Genomics and Pathophysiology: Dandruff as a Paradigm

9

Jun Xu, Teun Boekhout, Yvonne DeAngelis, Tom Dawson and Charles W. Saunders

Core Messages

› The recent sequencing of the genomes of dandruff-associated basidiomycetous yeasts, *Malassezia globosa* and *Malassezia restricta,* disclosed that the *M. globosa* genome is among the smallest for a free-living fungus. *M. globosa* produces a similar set of secreted hydrolases as the human pathogen *Candida albicans.* Although phylogenetically more closely related to the plant pathogen *Ustilago maydis*, *M. globosa* produces a different set of secreted hydrolases, which is a likely adaptation to the host niche and may be involved in pathogenicity. *M. globosa* is apparently missing several enzymes in fatty acid metabolism, including fatty acid synthase, Δ9 desaturase, and $\Delta^{2,3}$ enoyl CoA isomerase. The two former enzymes are apparently missing also in another skin microbe, *Corynebacterium jeikeium. M. globosa* has six lipase genes in each of two lipase families, which, compared with the lipases from a related fungus *U. maydis,* had undergone duplications since divergence from the *Ustilago*-containing lineage. There is also evidence for duplication of other *M. globosa* genes for secreted enzymes such as aspartyl proteases, phospholipases C, and acid sphingomyelinases. The *M. globosa* genome encodes proteins similar to all *Malassezia* allergens, the coding sequences of which have been isolated, and genes associated with mating, although mating has not yet been observed in *Malassezia.*

C. W. Saunders (✉)
Miami Valley Innovation Center, The Procter & Gamble Co, Cincinnati, OH 45253-8707, USA
e-mail: saunders.cw@pg.com

T. Boekhout et al. (eds.), *Malassezia and the Skin*
DOI: 10.1007/978-3-642-03616-3_9, © Springer Verlag Berlin Heidelberg 2010

9.1 Introduction

Malassezia and their role in dandruff and seborrheic dermatitis (SD) are of particular interest, as elucidation of pathogenetic mechanisms would instigate research into more effective treatments (Fig. 9.1).

Until recently, very little was known about the *Malassezia* genome. A recent report [1] on the genome sequences of *M. globosa* and *M. restricta* has brought new insights into the pathogenic capacity of these organisms. In this chapter, it is not our goal to repeat the results described in that paper and its 52-page supplementary material. Instead, we will provide additional perspective to the conclusions of that paper as well as some additional insights gleaned from the genome sequence.

Malassezia spp. are considered to have an essential role in dandruff and SD [2], although Koch's postulates have not been satisfied. The role of the fungus is complicated in that it is present essentially on all human scalps and yet many people lack symptoms, which is not unusual in diseases where microorganisms play a role. Disease-implicated microorganisms are present on many people without causing symptoms. Such microorganisms include *Staphylococcus aureus* [3], *Mycobacterium tuberculosis* [4], dermatophytes [5], and *C. albicans* [6]. In all of these cases, there are host attributes that determine whether the

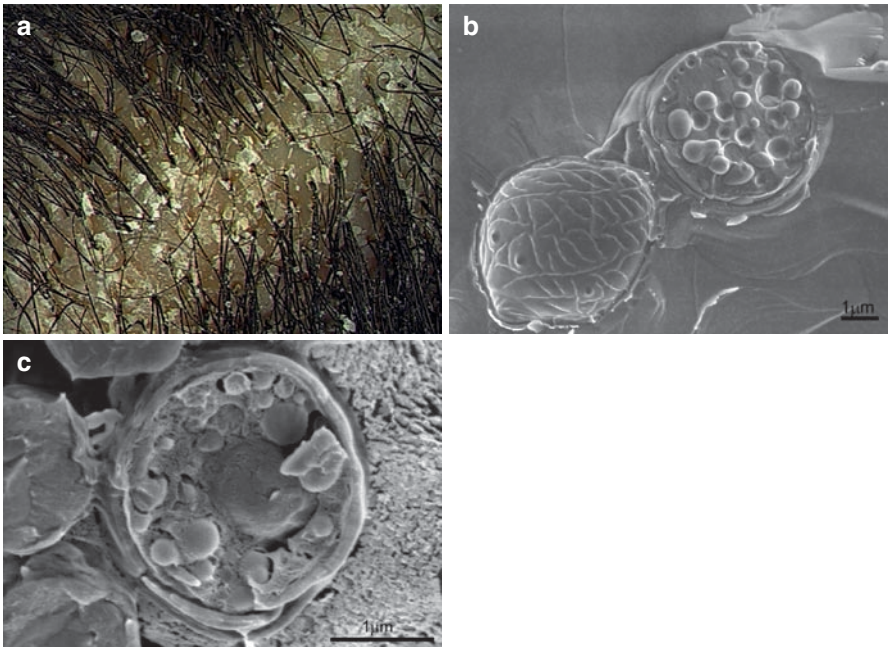

Fig. 9.1 Dandruff and *Malassezia*. (**a**) Scalp flaking known as dandruff. Note excessive flaking without overt irritation. Photo was taken with a Sony High-scope at 20× magnification. (**b**) Scanning electron micrograph of *Malassezia globosa* CBS strain 7966 after 14 days growth on LNA agar. Note the spiral structure in the cell wall (see also Chap. 2.3). (**c**) *Malassezia restricta* CBS strain 7877 grown under the same conditions as (**b**)

microorganism can cause disease. With dandruff, one host susceptibility component that has been described is sensitivity to oleic acid, which is a product of *Malassezia* lipase activity. Topically applied oleic acid can induce dandruff in dandruff-susceptible people only [7].

The role of *Malassezia* in dandruff and seborrheic dermatitis is supported by the efficacy of a variety of antifungal agents (i.e., pyrithione zinc, selenium sulfide, ciclopirox olamine, climbazole, and ketoconazole) [8] with no other obvious related biological activity. Upon treatment, the fungi are largely removed, followed by alleviation of symptoms [9]. With termination of treatment, the fungal numbers increase, and the symptoms often return [9]. While a different fungus may possibly contribute to these symptoms, no such fungus has been reported from the scalp. There are multiple hypotheses for the molecular mechanisms by which *Malassezia* cause dandruff and SD. One possibility is that lipase-mediated hydrolysis of sebum triglycerides leads to an excess of oleic acid that aggravates the scalp [10, 11]. Another possibility is that *Malassezia* produce tryptophan metabolites that are active on the aryl hydrocarbon receptor, thus producing inflammation [12]. An altered immune response has also been implicated (reviewed in Chap. 5).

