Effect of Domestication on the Spread of the [PIN+] Prion in *Saccharomyces cerevisiae*

Amy C. Kelly,* Ben Busby,[†] and Reed B. Wickner**¹

*Laboratory of Biochemistry and Genetics, National Institute of Diabetes, Digestive and Kidney Diseases, and [†]National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, Maryland 20892

ABSTRACT Prions (infectious proteins) cause fatal neurodegenerative diseases in mammals. In the yeast *Saccharomyces cerevisiae*, many toxic and lethal variants of the [PSI+] and [URE3] prions have been identified in laboratory strains, although some commonly studied variants do not seem to impair cell growth. Phylogenetic analysis has revealed four major clades of *S. cerevisiae* that share histories of two prion proteins and largely correspond to different ecological niches of yeast. The [PIN+] prion was most prevalent in commercialized niches, infrequent among wine/vineyard strains, and not observed in ancestral isolates. As previously reported, the [PSI+] and [URE3] prions are not found in any of these strains. Patterns of heterozygosity revealed genetic mosaicism and indicated extensive outcrossing among divergent strains in commercialized environments. In contrast, ancestral isolates were all homozygous and wine/vineyard strains were closely related to each other and largely homozygous. Cellular growth patterns were highly variable within and among clades, although ancestral isolates were the most efficient sporulators and domesticated strains showed greater tendencies for flocculation. [PIN+] infected strains had a significantly higher likelihood of polyploidy, showed a higher propensity for flocculation compared to uninfected strains, and had higher sporulation efficiencies compared to domesticated, uninfected strains. Extensive phenotypic variability among strains from different environments suggests that *S. cerevisiae* is a niche generalist and that most wild strains are able to switch from asexual to sexual and from unicellular to multicellular growth in response to environmental conditions. Our data suggest that outbreeding and multicellular growth patterns adapted for domesticated environments are ecological risk factors for the [PIN+] prion in wild yeast.

PRIONS cause a variety of transmissible spongiform encephalopathies (TSEs) in mammals, including scrapie in sheep, bovine spongiform encephalopathy (BSE) in cattle, chronic wasting disease (CWD) in cervids, and Creutzfeldt-Jakob disease (CJD) in humans. With all TSEs, endogenous prion protein (PrP) misfolds into an altered, usually protease-resistant, isoform, termed “PrPres” (“res” for protease resistant).

Several prions have been described in fungi (Wickner 1994), with many prion variants of the [PSI+] and [URE3] prions showing toxic or even lethal effects in *Saccharomyces cerevisiae* (McGlinchey *et al.* 2011). The [PSI+] prion of *S. cerevisiae* is formed from the cytosolic protein Sup35p. There are three distinct regions of Sup35p: a glutamine/

asparagine-rich N-terminal domain (amino acids 1–123) that is necessary and sufficient for prion formation (Ter-Avanesyan *et al.* 1994) and that facilitates deadenylation and decay of messenger RNA (Hosoda *et al.* 2003); a middle M domain (amino acids 124–253) of unknown function; and an essential C-terminal domain that functions during translation termination (reviewed by Wickner *et al.* 2004). The function of Rnq1p is not yet known, but the amyloid conformation of this protein, [PIN+], enhances the *de novo* formation of other yeast prions, including [PSI+] (hence [PIN+], for [PSI+] inducibility) (Derkatch *et al.* 1997, 2001).

The natural abundance of prions should reflect their rate of spread via mating, relative fitness effects, and spontaneous loss or gain of the prion state. Ecological factors could influence mating frequency and thus prion prevalence in the wild. Magwene *et al.* (2011) reported that strains from natural environments more readily undergo meiosis (under the tested conditions) compared to domesticated strains, whereas pseudohyphal growth was more readily induced in domesticated strains (Magwene *et al.* 2011). During meiosis wild yeast generate four haploid spores that can subsequently mate with

Copyright © 2014 by the Genetics Society of America
doi: 10.1534/genetics.114.165670

Manuscript received January 21, 2014; accepted for publication May 6, 2014;
published Early Online May 8, 2014.

Supporting information is available online at <http://www.genetics.org/lookup/suppl/doi:10.1534/genetics.114.165670/-DC1>.

¹Corresponding author: Laboratory of Biochemistry and Genetics, National Institute of Diabetes, Digestive and Kidney Diseases, National Institutes of Health, Room 207, 8 Center Dr., Bethesda, MD 20892. E-mail: wickner@helix.nih.gov

(1) cells from the same spore clone (homothallism), (2) cells from other spores within the ascus (intratetrad mating), or (3) spores from a different yeast clone (outcrossing) (Strathern *et al.* 1981). Studies conducted in *Saccharomyces paradoxus* have estimated the frequency of outcross mating at 1% of total matings and once per 10⁵ mitotic divisions (Tsai *et al.* 2008). For *S. cerevisiae*, estimates of mating from genomic data reported two outcrosses per 10⁵ mitotic doublings (Ruderfer *et al.* 2006), while another, more recent study reported one outcross per 100 mitotic divisions (Kelly *et al.* 2012).

Alternative to sporulation, nutrient deprivation can induce the formation of pseudohyphal filaments that allow yeast to invade solid media and forage for nutrients. This phenotype is highly variable, however, and even among genetically homogeneous strains only a subset of cells will undergo a pseudohyphal response to nutrient deprivation (Gimeno *et al.* 1992). Additionally, flocculation, characterized by increased cell–cell adhesion resulting in aggregates of vegetative cells, is often observed when sugars are depleted from the media (Guo *et al.* 2000). Enhanced cellular aggregation provides protection in harsh environments (Bruckner and Mosch 2011), and flocculent yeast strains are often utilized in beer fermentation and other industrialized settings (Verstrepen and Klis 2006).

The yeast *S. cerevisiae* has been isolated globally from a variety of natural substrates (fruit, tree bark, soil) (Sniegowski *et al.* 2002; Wang *et al.* 2012) and from environments closely associated with human activity (breweries, bakeries, vineyards) (Legras *et al.* 2007). We refer to strains isolated and adapted (or bred) for human use as domesticated strains. We include brewing, baking, vineyard, and clinical strains. We regard even “non-inoculated” wine strains isolated from vineyards as domesticated since they have been isolated from agricultural crops after at least one human use. Although there has been a recent surge in the number of articles examining yeast genomics, population genetics (Fay and Benavides 2005; Koufopanou *et al.* 2006; Liti *et al.* 2009), and reproduction (Ruderfer *et al.* 2006; Tsai *et al.* 2008; Kelly *et al.* 2012), few studies focus on the ecology of yeast in natural and domesticated environments. To improve our understanding of how ecological factors influence prion dynamics in yeast, we characterized genetic and phenotypic traits of wild isolates of *S. cerevisiae*, including [PIN+]–infected and uninfected strains. With these data we were able to identify ecological factors that contribute to [PIN+] transmission in wild fungi and suggest effects of human activity on the spread of [PIN+] in *S. cerevisiae*.

Materials and Methods

Genetic analysis

We amplified and sequenced *RNQ1*, *SUP35*, *TRP1*, and *YGL108C* loci from 75 wild *S. cerevisiae* isolates, 15 of which were previously shown to carry the [PIN+] prion (Nakayashiki *et al.* 2005; Kelly *et al.* 2012). The *RNQ1* and *SUP35* loci were

chosen primarily because these genes code for the most commonly studied yeast prions, but also because previous studies have shown that high polymorphism at *RNQ1* (Resende *et al.* 2003; Kelly *et al.* 2012) makes this locus a good candidate for population genetic studies (Kelly and Wickner 2013). The other two loci, *TRP1* and *YGL108C*, were chosen as controls. *TRP1*, like *SUP35*, is an essential, functional gene, and *YGL108C*, like *RNQ1*, is a nonessential open reading frame that codes for a protein of unknown function. When necessary, PCR products were cloned or spores were sequenced to resolve ambiguous sequencing reads from polyploids or strains heterozygous at more than one segregating site. We documented the location and frequency of SNPs, insertions, and deletions with CodonCode Aligner (version 3.7.1; CodonCode) using nucleotide sequences of laboratory strain S288C as reference.

We aligned nucleotide sequences of our strains with MAFFT (Kato *et al.* 2002) and inferred phylogeny from the four loci using a maximum-likelihood approach and the generalized time-reversible model of nucleotide evolution in FastTree (Price *et al.* 2009). Support for the tree was evaluated using 1000 bootstrap replicates to test for alternate topologies of each split (Price *et al.* 2010). The tree was rooted using homologous nucleotide sequences from *S. paradoxus* obtained from the Saccharomyces Genome Database.

DNAsp version 5.10 (Librado and Rozas 2009) was used to infer haplotypes for *RNQ1*. Haplotype inferences were confirmed by comparison with sequences from single clones or spores of heterozygous individuals. We used PHASE version 2.1 (Stephens and Scheet 2005) to calculate haplotype frequencies of [PIN+]–infected and uninfected strains and created a haplotype map in TCS version 2.1 (Clement *et al.* 2000) to visualize nucleotide substitutions and examine [PIN+] infections in haplotype groups. In all phylogenetic analyses gaps were considered missing data, and in haplotype analyses insertions and deletions were treated as additional alleles.

For molecular genetic analysis, we used the three most obvious natural splits to subdivide the phylogenetic tree into four clades, which identified groups with similar prion protein histories. To confirm clade membership for strains with low phylogenetic bootstrap values or discordant ecological origins (e.g., a wine strain that phylogenetically grouped with ancestral isolates), we used GeneClass version 2.0 (Piry *et al.* 2004) to perform assignment tests on diploid SNP genotypes. We employed allele-frequency-based methods to determine the probable origin of observed alleles and the likelihood of clade membership for each strain. Heterozygous ambiguity characters within nucleotide sequences are ignored by FastTree during phylogenetic analysis. However, both homozygous and heterozygous diploid genotypes were considered with assignment tests, which allowed us to further resolve clade membership. Methods for calculating likelihood scores for assignment tests are described in Piry *et al.* (2004) and in Supporting Information, File S1. Strains were assigned to the clade with the highest likelihood score.

Figure 1 Maximum-likelihood tree of *S. cerevisiae* nucleotide sequences. [PIN+] strains are highlighted in yellow. Branch lengths are proportional to the number of nucleotide substitutions per site, as indicated by the scale bar. Local bootstrap values are indicated by the numbers shown at branch nodes. The tree was rooted using homologous gene sequences from *S. paradoxus*.

Additionally, we examined outcrossing among clades by conducting assignment tests on single-locus haplotypes. For each strain, haplotypes for *SUP35* and *RNQ1* (*TRP1* and *YGL108C* were not included because low polymorphism at these loci prevented a powerful analysis) were assigned to the ecological clade with the highest likelihood score. To examine genetic mosaicism (Liti *et al.* 2009), we quantified the extent of genetic mixing at different regions of the genome based on each strain's assignment scores for *SUP35* and *RNQ1* haplotypes. We also created a mosaicism metric (*M*), which quantifies intragenic mosaicism, and was calculated as follows:

$$M_{\text{haplotype}} = \frac{[(1 - L1) * (1 - L2) * (1 - L3) * (1 - L4)]}{(1 - 0.25)^4},$$

where L1–L4 are assignment scores for each clade. The denominator serves to scale each score by maximum mosaicism (0.25 in each of the four clades) such that scores range from 0 (no mosaicism: all alleles have a high likelihood of originating from a single clade) to 1 (maximum mosaicism: alleles originated from all four clades). For diploid and polyploid strains, the highest haplotype mosaic score for *SUP35* and *RNQ1* was used for analysis.

