G3: Genes|Genomes|Genetics Early Online, published on January 8, 2018 as doi:10.1534/g3.117.300316

Genome report

The genomes of four *Meyerozyma caribbica* isolates and novel insights into the *Meyerozyma guilliermondii* species complex

AUTHORS

Leone De Marco[1,2], Sara Epis[3], Aida Capone[1], Elena Martin[3], Jovana Bozic[1], Elena Crotti[4], Irene Ricci[1], Davide Sassera[2]

AFFILIATIONS

[1]School of Bioscience and Veterinary Medicine, University of Camerino, Camerino, Italy

[2]Department of Biology and Biotechnology, University of Pavia, Pavia, Italy

[3]Department of Veterinary Science and Public Health, University of Milan, Milan, Italy

[4]Department of Food, Environmental and Nutritional Sciences (DeFENS), University of Milan, Milan Italy

DATA REFERENCE NUMBERS

Reads are available in the European Nucleotide Archives (https://www.ebi.ac.uk/ena) under the accession numbers ERX2126952-5. The assembled contigs of each genome are available at ENA under accession numbers GCA_900231965, GCA_900231995, GCA_900232055, and GCA_900232065.

RUNNING TITLE

Genomes of four *Meyerozyma caribbica*

KEYWORDS

Meyerozyma caribbica, Genome Report, *Drosophila suzukii*, *Culex quinquefasciatus*, *Anopheles stephensi*, *Aedes aegypti*

CORRESPONDING AUTHOR

Davide Sassera, Dipartimento di Biologia e Biotecnologie, Università degli Studi di Pavia, via Ferrata 9, 27100, Pavia. +39 0382 986028, davide.sassera@unipv.it

ABSTRACT

Yeasts of the *Meyerozyma guilliermondii* species complex are widespread in nature and can be isolated from a variety of sources, from the environment to arthropods to hospital patients. To date, the species complex comprises the thoroughly studied and versatile *M. guilliermondii*, and the hard to distinguish *M. caribbica*, and *Candida carpophila*. Here we report the whole-genome sequencing and *de novo* assembly of four *M. caribbica* isolates, identified with the most recent molecular techniques, derived from four Diptera species. The four novel assemblies present reduced fragmentation and comparable metrics (genome size, gene content) to the available genomes belonging to the species complex. We performed a phylogenomic analysis comprising all known members of the species complex, to investigate evolutionary relationships within this clade. Our results show a compact phylogenetic structure for the complex and indicate the presence of a sizable core set of genes. Furthermore, *M. caribbica*, despite a broad literature on the difficulties of discerning it from *M. guilliermondii*, seems to be more closely related to *C. carpophila*. Finally, we believe that there is evidence for considering these four genomes the first published for the species *M. caribbica*. Raw reads and assembled contigs have been made public to further the studies on these organisms.

INTRODUCTION

Meyerozyma caribbica (anamorph *Candida fermentati*) and *Meyerozyma guilliermondii* (anamorph *Candida guilliermondii*) are two closely related yeast species belonging to the *M. guilliermondii* species complex [Bai et al. 2000; Vaughan-Martini et al. 2005]. *M. guilliermondii* has been object of several studies, with a broad bibliography describing its multiple interesting properties and applications [Papon et al. 2013], and is extensively used in biotechnology in a variety of tasks. *M. guilliermondii* is employed in riboflavin production [Tanner Jr et al. 1945], bioconversion of xylose into xylitol [Zhou et al. 2010], and is a promising source of enzymes [Gong et al. 2007] and biofuel [Wang et al. 2012].

Moreover, it is considered a killer yeast, having a broad range anti-microbial activity against bacteria [Zhao et al. 2010], fungi [Coda et al. 2013], and even protozoa [Dantán-González et al. 2015]. This has led to its use as a biocontrol agent in agriculture and food industry [Wisniewski et al. 1991; Hashem et al. 2011]. Another killer yeast species, *Wickerhamomyces anomalus*, has been suggested as possible candidate for integrated vector control [Ricci et al. 2011; Martin et al, 2016]; interestingly, yeasts of the *M. guilliermondii* clade possess a similar antimicrobial activity and can be found in insect hosts as well, opening the possibility of envisioning similar approaches.