Among the *Malassezia* species, *M. restricta* and *M. globosa* are the most commonly found on the scalp [13–16]. Genome-sequencing efforts have been focused on the *M. globosa* type strain CBS 7966, as the tRFLP pattern of its rDNA region matches that of the *M. globosa* strain(s) most commonly found on human scalps (Gemmer and Dawson unpublished observations). A sevenfold (7×) sequence coverage of the *M. globosa* genome and a one-fold (1×) *M. restricta* type strain CBS 7877 genome sequence coverage were generated. In this chapter, we focus on the *M. globosa* genome, with some discussion on what was elucidated through the incomplete sequencing of the *M. restricta* genome.

9.2 Genome Size

At 8.9 Mb, the *M. globosa* genome is among the smallest for free-living fungi (Table 9.1).

The microsporidian, *Encephalitozoon cuniculi*, now classified as a fungus, contains a much smaller genome (2.9 Mb), but is incapable of growth outside animal cells [17]. Another obligate intracellular pathogen, *Pneumocystis carinii,* has a genome of only 7.0 Mb, not including the telomeric ends and centromeres (the *Pneumocystis* genome proj-

Table 9.1 Some properties of the *M. globosa* nuclear genome

Size (bp)	8.9 M
G+C content (%)	52
Protein-coding gene number	4,285
Percent coding	69
Percent genes with introns	27
Average gene length (bp)	1,484
Average coding length (bp)	1,447
tRNA genes	82

ect. http://pgp.cchmc.org). The free-living yeast *Eremothecium (Ashbya) gossypii* has a 9.2-Mb genome [18], a size similar to that of the *M. globosa* genome. Other sequenced yeast and fungal genomes are larger, with the largest at 80 Mb [1]. While most characterized bacterial genomes are smaller, some are larger, such as the 13 Mb *Sorangium cellulosum* [19] genome and the 9.7 Mb genome of *Rhodococcus* RHA1 [20]. The relatively small size of the *M. globosa* genome can be attributed to the small number of protein-coding genes ($n = 4{,}285$), the few introns (present in 27% of genes), and the few repetitive elements (comprising 0.78% of the genome). The number of *M. globosa* protein-coding genes is smaller than that of all other free-living fungi the genome of which has been sequenced. Among those bacteria the genome of which has been sequenced, most encode fewer genes than *M. globosa*. However, a larger number of proteins are encoded by some bacteria, such as *Streptomyces coelicolor* A3 (2) which encodes 7,825 proteins [21].

Small genomes are characteristic of organisms with specialized niches, with obligate intracellular parasites at the extreme. The genome size of *M. globosa*, relative to the genome size of its yeast and fungal relatives, suggests that *M. globosa* may have limited capacity to survive in diverse environments although this capacity is not as limited as in the intracellular microsporidia. The notion of a limited niche for *M. globosa* is consistent with the observation that *Malassezia* yeasts are found on the skin of warm-blooded animals but rarely in any other environment. This limited niche is reflected in the requirements for suitable culture media for growth *in vitro* (as described in Chap. 2.1).

9.3 Phylogenetic Relationship with Other Fungi

In a previous work using the sequence similarities of ribosomal RNA-encoding genes, *Malassezia* had been placed in a phylogenetic tree adjacent to plant pathogens such as *Ustilago maydis* [22, 23]. Further analysis, using a set of protein-coding genes [1], confirmed this placement. Most (75%) of the *M. globosa* genes show *U. maydis* genes as their best matches in the NCBI NR database (excluding the *M. restricta* sequences). Some of these plant pathogens, namely the smut fungi, have had a large impact through the ages on human food. Further study of these fungi should yield novel insights into the mechanisms by which these microorganisms interact with their hosts.

9.4 Lipid and Amino Acid Metabolism

Malassezia globosa requires lipid for growth in culture [24], an observation explained by the lack of a fatty acid synthase gene within the genome [1]. Most other characterized *Malassezia* spp., including *M. restricta*, also require lipid supplementation for growth. Genome sequencing suggests that *M. restricta* too, lacks a fatty acid synthase gene, as fragments of such a large gene would have been detected even with the 1× sequence cover-

age of that genome. Among free-living fungi with sequenced genomes, only these *Malassezia* species lack a fatty acid synthase gene. By contrast, fatty acid synthase genes are lacking in *Entamoeba histolytica* [25], *Encephalitozoon cuniculi* [17], and *Corynebacterium jeikeium* K411 [26].

In addition to its inability to synthesize fatty acids, *M. globosa* apparently lacks a Δ9 desaturase (EC 1.14.19.2) gene, the product of which places a double bond into common fatty acids, such as oleic acid. A Δ9 desaturase gene was not detected in the *M. restricta* genome either [1]. As Δ9 fatty acids, such as oleic acid, are commonly found on human skin [27], it is likely that the human host is satisfying *M. globosa* needs for such unsaturated fatty acids. In contrast, the genome of the close relative, *U. maydis* contains a Δ9 desaturase gene (UM00955.1).

The *M. globosa* genome encodes the enzymes capable of degrading saturated fatty acids via β-oxidation. However, there is an intriguing limitation. With fatty acids containing a *cis*-double bond, this must be modified to a *trans*-double bond before oxidizing that region of the fatty acid. Furthermore, if the double bond occurs at an odd-numbered carbon, then the double bond must be shifted to an even-numbered carbon. Both of these modifications must happen for β-oxidation of the common skin lipid, oleic acid. The *Saccharomyces cerevisiae ECI1*product, a $\Delta^{3,2}$-enoyl CoA isomerase (EC 5.3.3.8), catalyzes both reactions, and *ECI1*mutant *S. cerevisiae* strains are limited in growth on unsaturated fatty acids [28]. The *M. globosa* genome shows no indication of an *ECI1*-like gene, suggesting that *M. globosa* either (1) has limited capacity for β-oxidation of unsaturated fatty acids, (2) contains a $\Delta^{3,2}$ enoyl CoA isomerase that is different from *ECI1* and not found by similarity searches, or (3) uses a different biochemical pathway for unsaturated fatty acid oxidation. In addition, no evidence for an *ECI1*homolog was found in the *M. restricta* genome sequences. In contrast, a *S. cerevisiae ECI1*homolog (UM01599.1) was found in *U. maydis.* The most likely explanation is that, *M. globosa* and *M. restricta* have lost the *ECI1* and Δ9 desaturase gene homologs since *Malassezia* spp. last shared a common ancestor with *U. maydis,* although further biochemical experiments and genomic analysis will be needed to test this hypothesis. In that respect, it would be interesting to know which of these lipid metabolic enzymes are encoded by *M. pachydermatis*, the one known *Malassezia* species that, while lipophilic, does not require lipid for growth. Also, a comparison of the lipid metabolic capacity of *Cor. jeikeium* and *M. globosa* would be particularly interesting, as both organisms live on skin, lack a fatty acid synthase gene, and depend on exogenous lipids for growth in the laboratory. However, unlike *M. globosa*, *Cor. jeikeium* is probably capable of complete oxidation of oleic acid. The $\Delta^{3,2}$ enoyl CoA isomerase, apparently missing from *M. globosa,* is likely part of a multi-enzyme complex encoded by a homolog of *E. coli fadB* [26, 29].

