DOI: 10.1534/genetics.106.062166

Sequence Diversity, Reproductive Isolation and Species Concepts in Saccharomyces

Gianni Liti,* David B. H. Barton* and Edward J. Louis*,[1]

**Institute of Genetics, Queen's Medical Centre, University of Nottingham, Nottingham NG7 2UH, United Kingdom*

Manuscript received June 16, 2006
Accepted for publication August 4, 2006

ABSTRACT

Using the biological species definition, yeasts of the genus Saccharomyces *sensu stricto* comprise six species and one natural hybrid. Previous work has shown that reproductive isolation between the species is due primarily to sequence divergence acted upon by the mismatch repair system and not due to major gene differences or chromosomal rearrangements. Sequence divergence through mismatch repair has also been shown to cause partial reproductive isolation among populations within a species. We have surveyed sequence variation in populations of Saccharomyces *sensu stricto* yeasts and measured meiotic sterility in hybrids. This allows us to determine the divergence necessary to produce the reproductive isolation seen among species. Rather than a sharp transition from fertility to sterility, which may have been expected, we find a smooth monotonic relationship between diversity and reproductive isolation, even as far as the well-accepted designations of *S. paradoxus* and *S. cerevisiae* as distinct species. Furthermore, we show that one species of Saccharomyces—*S. cariocanus*—differs from a population of *S. paradoxus* by four translocations, but not by sequence. There is molecular evidence of recent introgression from *S. cerevisiae* into the European population of *S. paradoxus,* supporting the idea that in nature the boundary between these species is fuzzy.

THE concept of a species is central to the biological sciences (Mallet 1995; Coyne and Orr 1998), but the definition of a species is controversial. There are several sometimes conflicting definitions of species. The biological species concept (BSC) is based on patterns of breeding: species are considered to be units reproductively isolated from other such units, but within which interbreeding and genetic recombination reduce the possibility of divergence. Asexual reproduction in many organisms including fungi (Taylor *et al.* 2000), hybridization in plants (Rieseberg 1997) and fungi (Taylor *et al.* 2000), and lateral gene transfer between bacteria (Gogarten and Townsend 2005) create problems for the BSC. The genotypic cluster species concept (GCSC) can accommodate some gene flow between species (clusters) as long as their integrity remains such that they can be distinguished (Mallet 1995). Genotypic clusters can be at the subpopulation level as well as at species level. The phylogenetic species concept (PSC) defines the species as a distinct monophyletic group on the basis of evolutionary history and geographical extent. This concept faces the difficulties of incomplete knowledge and a somewhat arbitrary delineation of specific boundaries (Hudson and Coyne 2002).

Sequence and chromosomal evolution can produce genetic divergence and can reduce or prevent interbreeding, as required by these species concepts. Although these concepts do not always agree (Isaac *et al.* 2004), early analyses of Saccharomyces strains by interbreeding (Naumov 1987) (BSC) and DNA/DNA reassociation (Martini and Martini 1987; Vaughan Martini 1989) (PSC) were concordant in supporting the existence of three distinct species: *Saccharomyces cerevisiae*, *S. paradoxus*, and *S. bayanus*. These studies also confirmed the identification of one natural hybrid (*S. pastorianus*). Recent studies of reproductive compatibility have defined three new species within the complex on the basis of BSC: *S. cariocanus*, *S. mikatae*, and *S. kudriavzevii* (Naumov *et al.* 1995a,b, 2000).

S. cerevisiae is a well-known model organism associated with human activity (baking and brewing). Its closest relative is *S. paradoxus*, which is often isolated from Quercus (oak) trees and has little or no association with human activity. The two species show perfect synteny with no gross chromosomal rearrangements between them (Fischer *et al.* 2000; Kellis *et al.* 2003) and a level of sequence divergence of ~15% (Cliften *et al.* 2001). The different species within the *sensu stricto* complex can mate and generate viable hybrids, indicating an absence of prezygotic isolation. Coexistence of these species in similar habitats (Sniegowski *et al.* 2002) and the isolation of interspecific hybrids (de Barros Lopes *et al.*

Sequence data from this article have been deposited with the EMBL/GenBank Data Libraries under accession nos. AM296217–AM296448.

[1]*Corresponding author:* Queen's Medical Centre, University of Nottingham, Nottingham NG7 2UH, United Kingdom.
E-mail: ed.louis@nottingham.ac.uk

2002; Liti *et al.* 2005) show that interbreeding not only is possible but also does occur in the wild.

A postzygotic barrier exists between the *sensu stricto* species, resulting in few viable spores (≤1%) (Naumov 1987). Several experiments have shown that this postzygotic barrier is not due to genetic incompatibilities or to chromosomal mechanisms such as translocations, which are the two mechanisms accepted for intrinsic hybrid sterility (Coyne and Orr 2004). Many of the Saccharomyces *sensu stricto* species are collinear (Fischer *et al.* 2000; Kellis *et al.* 2003), ruling out major structural changes as causes of sterility in collinear hybrids. Restoration of collinearity between two of these species did result in the production of viable gametes (Delneri *et al.* 2003) but these derived fertile hybrids were highly aneuploid. Euploid, 2N hybrids between the collinear species exhibit the low fertility seen between the noncollinear species (Naumov 1987; Naumov *et al.* 1992, 1995a,b; Hunter *et al.* 1996). Major dominant genetic incompatibilities are ruled out by converting 2N hybrids into 4N allo-tetraploids (Greig *et al.* 2002). These allotetraploids have high levels of fertility and produce 2N hybrid gametes. The ability to replace individual *S. cerevisiae* chromosomes with *S. paradoxus* homologs in the lab reduces the likelihood that there are any recessive genetic incompatibilities maintaining reproductive isolation. Chromosome III has been replaced with no problems of viability (Chambers *et al.* 1996) and most of the rest of the *S. cerevisiae* chromosomes have successfully been replaced by *S. paradoxus* chromosomes (D. Greig, personal communication).

Instead of these two mechanisms, simple sequence divergence and the action of the mismatch repair system have been shown to account for hybrid sterility in Saccharomyces. In an *S. paradoxus* × *S. cerevisiae* hybrid only 1% of gametes are viable (Naumov 1987; Naumov *et al.* 1992; Hunter *et al.* 1996). These viable gametes are highly aneuploid and exhibit little or no recombination (Hunter *et al.* 1996). As mismatch repair is known to greatly reduce recombination between diverged genomes in bacteria (Rayssiguier *et al.* 1989), Hunter *et al.* (1996) tested whether its action on the sequence divergence between the yeast species could account for the sterility due to prevention of the homologous recombination necessary for proper meiotic segregation of the homologous chromosomes. In mismatch-repair-defective hybrids, fertility increased 10-fold, which correlated with an increase in homologous recombination and reduced aneuploidy. Further experiments on a single *S. paradoxus* chromosome in an *S. cerevisiae* background supported the hypothesis that sterility in these hybrids was mostly, if not entirely, due to the normal mismatch repair system acting on the sequence divergence (Chambers *et al.* 1996). This effect of mismatch repair on fertility was also seen in populations of the same species that have diverged and may play a role in the speciation process itself rather than simply maintaining the postzygotic barrier after speciation by some other mechanism (Greig *et al.* 2003). The question remains as to how much divergence results in sufficient reproductive isolation to designate a species.

