

The mycobiota of the sand fly *Phlebotomus perniciosus*: involvement of yeast symbionts in uric acid metabolism

Elena Martin¹, Ilaria Varotto Boccazzi¹, Leone De Marco², Gioia Bongiorno³, Matteo Montagna⁴, Luciano Sacchi⁵, Priscilla Mensah², Irene Ricci², Luigi Gradoni³, Claudio Bandi^{1,6}, Sara Epis^{1,6*}

¹Department of Biosciences, University of Milan, Milan, Italy

²School of Biosciences and Veterinary Medicine, University of Camerino, Camerino, Italy

³Unit of Vector-Borne Diseases, Istituto Superiore di Sanità, Rome, Italy

⁴Department of Agricultural and Environmental Sciences- Production, Landscape, Agroenergy, University of Milan, Milan, Italy

⁵Department of Biology and Biotechnology, University of Pavia, Pavia, Italy

⁶Pediatric Clinical Research Center Romeo and Enrica Invernizzi, Ospedale "Luigi Sacco", Milan, Italy

*Corresponding author: Sara Epis, Department of Biosciences, University of Milan, Via Celoria 26, 20133 Milan (Italy); Tel: +39 02 50314710; E-mail: sara.epis@unimi.it

Running title: yeast community of *Phlebotomus perniciosus*

Summary

The knowledge of the fungal mycobiota of arthropods, including the vectors of human and animal diseases, is still limited. Here, we investigated the mycobiota associated with the sand fly *Phlebotomus perniciosus*, the main vector of leishmaniasis in the western Mediterranean area, by a culture-dependent approach (microbiological analyses and sequencing of the 26S rRNA gene), internal transcribed spacer (ITS) rRNA amplicon-based next-generation sequencing, fluorescence in

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as an 'Accepted Article', doi: 10.1111/1462-2920.14044

This article is protected by copyright. All rights reserved.

situ hybridisation (FISH), and genome sequencing of the dominant yeast species. The dominant species was *Meyerozyma guilliermondii*, known for its biotechnological applications. We focused the attention on this yeast and we investigated its prevalence in adults, pupae and larvae of reared sand flies (overall prevalence: 57.5%) and of field-collected individuals (overall prevalence: 9%). Using whole-mount FISH and microscopic examination, we further showed that *M. guilliermondii* colonizes the midgut of females, males and larvae and the distal part of Malpighian tubules of female sand flies, suggesting a possible role in urate degradation. Finally, the sequencing and analysis of the genome of *M. guilliermondii* allowed to predict the complete uric acid degradation pathway, suggesting that the yeast could contribute to the removal of the excess of nitrogenous wastes after the blood meal of the insect host.

Introduction

In recent decades, the role of microbes (fungi and bacteria) as insect symbionts has drawn attention (Gibson and Hunter, 2010). One of the most promising fields of investigation is the interaction between yeasts and insects, due to the diversity and metabolic capacity of yeasts and their presence in almost every environment (e.g. Blackwell, 2017). However, the current knowledge of yeasts occurring in the gut, and thereby belonging to the intestinal microbiota of arthropod vectors, of medical and veterinary importance, remains still limited. Some publications report the presence of the yeasts in the gut of mosquitoes (Diptera: Culicidae). Gusmão *et al.* (2007 and 2010) identified yeasts of the genera *Pichia* and *Candida* in the *diverticulum* of the mosquito *Aedes aegypti*. Ricci *et al.* (2011a and b) observed the presence of *Wickerhamomyces anomalus* (also known as *Pichia anomala*) in the midgut of adult mosquitoes, of both sexes, belonging to different species (*Anopheles stephensi*, *Anopheles gambiae*, *Aedes albopictus* and *Ae. aegypti*); male gonads were also shown to be positive to yeast presence, as well as larvae and pupae. Recently, *W. anomalus* has also been isolated and phylogenetically characterized in laboratory-reared adults and larvae of *Phlebotomus perniciosus* (Diptera: Psychodidae), the main phlebotomine vector of human and canine leishmaniasis in the Mediterranean area (Martin *et al.*, 2016). Previous surveys on the fungal microflora in sand flies, using routine microbiological methods, identified fungal strains associated with different species of these vectors, collected in North-Western Iran (Akhoundi *et al.*, 2012). In addition, the yeast-like fungus *Pseudozyma* sp. has been isolated from wild-collected *Lutzomyia longipalpis*, a major *Leishmania* vector in central/south America (Sant'Anna *et al.*, 2014); this study also presents an interesting investigation on the interaction between *Leishmania* and component of the microbiota, in laboratory reared *Lu. longipalpis*.

Larvae of sand flies feed on the decomposing organic materials and can acquire a part of their microbiota (including bacteria and fungi), whereas as adults they likely acquire microorganisms through their daily feeding on natural sugar solutions, especially nectars, lymph and water from

plants. Even though yeasts do not appear to be as frequent/abundant as prokaryotes in the intestinal microbiota of sand flies, they can account for an important part of the microbiota biomass, as they have a cell volume 30- to 100-fold higher than bacteria (Gatesoupe, 2007). Intestinal yeasts are likely to interact with the other components of the microbiota, including pathogens, and are thus likely to influence the insect biology and its vector competence. Therefore, the development of new strategies for the control of sand fly-borne diseases should also consider the mycobiota of these insects.

The aim of our study was to investigate the diversity of the fungal microbiota that colonizes laboratory-reared and wild-collected *P. perniciosus* individuals, at different developmental stages, with the goal to identify yeast species, with the potential to be exploited for future development and applications in the control of leishmaniasis and other vector-borne diseases. *P. perniciosus* was selected as model of investigation because of its medical and veterinary importance, being the main vector, in the western Mediterranean area, of the protozoan parasite *Leishmania infantum* (Trypanosomatida: Trypanosomatidae), the causative agent of canine and human leishmaniases (Killick-Kendrick, 1990; Alten *et al*, 2016).

Results

Yeast isolation and characterization

Yeasts from adults and L4 larvae of laboratory reared sand flies *P. perniciosus* were isolated and identified at the species level by 26S rRNA gene amplification, sequencing and comparison with sequences in NCBI database. As reported in tables 1 and 2, we obtained a total of 112 yeast isolates derived from whole adults, from digestive apparatus of adults and from L4 larvae; table 1 shows that we were not able to obtain yeast isolation from all the processed samples.

The molecular identification of all the isolated yeasts is reported in table 2. The dominant yeast species was *Meyerozyma guilliermondii*, isolated from both sexes of whole adults and from the digestive apparatus of adults and L4 larvae. The yeast *Wickerhamomyces anomalus*, derived from both whole adults and from the digestive system of the L4 larvae, was also frequently isolated. Finally, the Basidiomycota yeasts *Trichosporon* sp. and *Rhodotorula mucilaginosa* were also isolated from the digestive apparatus of larvae and from whole females, respectively.

A total of eight species-specific 26S rRNA region sequences were submitted to GenBank (table 3). The analysis performed on the 26S rRNA sequences showed a variable sequence identity ranging from 99% to 100% (E-value: 0) with *Meyerozyma guilliermondii* KU729151 (strain ATCC 6260).

Amplicon-based analysis of fungal community

After removal of chimeras and low-quality base call, a total of 37,154 amplicons of the fungal ITS rRNA were obtained by larvae and adults of the sand fly. Applying a similarity threshold of 97%, these amplicons were clustered into a total of 295 operational taxonomic units (OTUs). In particular 103 OTUs were amplified from adult females, 114 OTUs from adult males and 142 OTUs from larvae. The taxonomic assignment is synthesized in fig. 1 and in table S1-S3. The fungal community of *P. perniciosus* is dominated by taxa belonging to the phylum Ascomycota (average $96.6\% \pm 3.1\%$). Analyzing the samples according with the developmental stages, Ascomycota predominates in both adults ($98.1\% \pm 2.4\%$ of the total retrieved OTUs) and larvae (93.5%). Zygomycota represents the second taxon in term of abundance in larvae (6.2%) and is roughly null in adults (0.04% in females, 0% in males); meanwhile Basidiomycota are more present in adult females (3.5%) than in adult males (0.3%) and larvae (0.2%). The most represented class in the analyzed sand flies is Saccharomycetes (49.4% in adults, in detail 53.7% in adult males and 45.2% in adult females, 33.4% in larvae). Sordariomycetes (40.9% in adults, in particular 45.5% in adult males and 36.2% in adult females, 29% in larvae) and Eurotiomycetes (0.3% in adults, in particular 0.2% in adult males and 0.3% in adult females, 27.5% in larvae) are also represented. A total of 14

genera were retrieved in the analyzed samples. Only four genera, namely *Candida*, *Engyodontium*, *Meyerozyma* and *Simplicillium*, are shared between the analyzed sand fly groups, being identified in both adult males, adult females and larvae. The great majority of amplicons in larvae belongs to the genera *Candida* (32.7%), *Aspergillus* (27.4%) and *Simplicillium* (24.7%). By contrast, the majority of amplicons obtained from adults (both males and females) were identified as *Meyerozyma* (44% in adult females, 53.6% in adult males) and as *Simplicillium* (34% in adult females, 45.4% in adult males). Nevertheless, the genera *Candida* and *Aspergillus* are poorly represented in adults ($0.2\% \pm 0.1\%$ and $0.1\% \pm 0.1\%$ respectively) and the genus *Meyerozyma* is low in larvae (0.6%). The obtained 295 OTUs are not equally distributed between the analyzed sand fly groups; in fact 37% of OTUs are exclusive of larvae (see Fig. S1).