Corynebacterium jeikeium may share with *M. globosa* a defect in oleic acid synthesis. There are two bacterial pathways for generation of the double bond in oleic acid. One is based on a Δ9 desaturase [30], and the second is based on 3-hydroxydecanoyl-ACP-dehydrogenase [31] that acts during fatty acid elongation. The latter process would not be expected in *Co. jeikeium* because of the absence of fatty acid synthesis in this organism. It is noteworthy that there was no evidence for a *Cor. jeikeium* homolog of a Δ9 desaturase gene, when the *S. cerevisiae OLE1* or the *Mycobacterium tuberculosis desA3* was used in

a similarity search [30], although a *desA3* homolog in the genome of the nonlipophilic *Cor. diptheriae* was identified. In addition, no evidence was found for a *Cor. jeikeium* homolog of 3- hydroxydecanoyl-ACP dehydrogenases, as in *E. coli fabA* or *fabZ*.

In contrast to lipid metabolism, where *M. globosa* takes advantage of the host provision of fatty acid sources, *M. globosa* is apparently self reliant for amino acids. The *M. globosa* genome appears to encode all of the enzymes necessary for synthesis of all twenty canonical amino acids. It was expected that amino acids should be available to the yeast through the skin, because the stratum corneum contains amino acids as a result of proteolysis of filaggrin [32]. Besides the lipid metabolic pathways, *Cor. jeikeium* also resembles *M. globosa* in its amino acid metabolic pathways, in that both organisms can apparently synthesize the twenty standard amino acids [26]. It would be interesting to know if the amino acid synthesis genes are transcriptionally active on the skin and whether there is a selective advantage for these microorganisms to synthesize amino acids.

9.5 Lipases

Lipases and phospholipase C may be used by *Malassezia* to generate fatty acids from sebum triglycerides, therefore compensating for the lack of fatty acid synthase. Beyond this nutritional role, lipases may contribute to dandruff symptoms (see above). A large number of gene copies are often interpreted as a sign that the corresponding proteins are of particular significance. Among enzymes encoded by the *M. globosa* genome, lipases and aspartyl proteases are represented by the highest gene copy number. The only characterized *M. globosa* lipase, the LIP1 product, hydrolyzes diglycerides and monoglycerides [10]. However, *M. globosa* cells contain a triglyceride-hydrolyzing activity, so other lipases must be active against triglycerides.

Malassezia globosa encodes twelve lipases that can be partitioned equally into either of two PFAM families and are depicted on a phylogenetic tree of lipases (Fig. 9.2). Our research group constructed the tree following a BLAST search in the NCBI NR database, using the amino acid sequence of each of these twelve lipases. Every lipase, that was one of the top fifteen matches to any of the *M. globosa* lipases, as well as any *U. maydis* and *M. restricta* lipases were included in the tree. Of course, many lipases were among the top matches to multiple *M. globosa* lipases. Although partial sequencing of the *M. restricta* genome [1] resulted in many gene fragments, it was, in four cases, possible to assemble sufficient lipase sequences to place these in the phylogenetic tree. However, some *M. restricta* lipases may be missing.

Among the lipases, one family is PFAM category Lipase 3 (PF01764). *U. maydis* encodes one lipase within this family, whereas *M. globosa* encodes six lipases in the family. As described previously [1], the *M. globosa* genome contains clusters of genes for secreted enzymes. One such cluster contains four consecutive lipase genes (MGL_0797 through MGL_0800), each potentially transcribed in the same direction. These lipases are more similar to each other than to any other protein. It is likely that this cluster arose from lipase gene duplication, after the separation of *Malassezia* and *Ustilago* from a common

