MOLECULAR ECOLOGY

Molecular Ecology (2014) 23, 921–932 doi: 10.1111/mec.12642

Global distribution, diversity hot spots and niche transitions of an astaxanthin-producing eukaryotic microbe

MÁRCIA DAVID-PALMA,* DIEGO LIBKIND† and JOSÉ PAULO SAMPAIO*
*Departamento de Ciências da Vida, Faculdade de Ciências e Tecnologia, Centro de Recursos Microbiológicos (CREM), Universidade Nova de Lisboa, 2829-516 Caparica, Portugal, †Laboratorio de Microbiología Aplicada y Biotecnología, Instituto de Investigaciones en Biodiversidad y Medioambiente (INIBIOMA), CONICET – UNComahue, Quintral 1250 (8400), Bariloche, Argentina

Abstract

Microbes establish very diverse but still poorly understood associations with other microscopic or macroscopic organisms that do not follow the more conventional modes of competition or mutualism. *Phaffia rhodozyma*, an orange-coloured yeast that produces the biotechnologically relevant carotenoid astaxanthin, exhibits a Holarctic association with birch trees in temperate forests that contrasts with the more recent finding of a South American population associated with *Nothofagus* (southern beech) and with stromata of its biotrophic fungal parasite *Cyttaria* spp. We investigated whether the association of *Phaffia* with *Nothofagus–Cyttaria* could be expanded to Australasia, the other region of the world where *Nothofagus* are endemic, studied the genetic structure of populations representing the known worldwide distribution of *Phaffia* and analysed the evolution of the association with tree hosts. The phylogenetic analysis revealed that *Phaffia* diversity in Australasia is much higher than in other regions of the globe and that two endemic and markedly divergent lineages seem to represent new species. The observed genetic diversity correlates with host tree genera rather than with geography, which suggests that adaptation to the different niches is driving population structure in this yeast. The high genetic diversity and endemism in Australasia indicate that the genus evolved in this region and that the association with *Nothofagus* is the ancestral tree association. Estimates of the divergence times of *Phaffia* lineages point to splits that are much more recent than the break-up of Gondwana, supporting that long-distance dispersal rather than vicariance is responsible for observed distribution of *P. rhodozyma*.

Keywords: astaxanthin, *Cyttaria*, *Nothofagus*, *Phaffia*, phylogeography, yeast

Received 2 January 2013; revision received 5 December 2013; accepted 13 December 2013

Introduction

Microorganisms establish an exceptionally vast array of interactions with either other microbes or macroscopic living beings. These diverse interactions range from mutual benefiting associations to cases of markedly detrimental outcomes to one of the partners. Not surprisingly, human, animal and plant diseases caused by microbes are the targets of most of the studies but a myriad of other types of strict or facultative associations adopted by microbes take place in most, if not all, of Earth's ecosystems. Fungi interact with living plants in a variety of ways, the most frequent being the formation of mycorrhizae, the mutualistic association between a fungus and the roots of a vascular plant, and the different kinds of bio- or necrotrophic plant parasitism (Taylor *et al.* 2004). However, other types of interactions that include, for example, exploring the phylloplane niche without apparent negative effects to the host (Fonseca & Inácio 2006) have remained much more obscure.

Correspondence: José Paulo Sampaio, Fax: 351 21 2948530; E-mail: jss@fct.unl.pt

© 2014 John Wiley & Sons Ltd

Phaffia rhodozyma (synonym *Xanthophyllomyces dendrorhous*) is a basidiomycete yeast that was originally found in exudates of *Betula* spp. and a few other deciduous trees of the genera *Alnus*, *Cornus* and *Fagus* (Phaff *et al.* 1972; Miller *et al.* 1976). There are a number of traits that suggest that tree exudates are not a transient habitat of *P. rhodozyma* but rather a niche to which this yeast is well adapted. First, *Phaffia* is able to ferment sugars and is Crabtree positive (Reynders *et al.* 1997), which means that glucose is fermented even under aerobic conditions. The ability to ferment sugars, extremely rare among basidiomycetous yeasts (Barnett *et al.* 2000), is likely to be advantageous in the moist and nutrient-rich tree exudate niche. Second, in *Phaffia*, sexual reproduction does not involve the transition of the unicellular stage to a filamentous one, which is at variance to what is known for all other basidiomycetous yeasts and might be related to the adaptive loss of filamentous structures normally related to the exploitation of solid substrates. Interestingly, the switch from asexual to sexual reproduction is triggered by polyols such as ribitol, which are compounds that are present on the tree exudates colonized by *P. rhodozyma* (Golubev 1995). *Phaffia* is an homothallic microorganism and sexual conjugation occurs between cells of a single strain, usually between a cell and its bud or, less frequently, between two independent cells (Golubev 1995), a process that fits the definition of haploid selfing (Billiard *et al.* 2012). Nevertheless, outcrossing is not excluded as conjugation between cells of distinct strains followed by the development of sexual structures has also been observed (Kucsera *et al.* 1998).

Phaffia produces astaxanthin, a carotenoid that protects the cells from oxidative stress like that caused by photogenerated reactive oxygen species (Schroeder & Johnson 1995). Astaxanthin was also found to have anti-inflammatory and neuroprotective properties (Naguib 2000; Lee *et al.* 2011), in addition to antioxidant activity. It is also used in aquaculture for pigmentation of fish and crustaceans, being the most expensive feed ingredient (Johnson & Schroeder 1995). Current research on *P. rhodozyma* focuses on the genetic manipulation of the pathway of astaxanthin biosynthesis and on improving final yields of this compound (Gassel *et al.* 2013; Loto *et al.* 2012). In the light of the biotechnological relevance of astaxanthin, its global market size is expected to be a quarter-billion dollar by 2015 (Schmidt *et al.* 2011). In addition, a photo-inducible UV-absorbing compound has been recently detected also in *P. rhodozyma*, which is known as mycosporine-glutaminol-glucoside (Libkind *et al.* 2011a).

The Holarctic distribution of *P. rhodozyma*, in predominant association with birch (*Betula*) in temperate forests in North America and Japan, was later extended to *Betula* spp. in Russia (Golubev *et al.* 1977). More recently, the range of *P. rhodozyma* was considerably expanded when we reported on a South American population associated with *Nothofagus* trees (southern beech) and stromata of its biotrophic fungal parasite *Cyttaria* spp. (Libkind *et al.* 2007). *Cyttaria* produces trunk and branch cancers solely on *Nothofagus*, and its fructifications (stromata) form annually in spring and can reach the size of golf balls. *Cyttaria* stromata have a distinctive yellow to orange colour and, when mature, a content of simple sugars of almost 10% (Lederkremer & Ranalli 1967). The discovery of the South American population that was geographically well separated from the previously known populations and the association with *Nothofagus–Cyttaria* increased the interest of *Phaffia* yeasts as an experimental model for the study of ecological associations, phylogeography and long-distance migration of saprobe microbes.

In this study, we investigated whether the association of *Phaffia* with *Nothofagus–Cyttaria* observed in South America could also be found in Australasia, the other region of the globe, reminiscent of the break-up of Gondwana, where the *Nothofagus–Cyttaria* system can be found (Peterson & Pfister 2010). Besides mapping the distribution of *Phaffia* yeasts, we wanted to assess the degree of genetic divergence within and between populations, to get insights into population structure at a continental scale and to analyse in detail the evolution of the association between *Phaffia* and trees of the genera *Nothofagus* and *Betula*.