The OTU core of the population, consisting of the OTUs that are in common between larvae, adult males and adult females, is composed of 10 taxonomic units, therefore representing 3.4% of the *P. perniciosus* panmycobiota. The OTUs in the core were assigned to *Engyodontium*, *Meira*, *Meyerozyma* and *Simplicillium*. Among them, one OTU in the core represents the 63% of all the reads obtained from adults and it is assigned to the genus *Meyerozyma*, confirming the results obtained using culture dependent methods. Another noteworthy OTU in the core is assigned to *Simplicillium*; this OTU represents 25% of all the reads from larvae and 8.8% of all the reads obtained from adults.

Prevalence of the yeast *M. guilliermondii* in sand fly populations

In order to evaluate the prevalence of *M. guilliermondii* (the dominant yeast species), 80 reared specimens (including males, females, L4 larvae and pupae) and 100 field collected sand flies, were screened with species-specific primers. Due to the difficulty to collect larvae and pupae of wild sand flies, the molecular screening of field collected insects have been performed only on adults. Results are summarized in table 4.

The prevalence of *M. guilliermondii*, evaluated by PCR, in the reared population was 65% (13/20) in males, 40% (8/20) in females, 90% (18/20) in larvae and 35% (7/20) in pupae, with an overall

prevalence of 57.5% (95% CI: 46.7-68). By contrast, the overall prevalence of the yeast in field collected populations was 9% (95% CI: 3.4-14,6) that is statistically lower than the prevalence in Istituto Superiore di Sanità (ISS) reared one ($\chi^2 = 49.268$, $P < 0.00001$). The analyzed wild individuals showed a prevalence of 12% (6/50) in males and 6% (3/50) in females.

The prevalences in different sites of collection were not equal, although there is no statistically significant difference (χ^2 , $P = 0.250$): in Frascati (RM) the prevalence was 5% (3/58), in Roma it was 10% (1/10) and in Vieste-Peschici (FG) the prevalence was 16% (5/32).

Microscopic analysis of fungi in sand flies and whole mount fluorescent *in situ* hybridization (FISH)

In order to confirm the presence of yeasts in digestive tract of sand flies, microscopic analyses were performed on males and females. Giemsa stained yeast cells and budding yeast cells were observed in 5/10 slide preparations made from digestive tracts and from the Malpighian tubules of *P. perniciosus* (2/5 from males, 3/5 from females). The fungi exhibited typical staining properties (as basophilic) and structure, as reported in fig. S2.

The presence of *M. guilliermondii* in adult sand flies and L4 larvae was further investigated with whole mount FISH. The specificity of the probe MGU26S410, designed for *M. guilliermondii* complex, was evaluated against the yeast strains of *R. mucilaginosa*, *W. anomalus*, *Trichosporon* sp. and *M. guilliermondii* isolated from sand flies in this work. The probe resulted highly specific for *M. guilliermondii* (Fig. S3). The probe MGU26S410 for *M. guilliermondii* complex localized yeasts in the midgut of males (Fig. S4), females (Fig. 2) and larvae (Fig. S5) and in the distal part of the Malpighian tubules of females (Fig. 2), where yeast cells were in active division. However, *M. guilliermondii* was not detected in all the analyzed specimens, confirming what observed with PCR screening or yeast isolation (Fig. S6). In addition, the signal obtained using the PF2 probe (the universal yeast probe) in Malpighian tubules of females is comparable in intensity and localization to the signal obtained with the specific probe for *M. guilliermondii* complex, indicating the highest concentration of yeast of *M. guilliermondii* complex in this compartment. The probe non-

MGU26S410 and the pre-treatment with RNase A (Fig. S7), utilized as negative controls, did not generate detectable fluorescence signals in the analyzed samples.

Phylogenomic analysis of a strain of *M. guilliermondii* isolated from *P. perniciosus*

After sequencing with an Illumina HiSeq 2500, 12,869,872 150bp paired-end reads were obtained. The reads were checked and resulted of good quality. The obtained assembly consists of 108 contigs and has an overall length of 10,666,172 base pairs (and 5,476 genes), comparable to the 10,609,954 base pairs (and 5,396 genes) of the only genome already published for *M. guilliermondii* (strain ATCC 6260) (table 5). In addition, the number of not shared genes is 21 for *M. guilliermondii* Pp, and four not shared genes for the *M. guilliermondii* ATCC 6260 (table 5). We retrieved 5,531 putative genes with an average length of 1,521.35 base pairs, composing most of the genome by totaling 8,414,591 nucleotides. Out of the total of putative genes, 82.73% were annotated. Starting from a comprehensive dataset of fungal genomes (table S4), OrthoMCL retrieved 190 shared orthologous clusters with a single copy in each genome. Each cluster of orthologs was multialigned with ClustalO and the resulting alignment was cleaned with Gblock. All cleaned alignments were concatenated to obtain the final input for the phylogenetic analysis, consisting of a total of 42,467 aminoacids. A maximum likelihood analysis performed with RaxML using 100 bootstrap pseudoreplicates resulted in the phylogenetic tree showed in fig. 3a. *M. guilliermondii* Pp is strongly supported to be a sister clade to *M. guilliermondii* ATCC 6260 with *Clavispora lusitaniae* as their closest relative. The phylogenetic picture is consistent with other phylogeny studies conducted on yeasts, with the fission yeast *Schizosaccharomyces pombe* as the outgroup and a clear separation between Saccharomycotina and Pezizomycotina species (Kurtzman, 2014; Spatafora *et al.*, 2006).

In order to perform the comparative genomics, species-specific and strain-specific clusters of orthologous genes were retrieved for *M. guilliermondii* Pp, *M. guilliermondii* ATCC 6260, and *C. lusitaniae* (Fig. 3). The biggest subset by far is composed of clusters shared between the three yeast species (4,242 clusters), whereas few clusters are strain specific. There is evidence for a substantial

subset of *Meyerozyma* specific subset composed of 793 clusters. A visualization of the GO terms assigned to the three genomes presents a homogeneous landscape, implying an overall high similarity of general gene functions. The genomic alignment (Fig. 3b) between *M. guilliermondii* Pp and *M. guilliermondii* ATCC 6260 shows that synteny is mainly conserved between the two, with the exception of an inversed translocation of about 70 kb located roughly at 3,365 kb into the *M. guilliermondii* Pp genome.

Bioinformatic prediction of uric acid degradation pathway and uricolytic activity of *M. guilliermondii*

All components of the uric acid degradation pathway, except for Allantoine racemase, were found in the set of predicted proteins of *M. guilliermondii* Pp. Each enzyme had a reciprocal blast hit with a percentage of identity higher than 56%. Specifically, Mguil2443_1 was identified as Urate oxidase (percentage of identity 64%), Mguil178_1 as HIU Hydrolase (63%), Mguil553_1 as OHCU decarboxylase (56%), Mguil3309_1 as Allantoinase (66%), Mguil2347_1 as Allantoicase (67%), Mguil178_1 as Ureidoglycolate hydrolase (63%). Uric acid degradation pathway of *M. guilliermondii* Pp is shown on fig. 4c.

After the prediction of uric acid degradation pathway, and following the particular localization of the yeast in Malpighian tubules, we evaluated the yeast capability of degrading uric acid; the uricolytic activity assays were performed on two strains for each isolated yeast species. All tested *R. mucilaginosa*, *W. anomalus*, *Trichosporon* sp. and *M. guilliermondii* yeast isolates were able to grow on modified YPU containing 4 g/l of uric acid.

However, only *M. guilliermondii* and *R. mucilaginosa* strains were able to degrade uric acid showing a clear halo (suggestive of uric acid utilization) surrounding the yeast (Fig. 4a and 4b). The halos obtained from *M. guilliermondii* strains were comparable with those produced by *R. mucilaginosa* strains, the yeast was able to grow and consume the uric acid as sole nitrogen source (Vera-Ponce de Leon *et al.*, 2016).