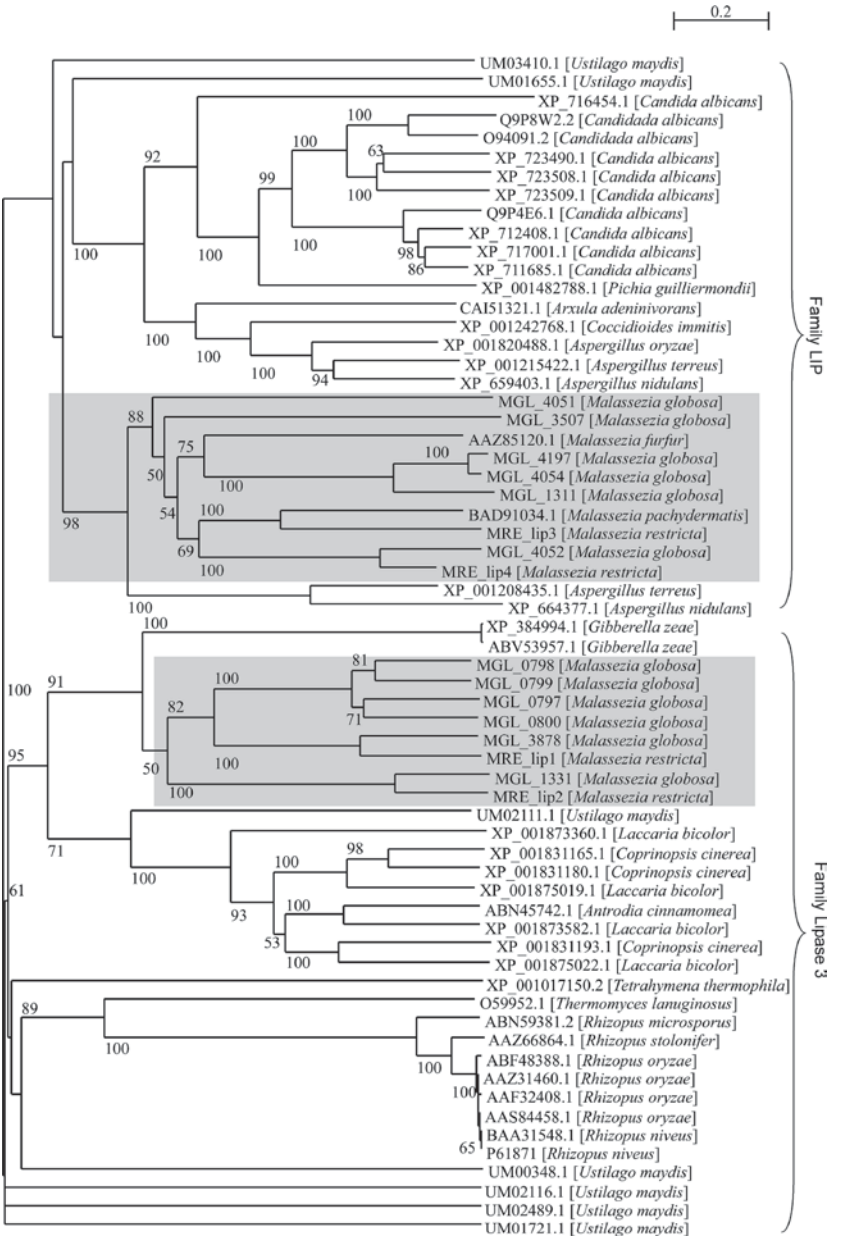

Fig. 9.2 Phylogenetic tree of lipases. To construct the tree, the protein sequences of twelve *M. globosa* lipases that belong to PFAM families LIP and Lipase_3 were searched against the NCBI NR database. The top fifteen matches to any of the *M. globosa* lipases were combined and duplicates removed. Partial open reading frames of putative *M. restricta* lipases were also included. A multiple sequence alignment was generated using MAFFT [33] L-INS-I model. The tree was constructed using PHYML as described previously [1]. The numbers shown on the tree are bootstrap values derived from 100 bootstrap datasets (only those that have a bootstrap value greater or equal than 50 were shown)

ancestor. The second family is the PFAM category LIP (PF03583), with one division containing well-characterized lipases from *Candida albicans* and one lipase from *U. maydis*. A second division contains lipases described from *M. furfur* [34] and *M. pachydermatis* [35], two lipases from *Aspergillus*, and six lipases from *M. globosa*. It is not clear if there is an *Ustilago* lipase that is homologous to the *Malassezia* lipases in this division. Three of these *M. globosa* lipases are encoded by a secreted enzyme gene cluster, with an aspartyl protease gene interspersed [1]. Unlike the case with the set of PFAM Lipase 3 lipase genes, the gene cluster of PFAM LIP members is not a set of most similar genes, making interpretation of its history more speculative. Although members of the LIP family clearly belong to the same family, they are not very similar to each other, with a protein sequence identity of about 30–45%. The nucleotide sequence has no apparent similarity to each other, which may be the reason why Brunke and Hube [34] were not able to identify additional lipase genes when using a *M. furfur* lipase gene as a probe in a southern blot analysis with the genomic DNA from seven species of *Malassezia*, although two hybridizing *M. pachydermatis* genome segments were detected. It is interesting that one *M. globosa* lipase, MGL_3507, appears to be about 110 amino acid residues longer at the C-terminus than the rest of the class. The extra coding sequence was confirmed by cDNA sequencing (DeAngelis and Dawson, unpublished observation). This extra region contains some low complexity amino acid sequence and is rich in serine, lysine and glycine. In the absence of experimental evidence and homology with other known protein sequences, we speculate that this region might be used to tether the lipase to cell wall.

Malassezia restricta has been reported to lack lipase activity in culture [10, 13]. However, the incomplete *M. restricta* genome sequence indicates the presence of at least two lipases in each of the LIP and the lipase 3 families, one of which was found among *M. restricta* secreted and cell wall-associated proteins [1]. The difficulty in detecting *M. restricta* lipase activity may be due to a combination of factors including culture optimization for lipase production and the choice of the appropriate lipase assay.

9.6 Aspartyl Proteases

Malassezia globosa encodes seventeen aspartyl protein genes, an uncommonly large number [1] (Fig. 9.3). Proteases could hydrolyze host proteins to supply nutrients, degrade host tissues, modify host cells to facilitate adhesion, or alter the immune response.

A phylogenetic tree of aspartyl proteases was created, using the methodology described for lipases; the only difference being that we have included in the tree, the top ten matches for each of the *M. globosa* aspartyl proteases. In one case, neighboring genes encoded aspartyl proteases with the highest similarity (MGL 3330 and 3331). In other cases, two *M. globosa* protease genes, despite being highly similar to each other, were not adjacent in the genome. As was the case with lipase genes, this pattern suggests gene duplication events more recent than the split from the common ancestor shared with *Ustilago*.