MATERIALS AND METHODS

Saccharomyces strains and crosses: Strains used in this study are listed in Table 1. Yeast crosses were obtained either by generating haploid versions of strains [by deleting the HO gene using the dominant drug resistance markers KAN and HYG (Goldstein and McCusker 1999)] or by crossing spores from *ura3* and *lys2* mutants [previously selected in 5-FOA and α-aminoadipate media, respectively (Naumov 1987)] in a minimal medium. The *HO* deletion was performed by PCR-mediated gene replacement and confirmed using the following primers (5′–3′): HO-DF, (AGAGCTTTTGAAGGTGAACC TGGTAGGTTAGACCCTAGGCGTAGAAC CGTACGCTGCA GGTCGAC); HO-DR, (GGCTCTTTTGTTGTACCGTCACCTA GCCATAGACCAAGCATCCAAGCCATATCGATGAATTCGAGC TCG); HO-A1, (GTTATGTGCGCAGATGGCTC); and HO-A4, (GTCACAGTAGCAGACATTCC).

Tetrad dissection: Sporulation was induced for 3–5 days at room temperature in 1% potassium acetate media and spores were dissected as previously described (Naumov *et al.* 1994). Some of the crosses involved strains differing in reciprocal and nonreciprocal translocations (Table 2). Each of the translocations studied induces the formation of multivalents during meiosis, causing a 50% loss of gametic viability if reciprocal and a 25% loss if not reciprocal.

Chromosome separation in Saccharomyces strains: Genomic DNA for CHEF analysis was prepared as previously described (Louis 1998). The program for separation of the chromosomes consisted of two blocks: block 1—20-hr, 60-sec switching time and block 2—12-hr, 90-sec switching time, using a CHEF–DRIII apparatus (Bio-Rad, Hercules, CA). Both blocks were run at a voltage of 6 V/cm, angle 120° in 0.045 M Tris–borate, 0.045 M boric acid, and 0.001 M EDTA (0.5× TBE) at 13°.

Yeast DNA purification, sequencing, and analysis: Genomic yeast DNA were extracted and purified as previously described (Louis and Borts 1995). DNA sequence was directly obtained from PCR products by Agowa (http://www.agowa.de). We sequenced *yKu70*, *yKu80*, *NEJ1*, *EST2*, *TLC1*, and *SPT4*. Genes were sequenced at two to six times coverage. The region indicated in Figure 4 was sequenced in a subset of strains. Sequences were filtered and assembled using the Staden Package (available at http://staden.sourceforge.net/). Synonymous and nonsynonymous substitutions were calculated using DnaSP software (http://www.ub.es/dnasp/; Rozas and Rozas 1999). Sequence comparison, alignment, and phylogenetic trees were obtained using DNAstar package software (EditSeq and MegAlign, http://www.dnastar.com). Sequence divergence is calculated by comparing sequence pairs in relation to the phylogeny reconstructed by MegAlign. The transfer of *S. cerevisiae* sequence to *S. paradoxus* CBS432 was identified when using ad hoc Python scripts to analyze the sequence data (Kellis *et al.* 2003) and mVISTA (http://genome.lbl.gov/vista/index.shtml).

The time estimation of gross chromosomal rearrangements (GCR) and geographic divergence reported in Table 1 is calculated using $t = K_s/(2 \times u \times g)$, where K_s is the synonymous substitution, u is the point mutation rate (1.84×10^{-10}), and g is the maximum number of generations per year according to Fay and Benavides (2005). The range of years takes into account a range from one (second value) to eight (first value) generations/day.

TABLE 1

List of Saccharomyces strains sequenced

OS	Strains	Source	Location	Notes
		S. cerevisiae		
1	DBVPG6763	Unknown	Unknown	Previous *S. boulardii*
3	DBVPG6765	Unknown	Unknown	Previous *S. boulardii*
17	SK1	Lab strain	United States	
60	DBVPG6044	Bili wine	West Africa	Previous *S. mangini*
84	DBVPG1788	Soil	Finland	
86	DBVPG3051	Grape must	Israel	
89	DBVPG1794	Soil	Finland	
91	DBVPG1373	Soil	Netherlands	
92	DBVPG1853	White Tecc	Ethiopia	
93	DBVPG1378	Grape must	Italy	
94	DBVPG1135	Cherries	Italy	
96	S288c[a]	Rotting fig (lab strain)	California	
97	Y55	Wine (lab strain)	France	
104	YPS128	Soil beneath	Pennsylvania	
		S. paradoxus		
5	DBVPG6565	Spoiled mayonnaise	Unknown	Previous *S. douglasii*
8	DG1768	Lab strain	Unknown	
26	N-17	Flux of Quercus	Russia	
31	DBVPG6303	Drosophila	California	
32	DBVPG6304	Drosophila	California	
33	DBVPG6037	Drosophila	California	
35	Q4.1	Flux of Quercus	London	
40	T21.4	Flux of Quercus	London	
76	N-43	Flux of Quercus	Far East	
77	N-44	Flux of Quercus	Far East	
98	CBS5829	Soil	Denmark	
114	YPS125	Flux of Quercus	Pennsylvania	
115	YPS138	Soil beneath	Pennsylvania	
116	YPS145	Soil beneath	Pennsylvania	
137	NBRC1804	Bark	Japan	
142	CBS432[a]	Sap exudate of tree	Russia	
		S. cariocanus		
20	UFRJ50791	Drosophila	Brazil	
21	UFRJ50816	Drosophila	Brazil	
		S. mikatae		
18	NBRC1815[a]	Soil	Japan	
19	NBRC1816	Decayed leaf	Japan	
127	NBRC 10995	Decayed leaf	Japan	
128	NBRC 10996	Decayed leaf	Japan	
130	NBRC 10998	Decayed leaf	Japan	
		S. kudriavzevii		
22	NBRC1802[a]	Decayed leaf	Japan	
23	NBRC1803	Decayed leaf	Japan	
122	NBRC 10990	Decayed leaf	Japan	
123	NBRC 10991	Decayed leaf	Japan	
		S. bayanus		
24	CBS7001[a]	Mesophylax	Spain	
99	VKMY361	Wine	Czech Republic	
100	VKMY508	Wine	Czech Republic	
101	NRRLY-969	Unknown	Unknown	
102	VKMY1146-6B	Grape berries	Russia	

Geographical origin and source have been included. CBS, Centraalbureau voor Schimmelcultures, Utrecht, The Netherlands; NBRC (former IFO), NITE Biological Resource Center, Chiba, Japan; DBVPG, Dipartimento Biologia Vegetale, Perugia, Italy; VKM, National Collection of Microorganisms, Moscow; UFRJ, Universidade Federal do Rio de Janeiro, Rio de Janeiro; NRRL, ARS Culture Collection, Peoria, Illinois; YPS, Yeast P. Sniegowski, University of Pennsylvania, Philadelphia; N-, strains isolated by G. Naumov, Moscow; Q and T, strains isolated by A. Burt, Imperial College, London. OS, other Saccharomyces.

[a] Indicates strains previously sequenced (see text).