Discussion

In this work, we investigated the mycobiota associated with the sand fly *P. perniciosus*, the main vector of canine and zoonotic visceral leishmaniasis. Culture-dependent analyses and ITS rRNA amplicon-based analysis show that the dominant yeast species is *M. guilliermondii*. This ascomycota yeast could be considered a “generalist yeast” that can form relationships with insects belonging to a variety insect taxa; in fact, it has been recovered from long-horned beetles (Nardon and Grenier, 1989), scarab beetles (Vishniac and Johnson, 1990), bark beetles (Rivera *et al.*, 2009), leaf beetle and crambid snout moths (Molnar *et al.*, 2008), fire ants (Ba *et al.*, 1995), adrenid bees and a mushroom-feeding fly (Zacchi and Vaughn-Martini, 2002), fishfly, dobsonfly and owlfly (Nguyen *et al.*, 2007), and several genera of mosquitoes (Gusmão *et al.*, 2010; Bozic *et al.*, 2017). Interestingly, it has never been recovered from sand flies. The prevalences recorded for *M. guilliermondii* in *P. perniciosus* indicate that this yeast is well distributed in the insectary population, in both sexes and in all analyzed life stages (with an overall prevalence of 57.5%). In wild-collected individuals, the yeast is less prevalent (overall: 9%) and less homogeneously distributed, with variability related also with the geographical location (Lazio region 7.5%; Puglia region: 16%). As for the diffusion of the yeast in the two sexes, we observed a higher prevalence in males, in both reared (65% in males, 40% in females) and field-collected individuals (12% in males and 6% in females).

The higher prevalence in lab-reared flies compared to wild collected ones could derive from different factors. First, it is reasonable to assume that laboratory rearing provides conditions that favors the horizontal transmission of the yeast among the insects, with cofeeding likely representing one of the main routes for transmission. Other ways for the horizontal transmission of the microbiota, that are known to occur in insects, are: trophallaxis (e.g. in termites; Nalepa *et al.*, 2001); venereal transfer (e.g. in mosquitoes; Damiani *et al.*, 2008), which includes the transmission of ejaculated components of the microbiota from males to females (e.g. in aphids; Moran and

Dunbar, 2006). Whether similar phenomena occur in sand flies is however unknown. Assuming that horizontal transmission is responsible for the higher prevalence of *M. guilliermondii* in the reared flies, we would conclude that this yeast is at least neutral toward the insect host, otherwise we would have observed some mortality or fitness reduction; indeed, we have evidence that *M. guilliermondii* is present in the examined colony since 2013 (unpublished observations). Second, we could also suggest that *M. guilliermondii* plays a beneficial role in laboratory condition. Indeed, yeasts are known to provide insects with several advantages, which include the provision of nutrients, the detoxification of harmful substances and the protection from biotic stresses (Gibson and Hunter, 2010; Janson *et al.*, 2008). It is reported that yeast cells can be a source of B vitamins, proteins and amino acids, that can be assimilated by the insects. We can thus hypothesize that, under laboratory conditions, *M. guilliermondii* contributes to the diet of reared sand flies with nutrients which are on the other side easily available in nature. Third, a founder-effect explanation could be proposed: the sand fly lab colony had been established from positive insects. Anyway, the maintenance of *M. guilliermondii* at a high prevalence, over a period of 30 years, would imply a neutral or beneficial effect of the yeast toward the insect. As for the higher prevalence of *M. guilliermondii* in males than in females, we could perhaps suggest that this difference derives from the different diet (males do not take a blood meal). The diet is indeed known to affect the composition of the microbiota, as shown in mosquitoes, in relation with the blood meal (Wang *et al.*, 2011).

Culture-independent ITS amplicon analysis of the fungal community in larvae and adults highlighted a core of 10 OTUs assigned to four genera of fungi (Fig. S1). Two genera that are worth to be considered are the already discussed yeast *Meyerozyma*, the predominant genus in adults, and the filamentous fungus *Simplicillium*. We could hypothesize that the genera that are in common between life stages and adults, represented in the core, can be transtadially transmitted. This hypothesis is also supported by the localization of the yeast *M. guilliermondii* in sand flies, i.e. in

the gut of adults and larvae, and also in the Malpighian tubules of females (Fig. 2). A previous, and so far unique, report of the presence yeasts in this excretory organ in dipteran was published in 1984 (Schlein *et al.*, 1985), where unidentified yeast-like cells (about 2-3 micrometers), without filaments or hyphae, were observed in the gut and in the Malpighian tubules of laboratory reared *P. papatasi* and *P. tobbi*.

The fact that *M. guilliermondii* can be housed in the Malpighian tubules of larvae and of adults opens the possibility that the yeast persistence could also be due to escape from the metamorphosis using the Malpighian tubules as a refuge. It is reported that in Diptera the larval Malpighian tubules remain intact during metamorphosis (Singh *et al.*, 2009) and tubular fluid including the microbes could then drain from Malpighian tubules towards the midgut.

In mosquitoes, the main role of Malpighian tubules after a blood meal switches from diuresis to detoxification and purine catabolism, and uric acid (an end product of purine catabolism) is accumulated in such structures within 24 h after a blood meal (Esquivel *et al.*, 2016). The physiology of Malpighian tubules in sand flies could be similar to that in mosquitoes, being both hematophagous nematoceran Diptera. The degradation of purines to uric acid is conserved in all organisms, but the resulting uric acid can either be excreted or further degraded and, in particular, Diptera excrete allantoin as the end product. Most fungi possess all the necessary nitrogen catabolic enzymes to completely degrade uric acid; the yeast *Saccharomyces cerevisiae*, where the pathway begins at the step of allantoin degradation, appears as an exception. Ammonia, the final product of uric acid degradation, is readily assimilated as a nitrogen source utilized by most bacteria and fungi (Lee *et al.*, 2013). In cockroaches, in which uric acid is accumulated in the fat body in periods of diet rich in proteins and decreases in times of shortness (Cochran *et al.*, 1979), the stored nitrogen can then be recycled thanks to the activity of his primary endosymbiont *Blattabacterium cuenoti* (Valovage and Brooks, 1979). In particular, Patiño-Navarrete *et al.* (2014) demonstrated that in the cockroach *Blattella germanica* the nitrogen recycling pathway is chimeric, with participation of enzymes from the host and the bacterial symbiont *B. cuenoti*.

Here, considering the detection of *M. guilliermondii* in Malpighian tubules of females of *P. perniciosus*, we can speculate that the yeast can play a role in uric acid degradation at the end of purine catabolism, in particular after the blood meal, explaining this unusual localization only in female individuals. The analysis of the genome of *M. guilliermondii* Pp. from *P. perniciosus*, supported this evidence: after the bioinformatic prediction of uric acid degradation pathway we demonstrated that all of the components of the uric acid degradation pathway, up to ammonia (except for Allantoine racemase), were found in the set of predicted proteins of *M. guilliermondii* Pp. The uricolytic activity of *M. guilliermondii* yeast strains isolated from sand flies in this work was also investigated determining the degradation halo in YPU medium (containing uric acid); we observed a clear halo surrounding yeast colonies, which is suggestive for uric acid degradation. We can thus propose that this yeast, associated with female *P. perniciosus* sand flies, could contribute to the degradation of urates, facilitating the removal of excesses of nitrogenous wastes in the Malpighian tubules after the blood meal. Ammonia could indeed be easily be drained towards the gut and, within the gut, ammonia could then be recycled by the resident microbial community, or eliminated with the fecal material.

Judging from genomic analysis, the genome of *M. guilliermondii* isolated from *P. perniciosus* is very similar to that of the available environmental strain (ATCC 6260) of the same species. A phylogenomic analysis performed on 190 genes shared between 15 yeast species placed the two *Meyerozyma* genomes as sister clades. Comparative genomics provides results coherent with those of phylogenomics, showing that, in fact, the two genomes have a comparable number of genes with overall similar general functions. Lastly, a genomic alignment points out that synteny is conserved, meaning that, apart from a minor translocation, blocks of similarity are present in the same order in the two genomes.

Finally, as reported in Papon *et al.* (2013), the yeast *M. guilliermondii* can display a large range of applications in various areas of fundamental and applied scientific research. Some *M. guilliermondii* strains have been shown to be useful as biological control agents due to their ability

to compete for nutrients and living space with pathogens. The presence of the yeast *M. guilliermondii* in phlebotominae sand flies could also play a protective role against other microorganisms, including transmitted pathogens (e.g. *Leishmania* spp. or viruses). Indeed, as in the case of other yeast species (e.g. *Wickerhamomyces anomalus*, symbiont of arthropod vectors, including sand flies), *M. guilliermondii* likely compete for nutrients or could be a source of enzymes, or could protect the host insect against infectious agents (Wang *et al.*, 2008; Cappelli *et al.*, 2014). As already demonstrated for bacteria, yeasts can also benefit insects by promoting the development of other mutualistic and beneficial organisms in the community, while decreasing the presence of insect pathogens (Zindel *et al.*, 2011).

In conclusion, although further studies will be required, the investigation of the mycobiota of *P. perniciosus* may have a great impact in a better understanding of some aspects of phlebotominae biology and towards the development of strategies aimed to reduce the vectorial capacity and/or inhibiting pathogen transmission using engineered yeast.