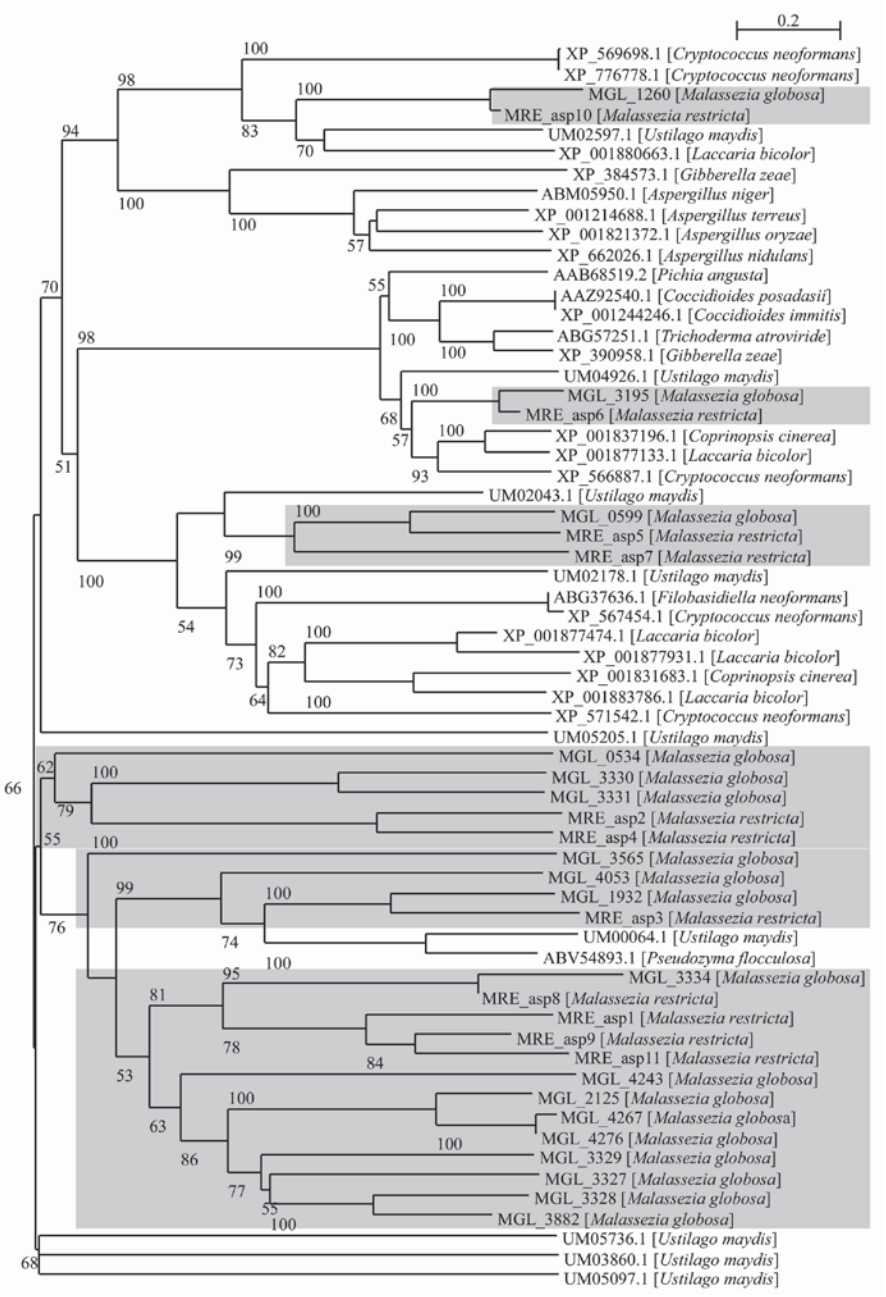

Fig. 9.3 Phylogenetic tree of aspartyl proteases. The tree was constructed using the same method as described in Fig. 9.2 except that only the top ten hits from the NCBI NR databases were included

9.7 Phospholipase C

We constructed a phylogenetic tree based on the twenty proteins that best matched each of the *M. globosa* phospholipases C. One branch of a tree contains only *Malassezia* genes, with six from *M. globosa* and five from *M. restricta.* The closest relatives are a set of ascomycete phospholipases, with a more distant set of bacterial phospholipases.

There are no basidiomycete phospholipases related closely enough to be shown (Fig. 9.4). Three of the *M. globosa* phospholipase genes are found adjacent to each other, and these three are each other's closest relatives within *M. globosa.* Each of these phospholipases has for its most similar protein a *M. restricta* phospholipase C. Our *M. restricta* genome sequence was not sufficiently complete to know if these genes are adjacent in the genome.

9.8 Acid Sphingomyelinases

A phylogenetic tree (not shown) indicates that one branch contains three adjacent *M. globosa* acid sphingomyelinase genes (MGL_1573, 1574, and 1575), as well as a fourth acid sphingomyelinase gene (MGL_568) and three *M. restricta* genes (MRE asm1, asm2, and asm3). Most similar to these *Malassezia* acid sphingomyelinase genes is a *U. maydis* gene, XP_758084.1. This suggests that the *Malassezia* lineage has undergone acid sphingomyelinase gene duplications since sharing a common ancestor with the *Ustilago* lineage.

9.9 Allergens

Multiple *Malassezia* proteins have been implicated as allergens contributing to atopic eczema [36]. Many of the putative allergens are produced in culture by *M. globosa*, as five (homologs of Mala s 1, f 2, s 6, s 8, s 9) were detected among secreted and cell wall-associated proteins [1]. Several cDNAs have been isolated that encode *Malassezia* proteins that react with human IgE [37–43]. Andersson et al. [44] used PCR to search for nine of the allergen gene homologs among seven *Malassezia* species, including *M. globosa*. Using *M. globosa* genomic DNA, they found PCR products matching Mala s 1, f 4, s 5, s 6, s 7, s 8, and s 9, but not Mala f 2 and f 3. With *M. globosa* mRNA, they only found the Mala s 6 PCR product.

Based on genome sequence, *M. globosa* encodes proteins similar to all of the putative *Malassezia* allergens. However, there is no simple one-to-one relationship among all of these potential allergens (Table 9.2). For example, Mala s 7 shows a weak match to two *M. globosa* proteins, MGL_0968 and MGL_2673. A single protein from *M. globosa,* MGL_4042, shows similarity to three previously described allergens, Mala f 2, f 3 and s 5, with the highest similarity to Mala f 2 (Table 9.2).