TABLE 2

Summary of crosses, percentage of spore viability, and number of total spores analyzed in this study

Crosses	Viable spores (%)	Spores dissected	Sequence divergence[a]	References
Collinear genomes				
Sc S288c × Sp CBS432	0.67	892	13.32	
Sc SK1 × Sc S288c	80.94	404	0.573	
Sc SK1 × Sc Y55	83.00	400	0.587	
Sc S288c × Sc Y55	83.33	288	0.15	
Sp DG1768 × Sp N-17	37.50	320	4.7	
Sp CBS432 × Sp YPS125	40.38	416	4.653	
Sp NBRC1804 × Sp YPS125	32.10	352	4.665	
Sp N44 × Sp YPS125	36.32	424	4.567	
Sp NBRC 1804 × Sp CBS432	73.47	392	1.196	
Sp N44 × Sp CBS432	77.13	328	1.156	
Sp N17 × Sp YPS138	34.11	516	4.631	
Sp N44 × Sp NBRC 1804	86.79	280	0.084	
Four reciprocal translocations, *S. paradoxus* × *S. cariocanus*				
Sp UFRJ50791 × Sp YPS138	4.81	208	0.292	
Sp UFRJ50816 × Sp YPS125	5.09	216	0.347	
One reciprocal translocation				
Sp N43 × Sp YPS125	25.56	360	4.528	
Sp N43 × Sp N44	44.90	412	0.15	
Sp N43 × Sp NBRC 1804	49.15	352	0.07	
Sp N43 × Sp CBS432	36.75	400	1.14	
One nonreciprocal translocation				
Sp N17 × Sp CBS5829	68.50	400	0.153	
Sp CBS432 × Sp CBS5829	71.50	200	0.097	
Sp N44 × Sp CBS5829	52.91	172	1.146	
Sp YPS138 × Sp CBS5829	24.38	160	4.594	
Previous data				
Sk NBRC1802 × Sk NBRC1803	38	208	4.586	Naumov *et al.* (1995b)
Sm NBRC1815 × Sm NBRC1816	44	152	0.613	Naumov *et al.* (1995b)
Sp UFRJ50791 × Sp UFRJ50816	95	236	0	Naumov *et al.* (1995a)
Sp YPS125 × Sp N-17	42	60	4.891	Sniegowski *et al.* (2002)
Sp YPS145 × Sp N-17	53	76	4.637	Sniegowski *et al.* (2002)
Sc Y55 × Sp N-17	1.2	852	13.369	Hunter *et al.* (1996)
Sc Y55 × Sc Y55	98.1	2176	0	Greig *et al.* (2003)

[a] Nucleotide diversity using the Jukes and Cantor correction.

RESULTS AND DISCUSSION

Genetic variation and gamete viability: In this report, we test the relationship between sequence divergence and meiotic sterility to determine if there is a level of divergence at which speciation occurs, that is, a level sufficient to result in reproductive isolation. We determined the degree of divergence by sequencing six nuclear genes (~10 kb) in 41 selected strains within the complex (Table 1). We sequenced *yKu70*, *yKu80*, *NEJ1*, *EST2*, *TLC1*, and *SPT4*. The proteins encoded by *yKu70* and *yKu80* form a strongly interacting heterodimer and are paradoxically involved in both telomere capping and DNA repair via nonhomologous end joining (NHEJ) (Dudasova *et al.* 2004). *NEJ1* is also involved in regulating NHEJ through ploidy and has a paradoxical role in end protection *vs.* double-strand break repair as well (Liti and Louis 2003). *EST2* and *TLC1* are part of the telomerase complex (Smogorzewska and de Lange 2004). *EST2* is the catalytic subunit of the reverse transcriptase and *TLC1* is the RNA template. Moreover, *TLC1* is not translated. Finally, *SPT4* is a transcription factor and has been recently shown to be part of the kinetocore complex (Crotti and Basrai 2004). The six genes are on different chromosomes except *NEJ1* and *EST2*, which are ~100 kbp apart on chromosome XII, and this allows linkage disequilibrium and recombination analyses.

These genes exhibit different levels of divergence and amounts of selective pressure (Figure 1). However, each

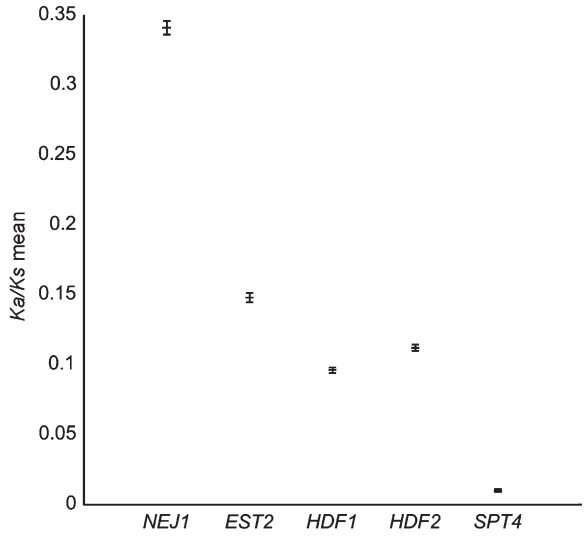

FIGURE 1.—Plot of K_a/K_s ratio of the genes sequenced in this study. The means and standard errors of K_a/K_s for the five coding genes are plotted. *TLC1* is the RNA template of telomerase and therefore cannot be included in the analysis.

of the six genes gave a similar phylogenetic topology (Figure 2) and were consistent with the genome-scale approach (ROKAS *et al.* 2003). A phylogenetic tree obtained from a concatenated analysis of four of these genes is shown in Figure 3 and is consistent with ITS1 phylogeny (LITI *et al.* 2005).

The sequences of *S. paradoxus* fall into three clades correlated with geography on a continental scale. We found three major subpopulations: European, Far Eastern (East Russia and Japan), and North American. Within these subpopulations, the levels of sequence divergence range from 0.08 to 0.21%. Between them, the divergence ranges from 1.5% (between European and Far Eastern populations) to 4.6% (between European or Far Eastern and North American populations; see Figure 3). The values of sequence divergence between *S. paradoxus* strains raise a question: using the PSC, at what level of divergence should monophyletic taxa be considered as subgroups, subspecies, or distinct species? With enough population data, the application of the GCSC may help with this distinction. A single lineage seems to occupy the whole North American continent. Different sets of strains, isolated in different places and at different times (Table 1) have similar sequences. The recently described species *S. cariocanus*, which has so far been found only in Brazil, appears in our analysis within the North American lineage of *S. paradoxus*. Strain NBRC1803 of *S. kudriavzevii* is highly divergent (4.6%) from the other three strains, despite being found geographically with the other isolates (Figure 3). Variation among *S. cerevisiae* strains, on the other hand, shows little in the way of geographical structure. Isolates from neighboring areas do not necessarily group together. Moreover, sequence divergence between locations was significantly lower in *S. cerevisiae* than in the other species, ranging from 0 to 0.6%. This is not really different from the phylogenetic studies seen in the analysis by AA *et al.* (2006) where they report differences between North American isolates (some of these are included in this study) and vineyard isolates, for example. The *S. cerevisiae* structures reported here and in AA *et al.* (2006) are not as distinct as those seen in *S. paradoxus*. The relative lack of diversity is probably a consequence of domestication (FAY and BENAVIDES 2005), could reflect movement through human activity, and is consistent with a more recent common ancestor, or founding population, for all extant *S. cerevisiae* strains (SNIEGOWSKI *et al.* 2002; AA *et al.* 2006).

Our data were used to assess the correlation between reproductive isolation and sequence divergence. The degree of divergence partially fills the gap between the extremes of ~1 and ~15% found within and between the species previously analyzed (NAUMOV 1987; GREIG *et al.* 2003). When the percentage of viable spores is plotted against sequence divergence (Figure 4A), several points lie outside a general linear trend. When chromosomal translocations are taken into account (see below), these data points move into the general mass of data (Figure 4B). The degree of sequence variation strongly correlates with meiotic sterility and several smooth, monotonic relationships can fit the data (including linear, exponential, etc.). The best fit is an exponential decay curve ($y = -12.75 + 104.52 \times e(-0.149x)$; $r = 0.98$, where x is sequence divergence and y is gamete viability). The linear and exponential curves are also very good fits (see Figure 4B). All combinations of crosses between the three major subpopulations in *S. paradoxus* exhibit partial reproductive isolation (Table 2 and Figure 4), in agreement with previous reports (SNIEGOWSKI *et al.* 2002; GREIG *et al.* 2003). Crosses within the same subpopulation exhibit a high frequency of viable spores (>~80%).

