MOLECULAR ECOLOGY

Molecular Ecology (2014) **23**, 2031–2045 doi: 10.1111/mec.12702

Population structure and reticulate evolution of *Saccharomyces eubayanus* and its lager-brewing hybrids

DAVID PERIS,* KAYLA SYLVESTER,* DIEGO LIBKIND,† PAULA GONÇALVES,‡ JOSÉ PAULO SAMPAIO,‡ WILLIAM G. ALEXANDER* and CHRIS TODD HITTINGER*

*Laboratory of Genetics, Genome Center of Wisconsin, Wisconsin Energy Institute, DOE Great Lakes Bioenergy Research Center, University of Wisconsin-Madison, Madison, WI 53706, USA, †Laboratorio de Microbiología Aplicada y Biotecnología, Instituto de Investigaciones en Biodiversidad y Medio-ambiente, INIBIOMA (CONICET-UNComahue), 8400 Bariloche, Argentina, ‡Centro de Recursos Microbiológicos, Departamento de Ciências da Vida, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, 2829-516 Caparica, Portugal

Abstract

Reticulate evolution can be a major driver of diversification into new niches, especially in disturbed habitats and at the edges of ranges. Industrial fermentation strains of yeast provide a window into these processes, but progress has been hampered by a limited understanding of the natural diversity and distribution of *Saccharomyces* species and populations. For example, lager beer is brewed with *Saccharomyces pastorianus*, an alloploid hybrid of *S. cerevisiae* and *S. eubayanus*, a species only recently discovered in Patagonia, Argentina. Here, we report that genetically diverse strains of *S. eubayanus* are readily isolated from Patagonia, demonstrating that the species is well established there. Analyses of multilocus sequence data strongly suggest that there are two diverse and highly differentiated Patagonian populations. The low nucleotide diversity found in the *S. eubayanus* moiety of hybrid European brewing strains suggests that their alleles were drawn from a small subpopulation that is closely related to one of the Patagonian populations. For the first time, we also report the rare isolation of *S. eubayanus* outside Patagonia, in Wisconsin, USA. In contrast to the clear population differentiation in Patagonia, the North American strains represent a recent and possibly transient admixture of the two Patagonian populations. These complex and varied reticulation events are not adequately captured by conventional phylogenetic methods and required analyses of Bayesian concordance factors and phylogenetic networks to accurately summarize and interpret. These findings show how genetically diverse eukaryotic microbes can produce rare but economically important hybrids with low genetic diversity when they migrate from their natural ecological context.

Keywords: admixture, hybridization, lager beer, phylogeography, *Saccharomyces eubayanus*, *Saccharomyces pastorianus*

Received 3 November 2013; revision received 17 February 2014; accepted 19 February 2014

Introduction

The process of hybridization between species and populations has long been known to have the potential to generate new varieties of plants and animals. Indeed, many crop species are recent or ancient interspecies hybrids, including wheat, maize, sugar cane, coffee, cotton and tobacco. Interspecies hybridization and admixture are less frequent in animals, but prominent examples have been described in insects, fishes, amphibians, reptiles and mammals [for a review, see Otto (2007)], including in primates (Zinner *et al.* 2011) and even suggested in ancient humans (Arnold 2008). Often, these types of reticulate evolutionary events can be beneficial in novel environments where the parental species or populations are not locally adapted (Verhoeven *et al.* 2011), but the creative potential of hybridization has been less thor-

Correspondence: Chris Todd Hittinger, Fax: +1 (608) 262-2976; E-mail: cthittinger@wisc.edu

oughly studied in eukaryotic microbes, in part due to the challenges of identifying the wild sources of the alleles found in hybrids.

The Saccharomycotina or hemiascomycete yeasts comprise a major eukaryotic subphylum with about 1000 described species, including several hybrids, especially in the *Saccharomyces* genus [for a review, see Morales & Dujon (2012)]. Unfortunately, little is known about the ecology, biogeography and population structure of most of the seven naturally occurring *Saccharomyces* species (Kurtzman *et al.* 2011; Hittinger 2013). Despite displaying predominantly vertical inheritance within species and lineages (Rokas *et al*. 2003; Dietrich *et al*. 2004; Holland *et al*. 2004; Yu *et al*. 2012), *Saccharomyces* yeasts provide examples of all of the major types of reticulation, including interspecies hybridization, mosaic lineages generated by admixture, introgression, horizontal gene transfer (HGT) and intragenic recombination (Liti *et al.* 2006, 2009; Novo *et al.* 2009; Dunn *et al*. 2012; Peris 2012; Peris *et al.* 2012a,c; Gladieux *et al*. 2014). Although the ecological forces favouring reticulation are not always well understood, interspecies hybrids have an advantage over parents in some industrial fermentation conditions, such as low-temperature wine-making and lager-brewing (Belloch *et al.* 2008; Gibson *et al.* 2013).

Reconstruction of the relationships of taxonomic groups that have undergone reticulate events requires a new layer of evolutionary thinking. Phylogenetic networks show considerable promise in aiding in the interpretation of conflicting phylogenetic signals (Bapteste *et al.* 2013). Using these network-based methods, incongruent data are visualized by connecting a taxon or clade with two or more distance-weighted edges to all of the lineages contributing to its evolution. Despite the potential of supernetworks, including application to the analysis of short internodes between *Saccharomyces* speciation events (Holland *et al.* 2004) and to the detection of recombination between *Saccharomyces* species in the mitochondrial-encoded gene *COX2* (Peris 2012; Peris *et al.* 2012a), their application to the study of the important biological processes of admixture and hybridization has been limited.

In the last several thousand years, humans domesticated multiple lineages of *S. cerevisiae* for wine-making, brewing and sake fermentation (Fay & Benavides 2005). Double and triple hybrids between *Saccharomyces* species have also been described in beer, wine, cider, dietary supplements and clinical samples (Masneuf *et al.* 1998; Le Jeune *et al.* 2007; González *et al.* 2008; Peris *et al.* 2012a). The lager-brewing yeast *S. pastorianus* is one of the best-known and most commercially important interspecies hybrids. Comparative genomic hybridization and DNA sequence data from the *S. cerevisiae* parents have convincingly established that at least two major groups of lager-brewing yeast, the Saaz and Frohberg lineages, resulted from two independent hybridization events between *S. cerevisiae* ale strains and *Saccharomyces eubayanus* (Dunn & Sherlock 2008; Libkind *et al.* 2011). Multiple independent hybridizations also appear to have given rise to *S. cerevisiae* × *S. kudriavzevii* hybrids (Erny *et al.* 2012; Peris *et al.* 2012b) and to *S. bayanus* triple hybrids containing genetic contributions from *S. cerevisiae*, *S. eubayanus* and *S. uvarum* (Libkind *et al.* 2011)[1].

The discovery of *S. eubayanus* in association with *Nothofagus* (southern beech) trees in Patagonia, Argentina, identified the second parental species of *S. pastorianus* hybrids and provided a model for their evolution (Libkind *et al.* 2011). Despite the high (99.56%) identity across the genome, key differences exist between the type strain of *S. eubayanus* and the *S. eubayanus* moiety found in domesticated *S. pastorianus*. Some differences, such as the inactivation of *SUL1* (a high-affinity sulfate permease), probably reflect the process of domestication (Libkind *et al.* 2011), but most sequence differences are expected to be neutral accumulated divergence or sites segregating within *S. eubayanus*. Broader surveys of *S. eubayanus* diversity are therefore necessary to determine which alleles from wild populations are most closely related to the alleles found in the interspecies hybrids present in the brewing environment and to infer which genetic changes occurred during domestication.

To better understand the complex reticulate evolution and domestication of hybrids containing *S. eubayanus* alleles, we launched a global effort to characterize the genetic diversity of *S. eubayanus* and its interspecies hybrids. Here, we combine population and phylogenetic supernetwork approaches to infer the genetic structure of *S. eubayanus* in nature and the history of its reticulation events. We also trace the relationships between wild and brewing strains in the context of hybridization and the exploration of new ecological niches.

Materials and methods

Yeast isolation and culture media

The complete yeast surveys will be described in more detail elsewhere, but *Saccharomyces eubayanus* was recovered from Patagonia using the 10 °C enrichment and isolation protocol of Sampaio and Gonçalves (2008). Outside Patagonia, this protocol and several other

[1]Most molecular geneticists study derivatives of CBS 7001, a pure European strain from the *S. uvarum* lineage of the *S. eubayanus*/*S. uvarum* species complex (Cliften *et al.* 2003, 2006; Kellis *et al.* 2003; Scannell *et al.* 2011; Caudy *et al.* 2013; Hittinger 2013).

protocols were deployed on samples from Europe, Asia, Oceania and North America. All non-Patagonian *S. eubayanus* strains came from a single site in North America and were enriched at 10 °C in synthetic complete media with 8% glucose as the sole carbon source (without ethanol). Representatives from more than 200 wild strains isolated in Patagonia were selected based on preliminary MSP-PCR fingerprinting data, which was performed as previously described (Libkind *et al.* 2011). Yeast strains used in this study (Table 1) were grown in YPD medium (2% glucose, 2% peptone and 1% yeast extract).

