Determinants of Divergent Adaptation and Dobzhansky-Muller Interaction in Experimental Yeast Populations

James B. Anderson,1,6, Jason Funt,2,3,6
Dawn Anne Thompson,3 Snehit Prabhu,3 Amanda Socha,3 Caroline Sirjusingh,1 Jeremy R. Dettman,1 Lucas Parreiras,1 David S. Guttman,4 Aviv Regev,2,3,5,7,* and Linda M. Kohn1,7
1Centre for the Analysis of Genome Function, Department of Cell and Systems Biology, and Department of Ecology and Evolutionary Biology, University of Toronto, Mississauga, Ontario L5L 1C6, Canada
2Department of Biology, Massachusetts Institute of Technology (MIT), Cambridge, MA 02140, USA
3Broad Institute of MIT and Harvard, 7 Cambridge Center, Cambridge, MA 02142, USA
4Centre for the Analysis of Genome Function and Evolution, University of Toronto, Toronto, Ontario M5S 3G5, Canada
5Howard Hughes Medical Institute, MIT, Cambridge, MA 02140, USA

Summary

Divergent adaptation can be associated with reproductive isolation in speciation [1]. We recently demonstrated the link between divergent adaptation and the onset of reproductive isolation in experimental populations of the yeast Saccharomyces cerevisiae evolved from a single progenitor in either a high-salt or a low-glucose environment [2]. Here, whole-genome resequencing and comparative genome hybridization of representatives of three populations revealed 17 mutations, six of which explained the adaptive increases in mitotic fitness. In two populations evolved in high salt, two different mutations occurred in the proton efflux pump gene PMA1 and the global transcriptional repressor gene CYC8; the ENS genes encoding sodium efflux pumps were overexpressed once through expansion of this gene cluster and once because of mutation in the regulator CYC8. In the population from low glucose, one mutation occurred in MDS3, which modulates growth at high pH, and one in MKT1, a global regulator of mRNAs encoding mitochondrial proteins, the latter recapitulating a naturally occurring variant. A Dobzhansky-Muller (DM) incompatibility between the evolved alleles of PMA1 and MKT1 strongly depressed fitness in the low-glucose environment. This DM interaction is the first reported between experimentally evolved alleles of known genes and shows how reproductive isolation can arise rapidly when divergent selection is strong.

Results

Incipient Speciation during Yeast Experimental Evolution in High Salt and Low Glucose

Divergent adaptation of populations may be associated with the evolution of reproductive isolation in two different ways: ecological isolation [3, 4] and Dobzhansky-Muller (DM) interaction [5]. Under ecological isolation, populations adapt to divergent environments through the accumulation of genetic changes that result in increased fitness. If formed, hybrid populations are genotypically intermediate and therefore suboptimally matched to any environment in which adaptation occurred. Reduced fitness in hybrids retards, if not prevents, gene flow between populations, contributing to speciation. With DM interaction, there is negative epistasis in hybrids among alleles that have never been tested together by natural selection. Ecological isolation and DM interaction can independently contribute to speciation.

Among fully fledged species, the majority of genes identified as components of DM interactions are unrelated to adaptation [6]. An exception is the DM interaction between a nuclear gene AEP2 in Saccharomyces bayanus and a mitochondrial gene OLI1 in S. cerevisiae [7]. It is unknown whether any of the DM incompatibilities identified to date among existing species drove the ancient speciation events.

To separate initial events from subsequent evolutionary change in extant species, we focused on the earliest mutations conferring adaptation and reproductive isolation in experimental populations of yeast under strongly divergent selection. We studied experimental populations of S. cerevisiae that evolved from a single progenitor (P) in either a high-salt (S) or a low-glucose (M) environment [2]. These populations were propagated as batch-transferred cultures, with population size fluctuating daily between 106 (“bottleneck size”) and 109 individuals. We then demonstrated that fitness reduction in hybrids in this system had origins both in ecological isolation and in DM interaction. Our study [2] required only 500 generations of divergent evolution from a common ancestor. This short time frame is in contrast to other studies of genes involved in speciation [5–10] and of isolating mechanisms among extant species [11–13].