Materials and methods

Sampling and isolation procedures

Yeast isolations from *Cyttaria* stromata were performed as previously described (Libkind *et al.* 2007). In parallel with the conventional malt and soytone yeast isolation medium (Libkind *et al.* 2007), a selective medium based on yeast nitrogen base (YNB; Difco) supplemented with trehalose (0.8% w/v) was also used (Tognetti *et al.* 2013). For isolations from *Nothofagus* leaves, approximately 10 leaves were placed in 50-mL polypropylene tubes with 40 mL of sterile water and incubated at 18–20 °C for 1 week. After this step, 100 μL of the suspension was used to inoculate the culture media, and the subsequent steps were performed in the same manner as in the isolations from *Cyttaria*. For preliminary identification of the isolates, the formation of sexual structures characteristic of *Phaffia* was checked microscopically using phase-contrast optics on cultures incubated on ribitol agar media (Golubev 1995) for 1 week at 18–20 °C.

© 2014 John Wiley & Sons Ltd

DNA extraction and sequencing

Genomic DNA was obtained as previously described (Sampaio *et al.* 2001). PCR conditions and primers used to amplify fragments of the genes *CRTI*, *CRTS*, *IDI*, *INV*, *L41*, *EPH1*, *GDHA* and rDNA are listed in Table S1 (Supporting information). For each gene, both strands of purified PCR products were sequenced. Gene fragments with heterozygous sites were inserted into the vector pJET1/blunt, and the cloning protocol of CloneJET™ PCR Cloning Kit from Fermentas was followed using DH5-α *Escherichia coli* cells. The obtained clones were sequenced to derive the two original sequences.

Phylogenetic analysis, population structure and ancestral state reconstruction

Sequences were aligned using CLUSTALW (Thompson *et al.* 1994) implemented in BIOEDIT 7.1 (Hall 1999) and also with T-Coffee (Notredame *et al.* 2000; Di Tommaso *et al.* 2011). Phylogenetic analyses using maximum parsimony were performed in MEGA5 (Tamura *et al.* 2011). The close-neighbor-interchange algorithm was used with search level 1 and random addition of sequences (10 replicates). Branch lengths were calculated using the average pathway method (Nei & Kumar 2000). Congruence of the seven individual gene trees was analysed in PAUP (Swofford 2003), using the partition homogeneity test performed with 1000 replicates. Congruence was considered to exist if $P > 0.05$.

Phylogenetic analyses using maximum likelihood were performed in RAxML (Stamatakis 2006). In this case, sequences were aligned using T-Coffee algorithm (Di Tommaso *et al.* 2011) as implemented in the T-Coffee Server (http://tcoffee.crg.cat/apps/tcoffee/index.html). Reliability of the ML trees was assessed by bootstrap analysis (Felsenstein 1985) including 1000 replications. The jModelTest (Posada 2008) was used to select the most appropriate model of nucleotide substitution under the Akaike information criterion (Akaike 1998) for single-gene concatenated phylogenies.

Phylogenetic networks were constructed using the neighbor-net model (Bryant & Moulton 2004). For the calculation of the pairwise homoplasy index (Phi statistics Φw), values of $P > 0.05$ were interpreted as indicating the absence of recombination (Bruen *et al.* 2006). Both analyses were conducted in SPLITSTREE 4 (Huson & Bryant 2006).

To investigate population structure, the multilocus sequence data set was analysed in STRUCTURE 2.3 (Pritchard *et al.* 2000; Falush *et al.* 2003), which implements a model-based clustering method to determine the optimal number of genetic clusters present in a population. A burn-in length of 100 000 steps followed by 10^6 steps of data collection was used. The ancestry model used was admixture with correlated allele frequency, and the number of populations (*K*) tested ranged from 1 to 10 (data set of 40 strains) and 1 to 8 (data set of 32 strains) with 20 runs per each *K*-value. The *ad hoc* measure of ΔK, which is based on the rate of change in the log probability of data between successive *K*-values (Evanno *et al.* 2005), was employed for estimating the number of clusters of *Phaffia rhodozyma*. Two distinct data sets were tested, one including 40 strains (populations A–F) and another with 32 strains (populations A–D). Ancestral state reconstruction regarding tree associations was performed in MESQUITE (http://mesquiteproject.org.). Associations with *Nothofagus*, *Betula* and *Cornus* trees were considered for the analysis. The phylogeny of the tree hosts was estimated with maximum likelihood and followed the procedures indicated above for this method.

Molecular dating

Using a primary calibration date taken from Coetzee *et al.* (2011), we calculated a secondary calibration date that was used to estimate the divergence times of the various *Phaffia* lineages in BEAST 1.7 (Drummond *et al.* 2012). We used the data set of Coetzee *et al.* (2011) to which *Phaffia* sequences were added. The first data set had sequences of the large subunit (LSU) region of the rDNA operon and corresponded to the Basidiomycota matrix of Coetzee *et al.* (2011) that included members of the Agaricomycotina (to which *Phaffia* belongs) and Ustilaginomycotina. The third main lineage of Basidiomycota, the Pucciniomycotina, was used as outgroup. We added five sequences of Cystofilobasidiales, comprising *Phaffia* and its sister genus *Cystofilobasidium*. Following Matheny *et al.* (2009) and Coetzee *et al.* (2011), we set the prior distribution for the separation of *Ustilago* from Agaricomycotina to a normal distribution with a mean of 430 Ma (standard deviation: 50). This allowed estimating the divergence dates between *Cystofilobasidium* and *Phaffia* as well as the divergence of the three *Phaffia* species. The estimates obtained with this data set were then used as calibration points for the second data set that combined LSU and internal transcribed spacer (ITS of the rDNA operon) and included representatives of all the *Phaffia* lineages. BEAUTI 1.7 (Drummond *et al.* 2012) was used to set prior distributions for the tree nodes, and a relaxed molecular clock with uncorrelated log-normal rate variation with a Yule speciation prior was used. The best fitting ML model for the data set was determined by MODELTEST 1.1 (Posada 2008). The remaining priors and operators were in default settings. A number of 2×10^8 generations of

© 2014 John Wiley & Sons Ltd

Markov chain Monte Carlo were used with sampling each 1000 steps. Statistical analyses of BEAST results were performed with TRACER 1.5 (http://tree.bio.ed.ac.uk/software/tracer/), data were summarized by TREEANNOTATOR 1.1 (http://beast.bio.ed.ac.uk/software/TreeAnnotator), and FIGTREE 1.3 (http://tree.bio.ed.ac.uk/software/figtree/) was used to draw the chronograms.

Results

Phaffia is present in Australasia

Field work was carried out in November and December of 2009 in *Nothofagus* woodlands infected with *Cyttaria* in Queensland and Tasmania (Australia) and in New Zealand's South Island (Fig. 1). Given that in Lamington National Park (Queensland), *Cyttaria* fruit bodies mature earlier than in the other locations due to the warmer climate, they were already absent when the collections were done. Consequently, in this collecting site, *Nothofagus* leaves were used instead. We were able to isolate *Phaffia* yeasts from all the localities surveyed and from a total of 32 samples that were collected, 18 yielded positive results (Table 1). Approximately one hundred Australasian *Phaffia* isolates were recovered and preliminarily identified by the detection of typical basidia. The success rate of isolation was considerable (Table 1), thus suggesting that these yeasts were not transient in the *Cyttaria–Nothofagus* system. The isolates chosen for further analysis were selected so as to minimize the possible inclusion of strains from the same population. Considering that our study had four sampling sites, 12 isolates were selected for further studies.