PCR amplification, sequencing and nucleotide sequences

Partial gene sequences were obtained for nine nuclear genes using primers and conditions described in Table S1 (Supporting Information): *DCR1* (*Sbay_13.48* following the Scannell *et al.* (2011) annotation of *S. uvarum* CBS 7001), *FSY1* (*LBYG08543* following the Nakao *et al.* (2009) annotation of *S. pastorianus* Weihenstephan 34/70), *FUN14*, *GDH1*, *HIS3*, *MET2*, *RIP1*, *URA3* and the *ITS* region of the *rDNA* locus (containing *ITS1*, *5.8S* and *ITS2*). Mitochondrial inheritance was assessed by amplifying and sequencing part of *COX2* (Belloch *et al.* 2000), which corresponds to positions 179–708 of the *S. cerevisiae* S288c *COX2* gene. We could not amplify yHCT96 *COX2* because it was a ρ^- petite (confirmed by its inability to grow with glycerol as the sole carbon source). Gene sequences were determined by colony-PCR and Sanger sequencing. Sequences were edited and assembled with STADEN Package version 1.7 (Staden *et al.* 2000). Sequences were deposited in GenBank under Accession nos. KF530330-KF530542 and KJ412200.

Nuclear gene sequences of the lager hybrid yeast *S. pastorianus* Weihenstephan 34/70 were obtained using the BLAST search tool (Altschul *et al.* 1990) against the *S. pastorianus* genome project ABPO00000000 (Nakao *et al.* 2009) and mtDNA genome sequence Accession no. EU852811.1 (Nakao *et al.* 2009). Gene sequence Accession nos. of the triple hybrid strains *S. cerevisiae* × *S. eubayanus* × *S. uvarum* (CBS 380, CBS 1546 and NBRC 1948) were previously described (Rainieri *et al.* 2008; Libkind *et al.* 2011; Peris 2012). For sequences that were heterozygous for *S. uvarum*/*S. eubayanus* alleles (annotated using IUPAC ambiguity codes in GenBank), we inferred both the *S. eubayanus* and *S. uvarum* alleles by comparing them with the reference strains FM1318 (yHCT76) and CBS 7001, respectively. All sequences for FM1318 and CBS 7001 were previously described (Libkind *et al.* 2011; Scannell *et al.* 2011), except the *ITS* region of CBS7001 and the *GDH1* and *COX2* genes of FM1318 and CBS 7001.

Multiple sequence alignments and individual gene trees

Gene sequences were aligned using CLUSTALW, as implemented in MEGA 5.1 (Tamura *et al.* 2011), and manually trimmed. Because *S. eubayanus* yeast strains were homozygous at the loci examined and *Saccharomyces* yeasts frequently autodiploidize and generally reproduce by clonal divisions (Tsai *et al.* 2008), we considered *S. eubayanus* to be haploid for subsequent analyses. We calculated Tajima's D (Tajima 1989) and Fu's Fs (Fu 1997) statistics in DNASP version 5 (Librado & Rozas 2009) to test for selection or unusual demography.

Individual phylogenetic trees were reconstructed using the maximum-likelihood (ML) method under the best-fit evolutionary model following the Bayesian Information Criterion (BIC), as implemented in MEGA 5.1 (Tamura *et al.* 2011). The *ITS* region was used to confirm species identification due to its status as a barcode gene. However, *ITS* was removed from downstream analyses due to the lack of variation within *S. eubayanus* and the presence of a recombinant (*S. cerevisiae* × *S. eubayanus*) sequence in the hybrid lager-brewing strain W34/70.

Recombinant-free sequence blocks were generated using IMGC (Woerner *et al.* 2007), removing blocks that violate the four-gamete test, such in *DCR1*, *FSY1*, *GDH1*, *MET2* and *URA3*. These recombinant-free sequences were concatenated into ~4 kb of nuclear sequence using FASCONCAT version 1.0 (Kück & Meusemann 2010). This recombinant-free alignment was used in the time-calibrated tree reconstruction and population size inferences because these methods assume no recombination.

Population structure

To delimit populations and infer the evolutionary history of the strains, we used the program STRUCTURE version 2.3.4 (Pritchard *et al.* 2000; Falush *et al.* 2003; Hubisz *et al.* 2009) after converting our FASTA file into STRUCTURE input format using SEQPHASE (Flot 2010). We assumed the admixture model and estimated the number of genetic clusters, K, testing from $K = 1$ to $K = 6$ subpopulations, and correlated allele frequencies with five parallel Markov chains run for all models of K with 200 000-iteration burn-ins and 500 000 iterations of sampling. STRUCTURE output data were used as input for STRUCTURE HARVESTER version 0.6 (Earl & vonHoldt 2012), which allowed us to compare the likelihood ratios associated with each K. Output data from STRUCTURE HARVESTER were visualized in CLUMPP version 1.1.2 (Jakobsson & Rosenberg 2007) and DISTRUCT version 1.1 (Rosenberg 2004). The fixation index (F_{ST}) was calculated

Table 1 List of strains used in this study

Name in Fig.	Species designation	Isolation region	Host	Substrate	Latitude	Longitude	Altitude	Temp °C*	Prcp† (mm)	Rad‡ (W/m²d)
W34/70§	*S. pastorianus*†††	Germany	Lager beer	Frohberg group	48.425555	11.731796	NA	Industrial	Industrial	Industrial
yHCT61¶	*S. eubayanus*	Nahuel Huapi, Argentina	*Nothofagus antarctica*	*Cyttaria harioti*	−41.36003	−71.5645	814	8.2	1304	153.42
yHCT62¶	*S. eubayanus*	Nahuel Huapi, Argentina	*Nothofagus dombeyi*	*Cyttaria harioti*	−41.35914	−71.51458	870	7.9	1178	153.42
yHCT63¶	*S. eubayanus*	Lanin, Argentina	*Nothofagus obliqua*	Bark	−40.149	−71.49942	815	8.0	1103	160.99
yHCT70¶	*S. eubayanus*	Villa Pehuenia, Argentina	*Nothofagus obliqua*	Bark	−38.83522	−71.26536	1340	8.4	1202	172.55
yHCT72¶	*S. eubayanus*	Villa Pehuenia, Argentina	*Nothofagus antarctica*	Bark	−38.83758	−71.22919	1487	7.4	1078	172.55
yHCT76**	*S. eubayanus*	Nahuel Huapi, Argentina	*Nothofagus dombeyi*	*Cyttaria harioti*	−41.35689	−71.51567	906	7.8	1166	153.42
CBS 1503††	*S. pastorianus*†††	Carlsberg brewery	Lager beer	Saaz group	56.035226	12.454834	NA	Industrial	Industrial	Industrial
yHCT88¶	*S. eubayanus*	Nahuel Huapi, Argentina	*Nothofagus pumilio*	Bark	−41.25015	−71.28339	1240	5.4	774	153.34
yHCT90¶	*S. eubayanus*	Nahuel Huapi, Argentina	*Nothofagus antarctica*	Bark	−41.24078	−71.18703	990	7.2	750	153.34
yHCT91¶	*S. eubayanus*	Nahuel Huapi, Argentina	*Nothofagus dombeyi*	*Cyttaria harioti*	−41.13165	−71.32996	782	8.6	991	153.34
yHCT92¶	*S. eubayanus*	Nahuel Huapi, Argentina	*Nothofagus antarctica*	*Cyttaria harioti*	−41.35059	−71.60139	815	7.4	1256	153.42
yHCT94¶	*S. eubayanus*	Nahuel Huapi, Argentina	*Nothofagus antarctica*	Bark	−41.38943	−71.49593	843	8.2	1188	153.34
yHCT96¶	*S. eubayanus*	Nahuel Huapi, Argentina	*Nothofagus antarctica*	Bark	−41.11494	−71.4426	812	8.4	1101	153.34
yHCT99¶	*S. eubayanus*	Nahuel Huapi, Argentina	*Nothofagus antarctica*	Bark	−41.24078	−71.18703	990	7.2	750	153.34
yHCT101¶	*S. eubayanus*	Nahuel Huapi, Argentina	*Nothofagus* sp.	Exudate	−41.11494	−71.4016	816	8.6	1081	153.34
yHCT104¶	*S. eubayanus*	Nahuel Huapi, Argentina	*Nothofagus antarctica*	Soil	−41.24078	−71.18703	990	7.2	750	153.34
yHCT105¶	*S. eubayanus*	Nahuel Huapi, Argentina	*Nothofagus pumilio*	Soil	−40.71897	−71.93669	1260	6.0	1512	157.66
yHCT107¶	*S. eubayanus*	Villa Pehuenia, Argentina	*Nothofagus pumilio*	Soil	−38.83064	−71.26536	1310	8.1	1178	172.55
yHCT114¶	*S. eubayanus*	Villa Pehuenia, Argentina	*Nothofagus obliqua*	Soil	−38.83522	−71.26536	1340	8.4	1202	172.55
yHKS210¶	*S. eubayanus*	Sheboygan, WI, USA	*Fagus grandifolia*	Bark	43.69611	−87.71778	187	7.7	800	128.27
yHKS211¶	*S. eubayanus*	Sheboygan, WI, USA	*Fagus grandifolia*	Bark	43.69611	−87.71778	187	7.7	800	128.27
yHKS212¶	*S. eubayanus*	Sheboygan, WI, USA	*Acer saccharum*	Bark	43.69611	−87.71778	187	7.7	800	128.27
CBS 7001‡‡	*S. uvarum*	Ávila, Spain		*Mesophylax adopersus*						
NBRC 1948§§	*S. bayanus*‡‡‡	Europe		Brewing contaminant	48.458352	8.232418	NA	Industrial	Industrial	Industrial
CBS 380¶¶	*S. bayanus*‡‡‡	Europe		Turbid beer	48.458352	8.232418	NA	Industrial	Industrial	Industrial
CBS 1546***	*S. bayanus*‡‡‡	Netherlands		Beer	52.106505	5.522459	NA	Industrial	Industrial	Industrial

*Current mean annual temperature.
†Annual mean precipitation.
‡Radiation.
Strain references: §Commercial; ¶This study; **(Libkind *et al.* 2011) (a monosporic derivative of the *S. eubayanus* type strain, CBS 12357T); ††(Groth *et al.* 1999); ‡‡(Scannell *et al.* 2011); §§(Rainieri *et al.* 2006); ¶¶(Kurtzman & Robnett 1991); ***(Montrocher *et al.* 1998).
†††*S. pastorianus* syn. *S. carlsbergensis* are interspecies hybrids between *S. cerevisiae* and *S. eubayanus* (Libkind *et al.* 2011).
‡‡‡*S. bayanus* are *S. eubayanus* × *S. uvarum* hybrids, some of which also have contributions from *S. cerevisiae* (Libkind *et al.* 2011).

between the STRUCTURE-inferred populations, and analysis of molecular variance (AMOVA) was performed in ARLEQUIN version 3.5 (Excoffier & Lischer 2010).