Molecular population genetic analysis based on segregating sites in the four sequenced loci was performed in DNAsp (Librado and Rozas 2009). Analyses were conducted on the

four genetic clades delineated from the tree and assignment tests and also with clades 2 and 3 combined into a single group. We calculated several metrics of nucleotide variation for the four clades: the frequency of synonymous and non-synonymous polymorphisms, nucleotide diversity (π , the average number of nucleotide differences per base between two strains), nucleotide polymorphism (θ , the total number of nucleotide differences, standardized by the number of samples), observed heterozygosity (H_o : the proportion of heterozygous individuals in the population), and expected heterozygosity [H_e : $(1 - \sum p_i^2)$, where p_i is the frequency of the i^{th} allele]. We tested for natural selection on nucleotide sequences by comparing nucleotide diversity (π) and nucleotide polymorphism (θ) with Tajima's *D*. Tajima's *D* was calculated as the standardized differences between π and θ (Librado and Rozas 2009). We also used GenA1Ex version 6.5 (Peakall and Smouse 2012) to calculate the number of private alleles per clade (alleles observed exclusively in a single clade). For all tests, *P*-values < 0.05 were considered significant.

We used GenePop version 4.2 (Rousset 2008) to calculate inbreeding coefficients (F_{IS}) and fixation indices (F_{ST}) for each clade based on diploid SNP genotypes. Additionally, we estimated the fraction of outcross matings given H_o and H_e according to Kelly *et al.* (2012).

Table 1 Origins of strains

Strain	Strain ID	Origin	Clade
CBS400	1	Sap of <i>Elaias guineensis</i> (palm wine), Ivory Coast, Africa	1
CBS405	2	Juice of <i>Osbeckia grandiflora</i> (bili wine), West Africa	1
CBS5287	3	Grapes, Russia	1
I4-36	4	Wild strain from insect gut in Louisiana, United States	1
I4-41	5	Wild strain from insect gut in Louisiana, United States	1
SD O4	6	North Carolina soil sample	1
SM12	7	North Carolina soil sample	1
Y12603	8	Fermented food, India	1
ATCC38554	9 ^a	Canned cherries	2
Fleishmann's	10	Baker's yeast	2
SAF Perfect Rise	11	Baker's yeast, Lesaffre Manufacturing, Belgium	2
Wyeast #2112XL	12	Brewery, lager, California	2
Y11878	13	Cane juice, Jamaica	2
Y12632	14	Top yeast, The Netherlands	2
Y12679	15	Tapai-ubi, fermented tapioca, Malaya	2
Y2416	16 ^a	Brewery, Europe	2
Y5997	17	Ragi, Java, Indonesia	2
Y7662	18	Wild strain, Pozol, Mexico	2
YB-3916	19 ^a	Clinical sputum, Oslo	2
YJM339	20 ^a	Clinical isolate, bile tube	2
YJM428	21	Clinical isolate, paracentesis fluid	2
WLP002	22	Brewery, ale, United Kingdom	2
Y976	23	Baker's yeast	2
YB2573	24 ^a	Fruit cocktail, refrigerated, Peoria, IL	2
WLP300	25	Wheat beer, Germany	3
Wyeast #1272	26	Brewery, ale, United States	3
ATCC38555	27	Pasteurized canned applesauce, Belgium	3
ATCC66349	28	Candied apple, Tokyo	3
ATCC96362	29	Wheat sourdough, Umbria, Italy	3
Boots Co. home beer	30 ^a	Brewery, Nottingham, United Kingdom	3
CBS6216	31 ^a	Tap water, Rotterdam, The Netherlands	3
MMRL1620	32	Fruit, Philippines	3
FPS449	33	Sourdough	3
Wyeast #1007	34	Brewery, ale, Germany	3
Y132	35	Distillery yeast, United States	3
Y1375	36 ^a	United States	3
Y382	37 ^a	Grain, Minnesota	3
Y977	38 ^a	Baker's yeast	3
YJM280	39	Clinical isolate, peritoneal fluid	3
YJM413	40	Clinical isolate, blood	3
ATCC1855	41 ^a	Unknown	4
ATCC2188	42 ^a	Unknown	4
ATCC32900	43 ^a	Ripe pear fruit, India	4
ATCC4127	44 ^a	Concord grapes, United States	4
ATCC76677	45 ^a	Fermenting fruit, Indonesia	4
CBS2087	46 ^a	Flower of lychee, China	4
CBS429	47 ^a	Champagne grapes, France	4
CBS4734	48	Sugar cane	4
CBS7957	49 ^a	Factory producing cassava flour, Brazil	4
Red Star	50 ^a	Dry wine	4
WLP565	51 ^a	Brewery, saison, Belgium	4
WLP705	52 ^a	Sake yeast	4
WLP715	53 ^a	Champagne grapes	4
Y11846	54	Alpechin, Spain	4
Y12617	55	Red wine, Spain	4
Y12629	56	Grape must, Spain	4
Y12637	57 ^a	Grape must, South Africa	4
Y12644	58 ^a	Wine, Spain	4
Y12649	59 ^a	Grape must, Italy	4
Y12656	60	Alpechin, Spain	4
Y12659	61 ^a	Human feces, Portugal	4
Y12660	62	Wine, Spain	4

(continued)

Table 1, continued

Strain	Strain ID	Origin	Clade
Y140	63 ^a	Wine, Switzerland	4
Y17732	64 ^a	Unknown	4
Y2034	65 ^a	Wine yeast, California	4
Y35	66 ^a	Fruit, <i>Ilex aquifolium</i> , Europe	4
Y629	67 ^a	Distillery yeast, United States	4
Y6677	68	Alpechín, Spain	4
Y6679	69 ^a	Alpechín, Spain	4
Y6680	70	Grape must, Russia	4
YB1191	71 ^a	Citrus juice, Southern Regional Research Laboratory, New Orleans	4
YB-4237	72	Grape must, Spain	4
YJM145	73	Clinical isolate, lung of AIDS patient	4
YJM320	74 ^a	Clinical isolate, blood	4
YJM326	75 ^a	Clinical isolate	4

^a Indicates strain with reference Rnq1p haplotype.

Phenotypic evaluation

To determine if cellular growth patterns correlate with the presence of the [PIN+] prion, we examined sporulation, flocculation, and pseudohyphal growth in several types of media. Three conditions for sporulation were examined: (1) potassium acetate agar media—0.17% yeast nitrogen base without amino acids, 1% potassium acetate, 50 μ M ammonium sulfate, and 2% agar (Magwene *et al.* 2011) at 23°; (2) McClary's media—0.1% glucose, 0.18% potassium chloride, 0.82% sodium acetate, 0.25% yeast extract, and 1.5% agar (McClary *et al.* 1959) at 23°; and (3) liquid potassium acetate with raffinose—0.3% potassium acetate and 0.02% raffinose (Abelson *et al.* 1991) at 30°. For growth in liquid media, cells were first grown to log phase in YPAD (1% yeast extract, 2% peptone, 2% glucose, 0.04% adenine) and then inoculated to a density of 5×10^5 cells/ml. For growth on agar media, cells were grown to a patch on YPAD and thinly spread onto agar plates. The number of tetrads, triads, and dyads out of 200 cells was counted after 7 days using a light microscope.

We evaluated pseudohyphal growth, cellular aggregation/flocculation, adhesion, and agar invasion by comparing yeast growth on YPAD (1% yeast extract, 2% peptone, 2% glucose, 0.04% adenine, 2% agar) to growth on SLAD [0.17% yeast nitrogen base without amino acids, 1% glucose, 50 μ M ammonium sulfate, and 2% agar (Gimeno *et al.* 1992)] and reduced glucose media (1% yeast extract, 2% peptone, 1% glucose, 0.04% adenine, and 2% agar). Using methods described in Zupan and Raspor (2008), cells were spotted onto agar media with an inoculation loop. After 7 days of growth at 30°, the proportion of cells growing pseudohyphally (elongated necks, unipolar budding, filaments) out of 200 cells on SLAD and reduced glucose media was examined using a light microscope. Additionally, we performed a plate-washing assay to test for adhesion (Guldal and Broach 2006) and subsequently assayed agar invasion by measuring the extent (in mm) of filament invasion in cross sections of agar.

We counted the number of cells in tightly clustered aggregates (flocs) out of 200 cells and subsequently performed

a flocculation test to determine if cellular aggregates observed under the microscope exhibited rapid sedimentation characteristic of flocculent strains (Soares and Mota 1997). A flocculation rate assay described by Smit *et al.* (1992) was modified for wild strains (Smit *et al.* 1992). Cells from SLAD plates (cellular aggregations were most apparent on SLAD media) were resuspended in deionized water to a final concentration of 5×10^7 cells/ml. Cells were vortexed for 5 sec, and 400 μ l of the cell suspension was immediately added to a 1.0-ml cuvette. Using a 400- μ l sample volume allowed us to measure the optical density (OD) slightly below the meniscus of the cell suspension with a Spectronic Genesys 2 spectrophotometer. The cuvette was inverted 10 times to induce flocculation (Smit *et al.* 1992), and OD was immediately recorded. Measurements were also recorded every minute for the first 5 min and again after 1 hr of sedimentation.

Statistical analysis

Statistical analyses were conducted in SAS (version 9.3; SAS Institute, Cary, NC). After examining the distributions of genetic and phenotypic variables and assessing data normality, we performed a nonparametric Kruskal–Wallis test on ranked values to identify genetic and phenotypic traits that differed among clades. The Mann–Whitney *U*-test was used to examine differences in phenotype and genetic traits between prion-infected and uninfected strains. Fisher's exact tests were performed to assess associations with prion status. All tests were considered significant at $P < 0.05$, although we also discuss results after applying a false discovery rate (Benjamini and Hochberg 1995) method to control for multiple significance testing. Details of statistical tests are outlined in the [Supporting Information, File S1](#).

Results

Genetic analysis

To examine genetic determinants of fungal prions and elucidate strains with similar prion protein histories, we characterized genetic variation at two prion loci and two

Table 2 Molecular population genetic analysis of *S. cerevisiae* nucleotide sequences

Clade ^a	<i>n</i>	% [PIN+]	% 2-μ DNA	No. of SNPs	No. of nonsynonymous SNPs	<i>H_O</i> ^b	π ^c	θ ^d	<i>H_e</i> ^e	<i>F_{IS}</i> ^f	% of outcross matings ^g
Ancestral	8	0	0	26	6	0.00	0.0026	0.0019	0.72	1.00	0
Commercial brewery	17	47	53	68	25	0.77	0.0035	0.0040	0.94	0.40	60
Commercial food	15	33	47	58	21	0.60	0.0032	0.0035	0.93	0.42	38
Wine/vineyard	35	6	66	70	31	0.49	0.0014	0.0035	0.96	0.60	26
Commercial clades	32	41	50	88	36	0.69	0.0038	0.0044	0.97	0.49	45

^a Clade membership was determined by phylogenetic analysis at four loci and assignment tests.