Experimental Procedures

Samples analyzed in this study

Laboratory-reared sand flies

A laboratory colony of *Phlebotomus perniciosus*, originated from samples collected in Madrid area (Spain) and established at the Institute of Health Carlos III of Madrid in 1983, was employed. The colony has been reared since June 2012 at the Unit of Vector-Borne Diseases of Istituto Superiore di Sanità, Rome, and at time of this study the sand flies were at the 27th generation. Adults were routinely maintained in thin mesh cages in thermostated cabinets at standard conditions of temperature, humidity and photoperiod ($28 \pm 1^\circ\text{C}$; RU 95-100%; light:dark period 7:17 hrs) and daily provided of 30% sugar solution (Maroli *et al.*, 1987). Larvae were grown in plastered jars at

same standard conditions, supplemented by sterilized larval food following the mass-rearing technique (Modi and Tesh, 1983), with strict regulation of food quantity and moisture.

Wild-caught sand flies

Sample collection was carried out in central and south Italy between 2012 and 2015, in the months of July and August, in three rural sites using five CDC miniature light traps equipped with a fine net cage for each site and placed near animal shelters. Site 1 was an agricultural area of Frascati commune (Rome province, central Italy) (N 41,838810-EO 12,700712) where available blood sources were horses, sheep and dogs. Site 2 consisted of a horse stable placed in Rome outskirts (N 41,936649-EO 12,367065). Site 3 was a farm hosting goats, sheep and dogs, and placed along a main road (SP 52) connecting two towns (Vieste and Peschici, Foggia province, south Italy) (N 343 41,901623-EO 16,138898). The sites were selected considering a variety of parameters consistent with *P. perniciosus* presence, and represented by different environmental conditions and biotic factors that may affect the sand fly microbiome.

All field collected specimens were morphologically identified to species level according to Theodor (1958) and Léger *et al.* (1983), by cutting the last abdominal segments in both sexes and the head in females, with subsequent clarification in chloro lactophenol and permanent mounting. The rest of the insect body was stored in absolute ethanol to preserve DNAs of both the host and the harbored microorganisms.

Yeast isolation

Laboratory reared sand flies were surface sterilized with ethanol and washed twice with sterile physiological solution (NaCl 0.9%) supplemented with detergent, in order to eliminate external contaminants and their hair. Ten whole females and 10 whole males were homogenized in 200 μ l of sterile NaCl 0.9% and diluted serially up to 10^{-3} . 100 μ l of each dilution were plated on YM agar medium (3 g/l yeast extract, 3 g/l malt extract, 5 g/l peptone, 10 g/l dextrose, 20 g/l agar, pH 6.2) and 100 μ l on PDA medium (Sigma Aldrich, St. Luis, USA), both supplemented with

chloramphenicol 100 mg/l to avoid bacterial growth. Fifteen females, 15 males and 10 L4 larvae, after surface cleaning, were aseptically subjected to dissection under a stereomicroscope (Leica M50, Leica Microsystems, Wetzlar, Germany) with sterile needles in order to separate the digestive system. Ten digestive systems from females, 10 digestive systems from males and all the digestive systems excised from larvae were homogenized in 100 µl NaCl 0.9% and plated individually directly on YM and on PDA agar medium. The digestive system of five females and the digestive system of five males were pooled for sex and homogenized in 200 µl NaCl 0.9% (each pool), diluted up to 10^{-2} and both the dilutions were plated on YM and on PDA medium. Plates were incubated at 28°C for 48-72 hours (Varotto Boccazzi *et al.*, 2017).

Characterization of yeast isolates

Growing yeast colonies from whole adult sand flies, digestive systems of adults and larvae were selected and collected based on colony morphology (Kurtzman and Fell, 2000), re-plated twice and the pure cultures of yeast strains, grown in liquid potato dextrose broth (PDB) medium (Sigma Aldrich, St. Luis, USA), were stored in 20% glycerol at – 80 °C.

DNAs from yeast isolates were extracted using the enzyme lyticase (Sigma Aldrich, St. Luis, USA) and DNeasy Blood and Tissue Kit (Qiagen, Hilden, Germany) following the supplementary protocol for purification of total DNA from yeasts. Total extracted DNAs were quantified by spectrophotometry (Nanodrop 2000, Thermo Scientific, Wilmington, USA).

In order to perform the characterization of the isolated strains, a PCR protocol for the amplification of 26S rRNA gene was carried out using primers NL1 (5'-GCATATCAATAAGCGGAGGAAAAG-3') and NL4 (5'-GGTCCGTGTTTCAAGACGG-3') (Ferreira *et al.*, 2010) in a total volume of 25 µl containing: 1X Green GoTaq Reaction Buffer, 0.2 mM dNTPs, 0.5 µM of each primer, 1.25 U GoTaq G2 (Promega, Madison, USA), 7-10 ng of template DNA following the thermal protocol reported in Ferreira *et al.* (2010). All amplicons, obtained from 26S rRNA gene amplification, were sequenced and the obtained sequences were corrected and subjected to BLAST analysis (<http://www.ncbi.nlm.nih.gov/blast>) in order to

compare our sequences with the available sequences of fungi in GenBank (nucleotide collection nr/nt; <http://www.ncbi.nlm.nih.gov/genbank/>). Two sequences for each yeast species isolated in this study were deposited in GenBank database.

Microscopic analysis of fungal infection in sand flies

Ten adult reared sand flies (five males and five females) were macerated in 10% KOH for 24 h at room temperature, dehydrated with steps in H₂O, 70% EtOH, 100% EtOH, finally xylol, following the protocol reported in Schlein and colleagues (1985). Digestive systems and Malpighian tubules from male and female sand flies were placed on microscopic slides in a drop of 1X PBS, smashed and smeared. The smears were fixed with methanol, stained with Giemsa and examined with a light microscope (Axioskop 2, Zeiss, Oberkochen, Germany).

ITS rRNA amplicon-based analysis of the fungal community of sand flies

ITS rRNA amplicon-based next-generation sequencing was carried out from pooled DNA of five reared specimens (in total, two pool of males, two pool of females and two pool of L4 larva) using primers ITS5 (5'-GGAAGTAAAAGTCGTAACAAGG-3') and ITS4 (5'-TCCTCCGCTTATTGATATGC-3') (White *et al.*, 1990). DNAs were extracted as reported below. After PCRs, amplicons were purified using Agencourt Ampure beads (Agencourt Bioscience Corporation, MA, USA) and then sequenced using Roche FLX 454 titanium. PCR reactions and amplicons sequencing were performed commercially by MR DNA (Shallowater, TX, USA). QIIME (Caporaso *et al.*, 2010) was used to remove pyrosequencing adaptors and low quality base calls (< 30 Phred score), perform the amplicon size selection (retained sequences between 200 and 400 bp) and purge the chimeric sequences. ITS2 gene sequences obtained by 454 pyrosequencing assays were deposited in European Nucleotide Archive with accession number PRJEB22022. High-quality 454 sequence reads were clustered into operational taxonomic units (OTUs) using UCLUST (Edgar, 2010) with a 97% sequence-identity threshold. The most abundant sequence of each OTU was selected as representative and used to build the overall OTU table. These OTUs were

taxonomically assigned by RDP Classifier (Wang *et al.*, 2007) using, as reference, the Warcup ITS fungal database (Deshpande *et al.*, 2016).

***M. guilliermondii*-specific primers design and PCR screening on wild-caught and laboratory-reared sand flies**

Before DNA extraction, 40 adults (20 males and 20 females), 20 pupae and 20 L4 larvae of *P. perniciosus*, reared in the insectary of ISS (Roma, Italy), and 100 field collected adults of *P. perniciosus* (50 males and 50 females), as reported in table 4, were carefully washed in sterile 1X PBS and dried. Samples were homogenized using sterile pestles in 200 µl of lysis buffer (1 M sorbitol, 0.1 M sodium EDTA and 200 units of lyticase (Sigma Aldrich, St. Luis, USA)) and incubated for 45 min at 30°C. Then, the extraction was carried out using the commercial kit DNeasy Blood and Tissue Kit (Qiagen, Hilden, Germany), following the protocol for purification of total DNA from animal tissues. DNAs were eluted in 25 µl of AE buffer. Each DNA was quantified by spectrophotometry (Nanodrop 1000, Thermo Scientific, Wilmington, DE, USA).

To investigate the prevalence of the yeast *M. guilliermondii*, a specific primer set mguITS_F108 (5'-CTTTGGTTTGGCCTAGAG-3') and mguITS_R481 (5'-ATAAACCTAATACATTGAGAGG-3'), that amplifies a fragment of 373 bp, was designed on ITS rRNA region. The amplification was carried out in a total volume of 25 µl containing: 1X Green GoTaq Reaction Buffer, 0.2 mM dNTPs, 0.5 µM of each primer, 1.25 U GoTaq G2 (Promega, USA) 7-10 ng of template DNA with the following thermal protocol: initial denaturation at 94°C for 7 min, followed by 35 cycles of 94°C for 45 sec, 56°C for 45 sec, 72°C for 45 sec and a final extension step of 72°C for 10 min. The PCR protocol was validated on three isolates of *M. guilliermondii* identified in this work by 26S rRNA gene sequencing.