Translocations and gamete viability: The new species *S. cariocanus* is interesting because, when the reductions in viability due to the four reciprocal translocations (16-fold; see MATERIALS AND METHODS) are accounted for, we find that crosses between *S. paradoxus* and *S. cariocanus* strains can produce a high corrected frequency of viable spores (77–81%; Table 2). Although these strains differ in chromosomal organization, they differ little in their sequences. Support for *S. cariocanus* as a member of the *S. paradoxus* species and in particular the American subpopulation comes from analysis of the viable spores. In hybrids between species, aneuploidy occurs at high frequencies due to the lack of homologous recombination (HUNTER *et al.* 1996). In the *S. cariocanus* × *S. paradoxus* hybrids we do not see any aneuploidy in the viable spores (Figure 5). Instead, we see the balanced chromosome sets expected for heterozygous translocations.

We analyzed the viability of spores in crosses between *S. paradoxus* strains with chromosomal rearrangements

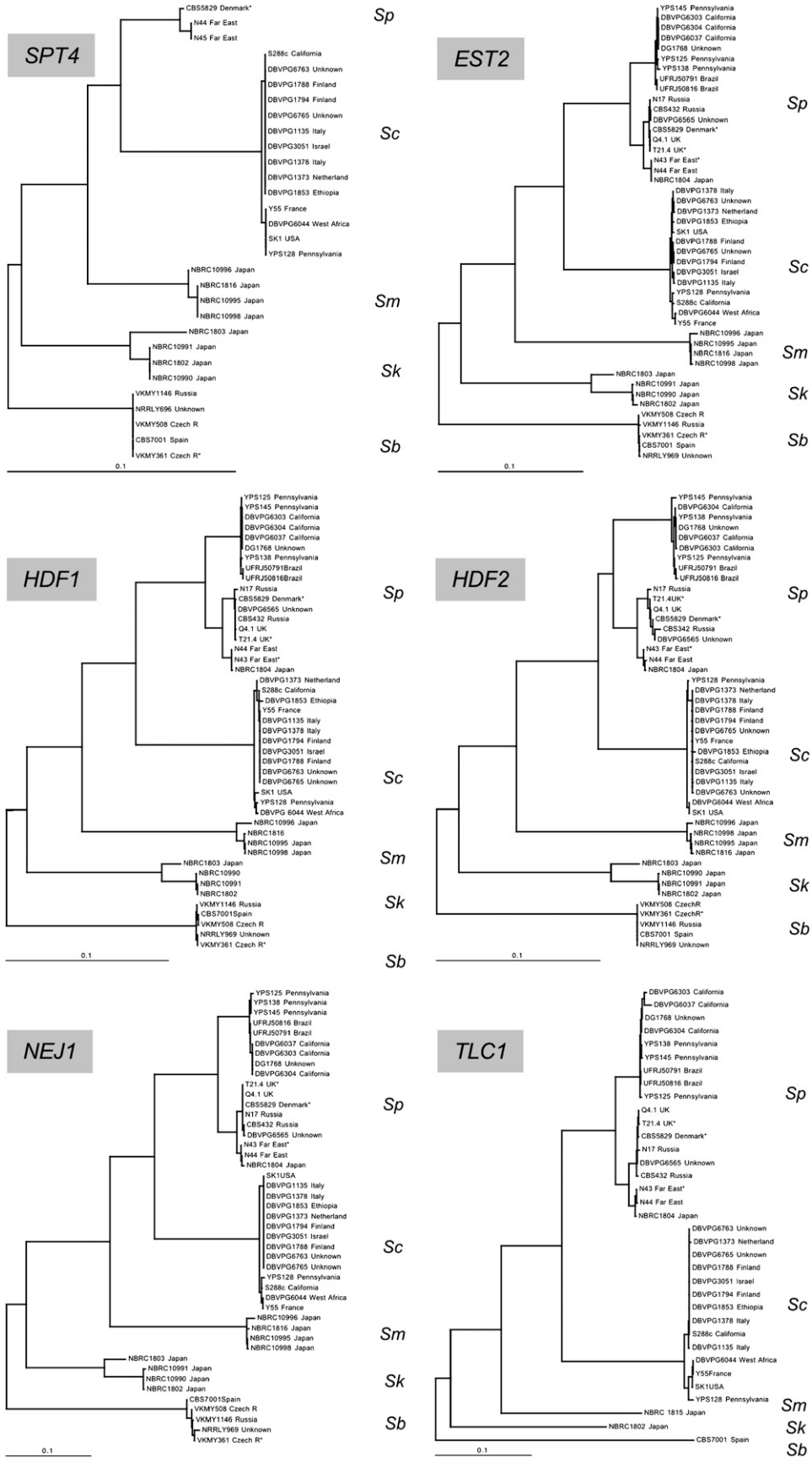

FIGURE 2.—Phylogenetic trees (neighbor-joining) of the six genes sequenced in this study. Sp, *S. paradoxus*; Sc, *S. cerevisiae*; Sm, *S. mikatae*; Sk, *S. kudriavzevii*; Sb, *S. bayanus*.