Although *C. guilliermondii,* the anamorph of *M. guilliermondii*, is considered safe and classified as a biosafety level 1 organism, it has been described as an occasional opportunistic pathogen in immunocompromised patients [Pflaller et al. 2006]. It is estimated to be the sixth most frequent nosocomial yeast [Pflaller et al. 2006], causing more than 11% of all episodes of systemic candidiasis [Girmenia et al. 2006].

The *Meyerozyma* species complex belongs to the Saccharomycotina CTG clade, a group of yeasts which has been thoroughly studied in the last 40 years, however its fine phylogenetic structure remains unclear. Taxonomy has been traditionally ruled by phenotypic (e.g. morphologic and metabolic) features, making it a challenging task in yeasts due to the paucity of discriminative morphological characters. More recently, molecular features (e.g. Single Nucleotide Polimorphysms in a single gene/sets of genes) have been widely adopted by taxonomists in general, and this shift clarified a number of phylogenetic relationships in all Kingdoms of life, including Fungi [Kurtzman, Cletus P. 1994]. Yeast taxonomy however remains a complicated matter, with multiple synonyms for each species and an exception to the "one species one name" rule concerning yeasts in different sexual stages (teleomorph/anamorph) [Taylor 2011].

Specifically, *M. guilliemondii*, formerly known as *Pichia guilliermondii*, has recently been placed into its own genus [Kurtzman 2010] and it is thought to form a species complex with close relatives *M. caribbica* and *Candida carpophila* (no known teleomorph) [Vaughan-Martini et al. 2005]. Furthermore, given its emerging pathogen status, it is important to be able to correctly identify yeasts belonging to this species complex, particularly *M. guilliermondii* and *M. caribbica*, the latter of which is less frequent and does not seem to present antibiotic resistance [Pfaller et al. 2006]. Since a morphological identification within the complex is impossible, several molecular protocols using microsatellites, internal transcribed spacer (ITS) polymorphisms, and ITS restriction fragment length polymorphism (RFLP) fingerprinting, have been developed [Romi et al. 2014, Merseguel et al. 2015 , Wrent et al. 2016].

We isolated four yeast strains from the gut of four different diptera species, namely *Drosophila suzukii*, *Culex quinquefasciatus*, *Anopheles stephensi*, and *Aedes aegypti*. First, we carefully identified them at the species level with RFLP fingerprinting of the ribosomal ITS, one of the most effective molecular protocols for the task [Romi et al. 2014]. We then performed whole genome sequencing and we *de novo* assembled the reads. The resulting genomes were employed for phylogenomic and comparative genomic analyses.

The objective of this study was twofold, first to examine the genomes of arthropod-associated yeasts, second to draw a comprehensive phylogenetic picture of the *M. guilliermondii* species complex exploiting whole genome data. To do so, we integrated our dataset with the reference genomes of the three members of the species complex and one close relative, *Clavispora lusitaniae*, completing it with an already published genome of *Phlebotomus*-associated *M. guilliermondii* [Martin et al. submitted].

Here we describe these four novel genomes in the context of the *M. guilliermondii* species complex, make them available to the public, and discuss the phylogenetic implications of our results.

METHODS

Yeast isolation and characterization

All arthropod samples employed in this study derive from insect colonies maintained at the University of Camerino and at the University of Torino. Diptera were maintained in cages at standard conditions of temperature, humidity and photoperiod, as previously reported (Ricci et al, 2011, Vacchini et al, 2017).

The yeast strains derived from the three mosquito species (isolates Clone C2, Clone 8, and Clone 1, respectively from *C. quinquefasciatus*, *An. stephensi,* and *Ae. aegypti*), were isolated following a published protocol (Bozic et al. 2017). Briefly, homogenised mosquitoes guts were pre-inoculated in YPD medium (1% yeast extract, 2% peptone, 2% glucose, 2% agar) and suspended in saline solution before plating on selective media 6.5% Sabouraud (Sabouraud powder prepared by the manufacturer, BD® Sabouraud Dextrose Agar) with Rifampicin 40µg/ml. Isolate AF2.6.P.231 was obtained from an adult individual of *D. suzukii*. After surface sterilization by washing once with ethanol and twice with deionized water, serial dilutions of the insect homogenate were plated on Potato Dextrose Agar (PDA). Once growth was visible, a colony was purified on solid PDA for three times and then conserved at -80°C. For identification, DNA was extracted from the isolate using boiling lysis (Marasco et al., 2012).