The *M. guilliermondii*-specific PCR was used to screen the prevalence of *M. guilliermondii* on the previously extracted DNAs from whole insects (wild-caught and laboratory-reared sand flies). The obtained amplicons were visualized on agarose gel. Only for negative samples, 1 µl of the amplification products was subjected to a second round of the same PCR.

Fluorescence *in situ* hybridization (FISH): probe design and evaluation of the specificity

A specific 26S rRNA oligonucleotide probe was designed to target yeasts of the complex *M. guilliermondii* associated with sand flies (MGU26S410_5'-FITC: 5'-GGCTCACAAAATATCGAGTCTG-3'). The MGU26S410 probe was designed on the basis of an alignment with sequences of a selection of yeast strains commonly isolated in arthropods derived from the NCBI Nucleotide database. The oligonucleotide sequence was further checked for its suitability with the software probeCheck ([http://www. microbial-ecology.net/probecheck](http://www.microbial-ecology.net/probecheck)) (Loy *et al.*, 2008). In addition, the probe PF2_5'-CY3: 5'-CTCTGGCTTCACCCTATTC-3' (Kempf *et al.*, 2000) was used as universal fungal probe and the probe non-MGU26S410, which is the complementary (antisense) sequence to the probe MGU26S410, was tested as a negative control for non-specific binding. As supplementary negative controls, males, females and larvae of sand flies were treated with RNase A (USB, Thermo Fisher Scientific, Paisley, United Kingdom), prior to incubation with probes. All FISH probes were purchased from Eurofins Genomics (Ebersberg, Germany). The specificity of the probe MGU26S410 was tested against the yeast strains of *R. mucilaginosa*, *W. anomalus*, *Trichosporon* sp. and *M. guilliermondii*, isolated from sand flies in this work, as previously described (Xufre *et al.*, 2006). Briefly, yeast cells were grown in PDB medium and harvested at the exponential growth phase, washed ones with 1X PBS and fixed for 4h with 4% paraformaldehyde at 4°C. Cells were hybridized in hybridization buffer (0.9 M sodium chloride, 0.01% w/v sodium dodecyl sulphate, 20 mM Tris–HCl) with 1.5 ng of each probe per microlitre, at 46°C for 2 h. After hybridization, cells were pelleted and washed with prewarmed hybridization buffer without probe for 30 min at 46 °C. The suspension was mixed with 200 µl 1× PBS and spotted onto microscope slides that were dried at 37°C and mounted with buffered glycerol 4% propyl gallate (80% glycerol, 4% n-propyl gallate, 0,02M Tris). The slides were examined using a laser-scanning confocal microscope SP8 (Leica, Wetzlar, Germany).

Whole mount fluorescence *in situ* hybridization (FISH)

Whole mount *in situ* hybridization was performed on reared *P. perniciosus*. Males, females and L4 larvae of *P. perniciosus* were cut near the last abdominal segment area utilizing a microclipper for dissection, fixed overnight in 4% paraformaldehyde at 4 °C and preserved in ethanol/PBS 1:1 at -20°C until processing. The samples were decolourized in 6% H₂O₂ for 30 minutes at room temperature and treated with 1.5 ng of the probe specific for *M. guilliermondii* complex MGU26S410_5'-FITC and of the universal yeast probe PF2_5'-CY3 (Kempf *et al.*, 2000) per microlitre in hybridization buffer (Xufre *et al.*, 2006) overnight at 42°C. After incubation, sand flies were washed with hybridization buffer without probe for 20 minutes at room temperature. Finally, 75 ng/ml of 4',6-diamidino-2-phenylindole (DAPI) were added for nuclei detection and incubated for 15 min at RT. The treated samples were mounted on slides with buffered glycerol 4% propyl gallate. The same protocol was applied for negative controls. The slides were observed using a laser-scanning confocal microscope SP8 (Leica, Wetzlar, Germany).

Phylogenomic analysis of a strain of *M. guilliermondii* isolated from *P. perniciosus*

DNA extraction and genome sequencing

We selected a yeast strain, isolated from the whole body of a reared male specimen of *P. perniciosus*, previously identified as *M. guilliermondii* by sequencing the 26S rRNA region, for the genome sequencing. The yeast strain, here referred to as *M. guilliermondii* Pp, was grown on YM agar medium for 30 hours, as previously described. 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). The assembled contigs are available on the European Nucleotide Archive under accession number GCA_900174495.1.

Reads assembly

After an assessment of the quality of the reads using FastQC (Andrews, 2010), the paired-end reads were assembled using the abyss-pe (Simpson *et al.*, 2009) *de novo*-assembler for short reads, setting all parameters at default values. The resulting contigs were filtered by length, retaining only contigs with a length greater than 1,000 nucleotides.

ORF calling

We performed yeast specific ORF calling using the software GeneMark (Besemer and Borodovsky, 2005) with an optional parameter for fungal sequences. The resulting putative genes were translated using transeq (Rice *et al.*, 2000; Goujon *et al.*, 2010). Putative genes were annotated using BLASTP (Camacho *et al.*, 2009) with a cut-off evalue of 0.000001 against a custom SwissProt database with the taxonomic filter “Diakarya” (Boeckmann *et al.*, 2003).

Phylogenetic study

Following a research on current literature, we identified a subset of yeast species representative of the Ascomycota phylum with a published genome sequence. We gathered 15 species (table S2), comprising the only published *M. guilliermondii* genome to date (strain ATCC 6260). Orthologous genes were retrieved using the OrthoMCL software (Li *et al.*, 2003). We selected from the resulting clusters of orthologs only genes present in a single copy in all species. A multiple sequence alignment was performed on each cluster of sequences using the ClustalO software (Sievers and Higgins, 2014). Alignments were cleaned with Gblock (Castresana, 2000), concatenated and used for a maximum likelihood phylogenetic analysis using raxml (Stamatakis, 2014) using 100 bootstrap.

Annotation and Comparative genomics

The genome of *M. guilliermondii* isolated from *P. perniciosus* (*M. guilliermondii* Pp) was compared with environmental *M. guilliermondii* ATCC 6260 and with *C. lusitaniae*, its closest relative on the phylogenetic tree. Species-specific and strain-specific cluster of orthologous genes were identified with OrthoMCL. Cluster of orthologs were manually curated to remove artifacts caused by repetitive elements. GO terms (Harris *et al.*, 2004) and putative domains were assigned to all three genomes with InterProScan5 (Jones *et al.*, 2014). The functional characterization of specific genes and genes shared between the three species was visualized using WEGO (Ye *et al.*, 2006). The genome of *M. guilliermondii* Pp was aligned to *M. guilliermondii* ATCC 6260 using

progressiveMauve (Darling *et al.*, 2010) to visualize large-scale evolutionary events such as rearrangements and inversions and to check the synteny of the two genomes.

Fungal uricolytic activity

The uricolytic activity of two strains of *R. mucilaginosa*, *W. anomalus*, *Trichosporon* sp. and *M. guilliermondii* isolated from sand flies in this work was tested on modified YPU medium (10 g/l yeast extract, 10 g/l peptone, 4 g/l uric acid (UA), 20 g/l agar). Yeast strains were maintained on YM agar until plating a single colony of each strain on YPU. After incubation at 26°C for 48 h, the degradation halo, suggesting uric acid utilization (Vera-Ponce de León *et al.*, 2016), was evaluated. The assay was performed three times for each tested isolate.

Bioinformatic analyses for prediction of uric acid degradation pathway of *M. guilliermondii*

Protein sequences of characterized Ascomycota uric acid catabolic enzymes were downloaded from Uniprot (UniProt Consortium, 2011). These proteins were queried against *M. guilliermondii* Pp predicted proteins using reciprocal BLASTP (Camacho *et al.*, 2009) searches.

Acknowledgments

The authors are grateful to Prof. Elena Crotti, Department of Food, Environmental and Evolutionary Sciences, University of Milan, Milan and Dr. Emanuela Clementi, Department of Biology and Biotechnology, University of Pavia, Pavia for their suggestions. This research was supported by the Italian Ministry of Education, University and Research FIR 2013 [RBFR136GFF (to SE)] and by the European Union Seventh Framework Programme ERC 2011 [FP7/2007-2013_FP7/2007-2011 under grant agreement no. 281222 (to IR)].

References

- Akhoundi, M., Bakhtiari, R., Guillard, T., Baghaei, A., Tolouei, R., Sereno, D., *et al.* (2012) Diversity of the bacterial and fungal microflora from the midgut and cuticle of phlebotomine sand flies collected in North-Western Iran. PLOS ONE 7(11): e50259.