Next-Generation Sequencing of Progenitor and Evolved Strains Identifies 17 Candidate Mutations

To identify the evolved mutations, we conducted whole-genome resequencing of single haploid representatives from two populations evolved in high salt (S2 and S6), one population evolved in low glucose (M8), and their common progenitor (P). The three evolved strains had increased fitness in the respective environments in which they evolved (see Figure S1 available online). We mapped all sequenced reads to the finished S. cerevisiae S288C genome and located mutations unique to each evolved strain (Supplemental Experimental Procedures).

Seventeen candidate mutations were confirmed by polymerase chain reaction, conventional sequencing, and comparative genome hybridization analysis (Table S1; Table S2). These included: in S2, nonsynonymous point mutations in the coding sequence of PMA1, GCD2, MET3, and LAP2, a point mutation in the intergenic region 3’ to SECl3 and PNP1, and an expansion of the ENS gene cluster; in S6, nonsynonymous point mutations in the PMA1 and CYC8 coding sequences, point mutations in the YBP2 and CYC8 coding sequences, and a contraction of the ASP3 gene cluster; and in M8, nonsynonymous mutations in the coding sequences of TIM11, RPH1, MDS3,
MKT1, and SGT1 and a synonymous mutation in UBI4. We note that two other studies have identified mutations in genome-wide screens from experimental yeast populations 

Assessing the Contribution of Each Evolved Allele to Fitness in the Adaptive Environment

To assess the contribution of these mutations to adaptation, we measured the fitness effects of each of the mutations unique to S2, S6, and M8 (Tables S1 and S3–S7) by monitoring culture density during growth (Supplemental Experimental Procedures). We compared the fitness of the progenitor (P) and evolved (S2, S6, and M8) strains, in both high-salt and low-glucose environments, to that of progeny genotyped for all five coding loci identified by sequencing, at 20 hr (A) and 30 hr (B) of growth on glucose. Data are available in Table S5. Light gray bars denote ancestral alleles; dark bars denote evolved alleles. Fitness of evolved parent is shown at the upper right corner. Significant differences are noted with a p value. All p values appear in Table S8. The measure of dispersion (error bars) was standard error.

**MKT1** and **SGT1** and a synonymous mutation in **UBI4**. We note that two other studies have identified mutations in genome-wide screens from experimental yeast populations

![Figure 1. Contribution of S2 and S6 Evolved Alleles to Fitness in High Salt](image1)

**Figure 1.** Contribution of S2 and S6 Evolved Alleles to Fitness in High Salt

Shown are fitness measurements (OD_{600}, mean and standard error, normalized to the progenitor value) for 48 offspring fully genotyped for all coding alleles segregating in the cross S2 × P (A, C; 5 loci) and in the cross S6 × P (B, D, 3 loci). Data are aggregated by specific alleles as marked (in each marked category, e.g., “PMA1-2,” the other alleles are segregating). Full data (including intergenic loci) are available in Tables S3 and S4.

(A and B) The bars represent the average fitness effect of each variant across all offspring. Light gray bars denote ancestral alleles; dark bars denote evolved alleles. Fitness of evolved parent is shown at the upper right corner. Significant differences are noted with a p value.

(C and D) Average pairwise effects of the two most advantageous mutations in each strain. Shown are the same data as in (A) and (B) but averaged for two-locus genotypes showing positive interaction. Superscript a denotes ancestral allele; superscript e denotes evolved allele. Interaction was tested by analysis of variance (ANOVA); all p values appear in Table S8.

The measure of dispersion (error bars) was standard error.