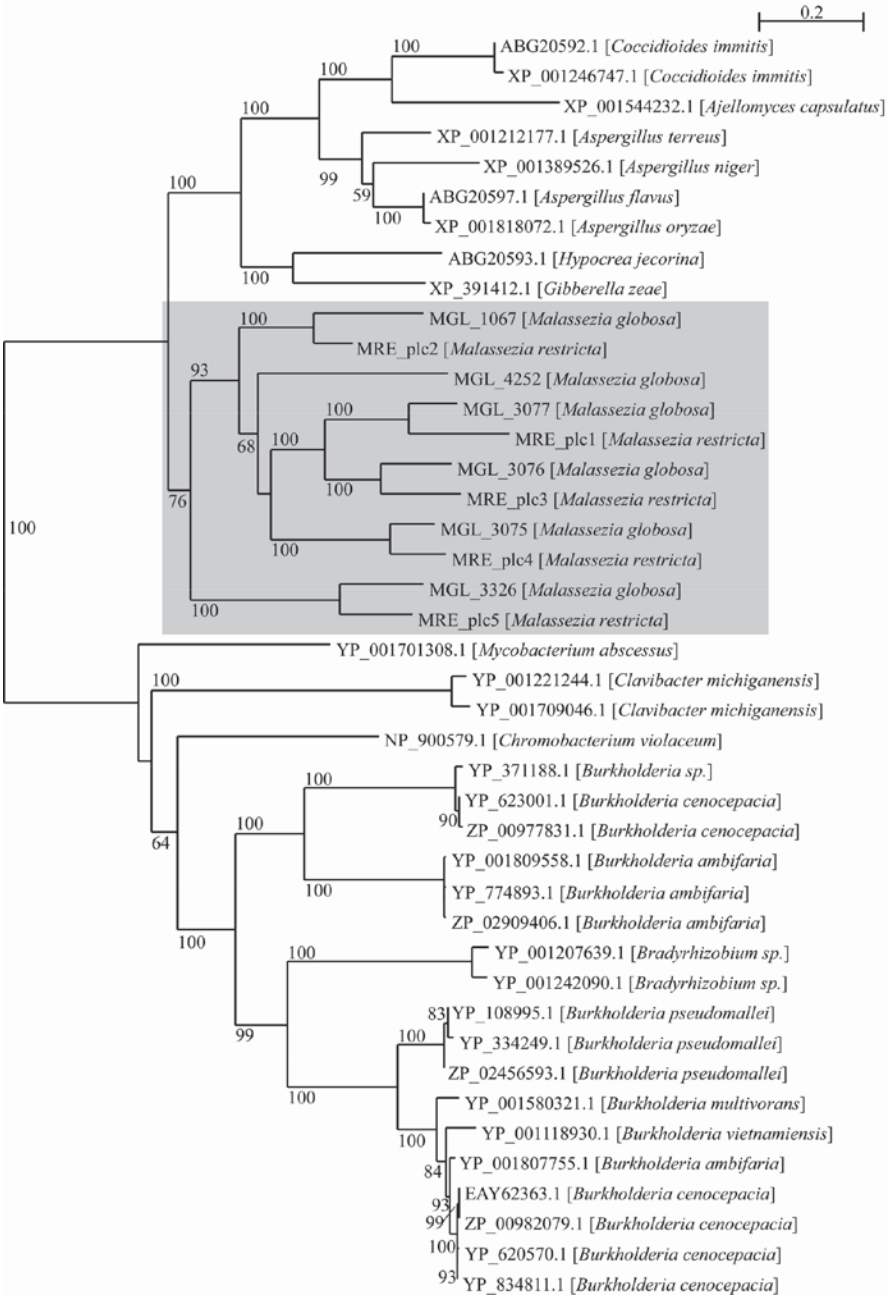

Fig. 9.4 Phylogenetic tree of phospholipases C. The tree was constructed using the same method as described in Fig. 9.2 except that only the top twenty hits from the NCBI NR databases were included

Table 9.2 Match of *M. globosa* proteins to *Malassezia* allergens described in the literature

Accession No.	Best matched *M. globosa* gene	Name	Blast bit-score	In *M. globosa* EST sequence database
AB011804	MGL_4042	Mala f 2	288	–
AB011805	MGL_4042	Mala f 3	199	–
AF084828	MGL_2703	Mala f 4	517	X
X96486	MGL_1303	Mala s 1	497	X
AJ011955	MGL_4042	Mala s 5	211	–
AJ011956	MGL_3612	Mala s 6	297	–
AJ011957	MGL_0968	Mala s 7	48.1	–
AJ011958	MGL_1304	Mala s 8	250	–
AJ011959	MGL_2179	Mala s 9	342	X
AJ428052	MGL_0201	Mala s 10	1,328	X
AJ548421	MGL_3190	Mala s 11	269	X
AJ871960	MGL_0750	Mala s 12	802	X
AJ937746	MGL_1781	Mala s 13	192	–

The accession number is the nucleotide sequence accession number. The bit-score is using the translated amino acid sequence of the allergens against the proteome of *M. globosa.* The match of Mala s 7 has a very low bit-score and only matches partially, but it is the best match to Mala s 7 in the NCBI nr database other than Mala s 7 itself

These three allergens were known to be similar [38] and thought to encode a peroxiredoxin. Similar genes are found in many other fungal species, including *U. maydis*, *Coprinopsis cinerea*, *Yarrowia lipolytica*, *Paracoccidioides brasiliensis*, *Laccaria bicolor*, *Aspergillus terreus*, and *Ajellomyces capsulatus*. The absence of the Mala f 2 and f 3 products in the PCR reactions of Andersson et al. [44] may be explained by the lack of a good match to MGL_4042 by the primers used for PCR.

There is some controversy over the presence of a Mala s 1 homolog in *Malassezia* other than *M. sympodialis.* Andersson et al. [44] found this homolog in four *Malassezia* species, including *M. globosa* 7966, whereas Gaitanis *et al.* [45] did not find a Mal s 1 homolog in several *Malassezia* species, including *M. globosa* 7966. The two groups collaborated on a letter [46] suggesting that choice of primers may have contributed to the discrepancy and that the Mala s 1 homologs may not closely match each other. Our alignment of Mala s 1 with its *M. globosa* homolog, MGL_1303, indicates that the two genes share about 70% identity in amino acid sequence. However, the nucleotide sequences are not highly conserved, and less than 1/5 of the nucleotide sequence can be aligned. We were not able to find a convincing match using the two sets of primers used by Gaitanis et al. [45] and Andersson et al. [44]. The lack of a good match to the PCR primers likely explains the difficulty in obtaining a PCR product for Mala s 1 in *M. globosa*. It is interesting that the *M. globosa* MGL_1303 gene (Mala s 1 homolog) and the MGL_1304 gene (Mala s 8 homolog) are located near each other in the chromosome, separated by only 605 base pairs.

In summary, genomic and proteomic analysis indicates that *M. globosa* encodes proteins similar to the allergens that have been identified in *M. furfur* and *M. sympodialis.*

Zargari et al. [47] have used Western blots to show that human IgE recognizes many proteins from *M. globosa*, and recognition of only some of these proteins is inhibited by competition with *M. sympodialis* protein. This implies that *M. globosa* produces many IgE-reactive epitopes that are not produced by *M. sympodialis.* Therefore, there may still be other *M. globosa* proteins that react with IgE and have not yet been characterized. More research is required to understand the role of these potential allergens in human disease.

9.10 Tryptophan Metabolites

In addition to the secreted enzymes that are likely to interact with the host, *Malassezia* also produce small molecules that may interact with the host. *Malassezia* are reported to produce tryptophan metabolites with a role in SD [12] and pityriasis versiolor [48]. Zuther et al. [48] used *U. maydis* as a model organism and demonstrated that this plant pathogen produced a similar set of tryptophan metabolites as does *M. furfur.* This analysis led to the discovery of two genes with a role in pigment production. The tryptophan aminotransferase Tam1p catalyzes the formation of indole pyruvate, which reacts to form pigments. We observed that *M. globosa* encodes a *TAM1* homolog, MGL_2601. Zuther et al. [48] also found that mutations in the sulphite reductase gene, *SIR1,* led to defective pigment production, *M. globosa* contains a *SIR1* homolog, MGL_0080.