^b *H_O*, observed heterozygosity calculated as number of heterozygotes/*n*.

^c π , nucleotide diversity; the average number of nucleotide differences per base between two strains.

^d θ , nucleotide polymorphism; the total number of nucleotide differences per base, standardized by the number of strains.

^e *H_e*, expected heterozygosity, calculated as $(1 - \sum p_i^2)$, where p_i is the frequency of the *i*th allele.

^f *F_{IS}*, the inbreeding coefficient, calculated according to Rousset (2008).

^g Percentage of outcross matings from heterozygosity, calculated according to Kelly *et al.* (2012).

control genes in *S. cerevisiae*. We identified 343 segregating sites (including insertions, deletions, and SNPs) and 140 SNPs of 4287 nucleotides sequenced from three genomic regions. Phylogenetic analysis of genetic data from two prion genes and two control genes revealed four major clades of *S. cerevisiae* (Figure 1). Rooting the tree with *S. paradoxus* sequences indicated that a clade of isolates sampled from insects, soil, plant material, and grapes (Table 1) was the most ancestral. In addition to this clade, there were two apparent clusters of commercial isolates (bakery, brewery, processed foods) and one cluster of wine/vineyard isolates (wine, grape must, etc.). Here we will refer to phylogenetic groupings as the following: clade 1: ancestral; clade 2: commercial brewery; clade 3: commercial food; and clade 4: wine/vineyard. Clades consisted of strains from similar ecological niches and for descriptive purposes were named based on the majority of strains in each clade.

To confirm clade membership for strains with low phylogenetic bootstrap values or discordant ecological origins, we performed assignment tests on diploid SNP genotypes. Overall, results from genotype assignment tests (which consider both alleles of heterozygotes) were consistent with the tree topology (which ignores heterozygous alleles). For the vast majority of strains, assignment scores into their respective phylogenetic clades were very high (>99.9%), indicating a high likelihood that the phylogenetic tree topology was consistent with population genetic structure. Assignment of SNP genotypes revealed that six strains shared SNPs with both commercial brewery and commercial food, but were assigned to a clade different from the tree predicted (Supporting Information, Table S1). This result was somewhat expected, given that ambiguous characters (*i.e.*, heterozygotes) are not considered during phylogenetic analysis, but would contribute to the likelihood assignment scores. For all remaining analyses the six strains that were assigned to a clade that differed from the one predicted by the tree were allotted to the clade with the highest likelihood assignment score.

Although some isolates possessed alleles from multiple clades, all clades were significantly differentiated geneti-

cally, with pairwise *F_{ST}* values ranging from 0.27 to 0.65. Each clade had between 3 and 13 private alleles that were shared among clade members but not detected in other clades.

Genetic and phenotypic characteristics of clades

We examined genetic and phenotypic characteristics of clades to identify ecological factors that influence [PIN+] spread. The prevalence of [PIN+] varied among clades (Table 2): ancestral isolates were uninfected, the two commercial clades had the highest prevalence, and the wine/vineyard clade had a relatively low prevalence (Fisher's exact $P < 0.001$). Moreover, the sexually transmitted 2-μm DNA plasmid was not detected in ancestral isolates, but was found in half of the commercial isolates and in two-thirds of the wine/vineyard isolates. Unlike the other three clades that contained a mixture of polyploids, heterozygotes, and homozygotes (Figure 1), ancestral isolates were all homozygous and were the least polymorphic (Table 2). Within the wine/vineyard clade, isolates had the lowest nucleotide diversity observed across all genomic regions (Table 2 and Figure S1; π , average number of pairwise differences between strains) and, compared to the commercial isolates, were substantially more inbred. Although nucleotide polymorphism in the two commercial clades was similar to that observed in the wine/vineyard clade, the commercial clades exhibited more heterozygosity (both observed *H_O* and expected *H_e*) and had a higher fraction of outcross matings.

To determine if [PIN+]-infected isolates exhibited evidence of prior mating, we used the number of heterozygous, shared SNPs per strain as an indicator of outcrossing for each strain. Individuals that are heterozygous for an allele that is commonly observed in the population (*i.e.*, shared SNPs) were likely produced by outcrossing (assuming that the probability of two independent mutational events occurring at the same nucleotide is low). This metric allowed us to distinguish between heterozygotes produced by outcrossing and those produced by asexual reproduction because single mutations accumulated during asexual growth would be observed only in a single strain. The proportion of

Figure 2 Variation in genetic characteristics among clades. Bars represent average values shown with standard deviations for each clade. P -values for the Kruskal–Wallis test are shown in upper right of each panel. (A) The number of heterozygous, shared SNPs per strain was averaged for each clade. (B and C) Intragenic haplotype mosaic scores were calculated for *RNQ1* and *SUP35*, and the highest score per strain was used to calculate an average for each clade.

heterozygous, shared SNPs per strain differed significantly among clades (Figure 2A; Kruskal–Wallis $P = 0.0003$). None of the ancestral isolates possessed heterozygous, shared alleles, whereas 60% of the commercial food isolates and 46% of the wine/vineyard isolates were heterozygous for at least one shared allele. All of the strains in the commercial brewery clade possessed heterozygous, shared alleles.

Commercial yeast are well known to be frequently polyploid (reviewed by Querol and Bond 2009). A total of 16 polyploid strains were identified. Thirteen were from commercial clades, and these strains had polyploid haplotypes at *RNQ1*, *SUP35*, or both loci. In contrast, three polyploid strains detected in the wine/vineyard clade were polyploid only at the *SUP35* locus. Polyploid strains had three to seven different haplotypes per locus (Figure 3), and over half would not sporulate or formed tetrads very infrequently. Of the seven polyploid strains that produced tetrads, spore viability was on average >40%, and all mated with *MATa* and/or *MATα* tester strains.

We examined interbreeding and therefore the propensity for [PIN+] spread among clades by performing assignment tests on diploid SNP profiles (from all four sequenced loci)

and haplotypes of *RNQ1* and *SUP35*. Many of the strains, particularly those from commercial clades, showed patterns of genetic mosaicism (Figure 3). Mosaic strains have mixed lineages and show variable patterns of ancestry for different regions of the genome. Nearly half (45%) of the strains showed genomic mosaicism in that at least one *RNQ1* or *SUP35* haplotype was assigned to a clade that differed from the one designated by diploid SNP profiles. Intragenic mosaicism, which we interpreted as gene flow among clades, was also apparent in single haplotypes with partial assignment to multiple clades (multicolored haplotypes Figure 3, B and C) and in strains with different haplotypes of the same locus assigned to different clades. Our metric of genetic mosaicism showed that mosaicism at *RNQ1* and *SUP35* differed among clades (Figure 2, B and C; Kruskal–Wallis $P = 0.0001$). At *RNQ1*, commercial clades had mosaic scores that were on average threefold higher than the wine/vineyard clade and 10-fold higher than the ancestral clade. At *SUP35*, the ancestral clade exhibited very little mosaicism, whereas the commercial food clade was highly mosaic compared to the commercial brewery and wine/vineyard clades.

Figure 3 Haplotype assignments into clades based on likelihood scores calculated for *SUP35* and *RNQ1*. Haplotypes were inferred using DNAsp (Librado and Rozas 2009) and confirmed by sequencing clones and spores of heterozygous strains. Bars represent the likelihood of haplotype assignment into clades. Solid bars indicate a high likelihood that haplotype alleles originated from a single clade, whereas multicolored bars indicate that alleles originated from multiple clades. [PIN+] strains are highlighted in yellow. Haplotypes for polyploids are grouped by brackets. *Strains with heterozygous spore clones indicating multiple copies of a single haplotype.

To determine if natural and artificial selection in yeast have influenced phenotypic responses to nutrient stress in domesticated environments, we examined differences in sporulation, pseudohyphal growth, and flocculation among clades. Of the three types of media tested, McClary's media proved far superior at inducing sporulation in wild strains. Eighty-six percent of the strains sporulated in 7 days and

93% of the strains sporulated after 14 days in McClary's media. In contrast, the majority of strains did not sporulate after 7 days in either solid or liquid potassium acetate media. Therefore, all subsequent analyses of sporulation report results from McClary's media. Furthermore, we chose to use the number of tetrads, rather than the total number of tetrads, triads, and dyads, as an indicator of sporulation efficiency, since the viability of triads and dyads is not certain (Safadi *et al.* 2010). With the exception of two polyploid strains, strains that had high sporulation efficiencies also had high spore survival and the spores mated frequently, indicating that sporulation efficiency was a good indicator of mating propensity.

Isolates from the ancestral clade were the best sporulators, as average sporulation efficiency was twofold higher compared to domesticated strains (Figure 4: Kruskal-Wallis: $P = 0.04$). Domesticated strains tended to aggregate more than ancestral isolates (Figure 4). In particular, the commercial food clade had more cells in flocs—approximately sevenfold higher than ancestral isolates (Kruskal-Wallis: $P = 0.018$).

Despite some marked differences between ancestral and domesticated isolates, in general, phenotypic characteristics were highly variable among strains (Figure 5). Furthermore, general tendencies for pseudohyphal growth in domesticated clades compared to ancestral isolates were not apparent, and there were no apparent differences among clades for adhesion and agar invasion. The vast majority of strains grew invasive filaments on SLAD and YPAD agar (85 and 83%, respectively), while less than half grew invasively on reduced glucose (Table S2). Adhesion was observed in 57% of strains on reduced glucose, 53% on YPAD, and 31% on SLAD. Although some strains showed strong preferences for pseudohyphal growth or sporulation, the vast majority of strains presented both asexual and sexual phenotypes on different media (Figure 5 and Figure 6, C and D). In fact, in a few of the strains pseudohyphal cells and flocs were observed directly adjacent to asci (Figure 6, A and B).

Risk factors for [PIN+] infection

Heterozygosity for the residue 129 M/V polymorphism in PrP in mammals is associated with resistance to prion disease (O'Rourke *et al.* 1997; Kelly *et al.* 2008; Mead *et al.* 2009), and previous reports in yeast have shown that certain polymorphisms in *SUP35* and *RNQ1* provide varying levels of protection against [PSI+] and [PIN+] transmission Resende *et al.* 2003; (Bateman and Wickner 2012; Kelly *et al.* 2012). To further investigate the effects of heterozygosity and outcrossing on prion infection, we examined genetic differences between [PIN+] and [pin-] strains. A disproportionate number of [PIN+] strains were polyploid (8 out of 15 [PIN+] vs. eight out of 60 [pin-]; Fisher's exact: $P = 0.018$), and [PIN+] strains had more heterozygous, shared SNPs than [pin-] strains (Figure 7A; Mann-Whitney U -test: $P = 0.0007$). Eighty percent of [PIN+] strains had at least one heterozygous, shared SNP, compared to 45% of [pin-] strains. Mosaic scores calculated

Figure 4 Phenotypic variation among clades. Bars represent average values shown with standard deviations for each clade. *P*-values for the Kruskal–Wallis test are shown in upper right of each panel. (A) The number of cells in aggregates after 7 days on SLAD was averaged for each clade. (B) The number of tetrads formed after 7 days on McClary’s media was averaged for each clade.

for *RNQ1*, were significantly higher for [PIN+] compared to [*pin*–] strains, though no difference in mosaic scores was observed at *SUP35* (Mann–Whitney *U*-test *P* = 0.121).