Genetic diversity

DNASP version 5 (Librado & Rozas 2009) was used to calculate genetic diversity statistics for each locus, such as the number of polymorphic sites (s), average number of differences between sequences (k), nucleotide diversity (π), number of haplotypes and haplotype diversity (Hd). Genetic diversity statistics were also calculated for each STRUCTURE-inferred population and between populations. The uncorrected and Tamura–Nei genetic distances were calculated within and between each STRUCTURE-inferred population using MEGA 5 (Tamura *et al.* 2011).

Divergence time reconstruction

To estimate divergence times, we first inferred the number of generations possible per year. *S. eubayanus* strains were grown in minimal media [6.7 g YNB with ammonium sulphate without amino acids (Amresco, USA)] + 2% glucose at 8 °C. These conditions were selected based on the average annual temperature of the Patagonian sampling sites and the likely rarity of rich conditions, such as YPD. OD_{595} was monitored in a BMG Labtech FLUOstar (BMG Labtech, USA). Background signal was removed using custom R scripts, and growth curve parameters were obtained using GCAT (http://www.glbrc.org/gcat-vm/). To test for growth rate differences between populations, a one-way analysis of variance (ANOVA) statistical test was performed using STATISTICA 7 (Hilbe 2007). To calibrate the molecular clock, we used the *S. cerevisiae* mutation rate of $0.33*10^{-9}$ substitutions/bp/generation (Lynch *et al.* 2008). Divergence times were obtained using a concatenated alignment of fourfold degenerate sites. Three independent runs of MCMC length 10^7 were performed in BEAST version 1.7.5 (Drummond & Rambaut 2007) with sampling every 1000 steps; convergence of posterior probabilities was monitored with TRACER version 1.5 (Rambaut & Drummond 2001). Convergence was confirmed when the estimated sample size (ESS) values were >300, and independent runs were combined using LOGCOMBINER from the BEAST package. To obtain the final tree, we used TREEANNOTATOR from the BEAST package. We discarded the first 10% of generations from each run as a burn-in. The calibrated tree with time divergences and 95% highest posterior density (HPD) of node age estimates was observed in FIGTREE version 1.3.1 (Rambaut & Drummond 2010).

Population differentiation: isolation-by-distance and isolation-by-ecology analyses

Tajima's D (Tajima 1989) and Fu's Fs (Fu 1997) were calculated, with 1000 permutations, for each STRUCTURE-inferred population using ARLEQUIN version 3.5 (Excoffier & Lischer 2010). In addition, extended Bayesian skyline plots (eBSPs) (Heled & Drummond 2008) were produced using BEAST, with a MCMC of length 10^6, sampling every 1000 steps, and three parallel runs that achieved ESS >300. eBSPs were represented using the script supplied in the eBSP tutorial (Heled 2010).

In order to study the possible mechanisms of population differentiation, we performed isolation-by-distance and isolation-by-ecology analyses. In addition to sampling information (e.g. host, substrate), GPS points for localities were entered into DIVA-GIS version 7.5 (Hijmans *et al.* 2001). We extracted current climates (BIO1: annual mean temperature, BIO12: annual mean precipitation) and last glacial maximum (BIO1, BIO12) grids from worldclim.org (Hijmans *et al.* 2005). Radiation grid (BIO20: annual mean radiation) was obtained from http://www.climond.org. Mantel tests (Sokal & Rohlf 1995) were performed in IBD WEB SERVICE version 3.23 (Jensen *et al.* 2005). Specifically, using 1000 permutations and the Rousset's distance measure (Rousset 1997), we tested for a correlation between genetic distance (F_{ST}), corrected by the Kimura 2-parameter model, and the geographical distance matrix generated using GEOGRAPHIC DISTANCE MATRIX GENERATOR version 1.2.3 (http://biodiversityinformatics.amnh.org/open_source/gdmg/index). Principal component analysis was also performed on ecological traits using the RGL package in the R statistical package (Adler & Murdoch 2009). The ecological dissimilarity matrix was calculated using the Euclidean distance method implemented in the ECODIST package of R (Sarah & Goslee 2007). Scatter plots and Pearson's correlation versus genetic distance were examined in STATISTICA 7 (Hilbe 2007).

Phylogenetic networks and supernetworks

A nexus file with the collection of ML trees of the nuclear genes (except *ITS*) was the input for SPLITSTREE 4 for super split network (supernetwork) reconstruction. This method was selected because some gene sequences were absent from the triple-hybrid brewing contaminants. Edges' weights were calculated using the tree size-weighted means option, which graphs the average genetic distance obtained from each tree (Huson *et al.* 2004). The NeighborNet (NN) method was employed for *COX2* phylogenetic network reconstruction in SPLITSTREE 4 (Huson & Bryant 2006). To test for recombinant sequences, we used RDPV4 (Martin *et al.* 2010).

Bayesian concordance analysis among gene trees

To provide an estimate of the level of concordance among individual phylogenetic gene trees, we performed Bayesian concordance analysis (BCA) (Ané *et al.* 2007). One of the useful descriptive statistics obtained from BCA is the clade concordance factor (CF), which describes the proportion of genes that contain a particular clade (Baum 2007). Two BCAs were performed, one in which the North American admixture strains were included and one in which the admixture strains were excluded. We reconstructed the individual phylogenetic trees using MrBayes version 3.2.1 (Ronquist *et al.* 2012). We selected the best-fit evolutionary model using MEGA 5.1 (Tamura *et al.* 2011). Two independent runs for each gene alignment were used with the default parameters. Chains were run for one million generations, sampling every 100 generations, for a total of 10 000 samples. We discarded 10% of generations as burn-in. CBS 7001 was used as the outgroup. In all cases, replicate analyses converged on the same posterior distribution, as observed using TRACER version 1.5 (Rambaut & Drummond 2001). We used the *mbsum* command, included in BUCKy version 1.4.2 (Larget *et al.* 2010), to combine the independent lists of tree topologies and posterior probabilities into one file for each gene. The combined file for each gene was the input for BUCKy version 1.4.2. Two replicate analyses were run for three different α values as priors (0.1, 1 and 10). $\alpha = 0$ indicates that all posterior distributions are represented by the same trees; $\alpha = \infty$ indicates that each gene has a distinct set of trees. We performed a MCMC of one million generations after a burn-in period of 100 000 generations. We applied this MCMC for the eight genes used in MRBAYES. CFs were calculated for all possible bipartitions in the 24- and 21-tip trees. From these CFs, primary concordance trees were reconstructed from the set of bipartitions with the highest overall CFs. In the supernetwork, we have provided the concordance results for key clades as the CF and its 95% credibility interval.

Results

Multilocus sequence diversity and relationships

To characterize the genetic diversity and phylogenetic relationships among wild *Saccharomyces eubayanus* and their domesticated hybrids, we sequenced portions of nine nuclear genes and one mitochondrial gene, resulting in a total of ~6.78 kbp for each strain. Summary statistics revealed no unusual signatures of selection (Table 2). Individual genes displayed variable levels of diversity and several alternative topologies (Fig. S1, Supporting Information). The *ITS* locus differentiated *S. uvarum* from *S. eubayanus* strains by a single base pair. *ITS* contained no polymorphisms within *S. eubayanus* (Fig. S1I, Supporting Information), so we excluded it from subsequent analyses. The gene with the highest genetic diversity (k, π, number of haplotypes and Hd) was the budding yeast *Dicer* (*DCR1*) gene, presumably because most strains contained premature stop codon(s) in the region sequenced, except for yHCT72, yHCT90, yHCT99 and yHCT114 (Fig. S1G, Supporting Information). Interestingly, hybrid brewing strains had particularly differentiated alleles of the subtelomeric *GDH1* and *FSY1* genes (Fig. S1C, E, Supporting Information),