**Assessing the Contribution of Each Evolved Allele to Fitness in the Adaptive Environment**

To assess the contribution of these mutations to adaptation, we measured the fitness effects of each of the mutations unique to S2, S6, and M8 (Tables S1 and S3–S7) by monitoring culture density during growth (Supplemental Experimental Procedures). We compared the fitness of the progenitor (P) and evolved (S2, S6, and M8) strains, in both high-salt and low-glucose environments, to that of progeny genotyped for all the identified mutations from crosses with the progenitor. The evolved allele of MKT1 (green) confers a benefit late in the growth cycle, relative to the ancestral genotype (black). Together, these two alleles produce a phenotype (red) that matches that of the M8 strain (dashed).

The measure of dispersion (error bars) was standard error.

**Figure 2. Contribution of M8 Evolved Alleles to Fitness in Low Glucose**

(A and B) Average fitness effect of each variant across the segregant offspring at log phase (20 hr) and postdiauxic shift (30 hr) during growth in low glucose. Shown are fitness measurements (OD_{600}, mean and standard error, normalized to the progenitor value) for 48 progeny from an M8 × P cross, fully genotyped for all five coding loci identified by sequencing, at 20 hr (A) and 30 hr (B) of growth on glucose. Data are aggregated by specific alleles as marked (in each marked category, e.g., “MKT1,” the other alleles are segregating). Full data are available in Table S5. Light gray bars denote ancestral alleles; dark bars denote evolved alleles. Fitness of evolved parent is shown at the upper right corner. Significant differences are noted with a p value. All p values appear in Table S8.

(C) Evolved alleles of **MDS3** and **MKT1** (**MDS3e** and **MKT1e**) account for the M8 phenotype. Shown are growth curves (OD_{600}) from three tetrads from each of two independent crosses segregating for **MDS3** and **MKT1** and from no other evolved alleles (based on full genotyping). The number of replicates for each time course varied between four and eight, reflecting independent assortment. The evolved allele of **MDS3** (green) confers a benefit early, whereas that of **MKT1** (blue) confers a benefit late in the growth cycle, relative to the ancestral genotype (black). Together, these two alleles produce a phenotype (red) that matches that of the M8 strain (dashed).

The measure of dispersion (error bars) was standard error.

**Recurrent Mutations in PMA1 and Phenocopy Mutations in ENA and CYC8 Contribute the Majority of the Observed Fitness Effects in High Salt**

Analysis of the 48 S2 × P progeny showed that the main adaptive determinants for the higher fitness of S2 in salt are
the ENA gene cluster expansion (mean fitness relative to progenitor: ENA1e segregants: 2.35; ENA1a segregants: 1.54, \( p < 0.008 \)) and the evolved allele of PMA1 (mean fitness relative to progenitor: PMA1e segregants: 3.03; PMA1a segregants: 1.16, \( p < 10^{-10} \)), with the PMA1 allele having a more pronounced effect (Figure 1A; Table S3). PMA1 encodes an essential ATP-driven proton pump responsible for maintaining the pH gradient across the cell membrane [16], and the ENA genes encode three paralogous ATP-driven sodium efflux pumps [17] (a similar ENA gene cluster expansion has been observed previously [18] with adaptation to high salt). ENA and PMA1 also had the only significant additive interaction (ANOVA, \( p < 10^{-4} \), Figure 1C), although this interaction was only marginally significant on a logarithmic scale (ANOVA of log(fitness), \( p < 0.07 \)). Nevertheless, the individual effects of the evolved alleles of ENA and PMA1 in increasing fitness act in an unreduced (noninterfering) manner when together in the same haploid genotype. This is consistent with a reduction of H\(^+\) efflux associated with the evolved allele of PMA1 and a greater Na\(^+\) efflux by the expanded ENA gene cluster. Together, the evolved allele of PMA1 and the ENA expansion conferred nearly the full fitness increase of the S2 haploid over the progenitor. Subsidiary minor effects of other mutations are summarized in Table S1.