Forty-seven different *Rnq1p* haplotypes were observed in the sampled population (Table 3). Haplotype mapping produced 11 separate networks (Figure 8). Fifteen of the haplotypes were not connected to any network due to high divergence (more than five mutational events from the most similar haplotype). Nonsynonymous changes most often involved glutamine and asparagines (red arrows in Figure 8), two amino acids that are often enriched in yeast prion domains. Interestingly, although glutamine and asparagines constitute 31% of the residues of *Rnq1p*, they were involved in 56% of nonsynonymous SNPs. In the majority of sites (10/12 glutamine/asparagine SNPs), charged residues, hydrophobic residues, or stop codons replaced the reference glutamine or asparagine. Nearly 75% of all haplotypes were observed in only one or two strains and were more or less evenly distributed among [PIN+]–infected and uninfected strains. Three protein haplotypes showed notable differences in frequency between [PIN+]–infected and uninfected strains (top three rows of Table 3): (1) reference, (2) 166 insertion GQ, Q360H, and (3) Δ167–170. Fisher’s exact tests revealed that the latter two haplotypes (Figure 8, A and E) were significantly associated with [PIN+] infection (*P* < 0.05) although, after correcting for multiple tests, only 166 insertion GQ, Q360H would be considered significant.

After comparing genetic and phenotypic characteristics among clades, we wanted to determine if there were risk factors that may have promoted the spread of [PIN+]

among domesticated isolates. For strains in the commercial beer, commercial food, and wine/vineyard clades, [PIN+] strains had significantly higher sporulation efficiencies (Figure 9D: Mann–Whitney *U*-test *P* = 0.021) than [*pin*–] strains. Cellular aggregation also tended to be more prevalent in [PIN+] strains. The number of cells in flocs was significantly higher in [PIN+] strains vs. [*pin*–] strains (Figure 9B: 1.4-fold difference, Mann–Whitney *U* test: *P* = 0.037). Flocculation tests revealed that [PIN+] strains sedimented faster than [*pin*–] strains when all strains were compared and also when only domesticated strains were examined (Figure 9, C and F: Mann–Whitney *U*-test: *P* = 0.016, *P* = 0.025, respectively). For each strain, the initial concentration was 5×10^7 cells/ml. Cells were allowed to sediment for 1 hr, and cells/ml was again measured. The result was that the cell concentration for [*pin*–] cells was higher than that of [PIN+] cells by 1×10^6 cells/ml. Thus, the [PIN+] cells sedimented more rapidly. Curing the [PIN+] prion did not change the results of sporulation, flocculation, or pseudohyphal growth tests.

Discussion

As animal husbandry practices have exacerbated scrapie, BSE (Kellar and Lees 2003), and CWD (Williams *et al.* 2002), domestication has apparently promoted the global spread of the [PIN+] prion. We observed strikingly different [PIN+] prevalences among four clades of *S. cerevisiae* derived largely from prion protein phylogenies. Neither [PIN+] nor the sexually transmitted 2-μm DNA plasmid

Figure 5 Pseudohyphal growth vs. sporulation for 72 wild *S. cerevisiae* strains. The percentage of cells with unipolar budding, elongated necks, or filaments out of 200 cells after 7 days on SLAD media was determined for each strain and plotted against the percentage of cells that were tetrads, triads, or dyads out of 200 cells after 7 days on McClary's sporulation media.

were detected in the ancestral clade, consisting of isolates sampled from soil, insects, fermented food, grapes, and African wine strains. The ancestral clade was the oldest lineage, showing the most similarity to *S. paradoxus*, a yeast that exists worldwide in natural environments, but has little or no association with humans (Sniegowski *et al.* 2002; Fay and Benavides 2005; Koufopanou *et al.* 2006; Sampaio and Goncalves 2008; Wang *et al.* 2012). Like ancient, wild strains of *S. cerevisiae* sampled in primeval forests of China (Wang *et al.* 2012), isolates from the ancestral clade showed the highest levels of inbreeding and reproductive isolation. Furthermore, strains from this clade had sporulation efficiencies that were similar to wild, undomesticated isolates sampled in forests and much higher than most domesticated isolates (Gerke *et al.* 2006; Magwene *et al.* 2011; Wang *et al.* 2012). Of course, sporulation is not sufficient to ensure frequent outcross mating. The degree of domestication of African wine strains is questionable, but previous work has shown that these strains are the progenitors of domesticated vineyard and sake strains (Fay and Benavides 2005).

Reports that [PSI+] is advantageous under certain conditions (Eaglestone *et al.* 1999; True and Lindquist 2000) have not been reproducible (True and Lindquist 2000; Namy *et al.* 2008). [PSI+] is reportedly a stress-inducible source of phenotypic variation that can promote environmental adaptation in yeast, although [PSI+] was detrimental under most of the reported prion-inducing stress conditions (Tyedmers *et al.* 2008). Moreover, the

[PSI+] inducibility was not reproducible (Kelly *et al.* 2012). If a yeast prion were beneficial for its host, natural selection, combined with the infectivity of all prions, would make it common in natural populations (Nakayashiki *et al.* 2005; Masel and Griswold 2009), like the [Het-s] prion found in 90% of wild isolates (Debets *et al.* 2012; reviewed by Saupe 2011). However, surveys from wild and industrial *Saccharomyces* isolates found that even the most mild variants of [PSI+], [URE3], and [PIN+] are quite rare in natural populations (Nakayashiki *et al.* 2005). One survey of 690 natural *S. cerevisiae* isolates detected [PIN+] in ~6% of samples and weak [PSI+] variants in ~1% (Halfmann *et al.* 2012), with all [PSI+] isolates obtained from wineries. In another study, [PSI+] was not detected in *S. paradoxus*, *Saccharomyces bayanus*, *Saccharomyces castellii*, *Saccharomyces dairensis*, *Saccharomyces exiguus*, *Saccharomyces unisporus*, *Saccharomyces kluyveri*, or *Pichia methanolica* isolates (Chernoff *et al.* 2000). The rarity of [PSI+], [URE3], and [PIN+] in nature suggests that they are not selectively favored (Masel and Griswold 2009), and it is inferred that even the mildest variants of the yeast prions [URE3], [PSI+], and [PIN+] are detrimental overall, decreasing the growth/survival of their hosts by at least 1% (Kelly *et al.* 2012).

Although rare in wild strains, [PSI+] and [PIN+] are found at frequencies greater than their frequency of generation, which is generally 10^{-5} or 10^{-6} . Either of two extreme scenarios could explain their expansion in the population. If [PSI+] or [PIN+] are sometimes an advantage to their hosts, mitotic expansion of prion-carrying cells would increase their occurrence, even if mating is too rare to produce any spread. At the other extreme, they may be deleterious to some degree, but spread horizontally by mating (= infection) at a rate faster than their detriment eliminates them. If prion-carrying cells became abundant for the prion's benefit to the cell, these cells would show no detectable bias toward a history of mating. If the spread is by mating, in spite of some prion toxicity, then prion-carrying strains would usually show evidence of prior outcross mating. We would expect a higher incidence of prion transmission among strains that favor sexual behavior, such as good sporulators. Additionally, multicellular phenotypes that increase contact among yeast cells might enhance opportunities for mating and prion exposure in wild populations.

Domestication has had a substantial effect on the spread of [PIN+] in natural populations of *S. cerevisiae*, most notably by increasing opportunities for outcross mating among genetically diverse strains. With the exception of a single strain of unknown origin, all of the [PIN+]-infected isolates originated from human-associated environments with at least 40% coming from factories or large-scale manufacturing plants. Commercialized settings, like modern breweries or bakeries, regularly utilize multiple yeast strains concurrently (e.g., blended strains) (White 2003; White and White 2013), and industrial improvement of yeast often involves hybridizing genetically diverse isolates (Anderson and

Figure 6 Diversity of phenotypes detected in *S. cerevisiae* strains. (A and B) Strain YB2573 exhibited a multitude of phenotypes on SLAD media. Arrows show multiple tetrads and a flocculent cell including a triad. (B) A tetrad adjacent to an elongated cell with an attached bud. (C and D) Strain YJM413 readily switched between pseudohyphal growth and meiosis when grown on McClary's sporulation media (C) and SLAD (D). (E) Large flocs observed in strain YJM320. (F) Cells with elongated necks and unipolar and radial buds.

Martin 1975; Spencer and Spencer 1983; Steensels *et al.* 2012). Commercial strains had complex patterns of genetic mosaicism, indicating that they are highly outcrossed. Furthermore, bakery and brewery strains had many more heterozygous, shared SNPs than did strains sampled from other niches (Figure S2), which shows that the gene pool within these environments is quite diverse. Others have shown that bread strains interbred with beer strains early on in the domestication process (Legras *et al.* 2007), consistent with historical reports that, in the 18th and 19th centuries, breweries provided most of the commercially available yeast (Frey 1930), and bakery strains were often mixed with industrial brewery isolates (Lesaffre 2013). By creating artificial mating opportunities, domestication has exacerbated [PIN+] spreading in *S. cerevisiae*. Wine yeasts were not often found infected with the [PIN+] prion, although the 2- μ m DNA plasmid, which is likewise spread only by outcross matings, is quite common in this group. Perhaps [PIN+] is somewhat more detrimental in this niche than in other settings.

In domesticating yeast, humans may have selected for traits that promote [PIN+] spread. Multicellular phenotypes, like flocculation and florulation, promote the removal of yeast following alcohol fermentation, and these traits

have been selected by brewers (Briggs *et al.* 2004; Verstrepen and Klis 2006; Bruckner and Mosch 2011). The ability of a strain to make contact and adhere to another yeast apparently evolved as a social response to stress, as flocs are protected from toxins like alcohol and antimicrobial compounds (Smukalla *et al.* 2008). Social behaviors that bring mixed populations into contact could conceivably enhance [PIN+] transmission among cells, particularly since flocculation seems to promote heterotypic interactions among yeast (Smukalla *et al.* 2008; Veelders *et al.* 2010).