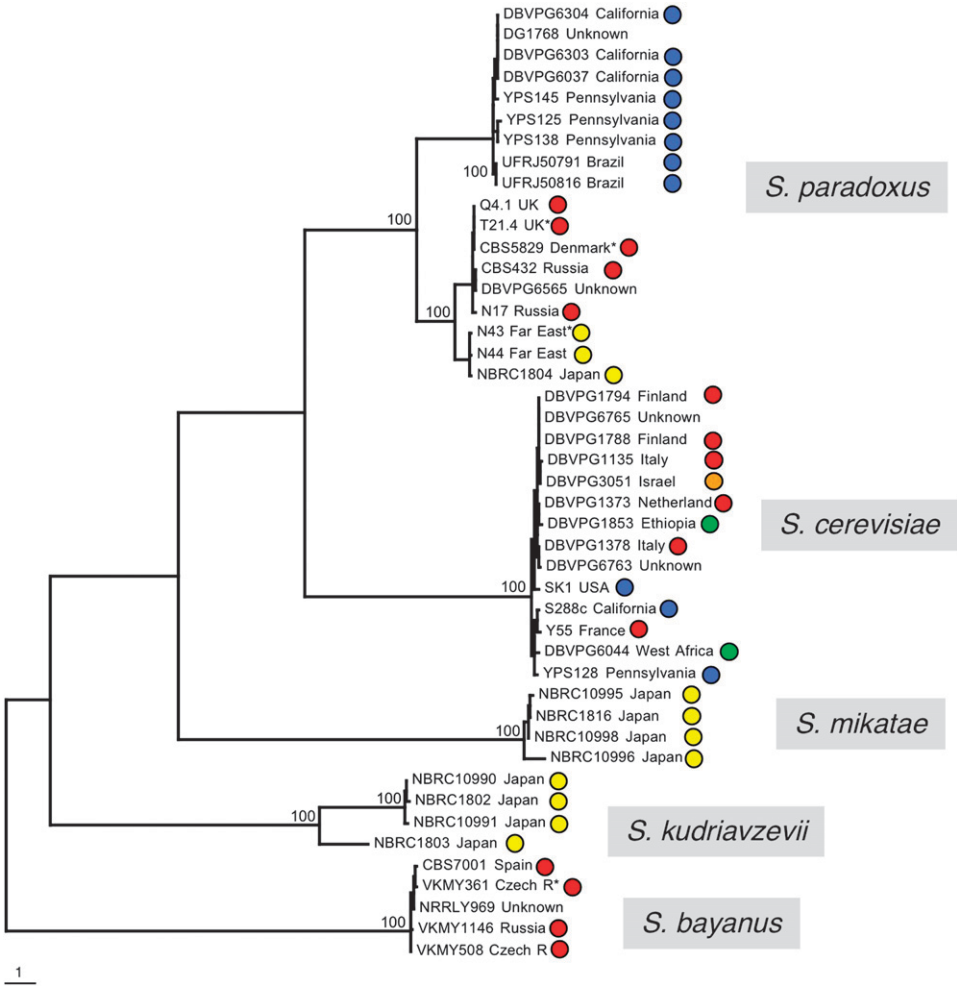

FIGURE 3.—Phylogenetic relationship among strains. Phylogenetic tree obtained from the concatenated analysis of *NEJ1*, *EST2*, *yKU70*, and *yKU80* (total length of 7.4 kbp). *TLC1* and *SPT4* were excluded from the analysis because they were missing from some strains. The tree was obtained using the neighbor-joining algorithm with 1000 bootstrap iterations. The tree was rooted at the *S. bayanus* branch according to ROKAS *et al.* (2003) and 100% of bootstrap value is indicated. The scale represents the number of nucleotide substitutions. An asterisk indicates strains with gross chromosomal rearrangements and colored dots indicate geographic origin (blue, North and South America; red, Europe; yellow, Far East; green, Africa; orange, Middle East).

and varying levels of sequence divergence. We used a Far Eastern isolate, N-43, carrying one reciprocal translocation (VIII t XV and XV t VIII, using the nomenclature of FISCHER *et al.* 2000) and a European strain, CBS5829, carrying one nonreciprocal translocation (XI t V). We crossed each of these with strains from all three geographic subpopulations of *S. paradoxus* (Figure 4). Reciprocal and nonreciprocal chromosomal translocations produce, respectively, 50 and 25% drops in gametic viability due to unbalanced segregation of essential genes on the translocated arms. The translocations are all large enough to contain at least one essential gene. The strains with chromosomal translocations do not show as large a sequence divergence compared to others within the same population, as might have been expected, thereby reinforcing the idea that chromosomal evolution has a minor impact on speciation in the Saccharomyces complex (*sensu stricto*) (FISCHER *et al.* 2000; DELNERI *et al.* 2003). A rapid accumulation of translocations, as happened in *S. cariocanus*, may influence the speciation process over a longer period as they clearly have an effect on fertility. At this point in time, however, *S. cariocanus* has not diverged much in sequence from the rest of the American *S. paradoxus*.

Evidence of introgression from *S. cerevisiae* into the European population of *S. paradoxus*: The smooth monotonic relationship between sequence divergence and gamete viability, rather than the appearance of a discontinuity or rapid decrease, makes the species boundary hard to define. Indeed, the lower, but not negligible, probability of successful breeding between diverged lineages may have strong influences on the population and genomic structures of the species concerned. Traces of interbreeding may still be present in the genomes of Saccharomyces. We searched for evidence of past interbreeding by comparing the available whole-genome sequences of *S. paradoxus* and *S. cerevisiae*. Pairwise comparisons of the two genomes show an average identity of 85% across chromosomes, a finding consistent with earlier analyses (KELLIS *et al.* 2003), except for one region that is aberrantly high. A subtelomeric segment on the left arm of chromosome XIV, 23 kb long, exhibits an anomalously high identity between *S. cerevisiae* S288C and *S. paradoxus* CBS432. The region extends between 15 and 38 kbp from the

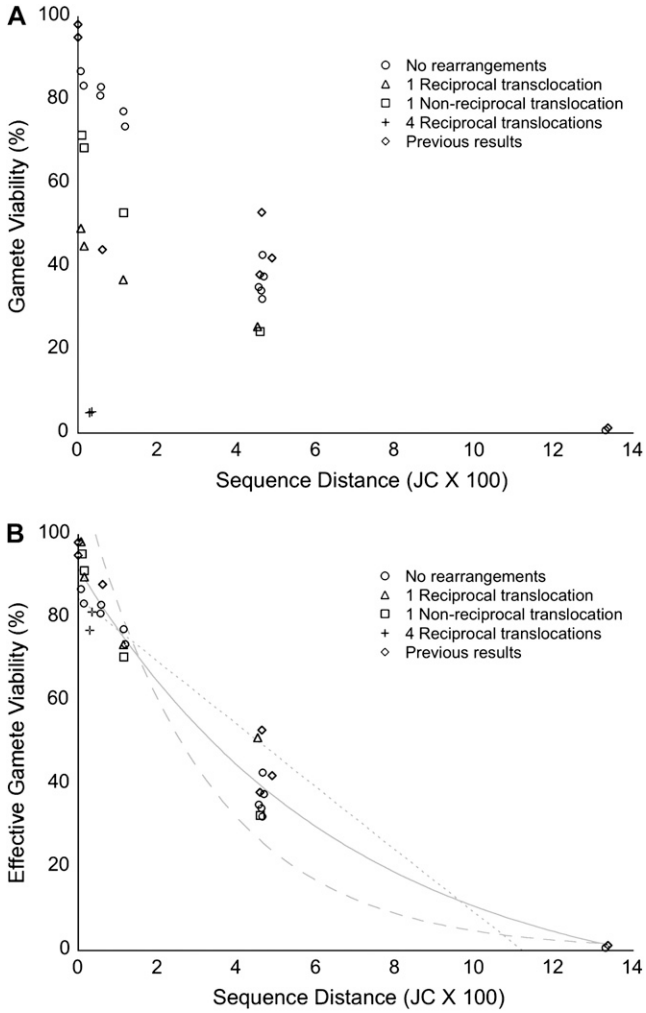

Figure 4.—Sequence divergence *vs.* fertility. Sequence divergence was plotted against gamete viability on a linear scale. In A, spore viability was not corrected for chromosomal translocations. In B, spore viability values were adjusted for the decrease in viability caused by translocations (see text). Various curves fit the data (dotted, linear; dashed, exponential; solid, exponential decay) with high correlation scores (r): linear (0.9347), exponential (0.9614), exponential decay (0.9805), and second order polynomial (0.9798; not shown). Log-transformed spore viabilities (not shown) resulted in a higher correlation (0.9518, linear curve fit) in agreement with Roberts and Cohan (1993).