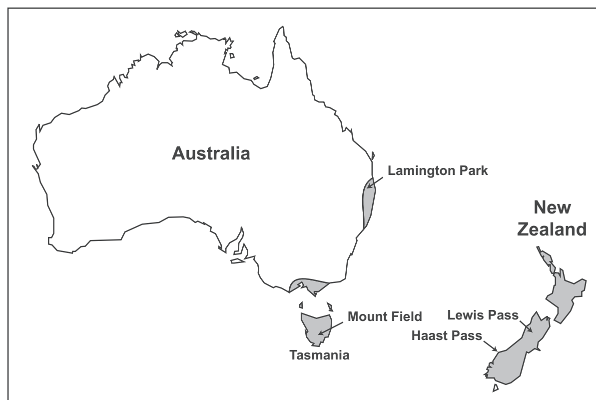

Fig. 1 Australasian map showing present-day *Nothofagus* distribution (shadowed areas) and collection localities (arrows).

Phylogeny and diversity

To study the phylogenetic relationships of the new isolates, seven genes (Table S1, Supporting information) were partially sequenced for a set of 40 strains representing the global geographical and ecological diversity currently known for this species (Table S2, Supporting information). Besides the newly found Australasian isolates, this strain data set included isolates from North and South America, Europe and Asia. First, individual maximum parsimony gene phylogenies were constructed, and their congruence was assessed with the partition homogeneity test implemented in PAUP. The null hypothesis of the absence of congruence between the seven data sets (corresponding to $P < 0.05$) could be rejected because in our study $P = 0.421$. Moreover, single-gene phylogenies constructed using both maximum parsimony and maximum likelihood were in general concordant, thus providing additional support for concatenation (Fig. S3, Supporting information). The concatenated alignment was used to generate the best maximum likelihood phylogenetic tree inferred in RaxML (Fig. 2; the equivalent maximum parsimony tree is shown in Fig. S3, Supporting information). Six main clades (A–F) were detected and found to be partially correlated with the distribution of *Phaffia* at a continental scale. While the South American isolates were grouped in clade A, the Australasian isolates were distributed in three divergent clades (B, E and F). The Japanese isolates were distributed in clades C and D, whereas all the North American/Northern European isolates were placed in clade C. Clade C was subdivided into an Asian/Alaskan subclade (C1) and a North American/North European subclade (C2). The divergence of clades E and F from each other and from

Table 1 Sampling localities, number and type of samples and success rates of isolation

Locality	Host system	Number of samples	Success rate of isolation (%)
Lamington National Park, Queensland, Australia	*N. mooreii* (leaves)	6	100
Mount Field National Park, Tasmania, Australia	*N. cunninghamii – C. gunnii*	13	30
Lewis Pass, New Zealand	*N. menziesii – C. nigra/C. gunnii*	4	50
Hasst Pass, New Zealand	*N. menziesii – C. nigra/C. gunnii*	9	65

C., Cyttaria; N., Nothofagus.

© 2014 John Wiley & Sons Ltd

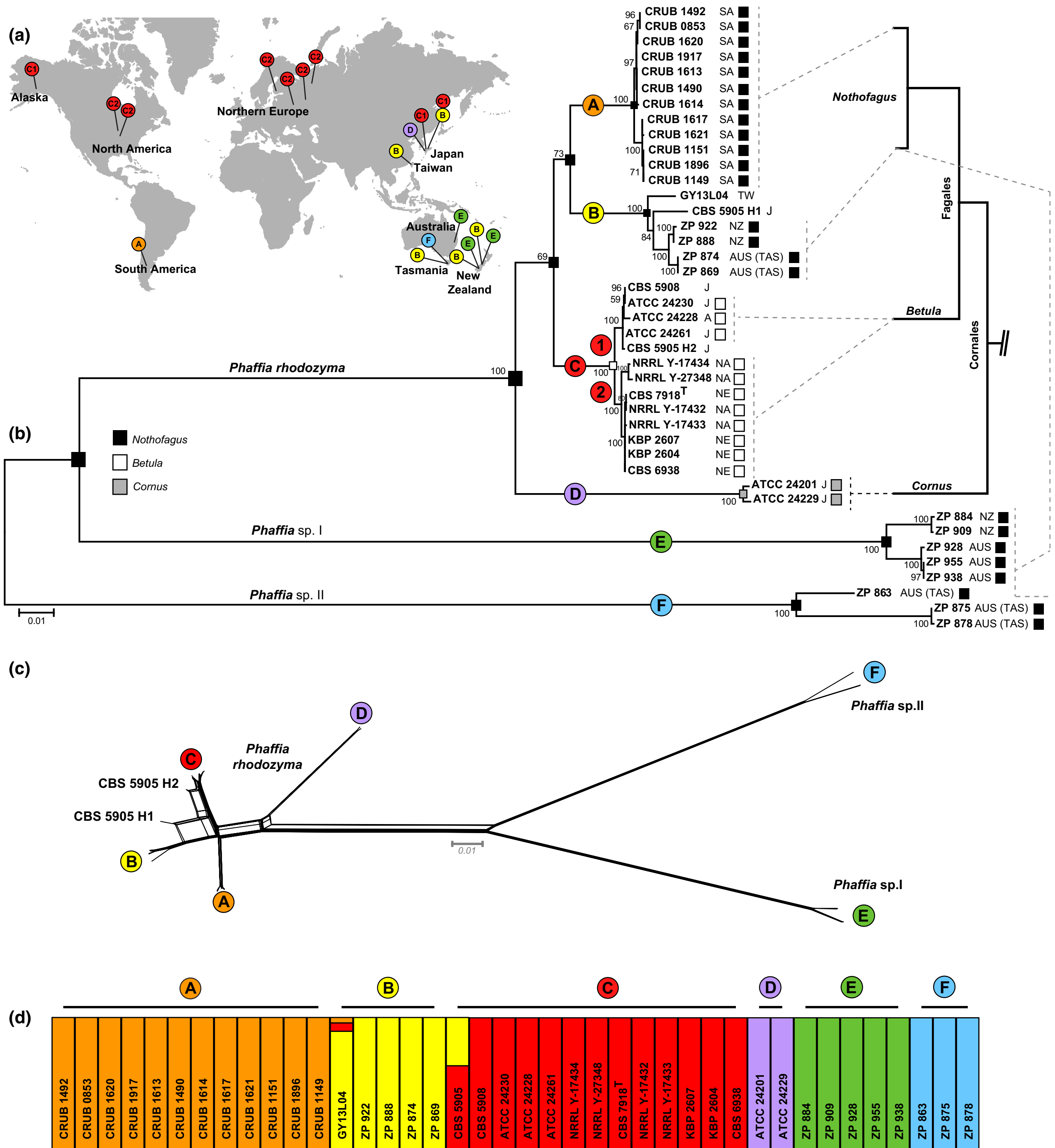

Fig. 2 Distribution, phylogeny, host tree association and phylogenetic network of *Phaffia*. (a) World map showing the distribution of populations, which are represented by capital letters. (b) Unrooted maximum likelihood phylogeny of *Phaffia* (left) based on a concatenated alignment of partial sequences of seven genes and host trees (right) based on internal transcribed spacer sequences. Bootstrap values (1000 replicates) higher than 50% are indicated. The geographical origin of the isolates is indicated after strain number (A, Alaska; AUS, Australia; J, Japan; NA, North America; NE, Northern Europe; NZ, New Zealand; SA, South America; TAS, Tasmania). Squares at internal and terminal tree nodes refer to ancestral state reconstruction (black squares, association with *Nothofagus*; open squares, association with *Betula*; grey squares, association with *Cornus*. (c) neighbor-net phylogenetic network of *Phaffia*. (d) Ancestry of *Phaffia* estimated with STRUCTURE). Each colour represents a distinct population. The vertical axis represents the probability of ancestry that each strain has to each of the inferred populations.