Genomic DNA was extracted from individual yeast samples grown in YPD medium, using JetFlex Genomic DNA Purification Kits (Genomed, Löhne, Germany). Quantity and quality of the recovered DNAs were checked by spectrophotometer and stored at − 20 °C. After incubation for 48h at 28°C, yeast colonies were subjected to PCR to amplify a polymorphic fragment of the 18S rRNA gene using oligos yeast-F1 and yeast-R1 (Ricci et al., 2011) or a region comprising the 5.8 rRNA gene and the two sideward regions, ITS1 and ITS2 using primers ITS1F and ITS4 (Manter and Vivanco, 2007). Subsequently the amplification products were sequenced and BLAST was used to characterize the isolated yeasts at the genus level.

Restriction Fragment Length Polymorphism (RFLP) was performed on all isolates to specifically discriminate between *M. guilliermondii* and *M. caribbica*, as previously described (Romi et al., 2014). Briefly PCR amplification using primers ITS1 (5′-TCCGTAGGTGAACCTGCGG-3′) and ITS4 (5′-TC CTCCGCTTATTGATATGC-3′) was carried out to amplify a polymorphic ITS fragment. The PCR product (4 µL) was digested with 5 U of TaqI (Promega, Madison, USA) in a 10 µL reaction volume at 65°C for 2 hours as per manufacturer's instructions. The restriction patterns were analyzed by electrophoresis of the 10 µL reaction volume on 2.0% (w/v) agarose gel.

Sequencing, assembly and annotation

Total DNA was sequenced by an external company (Mr. DNA, Shallowater, USA) in one run of 2x150 paired-end reads on a HiSeq-2500 platform (Illumina). Reads are available in the European Nucleotide Archives (https://www.ebi.ac.uk/ena) under the accession numbers ERX2126952-5.

Quality of the raw reads was assessed for each sample using FastQC [Simon 2010]. De novo assembly was performed using SPAdes version 3.8.2 [Bankevich et al. 2012] employing different k-mer lengths (21, 33, 55, 77, 99, 111), setting the --cov-cutoff parameter to auto, and using the --careful option. The assembled contigs of each genome are available at ENA under accession numbers GCA_900231965, GCA_900231995, GCA_900232055, GCA_900232065 respectively for *C. quinquefasciatus*, *An. stephensi, D. suzukii*, and *Ae. aegypti*. All genomes analyzed in this study, newly sequenced and reference ones, were annotated with the following procedure. First, gene calling was performed using GeneMark-ES Suite version 4.32 [Ter-Hovhannisyan et al. 2008] with the parameter --min_contig set to 10000 and using the --fungus option. Then, Clusters of Orthologous Groups (COGs) [Tatusov et al. 2000] were assigned to the obtained translated genes by the COGnitor software [Tatusov et al. 2000] .

Genomic analysis

The genomes of members of the *M. guilliermondii* species complex and of *C. lusitaniae* were retrieved from NCBI with the following GeneBank assembly accessions: *M. guilliermondii* (GCA_000149425.1 and GCA_900174495.1), *M. caribbica*

(GCA_000755205.1), *C. carpophila* (GCA_001599235.1), *C. lusitaniae* (GCA_000003835.1).

Orthogroups were inferred with Orthofinder 1.1.4 [Emms et al. 2015] from the predicted sets of proteins. Single Copy Orthogroups (SCO) were then selected, defined as orthogroups with exactly one protein in all samples, with a custom Python script. Each SCO was aligned using MUSCLE 3.8.31[Edgar et al. 2004] and consecutively tested each Multiple Sequence Alignment (MSA) for recombination using the software PhiPack [Bruen et al. 2005]. A SCO group was considered not having signs of recombination if it passed all three tests run by PhiPack. Then, non-recombinant SCO MSAs were polished with Gblocks 0.91b [Castresana 2000] and concatenated with a custom Python script. Finally, the concatenated alignment was used as input for RAxML version8.2.8 [Stamatakis 2014] under the PROTCAT approximation, using the LG substitution matrix and 100 bootstrap replicates. A comparative genomic approach was designed, integrating the functional and phylogenetic data obtained. Phylogenetic clades were analyzed for COG content using in-house Python scripts.