[PIN+] infection is significantly associated with polyploidy in wild yeast (Fisher's exact: $P = 0.018$). Of 16 polyploids detected in our population, 8 were [PIN+], and of these, nearly two-thirds were brewery or bakery isolates. We identified polyploids by the presence of more than two haplotypes, and all polyploids had numerous shared alleles, indicating that they resulted from multiple outcross matings/fusions of different strains. Fleischmann's baker's yeast, for example, possessed haplotype signatures of all four clades (Figure 3), which could have been acquired during successive matings. Haplotype maps, which depict clusters of individuals that have undergone recent genetic exchange, show that [PIN+]-infected polyploids like Fleischmann's, CBS6216, YB3916, and YB2573 have haplotypes that cluster

Figure 7 Genetic characteristics for [PIN+] and [pin-] strains. Bars represent average values shown with standard deviation. *P*-values for Mann-Whitney *U*-test are shown in upper right of each panel. (A) The number of heterozygous, shared SNPs was averaged for [PIN+] and [pin-]. (B and C) Haplotype mosaic scores were calculated for *RNQ1* and *SUP35*, and the highest score per strain was used to calculate an average for [PIN+] and [pin-] strains.

with multiple, genetically distinct groups (Figure 8). The enrichment of [PIN+] in these strains may reflect the fact that this prion is an infectious disease (Nakayashiki *et al.* 2005; McGlinchey *et al.* 2011; Kelly *et al.* 2012), and multiple outcross matings increase opportunities for infection. Interestingly, seven polyploid strains, of which six were [PIN+]-infected, formed tetrads with appreciable frequencies, and spores were viable and able to mate with α - and/or α -tester strains. In contrast, the majority of uninfected polyploid strains did not sporulate under the conditions tested. Recently, higher ploidy has been linked to toxicity resistance in baker's yeast harboring polyQ aggregates (Kaiser *et al.* 2013). While the effects of ploidy on prion toxicity have not been examined, it is possible that polyploid yeast are able to mitigate the detrimental growth/survival effects of [PIN+].

The prion domains from several other species have [PSI+]-forming ability in *S. cerevisiae*. (Chernoff *et al.* 2000; Santoso *et al.* 2000; Nakayashiki *et al.* 2001). The [URE3]-forming ability of Ure2p is not conserved, but rather appears sporadically among species related to *S. cerevisiae*. The Ure2p of *Kluyveromyces lactis* cannot form [URE3] in *S. cerevisiae* or in *K. lactis* itself (Safadi *et al.* 2011). In *Candida glabrata*, Ure2p cannot form [URE3] in either its own cells or in *S. cerevisiae* (Edskes *et al.* 2011; Edskes and Wickner 2013), although Ure2p of

Candida albicans can form the [URE3] prion in *C. glabrata* (Edskes and Wickner 2013). Yeast prion-forming domains are characteristically biased for glutamine and asparagine (Ter-Avanesyan *et al.* 1994; Masison and Wickner 1995; Derkatch *et al.* 2001; Toombs *et al.* 2012), whereas acidic, basic, and hydrophobic residues tend to be underrepresented (Ross and Toombs 2010). For [PIN+], deletions in any of the four glutamine-asparagine-rich repeat regions result in barriers to [PIN+] transmission (Kadnar *et al.* 2010; Kelly *et al.* 2012), whereas expanding these regions promotes [PIN+] formation (Vitrenko *et al.* 2007). Interestingly, the emergence of [PIN+] in our population coincided with the appearance of a two-amino-acid insertion within the GQ repeat region, and this allele was significantly associated with [PIN+] infection. Deletions that deter prion transmission ($\Delta 286-296$; triangle haplotypes in Figure 8) were common in the ancestral clade (38% of strains), and deletions of up to four repeat motifs in the glutamine-glycine repeat region (amino acids 153-172) were frequently detected in all clades. While the GQ-rich region does not create a transmission barrier when deleted, it does contribute to prion formation through interactions with glutamine-asparagine-rich regions (Kadnar *et al.* 2010). The repeat region of *Sup35p* is important for prion formation (Liu and Lindquist 1999; Parham *et al.* 2001; Osheroovich *et al.* 2004), not because of the repeat feature, but because of the amino acid composition (Ross *et al.* 2004; Ross *et al.* 2005; Toombs *et al.* 2011). The importance of repeats—as repeats—has not been tested for *Rnq1p*.

RNQ1, and in particular the prion domain, is highly polymorphic (Kelly *et al.* 2012). While fewer than one in three amino acids in *Rnq1p* are glutamine or asparagine, more than half of the nonsynonymous SNPs involved these residues, which suggests that not all amino acids mutated with equal probability. In 12/15 sites, glutamine or asparagine residues were replaced with charged or nonpolar residues that would be predicted to lower the prion-forming propensity of the protein (Ross and Toombs 2010) or with premature stop codons that completely abrogate prion transmission (Kelly *et al.* 2012). In fact, five different protein variants that abrogate prion transmission were observed in domesticated clades, including several premature stop codons that eliminate most of the prion domain. Given the dynamic changes to *Rnq1p* over time, our data seem to question the extent of evolutionary conservation of prion-forming ability.

The relatively low [PIN+] prevalence in the wine clade may be because they are largely asexual (Cubillos *et al.* 2009) or because the population experienced a bottleneck. Magwene *et al.* (2011) suggested that artificial selection in some yeast lineages has favored asexual growth over sporulation. Indeed, strains in the wine clade generally showed poor sporulation and a greater tendency for pseudohyphal growth than most strains. Heterozygosity was reduced in the wine clade compared to the other domesticated clades. Of the 70 SNPs detected in the wine clade, 30 were found in only one isolate. The number of strains

Table 3 Haplotype frequencies in [PIN+]-infected and uninfected strains

Haplotype description	No. of haplotypes	% in [pin-]	% in [PIN+]	Difference (%)	Strain ID (from Table 1)
Reference (S288C): NS	61	32.8	43.6	10.8	^a
166insGQ, Q360H**	13	5.2	15.4	10.2	10, 11, 17, 19, 23, 28, 33, 35
Δ167–170*	6	1.5	10.3	8.8	10, 11, 22, 24, 35
Δ165–166, Q360H, Q363H	8	6.0	0.0	6.0	4, 5, 32, 40
Q319R	7	5.2	0.0	5.2	50, 53, 55, 62, 72
166insGQ, Q360H, Q373P	2	0.0	5.1	5.1	29
A10S	2	0.0	5.1	5.1	12
166insGQ	6	4.5	0.0	4.5	54, 60, 68
D55E, Δ161–166, Q360H	6	4.5	0.0	4.5	6, 7, 8
G158S, Δ167–170, N238S, Δ286–296, Q360H, Q363H	4	3.0	0.0	3.0	1, 2
Δ159–160, Q360H	1	0.0	2.6	2.6	34
Q65STOP, Δ167–170	1	0.0	2.6	2.6	31
Q65STOP, Δ167–170, N253D	1	0.0	2.6	2.6	31
S58I, 158insDQ	1	0.0	2.6	2.6	34
Δ167–170, Q360H	1	0.0	2.6	2.6	10
Δ165–166, Q360H	3	2.2	0.0	2.2	16, 27
Q360H	6	3.0	5.1	2.1	29, 37, 38, 48
Δ163–172	2	0.7	2.6	1.8	37, 49
N377K	2	1.5	0.0	1.5	70
Y249STOP	2	1.5	0.0	1.5	47, 72
Δ167–170, N253D, N296S, N296S, Q313H	2	1.5	0.0	1.5	22
Δ167–168, Q360H	2	1.5	0.0	1.5	25
166insGQ, Q360H, Q363H	2	1.5	0.0	1.5	39
166insGQ, Δ297–307	2	1.5	0.0	1.5	56
Δ165–166, Δ167–168, Δ286–296, Q360H	2	1.5	0.0	1.5	3
Δ159–162	2	1.5	0.0	1.5	14
158insDQ	2	1.5	0.0	1.5	14
Δ156–161, Q193STOP	2	1.5	0.0	1.5	21
G59R, Q167STOP	2	1.5	0.0	1.5	13
D55E, Δ165–166, G294D, Q360H	2	1.5	0.0	1.5	26
D55E, A81T, 172insQGQ, Q360H	2	1.5	0.0	1.5	15
Δ165–166	1	0.7	0.0	0.7	26
D55E, Δ167–168, G294D, Q360H	1	0.7	0.0	0.7	22
D55E, Δ159–160	1	0.7	0.0	0.7	14
S323N, Q360H	1	0.7	0.0	0.7	18
Δ297–307	1	0.7	0.0	0.7	73
Q235STOP	1	0.7	0.0	0.7	48
Δ166–173, N253D, N296S, N296S, Q313H	1	0.7	0.0	0.7	22
Δ159–166	1	0.7	0.0	0.7	71
G59R	1	0.7	0.0	0.7	67
S58I	1	0.7	0.0	0.7	74
S58I, Δ297–307	1	0.7	0.0	0.7	73
D55E, Δ167–170, Q193STOP, N253D, N296S, N296S, Q313H	1	0.7	0.0	0.7	22
D55E, Δ165–172	1	0.7	0.0	0.7	26
D55E, Δ165–166, Q360H	1	0.7	0.0	0.7	26
D55E, Δ167–168, Q193STOP, G294D, Q360H	1	0.7	0.0	0.7	22
D55E, S58I, G59R, Q360H	1	0.7	0.0	0.7	18

^a Strains with reference Rnq1p haplotype are indicated in Table 1.

NS, not significant. * $P < 0.05$, ** $P < 0.001$ for Fisher's exact tests of association between haplotype and [PIN+] infection.

with a single SNP was more than twice as high as expected under a model of random mating (Figure S3), a trend not observed in the other clades. This pattern of genetic structure could also indicate a population bottleneck, which often occurs in vineyards following wine fermentation (Mortimer and Polsinelli 1999), with subsequent recovery that produces an excess of low-frequency alleles (Excoffier *et al.* 2009; Bisson 2012). The rare SNPs observed in the wine/vineyard clade are likely mutations acquired during

asexual growth rather than alleles inherited during outcrossing. Ultimately, strains in the wine/vineyard clade undergo relatively few outbred, sexual cycles and thus have a reduced likelihood of being infected with the sexually transmitted [PIN+] prion.

Genetic analyses permitted a thorough survey of heterozygosity and genetic mosaicism at two prion protein loci and two additional genomic regions. While our data allowed us to discern phylogeny of proteins with prion-

Figure 8 Haplotype map of *S. cerevisiae* RNQ1. Networks (A–K) representing haplotype connections, where nodes (black dots and shapes) represent a single mutation, were generated with TCS2.1. Each haplotype is represented with strain and haplotype identifiers (e.g., A or B), and colored according to the clade to which it was assigned; green, ancestral; orange, commercial beer; blue, commercial food; red, wine/vineyard. Haplotypes within each network are ordered top to bottom by increasing genetic divergence from *S. paradoxus*. Haplotype E (boxed area) was significantly associated with [PIN+] infection (Fisher's exact: $P < 0.001$). Haplotypes were not connected to networks when the number of mutational events was greater than five.

forming ability, we did not fully resolve genomic phylogeny, and it is unlikely that we detected the full extent of *S. cerevisiae* genetic mosaicism. Our results agree completely with the genome-wide study of Liti *et al.* (2009), except that we classified two African wine strains and four North American strains as sister taxa in a single clade, whereas Liti *et al.* (2009) placed these taxa in separate clades. None of these strains carry [PIN+] and additional analyses demonstrated that classifying African wine strains and North American isolates as separate clades would not affect the findings reported in our article. General trends reported for ancestral isolates remained consistent; strains in African wine and North American lineages tend to flocculate less and sporulate more than do the majority of strains in domesticated lineages. Phenotypes, however, were variable within these lineages. For example, three of four North American isolates had high sporulation efficiencies (>92% of domesticated strains) whereas a single North American strain had a somewhat lower sporulation efficiency (>72% of domesticated strains). Phenotypic variability among strains within a single phyloge-

netic lineage has been previously reported and suggests that environmental conditions can elicit different cellular growth responses in strains with the same genetic background (Yvert *et al.* 2013).