Table 2 Summary statistics for one mitochondrial and nine nuclear genes

Gene name	Systematic name	bp/bp*	*s*	*k*	π	#hap	Hd	Fs	Tajima's D
COX2	*Q0250*	530	23	5.545	0.01046 ± 0.00305	12	0.905 ± 0.039	−1.767	−0.45565
DCR1	*Sbay_13.48*	859/428	46	10.332	0.01207 ± 0.00114	16	0.968 ± 0.02	−2.409	−0.66639
FSY1	*LBYG08543**	1218/670	21	4.901	0.00403 ± 0.00066	12	0.905 ± 0.041	−2.062	−0.51242
FUN14	*YAL008W*	447	5	1.359	0.00304 ± 0.00032	6	0.779 ± 0.059	−1.127	0.01076
GDH1	*YOR375C*	677/481	23	6.075	0.00897 ± 0.00083	15	0.953 ± 0.025	−4.032	−0.24591
HIS3	*YOR202W*	537	8	1.936	0.00361 ± 0.00077	6	0.569 ± 0.114	−0.121	−0.34542
ITS	*ITS*†	693	0	0	0.0 ± 0.0	1	0.0 ± 0.0	N.A.	N.A.
MET2	*YNL277W*	513/383	9	2.561	0.00499 ± 0.0005	9	0.885 ± 0.038	−1.911	0.16682
RIP1	*YEL024W*	511	7	2.553	0.005 ± 0.00052	6	0.739 ± 0.079	0.692	1.09472
URA3	*YEL021W*	796/485	12	3.834	0.00483 ± 0.00032	7	0.767 ± 0.07	1.115	0.62113

bp: fragment length in base pairs; *bp: base pairs used in the concatenated alignment without recombinant segments that violate the four-gamete test; *s*: number of segregating sites; *k*: average number of differences between sequences; π: nucleotide diversity; #hap: number of haplotypes; Hd: haplotype diversity; Fs: Fu's Fs; Tajima's D (no values are statistically significant, $P < 0.05$).

*Located in the subtelomeric region of chromosome IV.

†Gene encoding portions of the internal transcribed spacer 1, 5.8S ribosomal RNA, internal transcribed spacer 2 and the 28S ribosomal RNA gene. Located on Chromosome XII. *ITS* sequences from lager-brewing strains were removed for this analysis because the W34/70 allele was recombinant.

which are known to play important roles during brewing in nitrogen (Godard *et al.* 2007) and fructose metabolism (Anjos *et al.* 2013), respectively. Although some Patagonian strains were subject to incomplete lineage sorting at specific loci, the placement of the North American strains was particularly variable.

Structure and admixture of two Patagonian populations

To infer the number of natural populations represented by our strain collection, we performed several simulations using the STRUCTURE software (Pritchard *et al.* 2000; Earl & vonHoldt 2012). These simulations consistently recovered two populations. ΔK, the rate of change in the log probability of data between successive cluster (K) values (Evanno *et al.* 2005), was highest when $K = 2$ ($\Delta K = 1164.5$). At higher K values, the ΔK value was not significantly different from zero (e.g. at $K = 3$, $\Delta K = 0.77$), and the results were stochastic. For example, $K = 3$ barplots varied radically between independent runs (Fig. 1B). These results led us to conclude that the data only support two populations.

Analysis of molecular variance (AMOVA) provided further support for strong structure in our data ($P < 10^{-4}$) with most of the genetic variation existing between the populations suggested by STRUCTURE (~73%) (Table 3). In addition to containing the type strain and the majority of wild strains of *S. eubayanus* from Patagonia, one of these populations also contained the Saaz and Frohberg lager-brewing strains, so we called it the 'Patagonia B (Lager)' population. We simply named the second population the 'Patagonia A' population. Interestingly, the North American strains appeared to be the result of admixture between the Patagonia A and Patagonia B (Lager) populations, having membership coefficients of 0.53 and 0.47 for the Patagonia A cluster and the Patagonia B (Lager) cluster, respectively (Fig. 1A).

The Patagonian populations are diverse and well differentiated

The distributions of single-nucleotide polymorphisms (SNPs) provided further support for the existence of two well-differentiated populations. The Patagonia A and Patagonia B (Lager) populations had 23 fixed and only four shared SNPs (Fig. 1C). The populations had 44 and 57 private SNPs, respectively. Similarly, analysing the lager-brewing strains and the wild populations separately revealed 41 fixed differences between the lager-brewing strains and the Patagonia A population. In contrast, there were only 15 fixed differences between the lager-brewing strains and the wild representatives of the Patagonia B (Lager) population, more

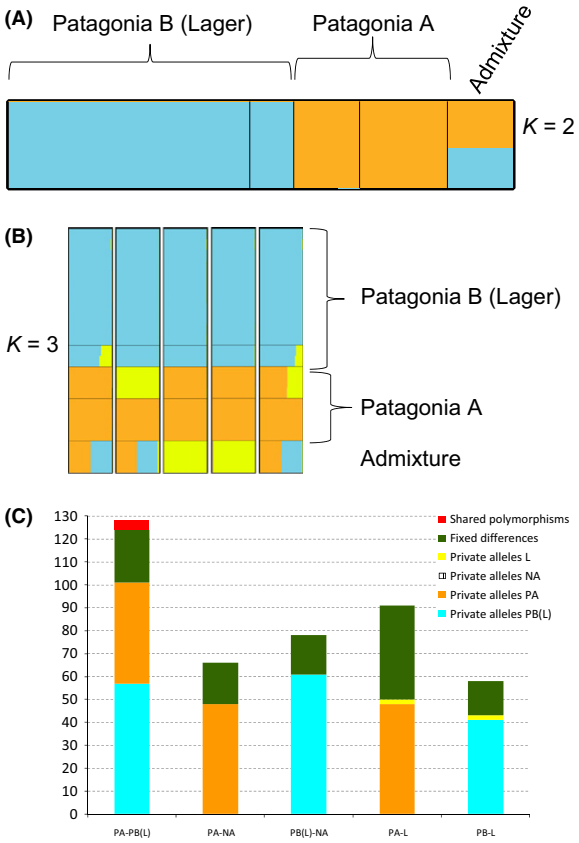

Fig. 1 Population structure and summary statistics of SNPs. (A) Inference of the genetic clusters (K) and composition of individuals by STRUCTURE. The most consistently supported number of genetic clusters/populations was $K = 2$ with a ΔK_2 value = 1164.5 ($\Delta K_3 = 0.77$). (B) Barplots for five independent $K = 3$ runs yielded variable, conflicting results. Each colour in (A) and (B) bar plots represents the cluster membership coefficients, and a mixture of colours suggests admixture. (C) Numbers of private segregating alleles, fixed differences and shared polymorphisms among SNPs found in pairwise comparison between populations or groups. PA, Patagonia A; PB(L), Patagonia B (Lager), NA, North America; L, *Saccharomyces eubayanus* moiety of *S. pastorianus* lager-brewing strains; PB, Patagonia B (Lager) population, excluding lager strains.

than a third of which were in *FSY1*. The North American strains had no private alleles and had nearly the same number of fixed differences when compared either to the Patagonia A population or to the Patagonia B (Lager) population (18 vs. 17, respectively; Fig. 1C), observations consistent with recent admixture.

Although the nucleotide diversity of the hybrid European lager-brewing strains was extremely low (Table S2D, Supporting Information $\pi = 0.0004$ with no variation at 6/9 nuclear *S. eubayanus* loci) and the admixed North American strains were identical at all genes examined, both *S. eubayanus* populations proved to be

Table 3 Analysis of molecular variance (AMOVA) of STRUCTURE-inferred populations

	Sum of d.f.	Variance squares	Variance components	Percentage components of variation
Among populations	1	301.598	31.865	73.27
Within populations	18	209.268	11.626	26.73
Total	19	510.866	43.491	

F_{ST} = 0.73268.
$P < 10^{-4}$.

remarkably diverse in Patagonia (Fig. 1C, Tables 4 and S2, Supporting Information). Extended Bayesian skyline plots (eBSPs) (Heled & Drummond 2008) imply that both natural populations of *S. eubayanus* have maintained a constant effective population size of around 20–30 million (Fig. S2, Supporting Information), suggesting that the Patagonian populations have been consistently large and diverse. The Patagonia B (Lager) effective population size may have decreased recently (Fig. S2B, Supporting Information), but this was likely driven by a strong lineage-specific bottleneck during the origin of hybrid lager-brewing strains.

The Patagonia A and Patagonia B (Lager) populations were highly divergent and differentiated from one another with a genetic divergence of 0.93% (Table 5) and a F_{ST} value of 0.73. To obtain a minimum estimate for when the Patagonian populations diverged, we applied an ultrametric molecular clock. We calibrated the molecular clock using the growth rate of *S. eubayanus* in minimal media at 8 °C (43.48 h/generation or 201.43 generations/year), a rate that did not differ between populations (unequal N HSD as post hoc test; Fig. S3, Supporting Information). This conservative calibration suggests that the *S. eubayanus* populations started to diverge at least 150 000 years ago (100–223 kybp, 95% HPD) (Fig. S4, Supporting Information). These results also imply that the *S. eubayanus* strains that hybridized with *S. cerevisiae* to form the *S. pastorianus* lager-brewing strains began to diverge from the wild Patagonia B (Lager) strains studied here at least several thousand years ago.

Table 5 Average pairwise genetic distances within and between STRUCTURE-inferred populations and the admixture group

	Patagonia A	Patagonia B (Lager)	North America
Patagonia A	**0.003830** **0.003846**	0.009262	0.006028
Patagonia B (Lager)	0.009334	**0.003294** **0.003305**	0.005743
North America	0.006059	0.005771	**0** **0**

Main diagonal (boldface): Top entry is the average pairwise distance within the population. Bottom entry is average Tamura–Nei-corrected distance within the population.
Rows: Average pairwise distance between two populations.
Columns: Average Tamura–Nei-corrected distance between two populations.