RESULTS AND DISCUSSION

Yeast isolation and characterization

Yeasts were isolated from four insect species and characterized at the species complex level using PCR and Sanger sequencing. 18S fragments were sequenced for the three yeasts isolated from mosquitoes, were identical and presented 99% sequence identity with an 18S gene belonging to *M. guilliermondii*, GenBank accession KX258468.1. For what concerns the yeast isolated from *D. suzukii*, a fragment of the ITS was sequenced, showing a 100% sequence identity with sequence KU216711.1 of *M. guillermondi*. These results allowed to identify the isolates at the genus level, however, in order to discriminate between *M. guilliermondii* and *M. caribbica*, a specific RFLP protocol [Romi et al. 2014] was performed, which clearly showed that all novel isolates exhibit the restriction fragment pattern typical of *M. caribbica*. The same RFLP protocol was performed on the yeast derived from *P. perniciosus* (characterized in Martin et al submitted) confirming its identification as *M. guilliermondii*.

Assembly and annotation

Raw reads were high quality for each of the samples analyzed. The four novel draft genomes present a reduced amount of fragmentation (from 43 to 144 contigs longer than 1000bp) and are sized coherently compared to the reference genomes of the *Meyerozyma* genus (table 1). It has to be noted that, while the two reference genomes were sequenced with a combination of long and short reads, we only employed a paired-end short reads library thus obtaining a larger amount of contigs. Nevertheless, GeneMark-ES, a self-training gene calling algorithm, predicted a comparable amount of genes (from 5111 to 5774) for all genomes analyzed (table 1), including the reference genomes. Additionally, COGnitor assigned a similar number of COGs (from 3288 to 3598) to a similar number of unique genes (from 3029 to 3312) for all sets of predicted proteins (table 1).

Species	Isolate	Genome size – bp	Contigs #	Genes #	Genes with COG
M. guilliermondii	ATCC6260	10609954	9	5401	3312
M. caribbica	MG20W	10609282	9	5390	3305
C. carpophila	JCM9396	10242926	10	5296	3219
C.lusitaniae	ATCC42720	12114892	9	5111	3029
M. guilliermondii	*P. perniciosus*	10642597	31	5487	3362
M. caribbica	*D. melanogaster* *	10387257	43	5367	3242
M. caribbica	*Cx. quinquefasciatus* *	10553449	51	5453	3301
M. caribbica	*An. stephensi* *	11040470	144	5774	3481
M. caribbica	*Ae. aegypti* *	10347015	48	5359	3237

Table 1. Assembly and annotation statistics of the four novel genomes (marked with an asterisk) and of the published genomes used for comparative analysis; contigs shorter than 1kbp were discarded.

Genomic analysis

All analyzed genomes show high similarity for what concerns the inferred orthogroups. Almost the totality of proteins (98.7%) was assigned to an orthogroup; moreover, out of a total 5371 orthogroups, 5142 had at least one protein in each genome of the *M. guilliermondii* species complex (all genomes analyzed except the outgroup *C. lusitaniae*), while 4050 had all species represented, indicating the presence of a strong core set of genes.

We retrieved 3408 SCOs which were then aligned with MUSCLE and tested for recombination with PhiPack. We retained 2147 non recombining SCOs and processed their MSAs with Gblocks. The polished MSAs were concatenated obtaining a 870,212 bp long alignment. We used this final MSA as input for RAxML, obtaining a phylogenetic tree with 100% bootstrap support for all branches (figure 1). Four clear clades can be seen in the tree: 1) *C. lusitaniae* as the outermost single species clade; 2) the reference genomes of *M. guilliermondii* and *M. caribbica* clustering together with the *M. guilliermondii* genome isolated from *P. perniciosus* in a clade with reduced branch lengths; 3) *C. carpophila* as a single species clade; 4) all novel *M. caribbica* genomes clustering together in a clade with reduced branch lengths, closer to the *C. carpophila* clade than to the reference *M. guilliermondii*/*M. caribbica* clade.

This is the first phylogenetic study attempting to describe the *M. guilliermondii* species complex employing Whole Genome Sequencing data. The genome of *M. guilliermondii* has already been included in a phylogenomic analysis in a study which, among other things, clarifies its position inside the CTG clade [Butler et al. 2009]. In this work we expand the genomic dataset of the *M. guilliermondii* species complex, including genomes of *M. caribbica* and *C. carpophila*. Literature on these two species is scarce compared to *M. guilliermondii* and past phylogenetic inferences were based on single or few genes and did not include all members of the species complex, leaving doubts about fine characterization [Kurtzman et al. 2010, 2013]. Considering these facts, it would have been presumptuous to make assumptions on the phylogenetic structure of the complex.