© 2014 John Wiley & Sons Ltd

the remaining clades was remarkable and ranged between 30% and 33% (nucleotide diversity calculated using the Jukes and Cantor correction), thus indicating a marked separation of these two groups. To better evaluate the divergence of lineages E and F, we constructed a broader phylogeny based on ITS sequences (comprising ITS 1, 5.8 S and ITS 2 of rRNA) that are used for fungal DNA barcoding (Schoch *et al*. 2012) and therefore allow the simultaneous assessment of sequence divergence of *Phaffia* lineages and of species of the sister genus *Cystofilobasidium* (Fig. S1, Supporting information). In this analysis, the divergence of lineages E and F from each other and from the remaining *Phaffia* lineages was confirmed. The number of nucleotide substitutions between the different species currently recognized in the genus *Cystofilobasidium* ranged between 11 and 26, whereas the nucleotide differences between the three main *Phaffia* clades (lineages A–D, E and F) varied between 25 and 36, representing therefore an equivalent or higher divergence than that observed for species of *Cystofilobasidium*. Therefore, the most comprehensive *Phaffia* strain data set analysed to date was resolved into three main clades that we tentatively assign the rank of species, primarily because of their degree of sequence divergence. Clades A–D represent *Phaffia rhodozyma,* and clades E and F represent two new Australasian species of *Phaffia*. The individual gene phylogenies (Fig. S3, Supporting information) validate a three-species phylogenetic concept *sensu* Taylor *et al*. (2000) as incongruences between tree topologies were only found within the three putative species, the delimitation of their respective clades being always maintained. Assessing levels of reproductive isolation in mating tests involving representatives of each lineage was not possible because sexual structures were readily produced by each strain alone. While the new species represented by clade E, *Phaffia* sp. I, had a broad distribution and was found in Lamington Park (Queensland) and in the two localities surveyed in New Zealand (Haast Pass and Lewis Pass), *Phaffia* sp. II (clade F) was found only in Tasmania.

Hybridization and recombination

The considerable genetic divergence encompassed by the six *Phaffia* lineages, the three-species hypothesis and the peculiar sexuality with predominance of a mode of reproduction close to haploid selfing led us to assess the relevance of reticulate events such as hybridization and recombination between lineages. For CBS 5905, isolated in Japan, we detected heterozygous sites in the sequences of three genes (*CRTS*, *EPH1* and *IDI*). For these cases, two haplotypes were separated by cloning and their sequences were used in the phylogenetic tree of Fig. 2. Haplotype 1 clustered in clade B that included the Australasian strains of *P. rhodozyma* and also a single strain from Taiwan, whereas haplotype 2 belonged to clade C, formed by strains from the Northern Hemisphere, mostly from *Betula*. As to the remaining genes, whereas the *INV* sequence belonged to clade B, the *CRTI*, *L41* and *GDHA* sequences clustered in clade C. For the seven concatenated genes used to build the phylogeny, the presence of different haplotypes was only observed for CBS 5905, and in the splits network based on the multilocus sequence data set depicted (Fig. 2), we observed a limited number of incompatible splits that did not affect the marked separation between lineages. The restricted reticulation involved clades B (Australasia and Taiwan), C (Europe, North America and Japan) and D (Japan) and suggested some sort of genetic contact between these groups. As expected, the hybrid nature of CBS 5905 contributed to the reticulation observed between clades B and C. The Patagonian population (clade A) and the divergent Australasian clades E and F were well isolated and had no reticulate events with the remaining populations. Besides CBS 5905, another case of possible heterozygosity was found for the ITS sequence of strain GY13L04. However, as multiple copies of the rDNA operon are normally present, it is possible, although not frequent, that more than one ITS version is present in the same organism. In the case of GY13L04, the two sequence versions belonged to the same clade (Fig. S1, Supporting information). For a statistical validation of the phylogenetic network approach, we calculated the pairwise homoplasy index (Bruen *et al*. 2006) within and between the populations of the core group. The results were concordant with the previous analysis, and the hypothesis of free recombination in this species could be rejected as recombination was only detected between populations B and C and solely when the two haplotypes of CBS 5905 were included in the analysis (Φ_w higher than 0.05 indicates the absence of recombination and was 2×10^{-15} when CBS 5905 was included and 0.7 when this strain was excluded). In summary, the reticulation observed for *Phaffia* did not cause a marked change in the topology of the phylogeny proposed in Fig. 2, thus suggesting limited roles for sexual recombination and hybridization between lineages. Finally, we used the quantitative clustering method implemented in STRUCTURE to analyse population memberships. Six ancestral clusters ($K = 6$) were inferred and found to be coincident with the six clades mentioned earlier (Fig. 2). The vast majority of strains were clean lineages, and only two strains (GY13L04 and CBS 5905) had mixed ancestries, which is in concordance with the phylogenetic analyses reported earlier. The method of Evanno *et al.* (2005) supported $K = 4$ and $K = 6$, and higher values of K

© 2014 John Wiley & Sons Ltd

failed to reveal additional populations (Fig. S2, Supporting information). For $K = 4$, the divergent Australasian clades E and F were maintained as clean lineages, and a mixed ancestry was suggested for clades B and D. The differences observed between the plots from $K = 4$ and $K = 6$ might be due to the extreme divergence of lineages E and F. Removing them from the analysis yielded four populations that were concordant with the phylogenetic analysis (Fig. S2, Supporting information).

Tree associations

Nothofagus is an emblematic genus for the study of biogeographical processes in the Southern Hemisphere related to the break-up of Gondwana (van Steenis 1971; Linder & Crisp 1995; Sanmartín & Ronquist 2004; Cook & Crisp 2005). This super-continent that once united Australia, New Zealand, Antarctica, Africa and South America fragmented gradually between 95 and 30 Ma (Sanmartín & Ronquist 2004). It is considered that Australia and New Zealand separated first (80 Ma), whereas Australia and South America were connected by Antarctica until 30 Ma (McLoughlin 2001). Likewise, the Tasman Sea, which separates Tasmania from New Zealand, reached its present size circa 65 Ma (Coleman 1980; Veevers *et al.* 1991). Molecular dating of extant *Nothofagus* lineages indicates that long-distance dispersal played a key role because, in general, the divergence times of several lineages are much younger than continental split (Cook & Crisp 2005; Knapp *et al.* 2005). Moreover, there is no instance of a *Nothofagus* or *Cyttaria* species common to both South America and Australasia or present both in Australia and New Zealand (Hill & Read 1991), (Peterson & Pfister 2010). The biogeography of *Phaffia* was analysed in the light of the chronology of the break-up of Gondwana and in comparison with what is known for southern beech and its fungal parasite. The results obtained for *Phaffia* are in sharp contrast with those of *Nothofagus* or *Cyttaria* because the pattern of allopatric isolation and consequent speciation at a continental scale was not observed. First, a single species, *P. rhodozyma*, appears to be associated with South American and Australasian *Nothofagus–Cyttaria* systems, and second, this species and *Phaffia* sp. I were found both in Australia (including Tasmania) and New Zealand. To explore this issue further, we used as template a recently published relaxed molecular clock estimate of the radiation of the plant pathogenic mushroom *Armillaria* (Coetzee *et al.* 2011) and included in an equivalent analysis, representatives of the six main lineages of *Phaffia*. The age of divergence of the three *Phaffia* species (clades A–D, E and F) was estimated to be between 13.3 and 20.7 Ma (Fig. 3, Table 2), which suggests that, even taking into account the considerable incertitude of these methodologies, speciation occurred after the break-up of Gondwana. Moreover, the estimation of the divergence between the South American and Australasian populations of *P. rhodozyma* (clades A and B) was 0.8 Ma, and in clade B, isolates from Tasmania and New Zealand show even more recent divergence which can be viewed as absence of a separation.