Previous studies have found that phenotypic differences among *S. cerevisiae* isolates do not necessarily correlate with ecological and geographical origin (Liti *et al.* 2009; Yvert *et al.* 2013). While we found trends in mating and multicellular growth within clades, the diverse array of phenotypes observed in wild and domesticated isolates likely reflects the plasticity of this species in response to stress and habitat variability (Kvitek *et al.* 2008; Liti *et al.* 2009). In agreement with Magwene *et al.* (2011), about half of our sampled isolates showed a preference for either asexual or sexual reproduction; however, the other half readily switched between sexual and asexual growth when exposed to different growth media. Phenotypic plasticity can be an advantage for yeast living in habitats where resources or growth conditions fluctuate (Halme *et al.* 2004; Avery 2006; Bishop *et al.* 2007; Yvert *et al.* 2013).

Figure 9 Phenotypic variation between [PIN+] and [pin-] strains. Bars represent average values shown with standard deviations for [PIN+] and [pin-] strains. *P*-values for Mann-Whitney *U*-test are shown in upper right of each panel. A (all strains) and D (only domesticated): The number of tetrads formed after 7 days on McClary's media was averaged for [PIN+] and [pin-] strains. B (all strains) and E (only domesticated): The number of cells in aggregates after 7 days on SLAD was averaged for [PIN+] and [pin-] strains. C (all strains) and F (only domesticated): Flocculation test measuring the change in OD₆₆₀ of cell suspensions after 1 hr. Differences in OD₆₆₀ were averaged for [PIN+] and [pin-] strains.

We observed statistical differences among clades in the efficiency of sporulation and flocculation, but great variability within clades, reflecting the well-known flexibility of *Saccharomyces* in dealing with its environment. The genetic variation between clades may reflect adaptation to the various environments represented, while the high heterozygosity and mosaicism of commercial strains may be due to the history of breeding efforts and the conditions of brewing/fermentation. We find a significant association of [PIN+] prion infection with commercial strains, polyploidy, and heterozygosity, and this prion is most likely to be found in strains bearing evidence of outcross mating

(heterozygosity for shared alleles) or of hybrid formation (polyploidy/mosaicism with genetic material from different clades).

Acknowledgments

We thank Frank Shewmaker for a thoughtful reading of the manuscript. This work was supported by the Intramural Program of the National Institute of Diabetes and Digestive and Kidney Diseases and in part by the Intramural Research Program of the National Library of Medicine, National Institutes of Health.

Literature Cited

- Abelson, J. N., M. I. Simon, C. Guthrie, and G. R. Fink, 1991 *Guide to Yeast Genetics and Molecular Biology*. Academic Press, Inc., San Diego.
- Anderson, E., and P. A. Martin, 1975 The sporulation and mating of brewing yeasts. *J. Inst. Brew.* 81: 242–247.
- Avery, S. V., 2006 Microbial cell individuality and the underlying sources of heterogeneity. *Nat. Rev. Microbiol.* 4: 577–587.
- Bateman, D. A., and R. B. Wickner, 2012 [PSI⁺] prion transmission barriers protect *Saccharomyces cerevisiae* from infection: intraspecies “species barriers.” *Genetics* 190: 569–579.
- Benjamini, Y., and Y. Hochberg, 1995 Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. B* 57: 289–300.
- Bishop, A. L., F. A. Rab, E. R. Sumner, and S. V. Avery, 2007 Phenotypic heterogeneity can enhance rare cell survival in stress-sensitive yeast populations. *Mol. Microbiol.* 63: 507–520.
- Bisson, L. F., 2012 Geographic origin and diversity of wine strains of *Saccharomyces*. *Am. J. Enol. Vitic.* 63: 165–176.
- Briggs, D. E., C. A. Boulton, and P. A. Brookes, 2004 *Brewing: Science and Practice*. CRC Press, Cleveland, OH; Boca Raton, FL.
- Bruckner, S., and H.-U. Mosch, 2011 Choosing the right lifestyle: adhesion and development in *Saccharomyces cerevisiae*. *FEMS Microbiol. Rev.* 36: 25–58.
- Chernoff, Y. O., A. P. Galkin, E. Lewitin, T. A. Chernova, G. P. Newnam *et al.*, 2000 Evolutionary conservation of prion-forming abilities of the yeast Sup35 protein. *Mol. Microbiol.* 35: 865–876.
- Clement, M., D. Posada, and K. A. Crandall, 2000 TCS: a computer program to estimate gene genealogies. *Mol. Ecol.* 9: 1657–1659.
- Cubillos, F. A., C. Vasquez, S. Faugeron, A. Ganga, and C. Martinez, 2009 Self fertilization is the main sexual reproduction mechanism in native wine yeast populations. *FEMS Microbiol. Ecol.* 67: 162–170.
- Debets, A. J., H. J. Dalstra, M. Slakhorst, B. Koopmanschap, R. F. Hoekstra *et al.*, 2012 High natural prevalence of a fungal prion. *Proc. Natl. Acad. Sci. USA* 109: 10432–10437.
- Derkatch, I. L., M. E. Bradley, P. Zhou, Y. O. Chernoff, and S. W. Liebman, 1997 Genetic and environmental factors affecting the de novo appearance of the [PSI⁺] prion in *Saccharomyces cerevisiae*. *Genetics* 147: 507–519.
- Derkatch, I. L., M. E. Bradley, J. Y. Hong, and S. W. Liebman, 2001 Prions affect the appearance of other prions: the story of [PIN⁺]. *Cell* 106: 171–182.
- Eaglestone, S. S., B. S. Cox, and M. F. Tuite, 1999 Translation termination efficiency can be regulated in *Saccharomyces cerevisiae* by environmental stress through a prion-mediated mechanism. *EMBO J.* 18: 1974–1981.
- Edskes, H. K., and R. B. Wickner, 2013 The [URE3] prion in *Candida*. *Eukaryot. Cell* 12: 551–558.
- Edskes, H. K., A. Engel, L. M. McCann, A. Brachmann, H. F. Tsai *et al.*, 2011 Prion-forming ability of Ure2 of yeasts is not evolutionarily conserved. *Genetics* 188: 81–90.
- Excoffier, L., M. Foll, and R. M. J. Petit, 2009 Genetic consequences of range expansions. *Annu. Rev. Ecol. Evol. Syst.* 40: 481–501.
- Fay, J. C., and J. A. Benavides, 2005 Evidence for domesticated and wild populations of *Saccharomyces cerevisiae*. *PLoS Genet.* 1: e5.
- Frey, C. N., 1930 History and development of the modern yeast industry. *Ind. Eng. Chem.* 22: 1154–1162.
- Gerke, J. P., C. T. L. Chen, and B. A. Cohen, 2006 Natural isolates of *Saccharomyces cerevisiae* display complex genetic variation in sporulation efficiency. *Genetics* 174: 985–997.
- Gimeno, C. J., P. O. Ljungdahl, C. A. Styles, and G. R. Fink, 1992 Unipolar cell divisions in the yeast *S. cerevisiae* lead to filamentous growth: regulation by starvation and RAS. *Cell* 68: 1077–1090.
- Guldal, C. G., and J. Broach, 2006 Assay for adhesion and agar invasion in *S. cerevisiae*. *J. Vis. Exp.* 8: 64.
- Guo, B., C. A. Styles, Q. Feng, and G. R. Fink, 2000 A *Saccharomyces* gene family involved in invasive growth, cell-cell adhesion, and mating. *Proc. Natl. Acad. Sci. USA* 97: 12158–12163.
- Halfmann, R., D. F. Jarosz, S. K. Jones, A. Chang, A. K. Lancaster *et al.*, 2012 Prions are a common mechanism for phenotypic inheritance in wild yeasts. *Nature* 482: 363–368.
- Halme, A., S. Bumgarner, C. Styles, and G. R. Fink, 2004 Genetic and epigenetic regulation of the FLO gene family generates cell-surface variation in yeast. *Cell* 116: 405–415.
- Hosoda, N., T. Kobayashi, N. Uchida, Y. Funakoshi, Y. Kikuchi *et al.*, 2003 Translation termination factor eRF3 mediates mRNA decay through the regulation of deadenylation. *J. Biol. Chem.* 278: 38287–38291.
- Kadnar, M. L., G. Artcov, and I. L. Derkatch, 2010 Distinct type of transmission barrier revealed by study of multiple prion determinants of Rnq1. *PLoS Genet.* 6: e1000824.
- Kaiser, C. J. O., S. W. Grotzinger, J. M. Eckl, K. Papsdorf, S. Jordan *et al.*, 2013 A network of genes connects polyglutamine toxicity to ploidy control in yeast. *Nat. Commun.* 4: 1571.
- Katoh, K., K. Misawa, K. I. Kuma, and T. Miyata, 2002 MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* 30: 3059–3066.
- Kellar, J. A., and V. W. Lees, 2003 Risk management of the transmissible spongiform encephalopathies in North America. *Rev. Sci. Tech.* 22: 201.
- Kelly, A. C., and R. B. Wickner, 2013 *Saccharomyces cerevisiae*: a sexy yeast with a prion problem. *Prion* 3: 215–220.
- Kelly, A. C., N. E. Mateus-Pinilla, J. Diffendorfer, E. Jewell, M. O. Ruiz *et al.*, 2008 Prion sequence polymorphisms and chronic wasting disease resistance in Illinois white-tailed deer (*Odocoileus virginianus*). *Prion* 2: 28–36.
- Kelly, A. C., F. P. Shewmaker, D. Kryndushkin, and R. B. Wickner, 2012 Sex, prions, and plasmids in yeast. *Proc. Natl. Acad. Sci. USA* 109: E2683–E2690.
- Koufopanou, V., J. Hughes, G. Bell, and A. Burt, 2006 The spatial scale of genetic differentiation in a model organism: the wild yeast *Saccharomyces paradoxus*. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 361: 1941–1946.
- Kvitek, D. J., J. L. Will, and A. P. Gasch, 2008 Variations in stress sensitivity and genomic expression in diverse *S. cerevisiae* isolates. *PLoS Genet.* 4: e1000223.
- Legras, J. L., D. Merdinoglu, J. Cornuet, and F. Karst, 2007 Bread, beer and wine: *Saccharomyces cerevisiae* diversity reflects human history. *Mol. Ecol.* 16: 2091–2102.
- Lesaffre, 2013 Lesaffre group history. Available at: <http://www.lesaffre.com/en/the-group/about-us/histoire.html>. Accessed: October 10, 2013.
- Librado, P., and J. Rozas, 2009 DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25: 1451–1452.
- Liti, G., D. M. Carter, A. M. Moses, J. Warringer, L. Parts *et al.*, 2009 Population genomics of domestic and wild yeasts. *Nature* 458: 337–341.
- Liu, J.-J., and S. Lindquist, 1999 Oligopeptide-repeat expansions modulate protein-only inheritance in yeast. *Nature* 400: 573–576.
- Magwene, P. M., O. Kayakci, J. A. Granek, J. M. Reininga, Z. Scholl *et al.*, 2011 Outcrossing, mitotic recombination, and life-history trade-offs shape genome evolution in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA* 108: 1987–1992.
- Masel, J., and C. K. Griswold, 2009 The strength of selection against the yeast prion [PSI⁺]. *Genetics* 181: 1057–1063.