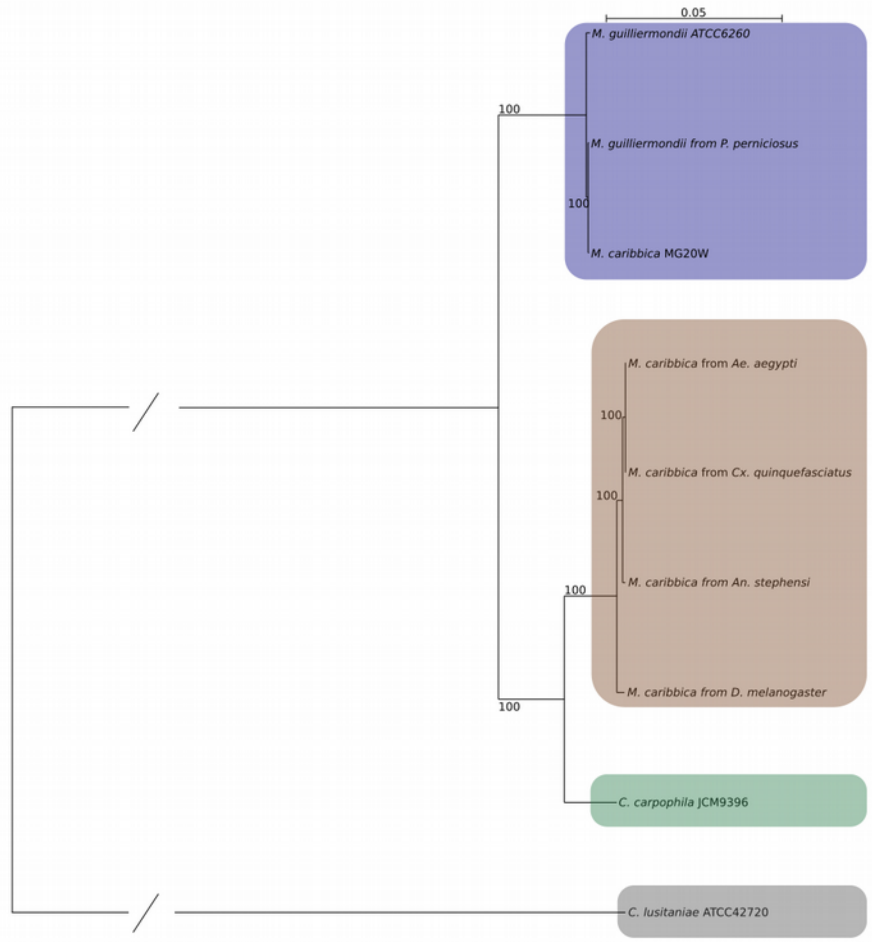

Figure 1. Maximum likelihood tree. Resulting clades are highlighted: *M. guilliermondii* (blue), *M. caribbica* (orange), *C. carpophila* (green), *C. lusitaniae* (grey).

Our phylogenetic tree has one glaring issue: the position of the reference genome of *M. caribbica* (figure 1). At first sight, its closeness to the genome of *M. guilliermondii* would not be suspicious since they share the genus name and they are notoriously difficult to differentiate, even with traditional molecular markers. The aforementioned difficulty to discern the two species, in our opinion, is the key to solve this phylogenetic enigma, considering two facts: 1) the reference genome of *M. caribbica* is sister group to the genome of *M. guilliermondii* isolated from *P. perniciosus*; 2) our genomes of *M. caribbica* isolated from arthropods form a distinct clade, closer to *C. carpophila*. If we accept the identification as *M. caribbica* for the isolate previously sequenced [Kim et al. 2015], it would seem that two distinct *M. caribbica* exist: one is phylogenetically indistinguishable from *M. guilliermondii* while the other forms a distinct clade, sister group to *C. carpophila*. We think that the most likely and parsimonious explanation is however that the reference genome of *M. caribbica* has been misidentified and actually belongs to *M. guilliermondii*.

As a follow-up to the phylogeny, the COG content of the four described clades was compared (figure 2). In total, 1093 COGs are shared between the four clades, constituting the vast majority of COGs assigned to each clade. This result confirms the evolutionary similarity of all analyzed genomes and a probable corresponding functional closeness. The outgroup, *C. lusitaniae*, holds the most unique COGs at 145 followed by the *M. caribbica* clade (128); the *M. guilliermondii* clade has a comparable number of unique COGs (107) while *C. carpophila* presents the least with 57. We think that this results confirm the phylogenetic distance between the *M. caribbica* clade and *M. guilliermondii* clade observed in the phylogenetic tree.