The *Betulaceae* are distributed in temperate regions of the Northern Hemisphere (Chen *et al.* 1999). It has been proposed that the Betulaceae arose in Central China 65 Ma, in the Late Cretaceous and Early Tertiary (Chen *et al.* 1999), long after the radiation of *Nothofagus*, and then colonized Eurasia and North America. At that time, migration to North America was possible from the east via the Bering land bridge and from the west via the Greenland North Atlantic land bridge (Tiffney 1985), and it is considered that *Betula* expansion followed these routes. With respect to tree associations in *Phaffia*, the ancestral state reconstruction performed in MESQUITE supported that the association with *Nothofagus* is the ancestral state and that transitions to other hosts occurred independently and at least twice, one to *Cornus* (clade D) and one to *Betula* (clade C) (Fig. 2). Moreover, according to the relaxed molecular clock estimates, the transition from *Nothofagus* to *Betula* occurred around 3.1 Ma (Fig. 3), therefore well after *Betula* colonized Eurasia and North America.

The ITS molecular phylogeny of *Phaffia* main tree hosts, *Nothofagus*, *Betula* and *Cornus*, was compared with that of *Phaffia* (Fig. 2 and Fig S4, Supporting information). As *Nothofagus* and *Betula* are more related and belong to the order Fagales, *Cornus* appears as outgroup. Interestingly, an equivalent topology was obtained for *Phaffia* lineages A–D that represent the species *P. rhodozyma*, the strains associated with *Cornus* being external to their *Nothofagus* and *Betula* relatives. It is also remarkable that the two *P. rhodozyma* lineages associated with *Nothofagus*, one South American (clade A) and the other Australasian (clade B), form a monophyllum that correlates with the two terminal nodes of the southern beech phylogeny representing the South American and Australasian *Nothofagus* (Fig. 2). Therefore, besides the dominance of the *Nothofagus* association, this comparison suggests a correlation between the phylogeny of the main lineages of *Phaffia* and that of the trees to which they are associated.

Discussion

Australasia as a diversity hot spot

The field of fungal phylogeography is less developed than that of animal or plant phylogeography, and few

© 2014 John Wiley & Sons Ltd

Table 2 Estimated divergence dates, node heights and confidence intervals of *Phaffia* lineages

Node	Date estimate	Median node height	95% HPD (L – H)
Phaffia rhodozyma			
Divergence of clades A and B	0.82	0.52	0–2.6
Divergence of clade C	3.11	2.76	0.5–6.4
Divergence of clade D	5.28	4.92	1.5–9.8
Phaffia sp. I	13.4	13.4	13.3–13.5
Phaffia sp. II	20.7	20.7	20.6–20.8

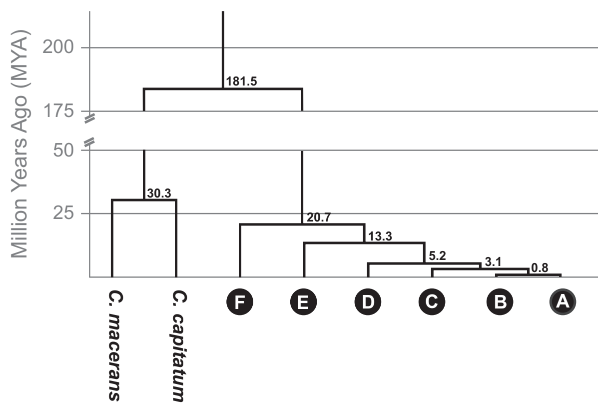

Fig. 3 Molecular dating of the divergence of *Phaffia* lineages and of its sister genus *Cystofilobasidium*, based on concatenated large subunit and internal transcribed spacer sequences. Estimated divergence times are indicated at the tree nodes.

studies have addressed the topic of the global distribution of fungi in space and through time (Thorsten *et al.* 2008). Among the several reasons that account for the poor development of evolutionary studies of fungal biogeography, the difficulty in delimiting and recognizing species, the high sampling effort associated with sound ecological studies and the puzzling vast distribution of some taxa over continents should be highlighted (Berbee & Taylor 2010). Second, the study of ecological phenomena in microbes is intrinsically complex because it involves multiple and elaborated interactions that can only be elucidated with detailed investigations carried out at the molecular level. In a previous study, we characterized a South American population of *Phaffia rhodozyma* in Patagonia and reported, for the first time, the occurrence of this species in the Southern Hemisphere (Libkind *et al.* 2007, 2011b). The Patagonian population represented a third lineage within the species, the other two being a clade that included several strains isolated from *Betula* in Finland, Russia and Alaska and a minor group of markedly divergent strains isolated in Japan from *Cornus brachypoda* (Libkind *et al.* 2007). In the present study, we investigated whether the *Nothofagus*–*Cyttaria* association previously detected in South America could be extended to Australasian *Nothofagus* infected with *Cyttaria*. Besides detecting *Phaffia* lineages in Australasia, we observed that those lineages represent the widest divergence so far reported for these yeasts. This divergence is illustrated by the presence in Australasia, but not in other regions, of three *Phaffia* species. Therefore, Australasia exhibits a much higher diversity of *Phaffia* yeasts than South America or the Holarctic region and can be therefore viewed as a diversity hot spot. New variation found in this restricted group of astaxanthin-producing yeasts might reveal stocks with increased biotechnological interest. In addition, gaining new insights into the ecology of these yeasts will foster our ability to understand the biological roles of astaxanthin in natural systems.