9.11 Mating

The *M. globosa* genome contains a mating locus, with pheromone receptor and bW and bE homeodomain genes, with some similarity to other basidiomycete mating-type loci [49]. This suggests that *M. globosa* may be capable of undergoing a sexual cycle. There are precedents for genome analysis providing the first indications of sexuality among fungi. With *C. albicans,* the discovery of mating- type genes was followed by the discovery of mating [50]. *Aspergillus oryzae* and *A. fumigatus* were thought to be asexual until the genome sequence indicated the presence of many genes associated with mating [51]. Perhaps, the existence of mating genes within the *M. globosa* genome will prompt the discovery of mating in these fungi. Mating can be important in the spread of virulence traits [52] and may be important in the distribution of pathogenic strains of *Malassezia.*

9.12 Adaptation to Animal Skin

As pointed out by Xu et al. [1], *M. globosa* and *C. albicans* are capable of living on animal skin and secrete a similar set of enzymes; lipases, aspartyl proteases, phospholipases, and

acid sphingomyelinases. In contrast, *U. maydis,* more closely related phylogenetically to *M. globosa,* secretes a different set of enzymes, including many glycosyl hydrolases. We propose that the sets of secreted enzymes by *M. globosa* and *C. albicans* represent an adaptation to life on animal skin. In this chapter, we have extended the previous observations by generating phylogenetic trees of the lipases, aspartyl proteases, phospholipases C, and acid sphingomyelinases, providing for the first time, evidence that *Malassezia* secreted enzyme genes have undergone duplications since sharing a common ancestor with *Ustilago.* As additional genomes of related organisms become available, it will be interesting to analyze the apparent losses and duplications of genes for secreted enzymes.

9.13 Conclusion

The *M. globosa* genome reveals duplications of genes for secreted enzymes, including lipases, aspartyl proteases, phospholipases C, and acid sphingomyelinases. These duplications likely arose after the divergence of *Malassezia* from its ancestors, shared with the plant pathogen *Ustilago.* These enzymes are also encoded by multiple genes in *C. albicans*, a human opportunistic pathogen found on skin, and it is therefore likely that these multicopy genes are a result of adaptation to animal skin.

Analysis of the *M. globosa* and *M. restricta* genomes discloses multiple pathways with potential roles in the pathogenesis of dandruff and seborrheic dermatitis, providing new intervention targets for the development of novel, more effective treatments.

Acknowledgments We thank the coauthors of our genome and proteome analysis publication [1] as well as all the individuals acknowledged in that publication. We especially highlight our appreciation to Pamela Trotter for her insights into lipid metabolism. We thank Julie Baker for the electron micrographs and Holly Krigbaum for the dandruff photograph.

References

1. Xu J, Saunders CW, Hu P et al (2007) Dandruff-associated *Malassezia* genomes reveal convergent and divergent virulence traits shared with plant and human fungal pathogens. Proc Natl Acad Sci USA 104:18730–18735
2. Ashbee HR, Evans EG (2002) Immunology of diseases associated with *Malassezia* species. Clin Microbiol Rev 15:21–57
3. Feng Y, Chen CJ, Su LH et al (2008) Evolution and pathogenesis of *Staphylococcus aureus*: lessons learned from genotyping and comparative genomics. FEMS Microbiol Rev 32:23–37
4. Zahrt TC (2003) Molecular mechanisms regulating persistent *Mycobacterium tuberculosis* infection. Microbes Infect 5:159–167
5. White TC, Oliver BG, Gräser Y et al (2008) Generating and testing molecular hypotheses in the dermatophytes. Eukaryot Cell 7:1238–1245
6. Calderone RA, Fonzi WA (2001) Virulence factors of *Candida albicans*. Trends Microbiol 9:327–335

7. Ro BI, Dawson TL (2005) The role of sebaceous gland activity and scalp microfloral metabolism in the etiology of seborrheic dermatitis and dandruff. J Investig Dermatol Symp Proc 10:194–197
8. Schwartz JR, Cardin CW, Dawson TL (2004) Dandruff and seborrheic dermatitis. In: Baran R, Maibach HI (eds) Textbook of cosmetic dermatology. Martin Dunitz, London, UK
9. Crespo Erchiga V, Delgado Florencio V (2002) *Malassezia* species in skin diseases. Curr Opin Infect Dis 15:133–142
10. DeAngelis YM, Saunders CW, Johnstone KR et al (2007) Isolation and expression of a *Malassezia globosa* Lipase Gene, LIP1. J Invest Dermatol 127:2138–2146
11. Troller JA (1971) Model system for the investigation of dandruff. J Soc Cosmet Chem 22:187–198
12. Gaitanis G, Magiatis P, Stathopoulou K et al (2008) AhR ligands, malassezin, and indolo[3, 2-b]carbazole are selectively produced by *Malassezia furfur* strains isolated from seborrheic dermatitis. J Invest Dermatol 128:1620–1625
13. Aspiroz C, Moreno LA, Rezusta A et al (1999) Differentiation of three biotypes of *Malassezia* species on human normal skin. correspondence with *M. globosa*, *M. sympodialis* and *M. restricta*. Mycopathologia 145:69–74
14. Gemmer CM, DeAngelis YM, Theelen B et al (2002) Fast, noninvasive method for molecular detection and differentiation of *Malassezia* yeast species on human skin and application of the method to dandruff microbiology. J Clin Microbiol 40:3350–3357
15. Midgley G (2000) The lipophilic yeasts: state of the art and prospects. Med Mycol 38(Suppl 1): 9–16
16. Paulino LC, Tseng CH, Blaser MJ (2008) Analysis of *Malassezia* microbiota in healthy superficial human skin and in psoriatic lesions by multiplex real-time PCR. FEMS Yeast Res 8: 460–471
17. Katinka MD, Duprat S, Cornillot E et al (2001) Genome sequence and gene compaction of the eukaryote parasite *Encephalitozoon cuniculi*. Nature 414:450–453
18. Dietrich FS, Voegeli S, Brachat S et al (2004) The *Ashbya gossypii* genome as a tool for mapping the ancient *Saccharomyces cerevisia*e genome. Science 304:304–307
19. Schneiker S, Perlova O, Kaiser O et al (2007) Complete genome sequence of the myxobacterium *Sorangium cellulosum*. Nat Biotechnol 25:1281–1289
20. McLeod MP, Warren RL, Hsiao WW et al (2006) The complete genome of *Rhodococcus* sp. RHA1 provides insights into a catabolic powerhouse. Proc Natl Acad Sci USA 103: 15582–15587
21. Bentley SD, Chater KF, Cerdeno-Tarraga AM et al (2002) Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). Nature 417:141–147
22. Begerow D, Bauer R, Boekhout T (2000) Phylogenetic placements of ustilaginomycetous anamorphs as deduced from nuclear LSU rDNA sequences. Mycol Res 104:53–60
23. Begerow D, Stoll M, Bauer R (2006) A phylogenetic hypothesis of Ustilaginomycotina based on multiple gene analyses and morphological data. Mycologia 98:906–916
24. Chen TA, Hill PB (2005) The biology of *Malassezia* organisms and their ability to induce immune responses and skin disease. Vet Dermatol 16:4–26
25. Loftus B, Anderson I, Davies R et al (2005) The genome of the protist parasite *Entamoeba histolytica*. Nature 433:865–868
26. Tauch A, Kaiser O, Hain T et al (2005) Complete genome sequence and analysis of the multiresistant nosocomial pathogen *Corynebacterium jeikeium* K411, a lipid-requiring bacterium of the human skin flora. J Bacteriol 187:4671–4682
27. Stewart ME, Steele WA, Downing DT (1989) Changes in the relative amounts of endogenous and exogenous fatty acids in sebaceous lipids during early adolescence. J Invest Dermatol 92:371–378
28. Gurvitz A, Mursula AM, Firzinger A et al (1998) Peroxisomal Delta3-cis-Delta2-trans-enoyl-CoA isomerase encoded by *ECI1* is required for growth of the yeast *Saccharomyces cerevisiae* on unsaturated fatty acids. J Biol Chem 273:31366–31374