Evidence for ecological and geographical differentiation among the strains from northwestern Patagonia was limited and equivocal. We found no evidence for isolation by distance (IBD) or isolation by ecology within populations (IBE) (Tables 1 and S3, Supporting Information). Two ecological traits (longitude and average annual precipitation) were marginally significant between populations ($P < 0.0215$ and $P < 0.0364$, Student's *t*-test).

Phylogenetic networks accurately summarize admixture and interspecies hybridization

In addition to the wild admixed or mosaic intraspecific hybrids of *S. eubayanus*, this species has contributed to several complex interspecies hybrids. To encapsulate these complex reticulation events, we performed a phylogenetic supernetwork reconstruction. This procedure clearly split the two natural species, *S. uvarum* and *S. eubayanus* (Fig. 2). Interspecies hybrids showed a wide range of contributions from *S. uvarum*, ranging from no detectable nuclear contributions for the *S. pastorianus* (*S. cerevisiae* × *S. eubayanus*) lager yeast hybrids W34/70 and CBS 1503, to a majority of alleles from *S. uvarum* in the *S. bayanus* triple hybrid CBS 380. Because the *S. eubayanus* alleles present in hybrid European

Table 4 Summary statistics for each STRUCTURE-inferred population and the admixture group

Clade	Sequences	#hap	Hd	π	Fs	Tajima's D	Effective pop size (Ne)
Patagonia A	7	7	1	0.00315	−1.194	0.70183	$3.36*10^7$
North America	3	1	0	0	N.A.	N.A.	N.A.
Patagonia B (Lager)	13	11	0.97	0.00198	−2.841	−0.60720	$1.93*10^7$

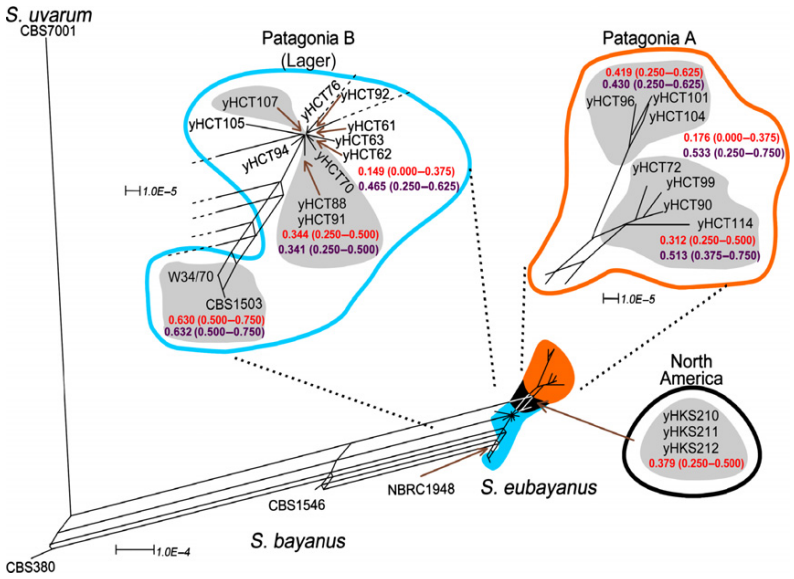

Fig. 2 A phylogenetic supernetwork captures reticulate evolutionary events. Phylogenetic supernetwork reconstructed using the maximum-likelihood (ML) trees of eight nuclear genes by the Z-closure method. Incongruent tree topologies are represented by nodes subtended by multiple edges. The scale bar represents the edges' weights inferred using the tree size-weighted means options, a measure similar to branch lengths in a phylogram. Orange and light blue shades correspond to the Patagonia A and Patagonia B (Lager) populations, respectively. The black shade corresponds to the admixed or mosaic North American strains. Grey shades highlight clades that were found only in the primary concordance tree obtained by BCA. Red numbers indicate the concordance factors (and 95% HPD) from BCA when the North American strains were included, while purple numbers show the values when the North American strains were excluded.

brewing strains were drawn from the Patagonia B (Lager) population or a closely related subpopulation, the supernetwork displays them along several close, nearly parallel edges with each interspecies hybrid strain's position determined primarily by the quantity of genetic contribution from *S. uvarum*. For example, the *S. bayanus* triple hybrids CBS 1546 and CBS 380 contain both full-length *S. eubayanus* and *S. uvarum* alleles, and they appear at intermediate locations between these two main groups with edges connecting them to both. For NBRC 1948, its position along the edge connecting it with *S. uvarum* is due entirely to *MET2* (Fig. S1F, Supporting Information), the only gene analysed that had a *S. uvarum* allele.

In contrast to the complex reticulate evolution in hybrid European brewing strains, the wild *S. eubayanus* strains branch into several well-supported nodes with few additional edges. Notably, the mosaic North American strains, which population genetic analyses had indicated were generated by the admixture of the Patagonia A and Patagonia B (Lager) populations, were placed at an intermediate position between the populations with edges connecting them to both. Importantly, the North American strains also have short but nonzero terminal edge lengths, which excludes both incomplete lineage sorting and laboratory contamination as the source of these mosaic strains.

To quantify the statistical support for the splits suggested by the supernetwork analyses, we performed BCA, which provides CFs or the proportion of genes that support the splits as clades in the primary concordance tree (Fig. 2). When the North American strains were included, low CFs were obtained for both the clade representing the Patagonia A and the Patagonia B (Lager) populations (0.176 and 0.149, respectively), indicating that only a handful of genes supported each population as a monophyletic clade. The exclusion of the mosaic North American strains increased the CFs to 0.533 and 0.465, respectively, demonstrating that admixture outside Patagonia is the main source of phylogenetic discordance among the wild strains of *S. eubayanus*.

Mitochondrial and nuclear intragenic recombination between species

To infer mitochondrial inheritance, we reconstructed a phylogenetic network using *COX2* gene sequences. This phylonetwork showed a unique cluster for most wild *S. eubayanus*, which we conclude corresponds to the

S. eubayanus COX2 allele. CBS 380 inherited a *S. uvarum COX2* allele, indicating the likely inheritance of *S. uvarum* mitochondria (Rainieri *et al.* 2008; Peris 2012). Phylonetwork analysis also suggested that there were two types of recombinant alleles with edges connecting them to both *S. eubayanus* and *S. uvarum* (Figs 3A and S1J, Supporting Information). The sites of interspecies recombination were found near a known recombination hotspot (Peris 2012; Peris *et al.* 2012a) and were readily identified by visual inspection (Fig. 3B) and formal analyses with RDP4 (Fig. S5, Supporting Information).

Surprisingly, we also detected recombination within several nuclear genes of the interspecies hybrids associated with brewing. For example, the ambiguous positions of some *FSY1* and *RIP1* alleles from triple hybrid strains (Fig. S1A, E, Supporting Information) were due to recombination between *S. uvarum* and *S. eubayanus* alleles. Specifically, the CBS 380 *S. eubayanus RIP1* allele and the *FSY1* alleles of CBS 380 and CBS 1546 are clear *S. uvarum*/*S. eubayanus* recombinants (Fig. S5C, D, Supporting Information). The *ITS* gene of the Frohberg lager strain W34/70 appears to be a *S. cerevisiae*/*S. eubayanus* recombinant allele (Fig. S1I, Supporting Information).

Discussion

Distribution of Saccharomyces eubayanus *and its hybrids*

The recent identification of *Saccharomyces eubayanus* as the non-*cerevisiae* parent of the alloploid lager-brewing yeast, *S. pastorianus* (Libkind *et al.* 2011), has allowed us to compare the natural genetic diversity of this species to the alleles present in brewing strains. Surprisingly, population genetic analyses suggest that there are two diverse and highly differentiated populations of *S. eubayanus* in Patagonia. Using a combination of Bayesian concordance factor and phylogenetic network analyses, we have conclusively demonstrated that *S. eubayanus* has been involved in three major types of reticulate evolution, predominantly outside Patagonia.