Reproductive modes and diversifying selection

The genetic diversity that we detected in *P. rhodozyma* correlates with the observed association with the tree genera, *Betula*, *Cornus* and *Nothofagus*, rather than with geography. Populations A and B, from South American and Australasian *Nothofagus*, respectively, are phylogenetically related in spite of being separated by more than 10 000 Km. In contrast, population C1 from *Betula* in Japan and Alaska is geographically close to the Australasian–Taiwanese population B from *Nothofagus*, or even in contact given the hybrid nature of CBS 5905, but the two groups show a genetic divergence higher than that observed between populations A and B (Fig. 2). Finally, the most divergent group within *P. rhodozyma* (population D, from *Cornus*) is from Japan, but clearly distinct from population C1 from *Betula* (and *Fagus*, see CBS 5908) also from Japan. A factor promoting this niche-dependent assortment of *Phaffia* genotypes could be diversifying selection. Kassen & Rainey (2004) have argued that diversifying selection can contribute to maintain diversity in nonuniform environments that are composed of distinct niches. The diversifying selection hypothesis proposes that different genotypes within a species may be favoured in each niche and, as a consequence, diversity within the species is promoted. Such a scenario has been recently confirmed for a flower-living yeast species for which plant microsite-dependent divergent selection was found to influence local genetic diversity (Herrera *et al.* 2011). *Phaffia* seems to represent an extreme case of diversifying selection in which distinct niches, represented by *Betula*, *Cornus* and *Nothofagus*, act as differential filters of genotypes and function as drivers of genetic diversity. Because of the disjunct distributions of the host trees, *Betula* and *Cornus* are typically

© 2014 John Wiley & Sons Ltd

Holarctic taxa (Xiang & Thomas 2008), whereas *Nothofagus* is restricted to the Southern Hemisphere; divergent selection in *P. rhodozyma* has implications not only for population biology (e.g. the sympatric occurrence of *Betula* and *Cornus* lineages in Japan), but also for biogeography and vicariance as distinct genetic lineages can be found in allopatry, like in the case of *Betula* and *Nothofagus* lineages.

At this stage, it is not clear how selection acting on genes relevant to niche adaptation would have contributed to the observed genetic diversity seen on genes that are not obvious candidates for such selective pressures. We hypothesize that the life cycle of *Phaffia* facilitates divergence between lineages, at the genome scale, for two main reasons. First, it is dominated by mitotic asexual reproduction, similar to what has been found in other fungal systems (Campbell *et al.* 2005; Billiard *et al.* 2012). Asexual reproduction in the form of bud formation is likely to promote selection for local adaptation to different host trees as recombination, which would break down advantageous allelic combinations, is absent. *Phaffia* sexual reproduction looks much less frequent than mitotic divisions and depends on the seasonal presence of chemical triggers originating from the host trees. Haploid selfing *sensu* Billiard *et al.* (2011) appears to be the most frequent mode of sexual reproduction and is accompanied by probably rare outcrossing events. Although changes in ploidy levels have not yet been confidently measured, conjugation between two cells of the same strain or even between a cell and its bud occurs immediately before the development of the basidium, the sexual structure that undergoes meiosis and forms the meiospores. Heterothalism (outcrossing) has also been shown to occur experimentally (Kucsera *et al.* 1998). The second reason is linked to assortative (nonrandom) mating and is likely to play a key role. Host factors, namely polyols present in tree fluxes, are known to trigger the development of the sexual cycle (Golubev 1995). Therefore, sexual reproduction only occurs in the different host trees that the different lineages are adapted to. Such habitat-dependent reproductive mode should promote the genetic differentiation observed for the various lineages (Giraud *et al.* 2010).

Origin and radiation of Phaffia

We have shown that Australasia has the highest diversity and the greatest level of endemism, which suggests that the genus *Phaffia* evolved originally in this region. The association with *Cyttaria–Nothofagus* seems to be an ancestral trait in the genus and is present in the three species known. Moreover, the distribution of *P. rhodozyma* has to be explained by long-distance dispersal rather than by vicariance. Given our estimate of the recent *Phaffia* radiation, long-distance dispersal is the most likely explanation for the presence of the same species in Australia and New Zealand and also in Australasia, South America and the Holarctic region. Within the long-distance dispersal scenario common to the other examples already mentioned (*Armillaria*, *Cyttaria* and *Nothofagus*), molecular dating suggested that the divergence of *Phaffia* lineages was considerably more recent than that of the other genera, even taking into account the level of incertitude associated with these approaches. The origin of *Cyttaria* was proposed to range between 112 and 148 Ma (Peterson & Pfister 2010), *Nothofagus* early divergence to be of 80 Ma (Knapp *et al.* 2005) and the deepest divergence of *Armillaria* was estimated to be 50 Ma (Coetzee *et al.* 2011). Therefore, taking into consideration that the divergence and possible migration of the populations of *P. rhodozyma* took place <5.5 Ma (Fig. 3), such a recent dispersal process could be the reason why speciation has not yet occurred, although the different populations are becoming gradually more genetically isolated from each other.

Betula and *Nothofagus* belong to the order Fagales, and *Nothofagus* is viewed as the ancestral lineage in the order (Li *et al.* 2004). Based on the fossil pollen record, it has been hypothesized that a Cretaceous Fagalean complex existed in the South-East Asian–Australian region (Hill 1992). Migrations southwards might have originated the *Nothofagus* group, whereas this complex might have given rise to the Betulaceae north of the equator, in South-East Asia. Our findings on *Phaffia* suggest that the three species originated in Australasia in association with *Nothofagus*, a radiation that started about 21 Ma. More recently, some lineages of *P. rhodozyma* shifted their ancestral association with *Cyttaria–Nothofagus* to associations with *Cornus* and *Betula*, while other lineages remained associated with *Nothofagus–Cyttaria*. The most recent divergence documented in our study is the one between the South American and Australasian populations that are associated with the *Nothofagus–Cyttaria* system (clades A and B). The topology and dating of the divergence of *Phaffia* lineages shown in Fig. 3 suggest that *P. rhodozyma* has been permanently associated with *Nothofagus* and that, at different time points, some populations become independently associated with *Cornus* and with *Betula* as supported by the ancestral state analysis performed in this study (Fig. 2). The relatively lower divergence between the Patagonian and the Australasian populations suggests that although geographically well apart, these two populations are not completely genetically isolated.

For the *Phaffia* lineage associated with *Betula* (clade C), a statistically well-supported subdivision was observed (Fig. 2). Subclade C1 corresponded to isolates found in Japan and western North America (Alaska),

© 2014 John Wiley & Sons Ltd

whereas subclade C2 corresponded to European and eastern North American isolates. The detection in North America of western and eastern lineages that encompass Japanese and European strains, respectively, suggests that colonization involved two independent migration routes and that the longitudinal separation provided by the Rocky Mountains might have contributed to maintain the eastern and western populations differentiated. Interestingly, the two possible migration routes of *Phaffia* proposed here correspond to the two *Betula* migration routes: one from Central Asia to western North America, across the Bering land bridge; and another one from Central Asia to Europe and then to eastern North America via the North Atlantic land bridge (Milne 2006).

Among the *Phaffia* lineages associated with *Nothofagus*, the genetic resemblance of the Patagonian and Australasian populations of *P. rhodozyma* contrasts with the differentiation observed for the Australasian lineages (*P. rhodozyma* and *Phaffia* sp. I and *Phaffia* sp. II). The Australasian lineages seem to be in apparent sympatry, and in three of the four localities surveyed, more than one *Phaffia* species was found (Table S2, Supporting information). One example is *Phaffia* sp. I that, in New Zealand, was isolated from mature *Cyttaria* stromata that had fallen on the ground, whereas in the same collection sites, *P. rhodozyma* was isolated from stromata still attached to *Nothofagus*. This suggests a process of ecological speciation and it would be interesting to know what the precise ecological niches of the three species are, as it is possible that they have evolved distinct ecological adaptations to avoid competitive exclusion. Another interesting question with relevance to the field of evolutionary microbial ecology is the identification of the factors that promoted the long-distance distribution and the colonization of new niches in *P. rhodozyma* but not in the other two species revealed in this work. However, before such questions can be addressed, basic aspects of *Phaffia* life cycle must be clarified, especially the precise identification of ploidy transitions during sexual cycle, the apparent haploid selfing reproductive mode and its coexistence with a heterothallic system.