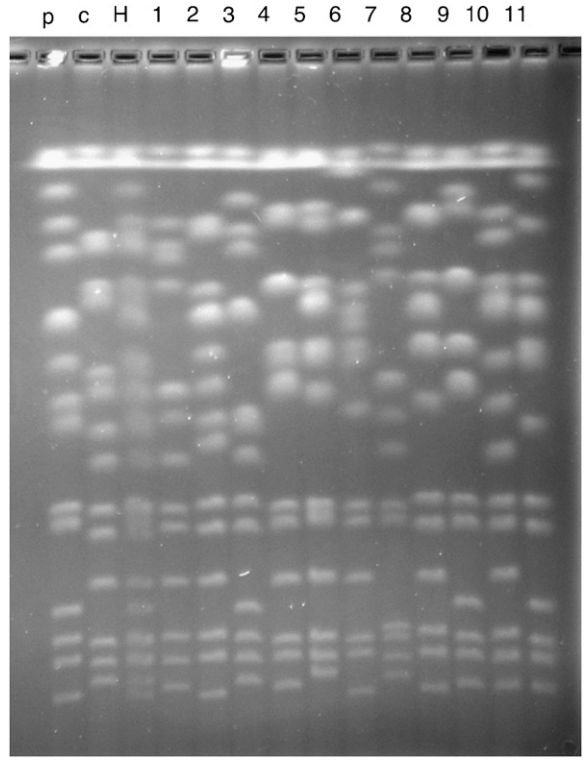

Figure 5.—Chromosome analysis in hybrid segregants. Chromosome separations using CHEF gel. Lanes: p, *S. paradoxus* YPS138; c, *S. cariocanus* UFRJ50816^T; H, cross *S. paradoxus* × *S. cariocanus*; 1–11, independent spores dissected from the hybrid. Chromosome aneuploidy, which appears as double-intensity bands, was not detected.

left telomere (in *S. cerevisiae*) and encompasses 12 open reading frames (Figure 6A). The average identity in this region (>99%) is significantly higher than in the rest of the genome. This could have been error in cloning or sequencing within the genome project, an error in handling the strains concerned, or a true introgression from one species into the other. To exclude a mistake in sequencing. we resequenced a small region within this segment in *S. paradoxus* and confirmed the original results (Figure 6, A and B). We then sequenced the same region in several isolates of *S. cerevisiae* and *S. paradoxus*. We found similar sequences in all strains of *S. cerevisiae* and all European strains of *S. paradoxus*, but not in the Far Eastern and American isolates (Figure 6B), which exhibit an average divergence from *S. cerevisiae* of 15%, in line with the rest of the genome. These results indicate that *S. cerevisiae* was the donor and *S. paradoxus* the recipient, either in a direct lateral transfer or through introgression after hybridization. They also indicate that the transfer took place after the split of the European and Far Eastern lineages. There is the possibility of a partial hybridization, where a single chromosome moves from one genome to the other, as in *kar1* mutants (Nilsson-Tillgren *et al.* 1980). However, we do not know how often such events may occur.

We favor introgression via hybridization for several reasons. Horizontal gene movements are quite rare in yeast (Dujon 2005; Liti and Louis 2005) and a large DNA segment is unlikely to be transferred. Furthermore, the segment involved is a homologous replacement and does not extend up to the telomeric repeats. *S. cerevisiae* and *S. paradoxus* can be found in the same localities in nature (Sniegowski *et al.* 2002) and therefore have the opportunity to hybridize. Naturally occurring hybrids between these species have been found (Liti *et al.* 2005). We propose a scenario where a spontaneous hybridization between *S. cerevisiae* and *S. paradoxus* resulted in a rare viable gamete, perhaps with a

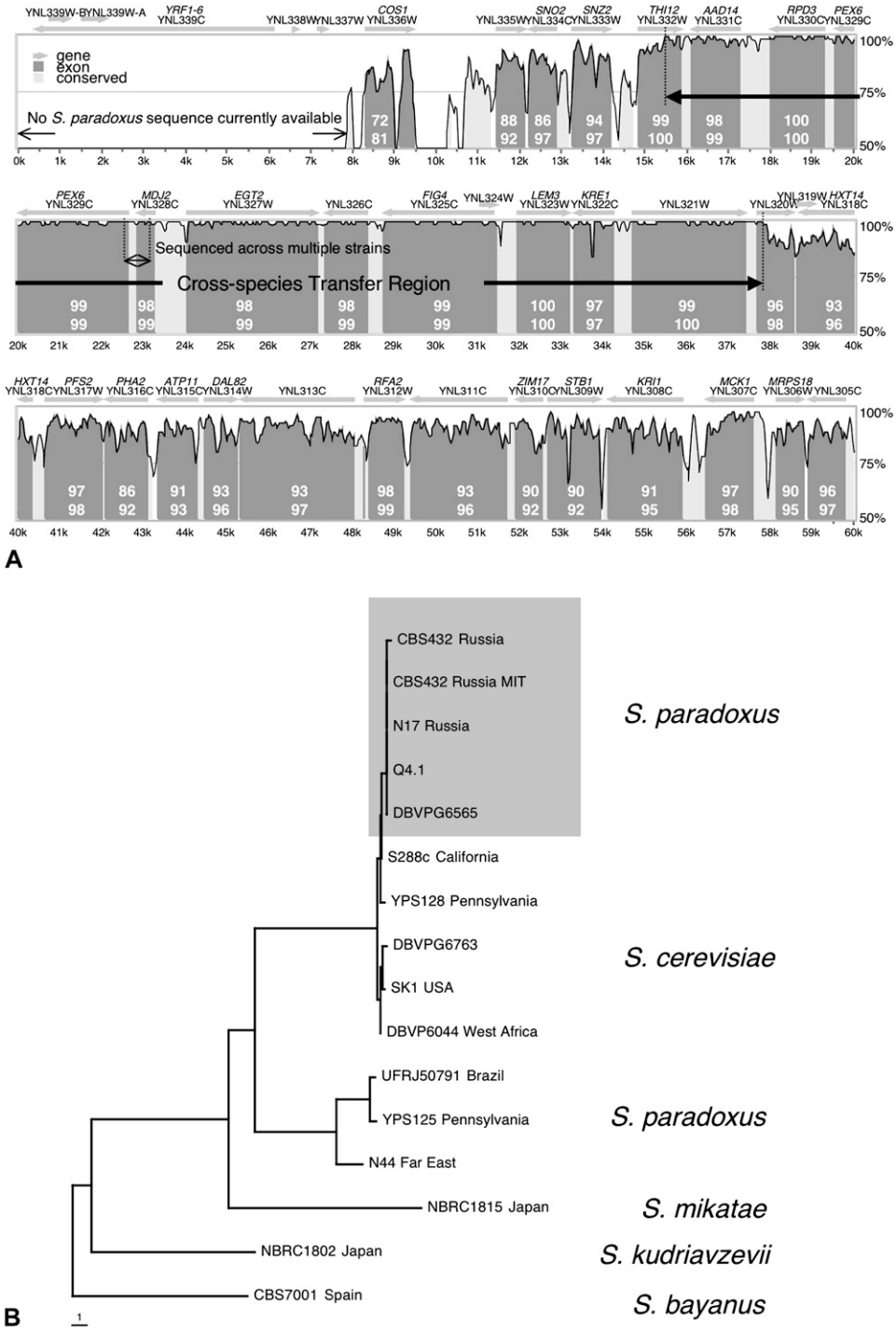

FIGURE 6.—Molecular evidence of introgression from *S. cerevisiae* into the European population of *S. paradoxus*. (A) VISTA alignment of *S. cerevisiae* S288C and *S. paradoxus* CBS432 genomes along the left subtelomeric region of chromosome XIV. The sequence identity is plotted along the pairwise comparison of the left subtelomere. The region showing high identity between *S. cerevisiae* and *S. paradoxus* is shown along with the small region resequenced and then checked in several isolates of both species. (B) Phylogenetic tree of various *S. cerevisiae* and *S. paradoxus* strains using sequence reads from the transferred subtelomeric region. *S. paradoxus* strains thought to be of European descent (shaded box) in this region are much more closely related to *S. cerevisiae* strains than to other geographical groups of *S. paradoxus*. The tree was produced by DNAstar using the neighbor-joining method.

majority of *S. paradoxus* sequences. Viable gametes could be composed of a range of chromosomes from either parent due to random segregation and therefore could have a majority of chromosomes from one parent or the other (G. LITI, unpublished results). The viable gametes, although aneuploid, will be nearly euploid as most of multiple aneuploidies are inviable (PARRY and COX 1970; LOUIS and HABER 1989).This could be followed by many backcrosses within the *S. paradoxus* parent population, leaving a single DNA segment from *S. cerevisiae*. To become fixed in the European population, this segment would probably require a selective advantage. A bottleneck in the ancestral European populations could also have fixed this transferred segment. The *KRE1* gene, which can confer resistance to killer toxins, seems at first sight to be a good candidate in this regard, but the different European isolates proved to have different resistances to killer toxins (results not shown). No cause for a selective advantage has so far been identified.