First, rare North American isolates of *S. eubayanus* originated through the recent admixture of the two Patagonian populations. Although the isolation of *S. eubayanus* was frequent (~47% of samples) across Patagonia (Libkind *et al.* 2011), we have only rarely (<1% of samples) isolated it in North America, so far from a single site. Second, after hybridizing with two distinct *S. cerevisiae* ale lineages, *S. eubayanus* has generated two distinct

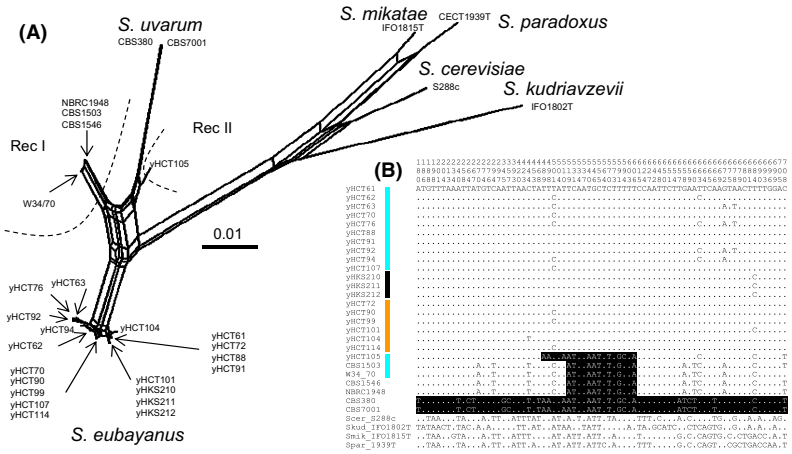

Fig. 3 Mitochondrial *COX2* reveals a history of interspecies recombination. (A) Phylogenetic Neighbor-Net network reconstructed from partial mitochondrial *COX2* gene sequences. Species-specific clusters are displayed using *COX2* gene sequences from the type or reference strains. Polymorphic sites for *COX2* gene sequences are displayed in (B). Black regions correspond to SNPs acquired from *S. uvarum*. RDP4 analysis (Fig. S5, Supporting Information) suggests that the lager strains (W34/70 and CBS 1503) and yHCT105 are both recombinant due to small insertions of *S. uvarum* sequence into the *Saccharomyces eubayanus* backbone. Note that *COX2* is highly polymorphic and prone to recombination due to endonuclease activity (Peris 2012; Peris *et al.* 2012a). In a previous study (Peris 2012), CBS 1546 and NBRC 1948 (CECT 11185) were found to share the same haplotype as CBS 1503 (CECT 1970), Haplotype 78, which is closely related to W34/70's Haplotype 93 (1-bp difference, Fig. 3B). Together with Haplotype 79, these haplotypes are enclosed in Haplogroup 6, which was previously considered to be the most plausible *S. eubayanus* allele. However, the intermediate positions between the alleles from *S. uvarum* and the wild strains of *S. eubayanus* suggest that Haplogroup 6 and yHCT105 represent recombinant versions of *COX2*. Orange and light blue bars mark strains included in Patagonia A and Patagonia B (Lager) populations, respectively. The black bar corresponds to the admixture strains.

lager-brewing lineages of *S. pastorianus* (Dunn & Sherlock 2008) that we have shown contain nearly identical *S. eubayanus* alleles. Third, we described clear evidence of intragenic recombination between *S. eubayanus*, *S. uvarum* and *S. cerevisiae* alleles within double and triple hybrid strains from the brewing environment. Thus, although reticulate evolution is rare in their natural ecological setting in Patagonia, *S. eubayanus* has participated in industrially important and genetically illuminating hybridization events in Europe and North America.

High genetic diversity suggests that Saccharomyces eubayanus *is well established in Patagonia*

Northwestern Patagonia in Argentina provides a rich natural habitat for *Saccharomyces* yeasts, including two diverse *S. eubayanus* populations and their sister species, *S. uvarum*, all of which exist in sympatry. One of these populations has a close affinity with hybrid strains associated with the European brewing industry, including the lager yeast hybrid *S. pastorianus* (*S. cerevisiae* × *S. eubayanus*). The second population was highly differentiated and ~1% divergent at the level of DNA sequence, a degree of divergence similar to pairs of allopatric populations of *S. paradoxus* and *S. kudriavzevii* on opposite sides of Eurasia (Liti *et al.* 2009; Hittinger *et al.* 2010). Moreover, the genetic divergence of the two *S. eubayanus* populations is greater than the pairwise divergence of any of the commonly studied *S. cerevisiae* strains from the *Saccharomyces* Genome Resequencing Project (Liti *et al.* 2009). The existence of multiple diverse populations of *S. eubayanus*, as well as the high frequency of isolation, demonstrates that it is well established in Patagonia. Given the high genetic diversity found in close proximity at the Patagonian sampling sites, further investigation of the ecological factors maintaining diversity and differentiation between and within populations of *S. eubayanus* is warranted. Its rare isolation from North America and its contribution to European hybrids suggest that, although *S. eubayanus* may be native to South America, it is not endemic or strictly exclusive to South America.

In contrast to the high genetic diversity in South America, the nucleotide diversity among the *S. eubayanus* moieties found in the Saaz and Frohberg lager-brewing strains was very low (0.04%), suggesting that alleles were drawn from a small and possibly transient subpopulation closely related to the *S. eubayanus* Patagonia B (Lager) population. The Saaz and Frohberg strains showed considerably higher divergence between their *S. cerevisiae* alleles (0.3%) (Dunn & Sherlock 2008), consistent with the nearly ubiquitous presence of diverse strains of *S. cerevisiae* in Europe. Even the highly polymorphic mitochondrial *S. eubayanus COX2* gene (Peris 2012) had low nucleotide diversity among lager-brewing strains, 0.085%.

Although fungal molecular clocks suffer from a sparse fossil record and heterotachy (Taylor & Berbee 2006), minimum estimates of divergence times have also been made using laboratory mutation and growth rates (Fay & Benavides 2005). Such calculations almost certainly underestimate divergence times due to suboptimal nutrient availability in nature. Our calibration in minimal media at 8 °C suggests divergence times of more than 150 kyr for the two Patagonian lineages of *S. eubayanus* and over 5 Myr for the origin of the *Saccharomyces* genus. Placing absolute dates on fungal branching events remains a serious challenge, but calibration by any of these methods implies that it is highly unlikely that any of the wild strains examined shares a common ancestor with *S. pastorianus* in the last few hundred years.

Local adaptation and the invasion of new niches

The success of reticulate evolutionary events, such as hybridization, admixture, introgression and HGT, depends on the ecological context in which they occur. When reticulate evolution happens in environments where parental strains are well adapted, the local adaption of the parents acts as a strong isolating force against reticulate evolution (Verhoeven *et al.* 2011). However, if the environment changes, new niches can become available where the acquisition of alleles by hybridization, admixture, introgression or HGT can be advantageous (Verhoeven *et al.* 2011; Abbott *et al.* 2013; Baltrus 2013). Environmental changes can be driven by geology, human or biological modification of habitats, or long-range dispersal to new locales (Vitousek *et al.* 1997; Kump 2008; Merow *et al.* 2011; Diffenbaugh & Field 2013).

In fungi, the ecological conditions that favour admixture or hybridization are still unknown, but hints of an association with novel habitats, disturbed environments and human activity are emerging. Genome sequencing projects have demonstrated several cases of hybridization, admixture, introgression, recombination and HGT (Brown *et al.* 1998; Liti *et al.* 2009; Novo *et al.* 2009; Schacherer *et al.* 2009; Dunn *et al.* 2012; Peris 2012; Peris *et al.* 2012a,b,c; Gladieux *et al.* 2014). The clearest cases are closely associated with human activity (Dunn & Sherlock 2008; Novo *et al.* 2009; Schacherer *et al.* 2009; Libkind *et al.* 2011; Dunn *et al.* 2012), but some fungal reticulation events appear to have occurred in nature (Liti *et al.* 2006; Doniger *et al.* 2008; Peris 2012), often in association with the acquisition of pathogenic capabilities or the adaptation to extreme environments (Gladieux *et al.* 2014). Several *Saccharomyces* species have also

been found in sympatry, but hybrids have only rarely been isolated from natural settings (Sniegowski *et al.* 2002; Sampaio & Gonçalves 2008; Libkind *et al.* 2011). No evidence of stable hybridization was found in the Patagonian location studied here.

In contrast, admixed *S. eubayanus* strains were isolated from novel tree genera (*Acer* and *Fagus*, instead of *Nothofagus*) in North America, while the interspecies hybrids provide an even clearer example of novel combinations of alleles exploiting the new brewing and winemaking niches created by humans. These observations suggest that local adaptation is often strong enough or ecological niches distinct enough that *Saccharomyces* hybrids are generally outcompeted, as usually occurs in animals and plants (Hatfield & Schluter 1996; Verhoeven *et al.* 2011). The means of long-range dispersal in *S. eubayanus* and other fungi remain speculative, but humans and the Central and Mississippi migratory bird flyways both provide plausible trans-hemisphere vectors (Francesca *et al.* 2012; Somveille *et al.* 2013).

Critically interpreting reticulate evolution

The reticulate evolution observed in *S. eubayanus* and its hybrids produced complex and sometimes contradictory phylogenetic signals. A population genetic framework was capable of capturing the population differentiation and admixture of the two wild populations of *S. eubayanus*, but phylogenetic networks provided an additional intuitive way to summarize admixture and more complex reticulate biological processes (Bapteste *et al.* 2013). Unfortunately, supernetworks do not currently have built-in statistical tests, and homoplasy can lead to misleading summaries if phylogenetic networks are not applied critically (Woolley *et al.* 2008). Combining phylogenetic network and Bayesian concordance factor analyses is an attractive approach that allows each gene to maintain an independent topology and model of evolution, while separately evaluating the statistical support that each gene lends to splits and clades.

This integrative approach allowed us to confidently visualize all three major types of reticulate evolution that had occurred in *S. eubayanus* and its brewing hybrids: admixture, interspecies hybridization and intragenic recombination between species. The vast majority of these reticulation events were associated with novel environments outside Patagonia, especially in the European brewing environment recently created by humans. In an era of global climate change, understanding the genetic consequences of even rare reticulation events between populations of eukaryotic microbes is likely to be increasingly important for human health and industry.