- Alten, B., Maia, C., Afonso, M.O., Campino, L., Jiménez, M., González, E., *et al.* (2016) Seasonal Dynamics of Phlebotomine Sand Fly Species Proven Vectors of Mediterranean Leishmaniasis Caused by *Leishmania infantum*. PLOS Neglected Tropical Diseases 10(2): e0004458.
- Andrews, S. (2010) FastQC: A quality control tool for high throughput sequence data. URL <http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc>.
- Ba, A.S., Guo, D.A., Norton, R.A., Phillips, S.A., and Nes, W.D. (1995) Developmental differences in the sterol composition of *Solenopsis invicta*. Arch Insect Biochem 29: 1-9.
- Besemer, J., and Borodovsky, M. (2005) GeneMark: web software for gene finding in prokaryotes, eukaryotes and viruses. Nucleic Acids Res 33(suppl 2): W451-W454.
- Blackwell, M. (2017) Made for Each Other: Ascomycete Yeasts and Insects. Microbiol Spectr 5(3).
- Boeckmann, B., Bairoch, A., Apweiler, R., Blatter, M.C., Estreicher, A., Gasteiger, E., *et al.* (2003) The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003. Nucleic Acids Res 31(1): 365-370.
- Bozic, J., Capone, A., Pediconi, D., Mensah, P., Cappelli, A., Valzano, M., *et al.* (2017) Mosquitoes can harbour yeasts of clinical significance and contribute to their environmental dissemination. Environ Microbiol Rep doi: 10.1111/1758-2229.12569.
- Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., and Madden, T.L. (2009) BLAST+: architecture and applications. BMC Bioinformatics 10: 421.
- Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., *et al.* (2010) QIIME allows analysis of high-throughput community sequencing data. Nat Methods 7: 335-336.
- Cappelli, A., Ulissi, U., Valzano, M., Damiani, C., Epis, S., Gabrielli, M.G., *et al.* (2014) A *Wickerhamomyces anomalus* killer strain in the malaria vector *Anopheles stephensi*. PLoS One. 9(5): e95988.

- Castresana, J. (2000) Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol Biol Evol* 17(4): 540-552.
- Cochran, D.G., Mullins, D.E., and Mullins, K.J. (1979) Cytological changes in the fat body of the American cockroach, *Periplaneta americana*, in relation to dietary nitrogen levels. *Ann Entomol Soc Am* 72: 197-205.
- Damiani, C., Ricci, I., Crotti, E., Rossi, P., Rizzi, A., Scuppa, P., *et al.* (2008) Paternal transmission of symbiotic bacteria in malaria vectors. *Curr Biol* 18(23):R1087-1088.
- Darling, A.E., Mau, B., and Perna, N.T. (2010) ProgressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. *PLoS one* 5(6): e11147.
- Deshpande, V., Wang, Q., Greenfield, P., Charleston, M., Porrás-Alfaro, A., Kuske, C.R., *et al.* (2016) Fungal identification using a Bayesian classifier and the Warcup training set of internal transcribed spacer sequences. *Mycologia* 108(1): 1-5.
- Edgar, R.C. (2010) Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26(19): 2460-2461.
- Esquivel, C.J., Cassone, B.J., and Piermarini P.M. (2016) A de novo transcriptome of the Malpighian tubules in non-blood-fed and blood-fed Asian tiger mosquitoes *Aedes albopictus*: insights into diuresis, detoxification, and blood meal processing. *PeerJ* 4: e1784.
- Ferreira, N., Belloch, C., Querol, A., Manzanares, P., Vallez, S., and Santos, A. (2010) Yeast microflora isolated from Brazilian cassava roots: taxonomical classification based on molecular identification. *Curr Microbiol* 60: 287-293.
- Gatesoupe, F.J. (2007) Live yeasts in the gut: natural occurrence, dietary introduction, and their effects on fish health and development. *Aquaculture* 267: 20-30.
- Gibson, C.M., and Hunter, M.S. (2010) Extraordinarily widespread and fantastically complex: comparative biology of endosymbiotic bacterial and fungal mutualists of insects. *Ecol Lett* 13(2): 223-234.

Goujon, M., McWilliam, H., Li, W., Valentin, F., Squizzato, S., Paern, J., and Lopez, R. (2010) A new bioinformatics analysis tools framework at EMBL-EBI. *Nucleic Acids Res* 38 Suppl: W695-W699

Gusmão D.S., Santos, A.V., Marini, D.C., Bacci, M.Jr., Berbert-Molina, M.A., and Lemos, F.J. (2010) Culture-dependent and culture-independent characterization of microorganisms associated with *Aedes aegypti* (Diptera: Culicidae) (L.) and dynamics of bacterial colonization in the midgut. *Acta Trop* 115(3): 275-281.

Gusmão, D.S., Santos, A.V., Marini, D.C., Russo Ede, S., Peixoto, A.M., Bacci, M.Jr., *et al.* (2007) First isolation of microorganisms from the gut diverticulum of *Aedes aegypti* (Diptera: Culicidae): new perspectives for an insect-bacteria association. *Mem Inst Oswaldo Cruz* 102(8): 919-924.

Harris, M.A., Clark, J., Ireland, A., Lomax, J., Ashburner, M., Foulger, R., *et al.* (2004) The Gene Ontology (GO) database and informatics resource. *Nucleic Acids Res* 32(suppl 1): D258-D261.

Janson, E.M., Stireman, J.O., Singer, M.S., and Abbot, P. (2008) Phytophagous insect-microbe mutualisms and adaptive evolutionary diversification. *Evolution* 62(5): 997-1012.

Jones, P., Binns, D., Chang, H.Y., Fraser, M., Li, W., McAnulla, C., *et al.* (2014) InterProScan 5: genome-scale protein function classification. *Bioinformatics* 30(9): 1236-1240.

Kempf, V.A.J., Trebesius, K., and Autenrieth, I.B. (2000) Fluorescent in situ hybridization allows rapid identification of microorganisms in blood cultures. *J Clin Microbiol* 38: 830-838.

Killick-Kendrick, R. (1990) Phlebotomine vectors of the leishmaniases: a review. *Med Vet Entomol* 4(1): 1-24.

Kurtzman, C.P. (2014) Use of gene sequence analyses and genome comparisons for yeast systematics. *Int J Syst Evol Microbiol* 64.2: 325-332.

Kurtzman, C.P., and Fell, J.W. (2000) *The yeasts; a taxonomic study*, 4th edn. Amsterdam, Netherlands: Elsevier Science.

- Lee, I.R., Yang, L., Sebetso, G., Allen, R., Doan, T.H., Blundell, R., *et al.* (2013) Characterization of the complete uric acid degradation pathway in the fungal pathogen *Cryptococcus neoformans*. PLoS One 8(5): e64292.
- Léger, N., Pesson, B., Madulo-Leblond, G., and Abonnenc, E. (1983) Sur la différenciation des femelles du sous-genre *Larroussius* Nitzulescu, 1931 (Diptera-Phlebotomidae) de la région méditerranéenne. Ann Parasitol Hum Comp 58: 611–623.
- Li, L., Stoeckert, C.J., and Roos, D.S. (2003) OrthoMCL: identification of ortholog groups for eukaryotic genomes. Genome Res 13(9): 2178-2189.
- Loy, A., Arnold, R., Tischler, P., Rattei, T., Wagner, M., and Horn, M. (2008) ProbeCheck - a central resource for evaluating oligonucleotide probe coverage and specificity. Environ Microbiol 10: 2894-2898.
- Maroli, M., Fiorentino, S., and Guandalini, E. (1987) Biology of a laboratory colony of *Phlebotomus perniciosus* (Diptera: Psychodidae). J Med Entomol 24(5): 547-551.
- Martin, E., Bongiorno, G., Giovati, L., Montagna, M., Crotti, E., Damiani, C., *et al.* (2016) Isolation of a *Wickerhamomyces anomalus* yeast strain from the sandfly *Phlebotomus perniciosus*, displaying the killer phenotype. Med Vet Entomo 30(1):101-106.
- Modi, G.B., and Tesh, R.B. (1983) A simple technique for mass rearing *Lutzomyia longipalpis* and *Phlebotomus papatasi* (Diptera: Psychodidae) in the laboratory. J Med Entomol 20(5): 568-569.
- Molnar, O., Wuczkowski, M., and Prillinger, H. (2008) Yeast biodiversity in the guts of several pests on maize; comparison of three methods: classical isolation, cloning and DGGE. Mycol Prog 7(2): 111-123.
- Moran, N.A., and Dunbar, H.E. (2006) Sexual acquisition of beneficial symbionts in aphids. Proc Natl Acad Sci USA 103(34): 12803-12806.
- Nalepa, C.A., Bignell, D.E. and Bandi, C. (2001) Detritivory, coprophagy, and the evolution of digestive mutualisms in Dictyoptera. Insectes Sociaux 48(3): 194-201.