- Masison, D. C., and R. B. Wickner, 1995 Prion-inducing domain of yeast Ure2p and protease resistance of Ure2p in prion-containing cells. *Science* 270: 93–95.
- McClary, D. O., W. L. Nulty, and G. R. Miller, 1959 Effect of potassium vs. sodium in the sporulation of *Saccharomyces*. *J. Bacteriol.* 78: 362–368.
- McGlinchey, R. P., D. Kryndushkin, and R. B. Wickner, 2011 Suicidal [PSI⁺] is a lethal yeast prion. *Proc. Natl. Acad. Sci. USA* 108: 5337–5341.
- Mead, S., J. Whitfield, M. Poulter, P. Shah, J. Uphill *et al.*, 2009 A novel protective prion protein variant that colocalizes with kuru exposure. *N. Engl. J. Med.* 361: 2056–2065.
- Medina, E. M., G. W. Jones, and D. A. Fitzpatrick, 2011 Reconstructing the fungal tree of life using phylogenomics and a preliminary investigation of the distribution of yeast prion-like proteins in the fungal kingdom. *J. Mol. Evol.* 73: 116–133.
- Mortimer, R., and M. Polsinelli, 1999 On the origins of wine yeast. *Res. Microbiol.* 150: 199–204.
- Nakayashiki, T., K. Ebihara, H. Bannai, and Y. Nakamura, 2001 Yeast [PSI⁺] “prions” that are crosstransmissible and susceptible beyond a species barrier through a quasi-prion state. *Mol. Cell* 7: 1121–1130.
- Nakayashiki, T., C. P. Kurtzman, H. K. Edskes, and R. B. Wickner, 2005 Yeast prions [URE3] and [PSI⁺] are diseases. *Proc. Natl. Acad. Sci. USA* 102: 10575–10580.
- Namy, O., A. Galopier, C. Martini, S. Matsufuji, C. Fabret *et al.*, 2008 Epigenetic control of polyamines by the prion [PSI⁺]. *Nat. Cell Biol.* 10: 1069–1075.
- O'Rourke, K. I., G. R. Holyoak, W. W. Clark, J. R. Mickelson, S. Wang *et al.*, 1997 PrP genotypes and experimental scrapie in orally inoculated Suffolk sheep in the United States. *J. Gen. Virol.* 78: 975–978.
- Oshervich, L. Z., B. S. Cox, M. F. Tuite, and J. S. Weissman, 2004 Dissection and design of yeast prions. *PLoS Biol.* 2: e86.
- Parham, S. N., C. G. Resende, and M. F. Tuite, 2001 Oligopeptide repeats in the yeast protein Sup35p stabilize intermolecular prion interactions. *EMBO J.* 20: 2111–2119.
- Peakall, R., and P. E. Smouse, 2012 GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics* 28: 2537–2539.
- Piry, S., A. Alapetite, J. M. Cornuet, D. Paetkau, L. Baudouin *et al.*, 2004 GENECLASS2: a software for genetic assignment and first-generation migrant detection. *J. Hered.* 95: 536–539.
- Price, M. N., P. S. Dehal, and A. P. Arkin, 2009 FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. *Mol. Biol. Evol.* 26: 1641–1650.
- Price, M. N., P. S. Dehal, and A. P. Arkin, 2010 FastTree 2: approximately maximum-likelihood trees for large alignments. *PLoS ONE* 5: e9490.
- Querol, A., and U. Bond, 2009 The complex and dynamic genomes of industrial yeasts. *FEMS Microbiol. Lett.* 293: 1–10.
- Resende, C. G., T. F. Outeiro, L. Sands, S. Lindquist, and M. F. Tuite, 2003 Prion protein gene polymorphisms in *Saccharomyces cerevisiae*. *Mol. Microbiol.* 49: 1005–1017.
- Ross, E. D., and J. A. Toombs, 2010 The effects of amino acid composition on yeast prion formation and prion domain interactions. *Prion* 4: 60–65.
- Ross, E. D., U. Baxa, and R. B. Wickner, 2004 Scrambled prion domains form prions and amyloid. *Mol. Cell. Biol.* 24: 7206–7213.
- Ross, E. D., H. K. Edskes, M. J. Terry, and R. B. Wickner, 2005 Primary sequence independence for prion formation. *Proc. Natl. Acad. Sci. USA* 102: 12825–12830.
- Rousset, F., 2008 Genepop007: a complete re-implementation of the genepop software for Windows and Linux. *Mol. Ecol. Res.* 8: 103–106.
- Ruderfer, D. M., S. C. Pratt, H. S. Seidel, and L. Kruglyak, 2006 Population genomic analysis of outcrossing and recombination in yeast. *Nat. Genet.* 38: 1077–1081.
- Safadi, R. A., M. Weiss-Gayet, J. Briolay, and M. Aigle, 2010 A polyploid population of *Saccharomyces cerevisiae* with separate sexes (dioecy). *FEMS Yeast Res.* 10: 757–768.
- Safadi, R. A., N. Talarek, N. Jacques, and M. Aigle, 2011 Yeast prions: Could they be exaptations? The URE2/[URE3] system in *Kluyveromyces lactis*. *FEMS Yeast Res.* 11: 151–153.
- Sampaio, J. P., and P. Goncalves, 2008 Natural populations of *Saccharomyces kudriavzevii* in Portugal are associated with oak bark and are sympatric with *S. cerevisiae* and *S. paradoxus*. *Appl. Environ. Microbiol.* 74: 2144–2152.
- Santos, A., P. Chien, L. Z. Oshervich, and J. S. Weissman, 2000 Molecular basis of a yeast prion species barrier. *Cell* 100: 277–288.
- Saupe, S., 2011 The [Het-s] prion of *Podospora anserina* and its role in heterokaryon incompatibility. *Semin. Cell Dev. Biol.* 22: 460–468.
- Smit, G., M. H. Straver, B. J. Lugtenberg, and J. W. Kijne, 1992 Flocculence of *Saccharomyces cerevisiae* cells is induced by nutrient limitation, with cell surface hydrophobicity as a major determinant. *Appl. Environ. Microbiol.* 58: 3709–3714.
- Smukalla, S., M. Caldara, N. Pochet, A. Beauvais, S. Guadagnini *et al.*, 2008 FLO1 is a variable green beard gene that drives biofilm-like cooperation in budding yeast. *Cell* 135: 726–737.
- Sniewowski, 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.
- Soares, E. V., and M. Mota, 1997 Quantification of yeast flocculation. *J. Inst. Brew.* 103: 93–98.
- Spencer, J. F. T., and D. M. Spencer, 1983 Genetic improvement of industrial yeasts. *Annu. Rev. Microbiol.* 37: 121–142.
- Steensels, J., T. Snoek, E. Meersman, M. Picca Nicolino, E. Aslan-koochi *et al.*, 2012 Selecting and generating superior yeasts for the brewing industry. *Cerevisia* 37: 63–67.
- Stephens, M., and P. Scheet, 2005 Accounting for decay of linkage disequilibrium in haplotype inference and missing-data imputation. *Am. J. Hum. Genet.* 76: 449–462.
- Strathern, J. N., E. W. Jones, and J. R. Broach, 1981 *The Molecular Biology of the Yeast Saccharomyces: A Life Cycle and Inheritance*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Ter-Avanesyan, M. D., A. R. Dagkesamanskaya, V. V. Kushnirov, and V. N. Smirnov, 1994 The SUP35 omnipotent suppressor gene is involved in the maintenance of the non-Mendelian determinant [psi⁺] in the yeast *Saccharomyces cerevisiae*. *Genetics* 137: 671–676.
- Toombs, J. A., N. M. Liss, K. R. Cobble, Z. Ben-Musa, and E. D. Ross, 2011 [PSI⁺] maintenance is dependent on the composition, not primary sequence, of the oligopeptide repeat domain. *PLoS ONE* 6: e21953.
- Toombs, J. A., M. Petri, K. R. Paul, G. Y. Kan, A. Ben-Hur *et al.*, 2012 De novo design of synthetic prion domains. *Proc. Natl. Acad. Sci. USA* 109: 6519–6524.
- True, H. L., and S. L. Lindquist, 2000 A yeast prion provides a mechanism for genetic variation and phenotypic diversity. *Nature* 407: 477–484.
- Tsai, I. J., D. Bensasson, A. Burt, and V. Koufopanou, 2008 Population genomics of the wild yeast *Saccharomyces paradoxus*: quantifying the life cycle. *Proc. Natl. Acad. Sci. USA* 105: 4957–4962.
- Tyedmers, J., M. L. Madariaga, and S. Lindquist, 2008 Prion switching in response to environmental stress. *PLoS Biol.* 6: e294.

- Veelders, M., S. Brackner, D. Ott, C. Unverzagt, H.-U. Masch *et al.*, 2010 Structural basis of flocculin-mediated social behavior in yeast. *Proc. Natl. Acad. Sci. USA* 107: 22511–22516.
- Verstrepen, K. J., and F. M. Klis, 2006 Flocculation, adhesion and biofilm formation in yeasts. *Mol. Microbiol.* 60: 5–15.
- Vitrenko, Y. A., M. E. Pavon, S. I. Stone, and S. W. Liebman, 2007 Propagation of the [PIN+] prion by fragments of Rnq1 fused to GFP. *Curr. Genet.* 51: 309–319.
- Wang, Q.-M., W.-Q. Liu, G. Liti, S.-A. Wang, and F.-Y. Bai, 2012 Surprisingly diverged populations of *Saccharomyces cerevisiae* in natural environments remote from human activity. *Mol. Ecol.* 21: 5404–5417.
- White, C., 2003 Using multiple yeast strains. Available at: <http://www.whitelabs.com/content/using-multiple-yeast-strains>. Accessed: October 10, 2013.
- White, L., and M. White, 2013 *White Labs Brewery Catalog*. White Labs, San Diego, CA.
- Wickner, R. B., 1994 [URE3] as an altered *URE2* protein: evidence for a prion analog in *S. cerevisiae*. *Science* 264: 566–569.
- Wickner, R. B., S. W. Liebman, and S. J. Saupe, 2004 Prions of yeast and filamentous fungi: [URE3], [PSI+], [PIN+], and [Het-s], pp. 305–372 in *Prion Biology and Diseases*, edited by S. B. Prusiner. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Williams, E. S., M. W. Miller, T. J. Kreeger, R. H. Kahn, and E. T. Thorne, 2002 Chronic wasting disease of deer and elk: a review with recommendations for management. *J. Wildl. Manag.* 66: 551–563.
- Yvert, G., S. Ohnuki, S. Nogami, Y. Imanaga, S. Fehrmann *et al.*, 2013 Single-cell phenomics reveals intra-species variation of phenotypic noise in yeast. *BMC Syst. Biol.* 7: 54.
- Zupan, J., and P. Raspor, 2008 Quantitative agar-invasion assay. *J. Microbiol. Methods* 73: 100–104.