TABLE 3

Time estimation of chromosomal rearrangements, geographic divergence, and speciation

Events	Years	Relative time
Chromosomal rearrangements		
One reciprocal translocation in N43	1000–8003	0.002
One nonreciprocal translocation in CBS5829	838–6700	0.002
One chromosomal amplification in Q4.1	1210–9678	0.003
Four reciprocal translocations in *S. cariocanus*	5164–41,319	0.012
Geographic subpopulations of *S. paradoxus*		
North/South American	5878–47,027	0.014
Far Eastern/European	27,355–218,847	0.063
Far Eastern–European/American	124,268–994,147	0.286
Speciation		
S. cerevisiae–S. paradoxus	434,084–3,472,677	1

Minimum age of common ancestor: The sequence divergence between populations and species can be used to estimate the minimum age since the most recent common ancestor. Using known mutation rates and estimates of generation times from studies in *S. cerevisiae* (see MATERIALS AND METHODS and FAY and BENAVIDES 2005), absolute dates can be estimated (Table 3). By assuming that the generation times and mutation rates are similar in all of the Saccharomyces *sensu stricto* isolates over time, we can also estimate the perhaps more informative relative age of populations. Using a range of potential generation rates, from one to eight/day on average, we can set the time of the most recent common ancestor of *S. cerevisiae* and *S. paradoxus* to between 0.4 and 3.4 MYA. The relative age of the common ancestor of the most diverse *S. cerevisiae* isolates is 2.7% of this, which is similar to the diversity seen within subpopulations of *S. paradoxus*. The relative age of the different subpopulations of *S. paradoxus* is much higher, ranging from 6.3% between Far Eastern and European isolates to 28.6% of the age of the species when comparing North American to Far Eastern or European populations (Table 3). This analysis can be applied to isolates with gross chromosomal rearrangements, which would be expected to have increased diversity if they were old and prevented gene flow with neighbors not harboring the same rearrangements. In every case, the relative ages of the GCRs containing isolates was on the order of 0.2–0.3% of the age of the species. The South American species *S. cariocanus* has a relative age, compared to the North American species *S. paradoxus*, of only 1.4% of the time to the most recent common ancestor, or between 5000 and 40,000 years ago. The four reciprocal translocations seen in this population must have occurred in a burst on an evolutionary time scale as predicted by FISCHER *et al.* (2000). However, we cannot infer if the translocations occurred simultaneously or sequentially.

Fuzzy boundaries in Saccharomyces species: Earlier results have shown that spontaneous hybrids can be found at reasonable frequencies (DE BARROS LOPES *et al.* 2002; LITI *et al.* 2005) and that the hybrid species *S. pastorianus* originated from several hybridizations rather than from one (LITI *et al.* 2005). Although lateral gene transfer has been rarely detected in Saccharomyces yeast, we have recently suggested that one of the yeast LTR–transposons, Ty2, has been transferred between *S. cerevisiae* and *S. mikatae* (LITI *et al.* 2005).

Overall, our data suggest a continuum of genetic differentiation: extant species may simply be the result of an extensive decay in sequence homology. A gradient of gamete viability is correlated with sequence divergence. These analyses are consistent with studies in Drosophila (COYNE and ORR 2004) and Neurospora (DETTMAN *et al.* 2003) where reproductive isolation (both pre- and postzygotic) increases over time (genetic distance), but in those studies the correlation appears weaker. The findings in Drosophila are consistent with the model of Bateson–Dobzhansky–Muller (BDM) (COYNE and ORR 2004) where two or more genetic incompatibilities account for the reproductive isolation. A special case of the BDM model acted in speciation of the hemi-ascomycete yeasts soon after the whole-genome duplication event (SCANNELL *et al.* 2006). Indeed, "speciation" genes have been identified (ORR *et al.* 2004). One might assume that the findings presented here are consistent with the BDM model of intrinsic postzygotic isolation with a large number of genes involved (COYNE and ORR 2004); however, there is no evidence of genetic incompatibilities here—only divergence acted upon by the normal process of mismatch repair. Any genetic incompatibilities between pairs of strains used in this study should result in fertilities inconsistent with the trend seen for sequence divergence. After accounting for the chromosomal incompatibilities of some crosses due to translocations, all

points fit the correlation with sequence divergence. In Neurospora, there was a great deal more variation in reproductive success over genetic distance than seen here with phylogenetically diverged strains still capable of successful reproduction (Dettman *et al.* 2003).

Evidence of frequent spontaneous hybridization, lateral gene transfer, and introgression suggests that species boundaries are fuzzy and that networks may be present in the phylogeny. Classical phylogenetic trees can give only a glimpse of true relationships. We suggest that a clear picture of the population structure of Saccharomyces requires modified phylogenies that can account for ongoing gene flow due to hybridization, introgression, or lateral transfer with evolving reproductive barriers. The GCSC concept may be the best for describing yeast species and its application at the population level up to the species and genus level may yield a better understanding of species and speciation in Saccharomyces. This view may be relevant to many other species.

We thank Bryan Clarke, Paul Sharp, John Brookfield, Jim Mallet, and Daniel Hartl for helpful discussions and The Wellcome Trust for funding.

LITERATURE CITED

Aa, E., J. P. Townsend, R. I. Adams, K. M. Nielsen and J. W. Taylor, 2006 Population structure and gene evolution in Saccharomyces cerevisiae. FEMS Yeast Res. **6:** 702–715.

Chambers, S. R., N. Hunter, E. J. Louis and R. H. Borts, 1996 The mismatch repair system reduces meiotic homeologous recombination and stimulates recombination-dependent chromosome loss. Mol. Cell. Biol. **16:** 6110–6120.

Cliften, P. F., L. W. Hillier, L. Fulton, T. Graves, T. Miner *et al.*, 2001 Surveying Saccharomyces genomes to identify functional elements by comparative DNA sequence analysis. Genome Res. **11:** 1175–1186.

Coyne, J. A., and H. A. Orr, 1998 The evolutionary genetics of speciation. Philos. Trans. R. Soc. Lond. B Biol. Sci. **353:** 287–305.

Coyne, J. A., and H. A. Orr, 2004 *Speciation.* Sinauer Associates, Sunderland, MA.

Crotti, L. B., and M. A. Basrai, 2004 Functional roles for evolutionarily conserved Spt4p at centromeres and heterochromatin in Saccharomyces cerevisiae. EMBO J. **23:** 1804–1814.

de Barros Lopes, M., J. R. Bellon, N. J. Shirley and P. F. Ganter, 2002 Evidence for multiple interspecific hybridization in Saccharomyces sensu stricto species. FEM Yeast Res. **1:** 323–331.

Delneri, D., I. Colson, S. Grammenoudi, I. N. Roberts, E. J. Louis *et al.*, 2003 Engineering evolution to study speciation in yeasts. Nature **422:** 68–72.

Dettman, J. R., D. J. Jacobson, E. Turner, A. Pringle and J. W. Taylor, 2003 Reproductive isolation and phylogenetic divergence in Neurospora: comparing methods of species recognition in a model eukaryote. Evolution Int. J. Org. Evolution **57:** 2721–2741.