- Nardon, P., and Grenier, A.M. (1989) Endosymbiosis in Coleoptera: biological, biochemical, and genetic aspects. In *Insect Endocytobiosis: Morphology, Physiology, Genetics, Evolution*. Schwemmler, W., and Gassner, G. (eds). Boca Raton: CRC Press, pp. 175-216.
- Nguyen, N.H., Suh, S.O., and Blackwell, M. (2007) Five novel *Candida* species in insect-associated yeast clades isolated from *Neuroptera* and other insects. *Mycologia* 99(6): 842-858.
- Papon, N., Savini, V., Lanoue, A., Simkin, A.J., Crèche, J., Giglioli-Guivarc'h, N., *et al.* (2013) *Candida guilliermondii*: biotechnological applications, perspectives for biological control, emerging clinical importance and recent advances in genetics. *Curr Genet* 59(3): 73-90.
- Patiño-Navarrete, R., Piulachs, M.D., Belles, X., Moya, A., Latorre, A., and Peretó, J. (2014) The cockroach *Blattella germanica* obtains nitrogen from uric acid through a metabolic pathway shared with its bacterial endosymbiont. *Biol Lett* 10(7): 20140407.
- Ricci, I., Damiani, C., Scuppa, P., Mosca, M., Crotti, E., Rossi, P., *et al.* (2011a) The yeast *Wickerhamomyces anomalus* (*Pichia anomala*) inhabits the midgut and reproductive system of the Asian malaria vector *Anopheles stephensi*. *Environ Microbiol* 13(4): 911-921.
- Ricci, I., Mosca, M., Valzano, M., Damiani, C., Scuppa, P., Rossi, P., *et al.* (2011b) Different mosquito species host *Wickerhamomyces anomalus* (*Pichia anomala*): perspectives on vector-borne diseases symbiotic control. *Antonie Van Leeuwenhoek* 99(1): 43-50.
- Rice, P., Longden, I., and Bleasby, A. (2000) EMBOSS: The European Molecular Biology Open Software Suite. *Trends Genet* 16(6): 276-277
- Rivera, F.N., Gonzalez, E., Gomez, Z., Lopez, N., Hernandez-Rodriguez, C., Berkov, A., and Zuniga, G. (2009) Gut-associated yeast in bark beetles of the genus *Dendroctonus* Erichson (Coleoptera: Curculionidae: Scolytinae). *Biol J Linn Soc* 98(2): 325-342.
- Sant'Anna, M.R., Diaz-Albiter, H., Aguiar-Martins, K., Al Salem, W.S., Cavalcante, R.R., Dillon, V.M., *et al.* (2014) Colonisation resistance in the sand fly gut: *Leishmania* protects *Lutzomyia longipalpis* from bacterial infection. *Parasit Vectors* 23(7): 329.

- Schlein, Y., Polacheck, I., and Yuval, B. (1985) Mycoses, bacterial infections and antibacterial activity in sand flies (Psychodidae) and their possible role in the transmission of leishmaniasis. *Parasitology* 90: 57-66.
- Sievers, F., and Higgins, D.G. (2014) Clustal Omega, accurate alignment of very large numbers of sequences. *Methods Mol Biol* 1079: 105-116.
- Simpson, J.T., Wong, K., Jackman, S.D., Schein, J.E., Jones, S.J., and Birol, I. (2009) ABySS: a parallel assembler for short read sequence data. *Genome Res* 19(6): 1117-1123.
- Singh, S.R., and Hou, S.X. (2009) Multipotent stem cells in the Malpighian tubules of adult *Drosophila melanogaster*. *J Exp Biol* 212: 413-423.
- Spatafora, J.W., Sung, G.H., Johnson, D., Hesse, C., O'Rourke, B., Serdani, M., *et al.* (2006) A five-gene phylogeny of Pezizomycotina. *Mycologia* 98(6):1018-1028.
- Stamatakis, A. (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30(9): 1312-1313.
- Theodor, O. (1958) Psychodidae–Phlebotomine. In *Die Fliegen der Palaearktischen Region*, 9c. Lindner, E. (ed). Stuttgart: E. Schweizerbart'sche Verlagsbuchhandlung, pp.1–55.
- UniProt Consortium (2011) Reorganizing the protein space at the Universal Protein Resource (UniProt). *Nucleic Acids Res* 40: D71-D75.
- Valovage, W.D., and Brooks, M.A. (1979) Uric acid quantities in the fat body of normal and aposymbiotic German cockroaches, *Blattella germanica*. *Ann Entomol Soc Am* 72: 687-689.
- Varotto Boccazzi, I., Ottoboni, M., Martin, E., Comandatore, F., Vallone, L., Spranghers, T., Eeckhout, M., Mereghetti, V., Pinotti, L., Epis, S. (2017) A survey of the mycobiota associated with larvae of the black soldier fly (*Hermetia illucens*) reared for feed production. *PLOS ONE*: 12(8):e0182533.
- Vera-Ponce de León, A., Sanchez-Flores, A., Rosenblueth, M., and Martínez-Romero, E. (2016) Fungal community associated with *Dactylopius* (Hemiptera: Coccoidea: Dactylopiidae) and its role in uric acid metabolism. *Front Microbiol* 7: 954.

- Vishniac, H.S., and Johnson, D.T. (1990) Development of a yeast flora in the adult green June beetle (*Cotinis nitida*, Scarabaeidae). *Mycologia* 82: 471-479.
- Wang, Q., Garrity, G.M., Tiedje, J.M., and Cole, J.R. (2007) Naïve Bayesian Classifier for Rapid Assignment of rRNA Sequences into the New Bacterial Taxonomy. *Appl Environ Microbiol* 73(16): 5261-5267.
- Wang, Y., Gilbreath, T.M. III, Kukutla, P., Yan, G., and Xu, J. (2011) Dynamic Gut Microbiome across Life History of the Malaria Mosquito *Anopheles gambiae* in Kenya. *PLOS ONE* 6(9): e24767.
- Wang, L., Yue, L., Chi, Z., and Wang, X. (2008) Marine killer yeasts active against a yeast strain pathogenic to crab *Portunus trituberculatus*. *Dis Aquat Organ* 80(3): 211-218.
- White, T.J., Bruns, T., Lee, S., and Taylor, J. (1990). Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In *PCR Protocols: A guide to Methods and Applications*. Innis, M.A., Gelfand, D.H., Sninsky, J.J and White, T.J. (eds). San Diego, USA: Academic Press, pp. 315-322.
- Xufre, A., Albergaria, H., Inácio J., Spencer-Martins, I., and Gírio, F. (2006) Application of fluorescence in situ hybridisation (FISH) to the analysis of yeast population dynamics in winery and laboratory grape must fermentations. *Int J Food Microbiol* 108: 376-384.
- Ye, J., Fang, L., Zheng, H., Zhang, Y., Chen, J., Zhang, Z., *et al.* (2006) WEGO: a web tool for plotting GO annotations. *Nucleic Acids Res* 34(suppl 2): W293-W297.
- Zacchi, I., and Vaughn-Martini, A. (2002) Yeasts associated with insects in agricultural areas of Perugia, Italy. *Annu Rev Microbiol* 52: 237-244.
- Zindel, R., Gottlieb, Y., and Aebi, A. (2011) Arthropod symbioses: a neglected parameter in pest- and disease-control programmes. *J Appl Ecol* 48: 864-872.

Tables

Table 1. Proportion between positive specimens of *P. perniciosus* out of all tested samples

	Adult males		Adult females		L4 larvae
	Whole body n=10	Digestive system n=10; pool n=1	Whole body n=10	Digestive system n=10; pool n=1	Digestive system n=10
Specimens positive for yeast isolation	2	3; pool:1	4	6; pool:1	6

Table 2. Identification of yeast isolates associated to *P. perniciosus* specimens.

Yeast species	Adult males			Adult females			L4 larvae
	Whole body	Digestive system	Digestive system pool	Whole body	Digestive system	Digestive system pool	Digestive system
<i>Meyerozyma guilliermondii</i>	4	3	2	26	41	2	8
<i>Rhodotorula mucilaginosa</i>	0	0	0	7	0	0	0
<i>Trichosporon</i> sp.	0	0	0	0	0	0	4
<i>Wickerhamomyces anomalus</i>	2	0	0	10	0	0	3
Total yeast isolates	6	3	2	43	41	2	15

Table 3. 26S rRNA sequence analysis and homology with GenBank sequences

Yeast species	GenBank acc. no.	26S D1/D2 sequence comparison		
		D1/D2 26S	Ident (%)	Species (GenBank acc. no.)
<i>Meyerozyma guilliermondii</i>	LT905498		99%	KY463383
<i>Meyerozyma guilliermondii</i>	LT905499		100%	KY463383
<i>Rhodotorula mucilaginosa</i>	LT905496		99%	LC176999
<i>Rhodotorula mucilaginosa</i>	LT905497		100%	KY296083

<i>Trichosporon</i> sp.	LT905494	98%	AF444740
<i>Trichosporon</i> sp.	LT905495	99%	AF444740
<i>Wickerhamomyces anomalus</i>	LN871207	100%	LT594899
<i>Wickerhamomyces anomalus</i>	LN871208	100%	KY296073

Table 4. Results of the molecular screening for the detection of *M. guilliermondii* DNA from *P. perniciosus* specimens.