Communicating editor: J. Lawrence

GENETICS

Supporting Information

<http://www.genetics.org/lookup/suppl/doi:10.1534/genetics.114.165670/-/DC1>

Effect of Domestication on the Spread of the [PIN+] Prion in *Saccharomyces cerevisiae*

Amy C. Kelly, Ben Busby, and Reed B. Wickner

File S1 Supporting Information

Description of Phylogenetic Analysis

To construct our phylogenetic tree, we employed FastTree (Price *et al.* 2009; Price *et al.* 2010) a method that builds approximate maximum-likelihood trees from large alignments. We applied a generalized-time reversible (GTR) model of nucleotide substitution to estimate distances between sequences (Lanave *et al.* 1984). This model is generalized in that it allows for a different rate of substitution between each nucleotide pair, and is time-reversible because substitution rates are symmetric (i.e. C changes into T at the same rate that T changes into C). A flexible time-reversible model was chosen to mitigate the potential effects of recombination during phylogenetic analysis. In a traditional maximum likelihood tree, tree space is searched, such that a tree that best explains the data can be defined (Felsenstein 1981). FastTree constructs an initial maximum likelihood tree where nearest neighbors are joined in a way that minimizes the sum of branch lengths (minimum evolution, where branch length correspond to the expected number of substitutions per site) (Price *et al.* 2009). To refine the initial topology, $2 \cdot \log(N)$ (N =the number of sequences being analyzed) rounds of nearest neighbor interchanges (NNIs) are used to compare the current branch configuration to alternate topologies where the two nearest neighbors are swapped and re-joined [eg. ((A,B), (C,D)) compared to ((A,C), (B,D)) and ((A,D), (B,C))]. With $2 \cdot \log(N)$ rounds of NNI, each node is tested in all possible positions on the tree. Additionally, for each subtree (or branch) two rounds of subtree-pruning-regrafting (SPRs) are performed such that a subtree is reinserted at the closest alternate locations [eg. ((A,B), (C,D), E) compared to ((A, (B,C)),D,E)] (Price *et al.* 2010). FastTree tests each NNI and SPR move and repositions the branches if alternate topologies further reduce the sum of branch lengths.

To calculate local support values for each branch, FastTree uses the Shimodaira-Hasegawa test (Shimodaira and Hasegawa 1999) on NNIs around the split. With this test, the sum of branch lengths for two NNI topologies are calculated and a log-likelihood test (standard goodness of fit test), specifically $2(l_1 - l_2)$, is used to determine which topology best explains the data. In this test, l_1 is the likelihood function for current configuration, and l_2 is the likelihood function for the NNI configuration being explored. The $2(l_1 - l_2)$ follows a χ^2 distribution (Anisimova and Gascuel 2006). To determine if the alternate branch topology resulted in a significant improvement in the likelihood of the tree, the calculated $2(l_1 - l_2)$ value is compared to a distribution of likelihood values generated using 1000 bootstrap replicates. For each bootstrap replicate, a single column in the alignment is altered, node distances are recalculated and a topology that minimizes the sum of branch lengths is obtained. Log-

likelihood values are calculated for each bootstrap replicate topology, and centered based on the mean log-likelihood for all replicate values (centering, or subtracting the mean from each value, is used to ensure that the resampled data represent appropriate topologies for the actual data). The distribution of $2(l_1-l_2)$ is generated by comparing each of the 1000 bootstrap replicate likelihoods to all other bootstrap replicate likelihoods calculated. The *P*-value reported at each split in the tree corresponds to the probability of obtaining a bootstrap $2(l_1-l_2) \leq$ the $2(l_1-l_2)$ calculated from the actual data (i.e. 0.9 means that the actual data generated a $2(l_1-l_2)$ value that was greater than 90% of the bootstrap replicate likelihoods generated).

Assignment Tests

Assignment tests were conducted on diploid genotypes and haplotypes based on observed allele frequencies in the four clades (Paetkau *et al.* 1995; Paetkau *et al.* 2004; Piry *et al.* 2004). For each individual, the likelihood of a single SNP occurring in a particular clade was estimated as $L=x_i^2$ for homozygotes and $L=2x_i x_j$ for heterozygotes, where *x* is the observed frequency of SNPs *i* and *j* in a given clade. To calculate the likelihood that an individual belonged to a given clade, likelihoods for each SNP in that clade were multiplied together.

Finally, to determine the most likely clade of origin for each individual, we calculated assignment scores according to Piry *et al.* (2004):

$$score_{i,l} = \frac{L_{i,l}}{\sum_{j=1}^k L_{i,j}}$$

Where *k* = the number of clades and *i* = the likelihood of an individual *i* in clade *l*. Here, $L_{i,l}$ is the likelihood value of the individual *i* in the population *l*.

Statistical Evaluation

Since genotypic and phenotypic variables contained outliers and were not normally distributed, we applied the nonparametric equivalent of a *t* test to examine differences between prion-infected and uninfected isolates. The Wilcoxon rank-sum test was performed on ranked values of all variables. Briefly, observations were ranked from highest to lowest, and

then the following test statistic (T) was computed based on the rank sum for group 1 (R_1) and the sample sizes of the two groups (n_1 and n_2):

$$T = \frac{\left[R_1 - \frac{n_1(n_1 + n_2 + 1)}{2} \right] - \frac{1}{2}}{\sqrt{\frac{n_1 n_2}{12} (n_1 + n_2 + 1)}}$$

The test statistic, T , follows a normal distribution and P -values were calculated based on comparison of the calculated T , with critical values (Rosner 2011). We rejected the null hypothesis of no differences between prion-infected and uninfected isolates at $P < 0.05$.

Similarly, hypotheses about genetic and phenotypic characteristics of clades were evaluated using the nonparametric alternative to the ANOVA, the Kruskal-Wallis Test. Similar to the Wilcoxon rank-sum test, the Kruskal-Wallis test pools observations from different groups (in this case, the four clades), ranks them and then compares the average ranks among groups by calculating the test statistic, H as follows:

$$H = \frac{12}{N(N+1)} \sum_{i=1}^k n_i (\bar{R}_i - \bar{\bar{R}})^2$$

where N is the total number of observations, n_i is the sample size of the i th group, k is the number of groups, $\bar{R}_i$ is the average rank in the i th group and $\bar{\bar{R}}$ is the average rank of all observations combined. The test statistic, H , follows a χ^2 distribution and P -values were calculated based on comparison of the calculated H , with critical values at $k-1$ degrees of freedom (Rosner 2011). For single comparisons, we rejected the null hypothesis of no differences among groups at $P < 0.05$, though we also applied a False Discovery Rate approach to minimize the probability of erroneously rejecting a null hypotheses when multiple tests are conducted. Briefly, significant P -values from multiple single tests were ranked and a corrected P -value (P_i) was calculated: $P_i = i / m$, where i is the rank of the original P -value and m is the total number of tests conducted. Single tests P -values $\leq P_i$ were considered significant (Benjamini *et al.* 2001; Benjamini and Hochberg 1995).

Table S1 Strains assigned to a different clade

Strain	Phylogenetic Clade	Assigned Clade	Likelihood Score Clade 2 (%)	Likelihood Score Clade 3 (%)
WLP300	2	3	0.3	99.7
WY1272	2	3	30.5	69.5
CBS6216	3	2	99.7	0.3
YB2573	3	2	98.7	1.3
WLP002	3	2	99.9	0.1
Y976	3	2	53.8	46.2

Table S2 Agar adhesion and invasion for 72 strains in SLAD, limiting glucose and YPAD media

	Adhesion on						
	SLAD	limiting glucose	YPAD	YPAD only	limiting* nutrient only	both limiting and YPAD	neither
Total # of strains	22	41	38	3	11	32	26
% of Total	31	57	53	4	15	44	36

	Invasion on						
	SLAD	limiting glucose	YPAD	YPAD only	limiting *nutrient only	both limiting and YPAD	neither
Total # of strains	61	34	60	8	7	52	5
% of Total	85	47	83	11	10	72	7

*Limiting nutrient only includes both limiting glucose and SLAD media.

Figure S1 Sliding window analysis of Tajima's D (upper) and nucleotide diversity (lower) for the four clades. Calculated using 75 *Saccharomyces cerevisiae* nucleotide sequences with 100 bp windows and 25-bp step size. * indicates a significant ($P < 0.05$) value for Tajima's D, with P-values calculated per (Tajima 1989). Significant Tajima's D values were observed for Clade 3, *RNQ1* N-domain bp 165-193, $D = -1.9$; Clade 1, *RNQ1* C-domain bp 1074-1161, $D = 2.0$; and Clade 2, *RNQ1* C-domain bp 1074-1161, $D = 2.6$.

Figure S2 Heterozygous, shared SNPs for sampled ecological niches. P -values for Kruskal-Wallis test shown in upper left.

Figure S3 Observed allele frequency distribution (red bars) for the four clades compared to the expected frequencies in a stable population (blue lines).

LITERATURE CITED

- Anisimova, M., and O. Gascuel, 2006 Approximate likelihood-ratio test for branches: A fast, accurate, and powerful alternative. *Systematic biology* **55**: 539-552.
- Benjamini, Y., D. Drai, G. Elmer, N. Kafkafi and I. Golani, 2001 Controlling the false discovery rate in behavior genetics research. *Behavioural brain research* **125**: 279-284.
- Benjamini, Y., and Y. Hochberg, 1995 Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*: 289-300.
- Felsenstein, J., 1981 Evolutionary trees from DNA sequences: a maximum likelihood approach. *J Mol Evol* **17**: 368-376.
- Lanave, C., G. Preparata, C. Saccone and G. Serio, 1984 A new method for calculating evolutionary substitution rates. *Journal of molecular evolution* **20**: 86-93.
- Paetkau, D., W. Calvert, I. Stirling and C. Strobeck, 1995 Microsatellite analysis of population structure in Canadian polar bears. *Molecular Ecology* **4**: 347-354.
- Paetkau, D., R. Slade, M. Burden and A. Estoup, 2004 Genetic assignment methods for the direct, real-time estimation of migration rate: a simulation-based exploration of accuracy and power. *Molecular Ecology* **13**: 55-65.
- Piry, S., A. Alapetite, J. M. Cornuet, D. Paetkau, L. Baudouin *et al.*, 2004 GENECLASS2: a software for genetic assignment and first-generation migrant detection. *Journal of heredity* **95**: 536-539.
- Price, M. N., P. S. Dehal and A. P. Arkin, 2009 FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. *Molecular biology and evolution* **26**: 1641-1650.
- Price, M. N., P. S. Dehal and A. P. Arkin, 2010 FastTree 2--approximately maximum-likelihood trees for large alignments. *PLoS One* **5**: e9490.
- Rosner, B. A., 2011 *Fundamentals of biostatistics*. Brooks/Cole, Boston, MA.
- Shimodaira, H., and M. Hasegawa, 1999 Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Molecular biology and evolution* **16**: 1114-1116.
- Tajima, F., 1989 Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**: 585-595.