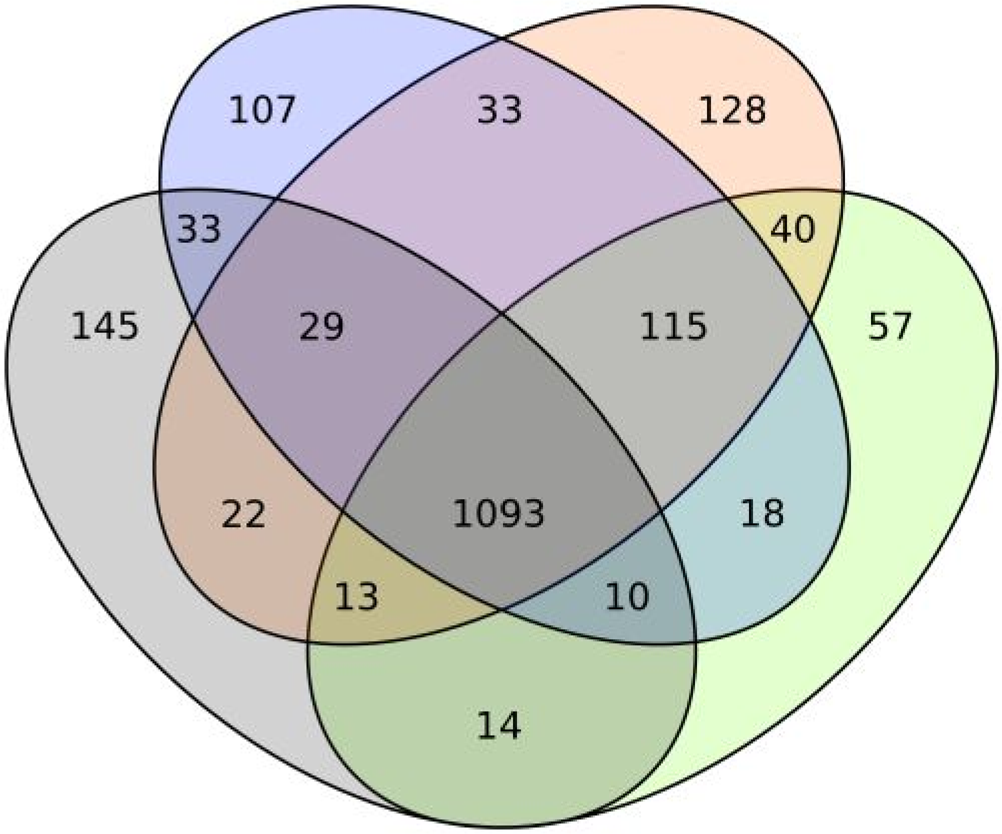

Figure 2. Venn diagram representing COG content in the clades resulting from the phylogenomic analysis: *M. guilliermondii* (blue), *M. caribbica* (orange), *C. carpophila* (green), *C. lusitaniae* (grey).

CONCLUSIONS

Yeasts of the *M. guilliermondii* species complex are relevant for multiple reasons. They are widely used in industry due to their useful properties, have recently emerged as nosocomial pathogens and due to their proprieties they can be envisioned as potential tools for the control of arthropod-borne diseases. Our study provides useful genomic resources and a detailed phylogenomic analysis of this species complex, providing novel insights into the evolutionary history of these yeasts. Additionally, we provide to the public four novel genomes belonging to *M. caribbica*, arguably the first of this species.

ACKNOWLEDGEMENTS

The authors are grateful to Alberto Alma (Università di Torino) for providing *D. suzukii* specimens. The research was supported by the European Union Seventh Framework Programme ERC 2011 [FP7/2007-2013_FP7/2007-2011 under grant agreement no. 281222 (to IR)] and by the Italian Ministry of Education, University and Research FIR 2013 [RBFR136GFF (to SE)].

LITERATURE CITED

Andrews, S. (2010). FastQC: a quality control tool for high throughput sequence data.

Bai, F. Y., Liang, H. Y., & Jia, J. H. (2000). Taxonomic relationships among the taxa in the *Candida guilliermondii* complex, as revealed by comparative electrophoretic karyotyping. International journal of systematic and evolutionary Microbiology, 50(1), 417-422.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., et al. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. Journal of computational biology, 19(5), 455-477.