Acknowledgements

This work was supported by FCT (Portugal) grants PTDC/BIA-BDE/73566/2006 and PEst-OE/BIA/UI0457/2011 (MP, JPS), SFRH/BD/81895/2011 (MP), ANPCYT (Argentina) grants PICT 2007-1745 and PICT 2012 – 1814 (DL). Ching-Fu Lee kindly provided strain GY13L04, and Celia Tognetti is acknowledged for her assistance in field and laboratory work in Patagonia. The following individuals and organizations are thankfully acknowledged for their help during preparation and execution of field work in Australia and New Zealand: Genevieve Gates and David Ratkowsky, University of Tasmania, AUS; Nigel Fechner, Queensland Herbarium, AUS; Peter Johnston, Landcare Research, NZ; Tony Young, The University of Queensland, AUS; Department of the Environment, Water, Heritage and the Arts, Australian Government; Department of Environment and Resource Management, Toowoomba Office, Queensland, AUS; Department of Conservation (West Coast Conservancy and Otago Conservancy), NZ.

References

Akaike H (1998) Information theory and an extension of the maximum likelihood principle. In: *Selected Papers of Hirotugu Akaike* (eds Parzen E, Tanabe K, Kitagawa G), pp. 199–213. Springer, New York City, New York.

Barnett J, Payne R, Yarrow D (2000) *Yeasts: Characteristics and Identification*, 3rd edn. Cambridge University Press, Cambridge.

Berbee ML, Taylor JW (2010) Dating the molecular clock in fungi – how close are we? *Fungal Biology Reviews*, **24**, 1–16.

Billiard S, López-Villavicencio M, Devier B *et al.* (2011) Having sex, yes, but with whom? Inferences from fungi on the evolution of anisogamy and mating types. *Biological Reviews*, **86**, 421–442.

Billiard S, LÓPez-Villavicencio M, Hood ME, Giraud T (2012) Sex, outcrossing and mating types: unsolved questions in fungi and beyond. *Journal of Evolutionary Biology*, **25**, 1020–1038.

Bruen TC, Philippe H, Bryant D (2006) A simple and robust statistical test for detecting the presence of recombination. *Genetics*, **172**, 2665–2681.

Bryant D, Moulton V (2004) Neighbor-Net: an agglomerative method for the construction of phylogenetic networks. *Molecular Biology and Evolution*, **21**, 255–265.

Campbell LT, Currie BJ, Krockenberger M *et al.* (2005) Clonality and recombination in genetically differentiated subgroups of *Cryptococcus gattii*. *Eukaryotic Cell*, **4**, 1403–1409.

Chen Z-D, Manchester SR, Sun H-Y (1999) Phylogeny and evolution of the Betulaceae as inferred from DNA sequences, morphology and paleobotany. *American Journal of Botany*, **86**, 1168–1181.

Coetzee MPA, Bloomer P, Wingfield MJ, Wingfield BD (2011) Paleogene radiation of a plant pathogenic mushroom. *PLoS ONE*, **6**, e28545.

Coleman PJ (1980) Plate tectonics background to biogeographic development in the southwest Pacific over the last 100 million years. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **31**, 105–121.

Cook LG, Crisp MD (2005) Not so ancient: the extant crown group of *Nothofagus* represents a post-Gondwanan radiation. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, **272**, 2535–2544.

Di Tommaso P, Moretti S, Xenarios I *et al.* (2011) T-Coffee: a web server for the multiple sequence alignment of protein and RNA sequences using structural information and homology extension. *Nucleic Acids Research*, **39**, W13–W17.

Drummond AJ, Suchard MA, Xie D, Rambaut A (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*, **29**, 1969–1973.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, **14**, 2611–2620.

© 2014 John Wiley & Sons Ltd

Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics*, **164**, 1567–1587.

Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. *Evolution*, **39**, 783–791.

Fonseca Á, Inácio J (2006) Phylloplane yeasts. In: *Biodiversity and Ecophysiology of Yeasts* (eds Péter G, Rosa C), pp. 263–301. Springer, Berlin.

Gassel S, Schewe H, Schmidt I, Schrader J, Sandmann G (2013) Multiple improvement of astaxanthin biosynthesis in *Xanthophyllomyces dendrorhous* by a combination of conventional mutagenesis and metabolic pathway engineering. *Biotechnology Letters*, **35**, 565–569.

Giraud T, Gladieux P, Gavrilets S (2010) Linking the emergence of fungal plant diseases with ecological speciation. *Trends in Ecology & Evolution*, **25**, 387–395.

Golubev W (1995) Perfect state of *Rhodomyces dendrorhous* (*Phaffia rhodozyma*). *Yeast*, **11**, 101–110.

Golubev WI, Bab'eva IP, Blagodatskaya VM, Reshetova IS (1977) Taxonomic study of yeasts isolated from spring sap flows of birch (*Betula verrucosa* Ehrh.). *Mikrobiologia*, **46**, 564–568.

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, **41**, 95–98.

Herrera C, Pozo MI, Bazaga P (2011) Clonality, genetic diversity and support for the diversifying selection hypothesis in natural populations of a flower-living yeast. *Molecular Ecology*, **20**, 4395–4407.

Hill RS (1992) *Nothofagus*: evolution from a southern perspective. *Trends in Ecology and Evolution*, **7**, 190–194.

Hill RS, Read J (1991) A revised infrageneric classification of *Nothofagus* (Fagaceae). *Botanical Journal of the Linnean Society*, **105**, 37–72.

Huson DH, Bryant D (2006) Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution*, **23**, 254–267.

Johnson E, Schroeder W (1995) Microbial carotenoids. In: *Downstream Processing Biosurfactants Carotenoids* (ed. Fiechter A), pp. 119–178. Springer, Berlin.

Kassen R, Rainey PB (2004) The ecology and genetics of microbial diversity. *Annual Review of Microbiology*, **58**, 207–231.

Knapp M, Stöckler K, Havell D *et al.* (2005) Relaxed molecular clock provides evidence for long-distance dispersal of *Nothofagus* (Southern Beech). *PLoS Biology*, **3**, e14.

Kucsera J, Pfeiffer I, Ferenczy L (1998) Homothallic life cycle in the diploid red yeast *Xanthophyllomyces dendrorhous* (*Phaffia rhodozyma*). *Antonie van Leeuwenhoek*, **73**, 163–168.

Lederkremer R, Ranalli M (1967) Hidratos de carbono en hongos de la Patagonia Argentina. Azúcares simples y polioles en *Cyttaria hariotii* Fischer. *Anales de la Asociación Química Argentina*, **55**, 199–203.

Lee D-H, Kim C-S, Lee YJ (2011) Astaxanthin protects against MPTP/MPP+-induced mitochondrial dysfunction and ROS production in vivo and in vitro. *Food and Chemical Toxicology*, **49**, 271–280.

Li RQ, Chen ZD, Lu AM *et al.* (2004) Phylogenetic relationships in Fagales based on DNA sequences from three genomes. *International Journal of Plant Sciences*, **165**, 311–324.

Libkind D, Ruffini A, van Broock M, Alves L, Sampaio J (2007) Biogeography, host specificity, and molecular phylogeny of the basidiomycetous yeast *Phaffia rhodozyma* and its sexual form, *Xanthophyllomyces dendrorhous*. *Applied and Environmental Microbiology*, **73**, 1120–1125.