Location of capture/origin	Sex or sample type	Number of <i>M. guilliermondii</i> positive samples/ n. of analyzed samples
Frascati (RM)	Male	1/21
N 41,838810-EO 12,700712	Female	2/37
Roma	Male	1/8
N 41,936649-EO 12,367065	Female	0/2
SP 52 Vieste-Peschici (FG)	Male	4/21
N 41,901623-EO 16,138898	Female	1/11
Insectary of ISS (Rome, Italy)	Male	13/20
	Female	8/20
	L4 larvae	18/20
	Pupae	7/20

Table 5. Statistics of the obtained assembly compared with the ATCC 6260 strain genome.

Assembly	Total length	Number of contigs	Contig length average	N50	Number of genes	Number of non shared genes
Abyss-pe	10,666,172 bp	108	98,760.85 bp	232,015	5,476	21
ATCC 6260	10,609,954 bp	9	1,178,883.78 bp	1,701,016	5,396	4

Figure legends

Figure 1. Fungal diversity associated with *P. perniciosus* analyzed by ITS rRNA amplicon-based sequencing. Only taxa represented more than 2% were considered. At the left of histograms are indicated sample stage/sex. Relative abundance of class and genera for each sample stage/sex are shown by histograms.

Figure 2. Localization of *M. guilliermondii* complex in a *P. perniciosus* female. (a) Schematic view of digestive apparatus of sand fly. (b) Cellular nuclei stained for cell viability with DAPI. (c) Staining with the universal yeast probe PF2 labeled with CY3. (d) Staining with FITC labeled probe, specific for *M. guilliermondii* complex MGU26S410. (e) Overlap between the two probes. (f-i) Higher magnification of the Malpighian tubule using the same probes described in figures b to e.

Figure 3. Phylogenomic analysis of a strain of *M. guilliermondii* isolated from *P. perniciosus*. (a) Phylogenetic yeast tree; Saccharomycotina species highlighted in lilac, Pezizomycotina species highlighted in blue lavender. (b) Genomic alignment obtained with Mauve; colors highlight regions sharing sequence homology between the two genomes

Figure 4. Uric acid degradation of *M. guilliermondii*. (a) Fungal uricolytic activity of the yeasts *R. mucilaginosa*, (b) *M. guilliermondii*, (c) *W. anomalus* and (d) the control plate only with modified YPU medium. (e) Uric acid degradation predicted pathway of *M. guilliermondii* Pp isolated from *P. perniciosus*.

Figure S1. Venn diagram showing the fungal OTUs (at 97%) shared by larvae, males and females of sand flies.

Figure S2. Budding yeasts stained with Giemsa in the digestive system of *P. perniciosus* visualized at 100X magnification. Arrows indicate multiple budding yeast-like cells.

Figure S3. Specificity of the probe MGU26S410 tested against the yeast strains of *R. mucilaginosa*, *W. anomalus*, *Trichosporon* sp. and *M. guilliermondii*, isolated from sand flies.

Figure S4. Detection of *M. guilliermondii* in a *P. perniciosus* male stained with the universal yeast probe PF2 labeled with CY3 (b) and the FITC labeled probe MGU26S410, specific for *M. guilliermondii* complex (c). Bright field (a). Overlap of the universal yeast probes (red) and the *M. guilliermondii* complex-specific probe (green) (d).

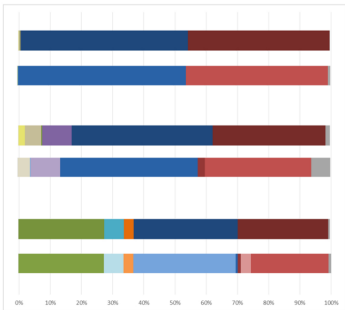
Figure S5. *M. guilliermondii* localization in a *P. perniciosus* larva after the hybridization with the universal yeast probe PF2 labeled with CY3 (b) and the FITC labeled probe MGU26S410, specific for *M. guilliermondii* complex (c). Bright field (a). Overlap of the universal yeast probes (red) and the *M. guilliermondii* complex-specific probe (green) (d).

Figure S6. *P. perniciosus* female sample, negative for *M. guilliermondii*. Staining with the universal yeast probe PF2 labeled with CY3 (a). Staining with FITC labeled probe, specific for *M. guilliermondii* complex MGU26S410 (b). Overlap between the two probes (c).

Figure S7. Negative controls of FISH staining. Hybridisation with the universal yeast probe PF2 and with the specific probe for *M. guilliermondii* complex MGU26S410 of a *P. perniciosus* specimen previously treated with Rnase (a). Hybridisation of a *P. perniciosus* specimen with the universal yeast probe PF2 (red) and with control probe non-MGU26S410 (b) .



♂

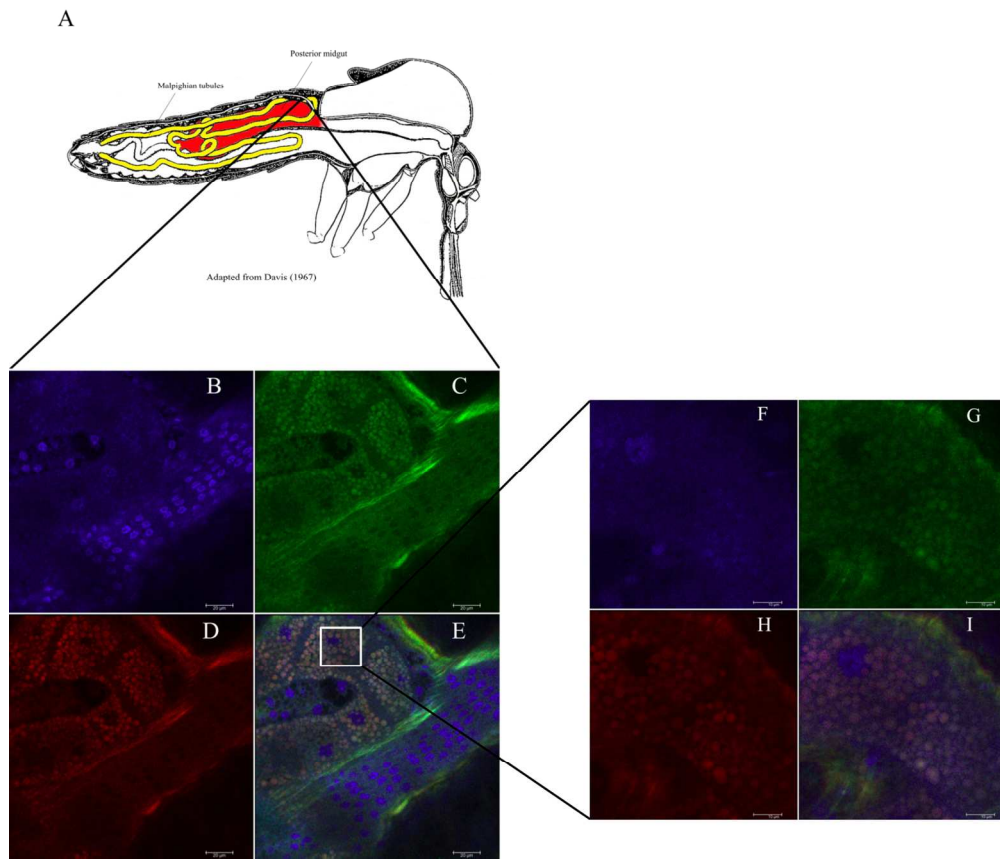


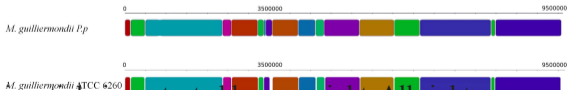
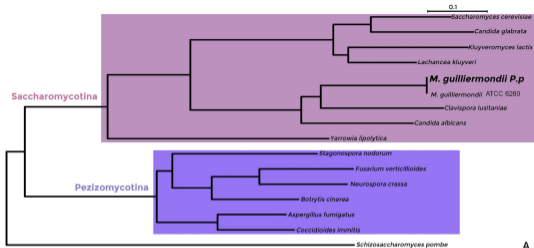
♀



Class Agaricomycetes Dothideomycetes Eurotiomycetes Leotiomycetes Mucoromycotina Orbiliomycetes Saccharomycetes Sordariomycetes Other
Pulsed-field gel electrophoresis of *Aspergillus fumigatus* and *Aspergillus nidulans* in the *Culex tritaeniorhynchus* mosquito
Meyerz, J. M., et al. (2015) *Journal of Invertebrate Pathology*, 144, 1-10
Pseudallescheria
Simplicillium

This article is protected by copyright. All rights reserved.





This article is protected by copyright. All rights reserved.