Bozic, J., Capone, A., Pediconi, D., Mensah, P., Cappelli, A., et al. (2017). Mosquitoes can harbour yeasts of clinical significance and contribute to their environmental dissemination. Environmental microbiology reports.

Bruen, T., & Bruen, T. (2005). PhiPack: PHI test and other tests of recombination. McGill University, Montreal, Quebec.

Butler, G., Rasmussen, M. D., Lin, M. F., Santos, M. A., Sakthikumar, S., et al. (2009). Evolution of pathogenicity and sexual reproduction in eight *Candida* genomes. Nature, 459(7247), 657-662.

Castresana, J. (2000). Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Molecular biology and evolution, 17(4), 540-552.

Coda, R., Rizzello, C. G., Di Cagno, R., Trani, A., Cardinali, G., et al. (2013). Antifungal activity of *Meyerozyma guilliermondii*: identification of active compounds synthesized during dough fermentation and their effect on long-term storage of wheat bread. Food microbiology, 33(2), 243-251.

Dantán-González, E., Quiroz-Castañeda, R. E., Cobaxin-Cárdenas, M., Valle-Hernández, J., Gama-Martínez, Y., et al. (2015). Impact of *Meyerozyma guilliermondii* isolated from chickens against *Eimeria sp*. protozoan, an in vitro analysis. BMC veterinary research, 11(1), 278.

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic acids research, 32(5), 1792-1797.

Emms, D. M., & Kelly, S. (2015). OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. Genome biology, 16(1), 157.

Ferreira, N., Belloch, C., Querol, A., Manzanares, P., Vallez, S., et al. (2010). Yeast microflora isolated from brazilian cassava roots: taxonomical classification based on molecular identification. Current microbiology, 60(4), 287-293.

Girmenia, C., Pizzarelli, G., Cristini, F., Barchiesi, F., Spreghini, E., et al. (2006). *Candida guilliermondii* fungemia in patients with hematologic malignancies. Journal of clinical microbiology, 44(7), 2458-2464.

Gong, F., Sheng, J., Chi, Z., & Li, J. (2007). Inulinase production by a marine yeast *Pichia guilliermondii* and inulin hydrolysis by the crude inulinase. Journal of industrial microbiology & biotechnology, 34(3), 179-185.

Hashem, M., & Abo-Elyousr, K. A. (2011). Management of the root-knot nematode *Meloidogyne incognita* on tomato with combinations of different biocontrol organisms. Crop Protection, 30(3), 285-292.

Kim, J. S., Baek, J. H., Park, N. H., & Kim, C. (2015). Complete genome sequence of halophilic yeast *Meyerozyma caribbica* MG20W isolated from rhizosphere soil. Genome announcements, 3(2), e00127-15.

Kurtzman, C. P. (1994). Molecular taxonomy of the yeasts. Yeast, 10(13), 1727-1740.

Kurtzman, C. P., & Suzuki, M. (2010). Phylogenetic analysis of ascomycete yeasts that form coenzyme Q-9 and the proposal of the new genera *Babjeviella*, *Meyerozyma*, *Millerozyma*, *Priceomyces*, and *Scheffersomyces*. Mycoscience, 51(1), 2-14.

Kurtzman, C. P., & Robnett, C. J. (2013). Relationships among genera of the *Saccharomycotina* (*Ascomycota*) from multigene phylogenetic analysis of type species. FEMS yeast research, 13(1), 23-33.

Manter, D. K., & Vivanco, J. M. (2007). Use of the ITS primers, ITS1F and ITS4, to characterize fungal abundance and diversity in mixed-template samples by qPCR and length heterogeneity analysis. Journal of Microbiological Methods, 71(1), 7-14.

Marasco, R., Rolli, E., Ettoumi, B., Vigani, G., Mapelli, F., et al. (2012). A drought resistance-promoting microbiome is selected by root system under desert farming. PLoS One, 7(10), e48479.

Martin, E., Bongiorno, G., Giovati, L., Montagna, M., Crotti, E., et al. (2016). Isolation of a *Wickerhamomyces anomalus* yeast strain from the sandfly *Phlebotomus perniciosus*, displaying the killer phenotype. Medical and veterinary entomology, 30(1), 101-106.

Martin, E., Varotto Boccazzi, I., De Marco, L., Bongiorno, G., Montagna, M., et al. The mycobiota of the sand fly Phlebotomus perniciosus: involvement of yeast 2 symbionts in uric acid metabolism. Submitted.