29. Yang SY, Li JM, He XY et al (1988) Evidence that the fadB gene of the fadAB operon of *Escherichia coli* encodes 3-hydroxyacyl-coenzyme A (CoA) epimerase, delta 3-cis-delta 2-trans-enoyl-CoA isomerase, and enoyl-CoA hydratase in addition to 3-hydroxyacyl-CoA dehydrogenase. J Bacteriol 170:2543–2548
30. Phetsuksiri B, Jackson M, Scherman H et al (2003) Unique mechanism of action of the thiourea drug isoxyl on *Mycobacterium tuberculosis*. J Biol Chem 278:53123–53130
31. Wang H, Cronan JE (2004) Functional replacement of the FabA and FabB proteins of *Escherichia coli* fatty acid synthesis by *Enterococcus faecalis* FabZ and FabF homologues. J Biol Chem 279:34489–34495
32. Rawlings AV, Harding CR (2004) Moisturization and skin barrier function. Dermatol Ther 17(Suppl 1):43–48
33. Katoh K, Kuma K, Toh H et al (2005) MAFFT version 5: improvement in accuracy of multiple sequence alignment. Nucleic Acids Res 33:511–518
34. Brunke S, Hube B (2006) MfLIP1, a gene encoding an extracellular lipase of the lipid-dependent fungus *Malassezia furfur*. Microbiology 152:547–554
35. Shibata N, Okanuma N, Hirai K et al (2006) Isolation, characterization and molecular cloning of a lipolytic enzyme secreted from *Malassezia pachydermatis*. FEMS Microbiol Lett 256: 137–144
36. Casagrande BF, Fluckiger S, Linder MT et al (2006) Sensitization to the yeast *Malassezia sympodialis* is specific for extrinsic and intrinsic atopic eczema. J Invest Dermatol 126: 2414–2421
37. Andersson A, Rasool O, Schmidt M et al (2004) Cloning, expression and characterization of two new IgE-binding proteins from the yeast *Malassezia sympodialis* with sequence similarities to heat shock proteins and manganese superoxide dismutase. Eur J Biochem 271: 1885–1894
38. Limacher A, Glaser AG, Meier C et al (2007) Cross-reactivity and 1.4-A crystal structure of *Malassezia sympodialis* thioredoxin (Mala s 13), a member of a new pan-allergen family. J Immunol 178:389–396
39. Lindborg M, Magnusson CG, Zargari A et al (1999) Selective cloning of allergens from the skin colonizing yeast *Malassezia furfur* by phage surface display technology. J Invest Dermatol 113: 156–161
40. Onishi Y, Kuroda M, Yasueda H et al (1999) Two-dimensional electrophoresis of *Malassezia* allergens for atopic dermatitis and isolation of Mal f 4 homologs with mitochondrial malate dehydrogenase. Eur J Biochem 261:148–154
41. Schmidt M, Zargari A, Holt P et al (1997) The complete cDNA sequence and expression of the first major allergenic protein of *Malassezia furfur*, Mal f 1. Eur J Biochem 246:181–185
42. Yasueda H, Hashida-Okado T, Saito A et al (1998) Identification and cloning of two novel allergens from the lipophilic yeast, *Malassezia furfur*. Biochem Biophys Res Commun 248: 240–244
43. Zargari A, Selander C, Rasool O et al (2007) Mala s 12 is a major allergen in patients with atopic eczema and has sequence similarities to the GMC oxidoreductase family. Allergy 62: 695–703
44. Andersson A, Scheynius A, Rasool O (2003) Detection of Mala f and Mala s allergen sequences within the genus *Malassezia*. Med Mycol 41:479–485
45. Gaitanis G, Menounos P, Katsambas A et al (2003) Detection and mutation screening of *Malassezia sympodialis* sequences coding for the Mal s 1 allergen implicated in atopic dermatitis. J Invest Dermatol 121:1559–1560
46. Gaitanis G, Velegraki A, Rasool O et al (2005) Clarifications on conflicting results published on the amplification of the Mala s 1 allergen gene sequences. J Invest Dermatol 124:479
47. Zargari A, Midgley G, Back O et al (2003) IgE-reactivity to seven *Malassezia* species. Allergy 58: 306–311

48. Zuther K, Mayser P, Hettwer U et al (2008) The tryptophan aminotransferase Tam1 catalyses the single biosynthetic step for tryptophan-dependent pigment synthesis in *Ustilago maydis*. Mol Microbiol 68:152–172
49. Bakkeren G, Kämper J, Schirawski J (2008) Sex in smut fungi: structure, function and evolution of mating-type complexes. Fungal Genet Biol 45(Suppl 1):S15–21
50. Miller MG, Johnson AD (2002) White-opaque switching in *Candida albicans* is controlled by mating-type locus homeodomain proteins and allows efficient mating. Cell 110:293–302
51. Galagan JE, Calvo SE, Cuomo C et al (2005) Sequencing of *Aspergillus nidulans* and comparative analysis with *A. fumigatus* and *A. oryzae*. Nature 438:1105–1115
52. Fraser JA, Giles SS, Wenink EC et al (2005) Same-sex mating and the origin of the Vancouver Island *Cryptococcus gattii* outbreak. Nature 437:1360–1364