S6 revealed a pattern of adaptation remarkably parallel to that of S2 (Figure 1B; Table S4). A mutation in PMA1 distinct from that in S2 and another in CYC8, a general transcriptional repressor that acts together with TUP1, each conferred large gains in fitness (mean fitness relative to progenitor: PMA1e segregants: 2.40; PMA1a segregants: 1.64, \( p < 0.002 \); CYC8e segregants: 2.68; CYC8a segregants: 1.39, \( p < 10^{-3} \)). A pairwise interaction between PMA1 and CYC8 (Figure 2D) was positive and marginally significant on an additive scale (ANOVA, \( p < 0.0074 \), significance threshold of \( p = 0.0083 \) with six comparisons), but not on a logarithmic scale (\( p < 0.023 \), significance threshold of \( p = 0.0083 \) with six comparisons). The fitness effects of the evolved alleles of PMA1 and CYC8 are noninterfering when together in the same haploid genotype. The growth defect of S6 (Figures S1A and S1B) was due to the mutation in PMA1; all genotyped strains with the evolved allele grew poorly in YPD and in low glucose (Figure S1G).

The cluster of genes whose expression is specifically induced in S6 (Figure 3B) is enriched for targets of the Tup1-Cyc8 complex (140 common genes between S6 and YPD targets and 240 genes in the S6 upregulated cluster out of 5728 genes in array, \( p < 1.5 \times 10^{-58} \)), suggesting that the evolved CYC8 allele encodes a less potent transcriptional repressor than the ancestral allele. Furthermore, these genes—repressed by Tup1-Cyc8 in YPD [19] and specifically induced in S6—are enriched for known genes induced in the osmotic stress response [20] (53 common genes between 259 osmotic stress response genes and 240 genes in the S6 upregulated cluster out of 5728 genes in array, \( p < 1.52 \times 10^{-23} \)). Among the Tup1-Cyc8 target genes that are derepressed in S6 are the glycerol biosynthesis enzyme HOR2 (important for high salt tolerance) and the ENA1 and ENA2 genes, phenocopying the effect of the genetic expansion of the ENA cluster in S2.

Mutations in MKT1 and MDS3 Contribute to Increased Fitness in Distinct Growth Phases in Low Glucose

The contribution of the M8 evolved alleles to increased fitness and reproductive isolation in low glucose depended on growth
phase (Figure 2; Table S5). At 20 hr, when the cultures were growing exponentially by fermentation, only the MDS3 allele conferred a significant fitness advantage (mean fitness relative to progenitor: MDS3e segregants: 1.3; MDS3a segregants: 0.99, p < 0.003) among the M8 × P offspring (Figure 2A), and there were no significant allele interactions. MDS3 is necessary for growth under alkaline conditions [21], consistent with the fitness benefit it conferred when culture pH was highest (near neutrality). In contrast, the evolved allele of MKT1—a major regulator of the mRNAs encoding mitochondrial proteins [22]—conferred a fitness disadvantage at this phase (mean fitness relative to progenitor: MKT1e segregants: 0.83; MKT1a segregants: 1.36, p < 10^{-4})]. The effect of each of these alleles was reversed after the diauxic shift from fermentation to respiration (30 hr, Figure 2B), when the evolved MDS3 allele conferred a fitness disadvantage (mean fitness relative to progenitor: MDS3e segregants: 0.82; MDS3a segregants: 1.12, p < 10^{-4}) and the evolved MKT1 allele was nearly neutral (mean fitness relative to progenitor: MKT1e segregants: 1.00; MKT1a: 0.97).