Merseguel, K. B., Nishikaku, A. S., Rodrigues, A. M., Padovan, A. C., e Ferreira, R. C., et al. (2015). Genetic diversity of medically important and emerging *Candida* species causing invasive infection. BMC infectious diseases, 15(1), 57.

Papon, N., Savini, V., Lanoue, A., Simkin, A. J., Crèche, J., et al. (2013). *Candida guilliermondii*: biotechnological applications, perspectives for biological control, emerging clinical importance and recent advances in genetics. Current genetics, 59(3), 73-90.

Pfaller, M. A., Diekema, D. J., Mendez, M., Kibbler, C., Erzsebet, P., et al. (2006). *Candida guilliermondii*, an opportunistic fungal pathogen with decreased susceptibility to fluconazole: geographic and temporal trends from the ARTEMIS DISK antifungal surveillance program. Journal of clinical microbiology, 44(10), 3551-3556.

Ricci, I., Mosca, M., Valzano, M., Damiani, C., Scuppa, P., et al. (2011). Different mosquito species host *Wickerhamomyces anomalus* (*Pichia anomala*): perspectives on vector-borne diseases symbiotic control. Antonie Van Leeuwenhoek, 99(1), 43-50.

Romi, W., Keisam, S., Ahmed, G., & Jeyaram, K. (2014). Reliable differentiation of *Meyerozyma guilliermondii* from *Meyerozyma caribbica* by internal transcribed spacer restriction fingerprinting. BMC microbiology, 14(1), 52.

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics, 30(9), 1312-1313.

Tanner Jr, F., Vojnovich, C., & Van Lanen, J. M. (1945). Riboflavin production by Candida species. Science (Washington), 101, 180-161.

Tatusov, R. L., Galperin, M. Y., Natale, D. A., & Koonin, E. V. (2000). The COG database: a tool for genome-scale analysis of protein functions and evolution. Nucleic acids research, 28(1), 33-36.

Taylor, J. W. (2011). One fungus= one name: DNA and fungal nomenclature twenty years after PCR. IMA fungus, 2(2), 113-120.

Ter-Hovhannisyan, V., Lomsadze, A., Chernoff, Y. O., & Borodovsky, M. (2008). Gene prediction in novel fungal genomes using an ab initio algorithm with unsupervised training. Genome research, 18(12), 1979-1990.

Vacchini, V., Gonella, E., Crotti, E., Prosdocimi, E. M., Mazzetto, F., et al. (2017). Bacterial diversity shift determined by different diets in the gut of the spotted wing fly *Drosophila suzuki*i is primarily reflected on acetic acid bacteria. Environmental microbiology reports, 9(2), 91-103.

Vaughan-Martini, A., Kurtzman, C. P., Meyer, S. A., & O'Neill, E. B. (2005). Two new species in the *Pichia guilliermondii* clade: *Pichia caribbica sp. nov*., the ascosporic state of *Candida fermentati*, and *Candida carpophila comb. nov*. FEMS yeast research, 5(4-5), 463-469.

Wang, G. Y., Chi, Z., Song, B., Wang, Z. P., & Chi, Z. M. (2012). High level lipid production by a novel inulinase-producing yeast *Pichia guilliermondii* Pcla22. Bioresource technology, 124, 77-82.

Wisniewski, M., Biles, C., Droby, S., McLaughlin, R., Wilson, C., et al. (1991). Mode of action of the postharvest biocontrol yeast, *Pichia guilliermondii*. I. Characterization of attachment to *Botrytis cinerea*. Physiological and Molecular Plant Pathology, 39(4), 245-258.

Wrent, P., Rivas, E. M., Peinado, J. M., & de Silóniz, M. I. (2016). Development of an affordable typing method for *Meyerozyma guilliermondii* using microsatellite markers. International journal of food microbiology, 217, 1-6.

Zhao, J., Mou, Y., Shan, T., Li, Y., Zhou, L., et al. (2010). Antimicrobial metabolites from the endophytic fungus *Pichia guilliermondii* isolated from *Paris polyphylla* var. yunnanensis. Molecules, 15(11), 7961-7970.

Zou, Y. Z., Qi, K., Chen, X., Miao, X. L., & Zhong, J. J. (2010). Favorable effect of very low initial k L a value on xylitol production from xylose by a self-isolated strain of *Pichia guilliermondii*. Journal of bioscience and bioengineering, 109(2), 149-152.