).

Although 64.7% of the 5,789 contigs had the highest similarity with proteins of the Pezizomycotina, the percentage is very different between functionally different proteins. For example, among the 17 contigs that had the highest similarity with DNA polymerase subunits, 14 contigs (82.4%) had similarity with those of Pezizomycotina (Fig. 2). On the other hand, among the 30 contigs that had the highest similarity with RNA poly-

merase subunits, 19 contigs (63.3%) had similarity with those of Taphrinomycotina and only 7 contigs (23.3%) had similarity with those of Pezizomycotina (Fig. 2). In addition, among the 49 contigs that had the highest similarity with ribosomal proteins, 22 contigs (44.9%) had similarity with those of Pezizomycotina (Fig. 2). Those results strongly suggest that the *Saitoella* genome consists of evolutionarily heterogeneous genes.

Acknowledgment

This work was supported in part by a research grant from the Institute for Fermentation, Osaka (IFO), Japan to H. N.

References

- DiGuistini, S., Liao, N. Y., Platt, D., Robertson, G., Seidel, M., Chan, S. K., Docking, T. R., Birol, I., Holt, R. A., Hirst, M., Mardis, E., Marra, M. A., Hamelin, R. C., Bohlmann, J., Breuil, C., and Jones, S. J. M. (2009) De novo genome sequence assembly of a filamentous fungus using Sanger, 454 and Illumina sequence data. *Genome Biol.*, **10**, R94.
- Goto, S., Sugiyama, J., Hamamoto, M., and Komagata, K. (1987) *Saitoella*, a new anamorph genus in the Cryptococcaceae to accommodate two Himalayan yeast isolates formerly identified as *Rhodotorula glutinis*. *J. Gen. Appl. Microbiol.*, **33**, 75–85.
- Hibbett, D. S., Binder, M., Bischoff, J. F., Blackwell, M., Cannon, P. F., Eriksson, O. E., Huhndorf, S., James, T., Kirk, P. M.,

- Lücking, R., Thorsten Lumbsch, H., Lutzoni, F., Matheny, P. B., McLaughlin, D. J., Powell, M. J., Redhead, S., Schoch, C. L., Spatafora, J. W., Stalpers, J. A., Vilgalys, R., Aime, M. C., Aptroot, A., Bauer, R., Begerow, D., Benny, G. L., Castlebury, L. A., Crous, P. W., Dai, Y. C., Gams, W., Geiser, D. M., Griffith, G. W., Gueidan, C., Hawksworth, D. L., Hestmark, G., Hosaka, K., Humber, R. A., Hyde, K. D., Ironside, J. E., Köljal, U., Kurtzman, C. P., Larsson, K. H., Lichtwardt, R., Longcore, J., Miadlikowska, J., Miller, A., Moncalvo, J. M., Mozley-Standridge, S., Oberwinkler, F., Parmasto, E., Reeb, V., Rogers, J. D., Roux, C., Ryvarden, L., Sampaio, J. P., Schüßler, A., Sugiyama, J., Thorn, R. G., Tibell, L., Untereiner, W. A., Walker, C., Wang, Z., Weir, A., Weiss, M., White, M. M., Winka, K., Yao, Y. J., and Zhang, N. (2007) A higher-level phylogenetic classification of the Fungi. *Mycol. Res.*, **111**, 509–547.
- Liu, Y., Leigh, J. W., Brinkmann, H., Cushion, M. T., Rodriguez-Ezpeleta, N., Philippe, H., and Lang, B. F. (2009) Phylogenomic analyses support the monophyly of Taphrinomycotina, including *Schizosaccharomyces* fission yeasts. *Mol. Biol. Evol.*, **26**, 27–34.
- Martinez, D., Challacombe, J., Morgenstern, I., Hibbett, D., Schmoll, M., Kubicek, C. P., Ferreira, P., Ruiz-Duenas, F. J., Martinez, A. T., Kersten, P., Hammel, K. E., Vanden Wymelenberg, A., Gaskell, J., Lindquist, E., Sabat, G., Bonduant, S. S., Larrondo, L. F., Canessa, P., Vicuna, R., Yadav, J., Doddapaneni, H., Subramanian, V., Pisabarro, A. G., Lavín, J. L., Oguiza, J. A., Master, E., Henrissat, B., Coutinho, P. M., Harris, P., Magnuson, J. K., Baker, S. E., Bruno, K., Kenealy, W., Hoegger, P. J., Kües, U., Ramaiya, P., Lucas, S., Salamov, A., Shapiro, H., Tu, H., Chee, C. L., Misra, M., Xie, G., Teter, S., Yaver, D., James, T., Mokrejs, M., Pospisek, M., Grigoriev, I. V., Brettin, T., Rokhsar, D., Berka, R., and Cullen, D. (2009) Genome, transcriptome, and secretome analysis of wood decay fungus *Postia placenta* supports unique mechanisms of lignocellulose conversion. *Proc. Natl. Acad. Sci. USA*, **106**, 1954–1959.
- Nishida, H., Nagatsuka, Y., and Sugiyama, J. (2011) Draft genome sequencing of the enigmatic basidiomycete *Mixia osmundae*. *J. Gen. Appl. Microbiol.*, **57**, 63–67.
- Nishida, H. and Sugiyama, J. (1993) Phylogenetic relationships among *Taphrina*, *Saitoella*, and other higher fungi. *Mol. Biol. Evol.*, **10**, 431–436.
- Nishida, H. and Sugiyama, J. (1994) Archiascomycetes: Detection of a major new lineage within the Ascomycota. *Mycoscience*, **35**, 361–366.
- Nowrousian, M., Stajich, J. E., Chu, M., Engh, I., Espagne, E., Halliday, K., Kamerewerd, J., Kempken, F., Knab, B., Kuo, H.-C., Osiewacz, H. D., Pöggeler, S., Read, N. D., Seiler, S., Smith, K. M., Zickler, D., Kück, U., and Freitag, M. (2010) De novo assembly of a 40 Mb eukaryotic genome from short sequence reads: *Sordaria macrospora*, a model organism for fungal morphogenesis. *PLoS Genet.*, **6**, e1000891.
- Sugiyama, J., Fukagawa, M., Chiu, S.-W., and Komagata, K. (1985) Cellular carbohydrate composition, DNA base composition, ubiquinone systems, and Diazonium Blue B color test in the genera *Rhodospiridium*, *Leucosporidium*, *Rhodotorula* and related basidiomycetous yeasts. *J. Gen. Appl. Microbiol.*, **31**, 519–550.
- Sugiyama, J., Hosaka, K., and Suh, S.-O. (2006) Early diverging Ascomycota: Phylogenetic divergence and related evolutionary enigmas. *Mycologia*, **98**, 996–1005.