Acknowledgements

We thank David A. Baum and Bret A. Payseur for critical comments on the manuscript; Amanda B. Hulfachor for artwork; and the administration of the Patagonian National Parks for sampling permits. Funding statement for D.P., K.S., W.G.A. and C.T.H.: This material is based upon work supported by the National Science Foundation under Grant No. DEB-1253634 and funded in part by the DOE Great Lakes Bioenergy Research Center (DOE Office of Science BER DE-FC02-07ER64494). D.L.: ANPCyT PICT 1814 and UNComahue B171 projects (Argentina). J.P.S. and P.G.: Grants PTDC/AGR-ALI/118590/2010 and PTDC/BIA-EVF/118618/2010 FCT (Portugal).

References

Abbott R, Albach D, Ansell S *et al.* (2013) Hybridization and speciation. *Journal of Evolutionary Biology*, **26**, 229–246.

Adler D, Murdoch D (2009) rgl: 3D visualization device system (Open GL). Available from http://CRAN.R-project.org/package=rgl.5

Altschul S, Gish W, Miller W, Myers E, Lipman D (1990) Basic local alignment search tool. *Journal of Molecular Biology*, **215**, 403–410.

Ané C, Larget B, Baum DA, Smith SD, Rokas A (2007) Bayesian estimation of concordance among gene trees. *Molecular Biology and Evolution*, **24**, 412–426.

Anjos J, Rodrigues de Sousa H, Roca C *et al.* (2013) Fsy1, the sole hexose-proton transporter characterized in *Saccharomyces* yeasts, exhibits a variable fructose:H+ stoichiometry. *Biochimica et Biophysica Acta (BBA) – Biomembranes*, **1828**, 201–207.

Arnold ML (2008) *Reticulate Evolution and Humans: Origins and Ecology*, 1st edn. Oxford University Press, New York.

Baltrus DA (2013) Exploring the costs of horizontal gene transfer. *Trends in Ecology & Evolution*, **28**, 489–495.

Bapteste E, van Iersel L, Janke A *et al.* (2013) Networks: expanding evolutionary thinking. *Trends in Genetics*, **29**, 439–441.

Baum DA (2007) Concordance trees, concordance factors, and the exploration of reticulate genealogy. *Taxon*, **56**, 417–426.

Belloch C, Querol A, Garcia MD, Barrio E (2000) Phylogeny of the genus *Kluyveromyces* inferred from the mitochondrial cytochrome-c oxidase II gene. *International Journal of Systematic and Evolutionary Microbiology*, **50**, 405–416.

Belloch C, Orlic S, Barrio E, Querol A (2008) Fermentative stress adaptation of hybrids within the *Saccharomyces sensu stricto* complex. *International Journal of Food Microbiology*, **122**, 188–195.

Brown CJ, Todd KM, Rosenzweig RF (1998) Multiple duplications of yeast hexose transport genes in response to selection in a glucose-limited environment. *Molecular Biology and Evolution*, **15**, 931–942.

Caudy AA, Guan Y, Jia Y *et al.* (2013) A new system for comparative functional genomics of *Saccharomyces* yeasts. *Genetics*, **195**, 275–287.

Cliften P, Sudarsanam P, Desikan A *et al.* (2003) Finding functional features in *Saccharomyces* genomes by phylogenetic footprinting. *Science*, **301**, 71–76.

Cliften PF, Fulton RS, Wilson RK, Johnston M (2006) After the duplication: gene loss and adaptation in *Saccharomyces* genomes. *Genetics*, **172**, 863–872.

Dietrich FS, Voegeli S, Brachat S *et al.* (2004) The *Ashbya gossypii* genome as a tool for mapping the ancient *Saccharomyces cerevisiae* genome. *Science*, **304**, 304–307.

Diffenbaugh NS, Field CB (2013) Changes in ecologically critical terrestrial climate conditions. *Science*, **341**, 486–492.

Doniger SW, Kim HS, Swain D *et al.* (2008) A catalog of neutral and deleterious polymorphism in yeast. *PLoS Genetics*, **4**, e1000183.

Drummond A, Rambaut A (2007) BEAST: bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, **7**, 214.

Dunn B, Sherlock G (2008) Reconstruction of the genome origins and evolution of the hybrid lager yeast *Saccharomyces pastorianus*. *Genome Research*, **18**, 1610–1623.

Dunn B, Richter C, Kvitek DJ, Pugh T, Sherlock G (2012) Analysis of the *Saccharomyces cerevisiae* pan-genome reveals a pool of copy number variants distributed in diverse yeast strains from differing industrial environments. *Genome Research*, **22**, 908–924.

Earl D, vonHoldt B (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetic Resources*, **4**, 359–361.

Erny C, Raoult P, Alais A *et al.* (2012) Ecological success of a group of *Saccharomyces cerevisiae/Saccharomyces kudriavzevii* hybrids in the Northern European wine making environment. *Applied and Environmental Microbiology*, **78**, 3256–3265.

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.

Excoffier L, Lischer H (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, **10**, 564–567.

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.

Fay JC, Benavides JA (2005) Evidence for domesticated and wild populations of *Saccharomyces cerevisiae*. *PLoS Genetics*, **1**, e5.

Flot JF (2010) SeqPHASE: a web tool for interconverting phase input/output files and fasta sequence alignments. *Molecular Ecology Resources*, **10**, 162–166.

Francesca N, Canale DE, Settanni L, Moschetti G (2012) Dissemination of wine-related yeasts by migratory birds. *Environmental Microbiology Reports*, **4**, 105–112.

Fu YX (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, **147**, 915–925.

Gibson BR, Storgårds E, Krogerus K, Vidgren V (2013) Comparative physiology and fermentation performance of Saaz and Frohberg lager yeast strains and the parental species *Saccharomyces eubayanus*. *Yeast*, **30**, 255–266.

Gladieux P, Ropars J, Badouin H *et al.* (2014) Fungal evolutionary genomics provides insight into the mechanisms of adaptive divergence in eukaryotes. *Molecular Ecology*, **23**, 753–773.

Godard P, Urrestarazu A, Vissers S *et al.* (2007) Effect of 21 different nitrogen sources on global gene expression in the yeast *Saccharomyces cerevisiae*. *Molecular and Cellular Biology*, **27**, 3065–3086.

González SS, Barrio E, Querol A (2008) Molecular characterization of new natural hybrids between *S. cerevisiae* and *S. kudriavzevii* from brewing. *Applied and Environmental Microbiology*, **74**, 2314–2320.

Groth C, Hansen J, Piskur J (1999) A natural chimeric yeast containing genetic material from three species. *International Journal of Systematic Bacteriology*, **49**, 1933–1938.

Hatfield T, Schluter D (1996) A test for sexual selection on hybrids of two sympatric Sticklebacks. *Evolution*, **50**, 2429–2434.

Heled J (2010) *Extended Bayesian Skyline Plots Tutorial*. (http://beast.bio.ed.ac.uk/Tutorials). Accessed on 20/09/2013.

Heled J, Drummond A (2008) Bayesian inference of population size history from multiple loci. *BMC Evolutionary Biology*, **8**, 289.

Hijmans R, Guarino L, Cruz M, Rojas E (2001) Computer tools for spatial analysis of plant genetic resources data: 1. DIVA-GIS. *Plant Genetic Resources Newsletter*, **127**, 15–19.

Hijmans R, Cameron S, Parra J, Jones P, Jarvis A (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, **25**, 1965–1978.

Hilbe J (2007) STATISTICA 7: an overview. In: *The American Statistician* (ed. Christensen R) pp. 91–94. Taylor & Francis, London.

Hittinger CT (2013) Saccharomyces diversity and evolution: a budding model genus. *Trends in Genetics*, **29**, 309–317.

Hittinger CT, Gonçalves P, Sampaio JP *et al.* (2010) Remarkably ancient balanced polymorphisms in a multi-locus gene network. *Nature*, **464**, 54–58.

Holland BR, Huber KT, Moulton V, Lockhart PJ (2004) Using consensus networks to visualize contradictory evidence for species phylogeny. *Molecular Biology and Evolution*, **21**, 1459–1461.

Hubisz M, Falush D, Stephens M, Pritchard JK (2009) Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources*, **9**, 1322–1332.

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

Huson DH, Tobias D, Tobias K, Steel MA (2004) Phylogenetic Super-Networks from partial trees. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, **1**, 151–158.

Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, **23**, 1801–1806.

Jensen J, Bohonak A, Kelley S (2005) Isolation by distance, web service. *BMC Genetics*, **6**, 13.

Kellis M, Patterson N, Endrizzi M, Birren B, Lander ES (2003) Sequencing and comparison of yeast species to identify genes and regulatory elements. *Nature*, **423**, 241–254.

Kück P, Meusemann K (2010) *FASconCAT, Version 1.0* (http://zfmk.de/web/Forschung/Abteilungen/AG_Wgele/Software/index.en.html) Accessed on 20/09/2013.

Kump LR (2008) The rise of atmospheric oxygen. *Nature*, **451**, 277–278.

Kurtzman CP, Robnett CJ (1991) Phylogenetic relationships among species of *Saccharomyces*, *Schizosaccharomyces*, *Debaryomyces* and *Schwanniomyces* determined from partial ribosomal RNA sequences. *Yeast*, **7**, 61–72.

Kurtzman CP, Fell JW, Boekhout T (2011) *The Yeasts: A Taxonomic Study*, 5th edn. Elsevier, Amsterdam.

Larget BR, Kotha SK, Dewey CN, Ané C (2010) BUCKy: gene tree/species tree reconciliation with bayesian concordance analysis. *Bioinformatics*, **26**, 2910–2911.