Libkind D, Moliné M, van Broock M (2011a) Production of the UVB-absorbing compound mycosporine–glutaminol–glucoside by *Xanthophyllomyces dendrorhous* (*Phaffia rhodozyma*). *FEMS Yeast Research*, **11**, 52–59.

Libkind D, Tognetti C, Ruffini A, Sampaio J, van Broock M (2011b) *Xanthophyllomyces dendrorhous* (*Phaffia rhodozyma*) on stromata of *Cyttaria hariotii* in northwestern Patagonian Nothofagus forests. *Revista Argentina de Microbiología*, **43**, 198–202.

Linder HP, Crisp MD (1995) *Nothofagus* and Pacific biogeography. *Cladistics*, **11**, 5–32.

Loto I, Gutierrez M, Barahona S *et al.* (2012) Enhancement of carotenoid production by disrupting the C22-sterol desaturase gene (CYP61) in *Xanthophyllomyces dendrorhous*. *BMC Microbiology*, **12**, 235.

Matheny PB, Aime MC, Bougher NL *et al.* (2009) Out of the Palaeotropics? Historical biogeography and diversification of the cosmopolitan ectomycorrhizal mushroom family *Inocybaceae*. *Journal of Biogeography*, **36**, 577–592.

McLoughlin S (2001) The breakup history of Gondwana and its impact on pre-Cenozoic floristic provincialism. *Australian Journal of Botany*, **49**, 271–300.

Miller M, Yoneyama M, Soneda M (1976) *Phaffia*, a new yeast genus in the *Deuteromycotina* (*Blastomycetes*). *International Journal of Systematic Bacteriology*, **26**, 286–291.

Milne RI (2006) Northern Hemisphere plant disjunctions: a window on Tertiary land bridges and climate change? *Annals of Botany*, **98**, 465–472.

Naguib YMA (2000) Antioxidant activities of astaxanthin and related carotenoids. *Journal of Agricultural and Food Chemistry*, **48**, 1150–1154.

Nei M, Kumar S (2000) *Molecular Evolution and Phylogenetics*. Oxford University Press, New York City, New York.

Notredame C, Higgins DG, Heringa J (2000) T-coffee: a novel method for fast and accurate multiple sequence alignment. *Journal of Molecular Biology*, **302**, 205–217.

Peterson KR, Pfister DH (2010) Phylogeny of *Cyttaria* inferred from nuclear and mitochondrial sequence and morphological data. *Mycologia*, **102**, 1398–1416.

Phaff H, Miller M, Yoneyama M, Soneda M (1972) A comparative study of the yeast florae associated with trees on the Japanese islands and on the west coast of North America. In: *Fourth International Fermentation Symposium Proceedings: Fermentation Technology Today, Society of Fermentation Technology* (ed. Terui G), pp. 759–774. Osaka, Japan.

Posada D (2008) jModelTest: phylogenetic model averaging. *Molecular Biology and Evolution*, **25**, 1253–1256.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.

Reynders MB, Rawlings DE, Harrison STL (1997) Demonstration of the Crabtree effect in *Phaffia rhodozyma* during continuous and fed-batch cultivation. *Biotechnology Letters*, **19**, 549–552.

Sampaio JP, Gadanho M, Santos S *et al.* (2001) Polyphasic taxonomy of the basidiomycetous yeast genus *Rhodosporidium*: *Rhodosporidium kratochvilovae* and related anamorphic species. *International Journal of Systematic and Evolutionary Microbiology*, **51**, 687–697.

© 2014 John Wiley & Sons Ltd

Sanmartín I, Ronquist F (2004) Southern Hemisphere biogeography inferred by event-based models: plant versus Animal patterns. *Systematic Biology*, **53**, 216–243.

Schmidt I, Schewe H, Gassel S *et al*. (2011) Biotechnological production of astaxanthin with *Phaffia rhodozyma/Xanthophyllomyces dendrorhous*. *Applied Microbiology and Biotechnology*, **89**, 555–571.

Schoch CL, Seifert KA, Huhndorf S *et al*. (2012) Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. *Proceedings of the National Academy of Sciences*, **109**, 6241–6246.

Schroeder WA, Johnson EA (1995) Singlet oxygen and peroxyl radicals regulate carotenoid biosynthesis in *Phaffia rhodozyma*. *Journal of Biological Chemistry*, **270**, 18374–18379.

Stamatakis A (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics*, **22**, 2688–2690.

van Steenis C (1971) *Nothofagus*, key genus of plant geography, in time and space, living and fossil, ecology and phylogeny. *Blumea*, **19**, 65–98.

Swofford D (2003) *PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods)*. Version 4. Sinauer Associates, Sunderland, MA.

Tamura K, Peterson D, Peterson N *et al*. (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution*, **28**, 2731–2739.

Taylor JW, Jacobson DJ, Kroken S, *et al*. (2000) Phylogenetic species recognition and species concepts in fungi. *Fungal Genetics and Biology*, **31**, 21–32.

Taylor J, Spatafora J, O'Donnell K *et al*. (2004) The relationship of fungi. In: *Assembling the Tree of Life* (eds Cracraft J, Donoghue M), pp. 171–194. Oxford University Press, Oxford.

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, **22**, 4673–4680.

Thorsten LH, Buchanan PK, May TW, Mueller GM (2008) Phylogeography and biogeography of fungi. *Mycological Research*, **112**, 423–424.

Tiffney B (1985) The Eocene North Atlantic land bridge: its importance in tertiary and modern phytogeography of the Northern Hemisphere. *Journal of the Arnold Arboretum*, **66**, 243–273.

Tognetti C, Moliné M, van Broock M, Libkind D (2013) Favored isolation and rapid identification of *Xanthophyllomyces dendrorhous* (*Phaffia rhodozyma*) from environmental samples. *Journal of Basic Microbiology*, **53**, 766–772.

Veevers JJ, Powell CM, Roots SR (1991) Review of seafloor spreading around Australia. I. synthesis of the patterns of spreading. *Australian Journal of Earth Sciences*, **38**, 373–389.

Xiang Q, Thomas D (2008) Tracking character evolution and biogeographic history through time in Cornaceae – does choice of methods matter? *Journal of Systematics and Evolution*, **46**, 349–374.

D.L. and J.P.S. conceived and designed the research; M.D.-P., D.L. and J.P.S. performed the experiments and analysed the data; M.D.-P. and J.P.S. wrote the manuscript with advice from D.L.

Data accessibility

DNA sequences deposited at GenBank with accession numbers JN636878–JN637130.

Final DNA matrix for multilocus and ITS data and corresponding phylogenetic trees are deposited in TreeBASE: http://purl.org/phylo/treebase/phylows/study/TB2:S15043.

Supporting information

Additional supporting information may be found in the online version of this article.

Fig. S1. ITS phylogeny (a) maximum parsimony and (b) maximum likelihood of *Phaffia* and *Cystofilobasidium* rooted with *Filobasidella depauperata*.

Fig. S2. Ancestry of *Phaffia* estimated with STRUCTURE.

Fig. S3. Individual gene phylogenies by both maximum parsimony (a–h) and maximum likelihood (i–o) methods, next to each phylogeny, the corresponding gene is identified.

Fig. S4. Host tree maximum likelihood phylogeny (model GTR+G+I).

Table S1 List of genes and genetic regions amplified with respective PCR conditions and GenBank accession numbers of the resulting sequences produced in this study.

Table S2 List of strains used in the study and relevant information.

© 2014 John Wiley & Sons Ltd