Dudasova, Z., A. Dudas and M. Chovanec, 2004 Non-homologous end-joining factors of Saccharomyces cerevisiae. FEMS Microbiol. Rev. **28:** 581–601.

Dujon, B., 2005 Hemiascomycetous yeasts at the forefront of comparative genomics. Curr. Opin. Genet. Dev. **15:** 614–620.

Fay, J. C., and J. A. Benavides, 2005 Evidence for domesticated and wild populations of Saccharomyces cerevisiae. PLoS Genet. **1:** e5.

Fischer, G., S. A. James, I. N. Roberts, S. G. Oliver and E. J. Louis, 2000 Chromosomal evolution in Saccharomyces. Nature **405:** 451–454.

Gogarten, J. P., and J. P. Townsend, 2005 Horizontal gene transfer, genome innovation and evolution. Nat. Rev. Microbiol. **3:** 679–687.

Goldstein, A. L., and J. H. McCusker, 1999 Three new dominant drug resistance cassettes for gene disruption in Saccharomyces cerevisiae. Yeast **15:** 1541–1553.

Greig, D., R. H. Borts, E. J. Louis and M. Travisano, 2002 Epistasis and hybrid sterility in Saccharomyces. Proc. R. Soc. Lond. B Biol. Sci. **269:** 1167–1171.

Greig, D., M. Travisano, E. J. Louis and R. H. Borts, 2003 A role for the mismatch repair system during incipient speciation in Saccharomyces. J. Evol. Biol. **16:** 429–437.

Hudson, R. R., and J. A. Coyne, 2002 Mathematical consequences of the genealogical species concept. Evolution Int. J. Org. Evolution **56:** 1557–1565.

Hunter, N., S. R. Chambers, E. J. Louis and R. H. Borts, 1996 The mismatch repair system contributes to meiotic sterility in an interspecific yeast hybrid. EMBO J. **15:** 1726–1733.

Isaac, N. J. B., J. Mallet and G. M. Mace, 2004 Taxonomic inflation: its influence on macroecology and conservation. Trends Ecol. Evol. **19:** 464–469.

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

Liti, G., and E. J. Louis, 2003 NEJ1 prevents NHEJ-dependent telomere fusions in yeast without telomerase. Mol. Cell **11:** 1373–1378.

Liti, G., and E. J. Louis, 2005 Yeast evolution and comparative genomics. Annu. Rev. Microbiol. **59:** 135–153.

Liti, G., A. Peruffo, S. A. James, I. N. Roberts and E. J. Louis, 2005 Inferences of evolutionary relationships from a population survey of LTR-retrotransposons and telomeric-associated sequences in the Saccharomyces sensu stricto complex. Yeast **22:** 177–192.

Louis, E. J., 1998 Whole chromosome analysis, pp. 15–31 in *Yeast Gene Analysis*, edited by A. J. P. Brown and M. F. Tuite. Academic Press, London.

Louis, E. J., and R. H. Borts, 1995 A complete set of marked telomeres in *Saccharomyces cerevisiae* for physical mapping and cloning. Genetics **139:** 125–136.

Louis, E. J., and J. E. Haber, 1989 Nonrecombinant meiosis I nondisjunction in *Saccharomyces cerevisiae* induced by tRNA ochre suppressors. Genetics **123:** 81–95.

Mallet, J., 1995 A species definition for the modern synthesis. Trends Ecol. Evol. **10:** 294–299.

Martini, A. V., and A. Martini, 1987 Three newly delimited species of Saccharomyces sensu stricto. Antonie Van Leeuwenhoek **53:** 77–84.

Naumov, G. I., 1987 Genetic basis for classification and identification of the ascomycetous yeasts. Stud. Mycol. **30:** 469–475.

Naumov, G. I., E. S. Naumova, R. A. Lantto, E. J. Louis and M. Korhola, 1992 Genetic homology between Saccharomyces cerevisiae and its sibling species S. paradoxus and S. bayanus: electrophoretic karyotypes. Yeast **8:** 599–612.

Naumov, G. I., T. A. Nikonenko and V. I. Kondrat'eva, 1994 Taxonomic identification of Saccharomyces from yeast genetic stock centers of the University of California. Genetika **30:** 45–48.

Naumov, G. I., E. S. Naumova, A. N. Hagler, L. C. Mendonca-Hagler and E. J. Louis, 1995a A new genetically isolated population of the Saccharomyces sensu stricto complex from Brazil. Antonie Van Leeuwenhoek **67:** 351–355.

Naumov, G. I., E. S. Naumova and E. J. Louis, 1995b Two new genetically isolated populations of the Saccharomyces sensu stricto complex from Japan. J. Gen. App. Microbiol. **41:** 499–505.

Naumov, G. I., S. A. James, E. S. Naumova, E. J. Louis and I. N. Roberts, 2000 Three new species in the Saccharomyces sensu stricto complex: Saccharomyces cariocanus, Saccharomyces kudriavzevii and Saccharomyces mikatae. Int. J. Syst. Evol. Microbiol. **50** (Pt. 5): 1931–1942.

Nilsson-Tillgren, T., J. G. L. Petersen, S. Holmberg and M. C. Kielland-Brandt, 1980 Transfer of chromosome III during Kar mediated cytoduction in yeast Carlsberg Res. Commun. **45:** 113–117.

Orr, H. A., J. P. Masly and D. C. Presgraves, 2004 Speciation genes. Curr. Opin. Genet. Dev. **14:** 675–679.

Parry, E. M., and B. S. Cox, 1970 The tolerance of aneuploidy in yeast. Genet. Res. **16:** 333–340.

Rayssiguier, C., D. S. Thaler and M. Radman, 1989 The barrier to recombination between Escherichia coli and Salmonella

typhimurium is disrupted in mismatch-repair mutants. Nature **342:** 396–401.

Rieseberg, L., 1997 Hybrid origins of plant species. Annu. Rev. Ecol. Syst. **28:** 359–389.

Roberts, M. S., and F. M. Cohan, 1993 The effect of DNA sequence divergence on sexual isolation in Bacillus. Genetics **134:** 401–408.

Rokas, A., B. L. Williams, N. King and S. B. Carroll, 2003 Genome-scale approaches to resolving incongruence in molecular phylogenies. Nature **425:** 798–804.

Rozas, J., and R. Rozas, 1999 DnaSP version 3: an integrated program for molecular population genetics and molecular evolution analysis. Bioinformatics **15:** 174–175.

Scannell, D. R., K. P. Byrne, J. L. Gordon, S. Wong and K. H. Wolfe, 2006 Multiple rounds of speciation associated with reciprocal gene loss in polyploid yeasts. Nature **440:** 341–345.

Smogorzewska, A., and T. de Lange, 2004 Regulation of telomerase by telomeric proteins. Annu. Rev. Biochem. **73:** 177–208.

Sniegowski, P. D., P. G. Dombrowski and E. Fingerman, 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 Res. **1:** 299–306.

Taylor, J. W., D. J. Jacobson, S. Kroken, T. Kasuga, D. M. Geiser *et al.*, 2000 Phylogenetic species recognition and species concepts in fungi. Fungal Genet. Biol. **31:** 21–32.

Vaughan Martini, A., 1989 Saccharomyces paradoxus comb. nov., a newly separated species of the Saccharomyces sensu stricto complex based upon nDNA/nDNA homologies. Syst. Appl. Microbiol. **12:** 179–182.

Communicating editor: D. Fw. Voytas