Le Jeune C, Lollier M, Demuyter C *et al.* (2007) Characterization of natural hybrids of *Saccharomyces cerevisiae* and *Saccharomyces bayanus* var. *uvarum*. *FEMS Yeast Research*, **7**, 540–549.

Libkind D, Hittinger CT, Valério E *et al.* (2011) Microbe domestication and the identification of the wild genetic stock of lager-brewing yeast. *Proceedings of the National Academy of Sciences of the United States of America*, **108**, 14539–14544.

Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, **25**, 1451–1452.

Liti G, Barton DB, Louis EJ (2006) Sequence diversity, reproductive isolation and species concepts in *Saccharomyces*. *Genetics*, **174**, 839–850.

Liti G, Carter DM, Moses AM *et al.* (2009) Population genomics of domestic and wild yeasts. *Nature*, **458**, 337–341.

Lynch M, Sung W, Morris K *et al.* (2008) A genome-wide view of the spectrum of spontaneous mutations in yeast. *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 9272–9277.

Martin DP, Lemey P, Lott M *et al.* (2010) RDP3: a flexible and fast computer program for analyzing recombination. *Bioinformatics*, **26**, 2462–2463.

Masneuf I, Hansen J, Groth C, Piskur J, Dubourdieu D (1998) New hybrids between *Saccharomyces Sensu Stricto* yeast species found among wine and cider production strains. *Applied and Environmental Microbiology*, **64**, 3887–3892.

Merow C, LaFleur N, John AS, Adam M, Rubega M (2011) Developing dynamic mechanistic species distribution models: predicting bird-mediated spread of invasive plants across Northeastern North America. *The American Naturalist*, **178**, 30–43.

Montrocher R, Verner MC, Briolay J, Gautier C, Marmeisse R (1998) Phylogenetic analysis of the *Saccharomyces cerevisiae* group based on polymorphisms of rDNA spacer sequences. *International Journal of Systematic Bacteriology*, **48**, 295–303.

Morales L, Dujon B (2012) Evolutionary role of interspecies hybridization and genetic exchanges in yeasts. *Microbiology and Molecular Biology Reviews*, **76**, 721–739.

Nakao Y, Kanamori T, Itoh T *et al.* (2009) Genome sequence of the lager brewing yeast, an interspecies hybrid. *DNA Research*, **16**, 115–129.

Novo M, Bigey F, Beyne E *et al.* (2009) Eukaryote-to-eukaryote gene transfer events revealed by the genome sequence of the wine yeast *Saccharomyces cerevisiae* EC1118. *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 16333–16338.

Otto SP (2007) The evolutionary consequences of polyploidy. *Cell*, **131**, 452–462.

Peris D (2012) *Genome Characterization of Natural Saccharomyces hybrids of Biotechnological Interest*. University of Valencia. (http://roderic.uv.es/handle/10550/24743). Accessed on 20/9/2013.

Peris D, Belloch C, Lopandic K *et al.* (2012a) The molecular characterization of new types of *S. cerevisiae* × *S. kudriavzevii* hybrid yeasts unveils a high genetic diversity. *Yeast*, **29**, 81–91.

Peris D, Lopes CA, Arias A, Barrio E (2012b) Reconstruction of the evolutionary history of *Saccharomyces cerevisiae* × *S. kudriavzevii* hybrids based on multilocus sequence analysis. *PLoS One*, **7**, e45527.

Peris D, Lopes CA, Belloch C, Querol A, Barrio E (2012c) Comparative genomics among *Saccharomyces cerevisiae* × *Saccharomyces kudriavzevii* natural hybrid strains isolated from wine and beer reveals different origins. *BMC Genomics*, **13**, 407.

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

Rainieri S, Kodama Y, Kaneko Y *et al.* (2006) Pure and mixed genetic lines of *Saccharomyces bayanus* and *Saccharomyces pastorianus* and their contribution to the lager brewing strain genome. *Applied and Environmental Microbiology*, **72**, 3968–3974.

Rainieri S, Kodama Y, Nakao Y, Pulvirenti A, Giudici P (2008) The inheritance of mtDNA in lager brewing strains. *FEMS Yeast Research*, **8**, 586–596.

Rambaut A, Drummond A. 2001. *Tracer v1.4. Molecular Evolution, Phylogenetics and Epidemiology*. (http://beast.bio.ed.ac.uk/Tracer). Accessed on 20/9/2013.

Rambaut A, Drummond AJ. 2010. *FigTree v1.3.1. Institute of Evolutionary Biology*, University of Edinburgh: Edinburgh, UK. (http://tree.bio.ed.ac.uk/software/figtree/). Accessed on 20/9/2013.

Rokas A, Williams BL, King N, Carroll SB (2003) Genome-scale approaches to resolving incongruence in molecular phylogenies. *Nature*, **425**, 798–804.

Ronquist F, Teslenko M, van der Mark P *et al.* (2012) MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology*, **61**, 539–542.

Rosenberg NA (2004) DISTRUCT: a program for the graphical display of population structure. *Molecular Ecology Notes*, **4**, 137–138.

Rousset F (1997) Genetic differentiation and estimation of gene flow from F-Statistics under isolation by distance. *Genetics*, **145**, 1219–1228.

Sampaio JP, Gonçalves P (2008) Natural populations of *Saccharomyces kudriavzevii* in Portugal are associated with oak bark and are sympatric with *S. cerevisiae* and *S. paradoxus*. *Applied and Environmental Microbiology*, **74**, 2144–2152.

Sarah C, Goslee D (2007) The ecodist package for dissimilarity-based analysis of ecological data. *Journal of Statistical Software*, **22**, 1–19.

Scannell DR, Zill OA, Rokas A *et al.* (2011) The awesome power of yeast evolutionary genetics: new genome sequences and strain resources for the *Saccharomyces sensu stricto* genus. *G3: Genes, Genomes, Genetics*, **1**, 11–25.

Schacherer J, Shapiro JA, Ruderfer DM, Kruglyak L (2009) Comprehensive polymorphism survey elucidates population structure of *Saccharomyces cerevisiae*. *Nature*, **458**, 342–345.

Sniegowski PD, Dombrowski PG, Fingerman E (2002) *Saccharomyces cerevisiae* and *Saccharomyces paradoxus* coexist in a natural woodland site in North America and display different levels of reproductive isolation from European conspecifics. *FEMS Yeast Research*, **1**, 299–306.

Sokal R, Rohlf F (1995) *Biometry*, 3rd edn. Freeman, New York.

Somveille M, Manica A, Butchart SHM, Rodrigues ASL (2013) Mapping global diversity patterns for migratory birds. *PLoS One*, **8**, e70907.

Staden R, Beal KF, Bonfield JK (2000) The Staden package, 1998. In: *Methods in Molecular Biology* (ed. Clifton N), pp. 115–130.

Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, **123**, 585–595.

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, Berbee ML (2006) Dating divergences in the Fungal Tree of Life: review and new analyses. *Mycologia*, **98**, 838–849.

Tsai IJ, Bensasson D, Burt A, Koufopanou V (2008) Population genomics of the wild yeast *Saccharomyces paradoxus*: quantifying the life cycle. *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 4957–4962.

Verhoeven KJF, Macel M, Wolfe LM, Biere A (2011) Population admixture, biological invasions and the balance between local adaptation and inbreeding depression. *Proceedings of the Royal Society B: Biological Sciences*, **278**, 2–8.

Vitousek PM, Mooney HA, Lubchenco J, Melillo JM (1997) Human domination of Earth's ecosystems. *Science*, **277**, 494–499.

Woerner AE, Cox MP, Hammer MF (2007) Recombination-filtered genomic datasets by information maximization. *Bioinformatics*, **23**, 1851–1853.

Woolley SM, Posada D, Crandall KA (2008) A comparison of phylogenetic network methods using computer simulation. *PLoS One*, **3**, e1913.

Yu Y, Degnan JH, Nakhleh L (2012) The probability of a gene tree topology within a phylogenetic network with applications to hybridization detection. *PLoS Genetics*, **8**, e1002660.

Zinner D, Arnold ML, Roos C (2011) The strange blood: natural hybridization in primates. *Evolutionary Anthropology: Issues, News, and Reviews*, **20**, 96–103.

D.P., D.L., P.G., J.P.S., W.G.A. and C.T.H. conceived and designed this study. K.S. and D.L. isolated and identified the strains from North America and South America, respectively. D.P. generated and analysed the data. D.P., D.L. and C.T.H. wrote the manuscript.

Data accessibility

Gene sequences are available in GenBank under Accession nos KF530330-KF530542 and KJ412200. Input and output files from the software used in this study, as well as phylogenetic trees, networks, and the alignments, were deposited in the Data Dryad repository under doi: 10.5061/dryad.153b8.

Supporting information

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

Fig. S1 Individual gene trees.

Fig. S2 Extended Bayesian Skyline plots.

Fig. S3 Population and admixture growth rate at 8 °C.

Fig. S4 Time-calibrated phylogenetic tree.

Fig. S5 RDP4 analyses.

Table S1 PCR primers and conditions.

Table S2 (A) Summary statistics for the Patagonia A population. (B) Summary statistics for the Patagonia B (Lager) population. (C) Summary statistics for the *Saccharomyces eubayanus* moiety in all hybrids. (D) Summary statistics for the *Saccharomyces eubayanus* moiety in *S. pastorianus*.

Table S3 Strains used in this study and geographical and ecological factors associated with them.