To explore the stage-dependent effects of MDS3 and MKT1, we used 24 genotyped offspring of two crosses (three tetrads from each cross), segregating only for the evolved and ancestral alleles of MDS3 and MKT1 and for no other evolved SNPs. The evolved allele of MKT1 alone showed no fitness deficit relative to the progenitor in early time points (Figure 2C) but had a strong increase in fitness late in the growth cycle. This is in contrast to the aggregate effect of MKT1 in the presence of other segregating SNPs (Figures 2A and 2B), where we found a fitness deficit early and near neutrality late. Nevertheless, in both experiments, the effect of MKT1e had the same directionality: it performs better late in the growth cycle than early. The evolved allele of MDS3 showed the opposite directionality, performing better early than late. Importantly, genotypes carrying only the evolved alleles of both MDS3 and MKT1 closely approximated the growth curve of the M8 haploid strain, accounting for the adaptation observed in low glucose (Figure 2C).

A competitive fitness assay over a 24 hr period provided a third, independent measure of the individual fitness effects in low glucose of the evolved alleles of MDS3 and MKT1. This period matched the daily batch-culture regimen in the original 500 generation experiment [2], which included both fermentative and respiratory energy production. Each mutation conferred a fitness advantage over the progenitor alleles (MDS3, 1.25 ± 0.1 standard error [SE] [n = 9]; MKT1, 1.10 ± 0.2 SE [n = 6]). We conclude that our experimental regimen selected for alleles conferring advantages at distinct phases of the yeast growth cycle.

Finally, the evolved alleles of the mitochondrial protein TIM11 and the chromatin modifier gene RPH1 conferred smaller, nonsignificant growth increases at 30 hr (post shift; Figure 2B; Tables S5 and S8). This effect is consistent with the role of the RPH1 paralog in regulating gene expression post diauxic shift [23]. However, the evolved RPH1 allele was not essential to reconstitute the full M8 phenotype.

The MKT1 Allele Reverted to a Wild Allele during Experimental Evolution

The evolved MKT1 allele of M8 is identical to the allele (89G) observed in strains of S. cerevisiae of diverse environmental origin and of S. paradoxus [24], leading to a nonconservative amino acid change from aspartate (P) to glycine (M8). MKT1 encodes a major component in the interaction between Puf3, a sequence-specific RNA-binding protein targeting mRNAs involved in mitochondrial function, and P bodies, which control sequestration and expression of certain mRNAs [22].

The cluster of genes of elevated expression in M8 strains (Figure 3C) is highly enriched for mitochondrial genes (62 common genes between 568 mitochondrial genes and 90 genes in the M8 upregulated cluster out of 5728 genes in array, p < 2.7 × 10^{-41}), including aerobic respiration genes (10 common genes between 64 aerobic respiration genes and 90 genes in the M8 upregulated cluster out of 5728 genes in array, p < 9.7 × 10^{-79}). Furthermore, the M8 cluster includes genes more highly expressed in the vineyard strain RM-11 than the lab strain BY (Figure 3C, bottom). The expression quantitative trait loci for these genes were previously found to be linked to the MKT1 allele that segregates in the BY × RM-11 cross [22].

Taken together, the data suggest a past mutation from the allele (89G) uniformly present in wild strains to that of the laboratory standard (89A), carried by our P strain, followed by an exact reversion of that mutation at some point during the 500 generations of evolution from P to M8. Thus, the progenitor (P) laboratory reference strain carries a less potent form of MKT1, with lower expression of target genes, strongly selected for in lab experiments focusing on early or mid log phase cells in which the wild allele (here the “evolved MKT1”) confers a growth disadvantage. In contrast, the low-glucose selection regimen on a 24 hr batch-transfer cycle used in this study may more closely approximate natural conditions in which growth more often approaches stasis, a condition that would favor the reversion to the naturally occurring 89G allele, and corresponding higher expression of gene targets.

A DM Interaction between PMA1 and MKT1

We next tested for the presence of DM interactions, defined as genetic incompatibilities between alleles independently evolved in the two environments. We measured the fitness, in the two selective environments, of 96 offspring from 24 tetrads from the S2 × M8 and S6 × M8 crosses (Figures S1 and S2). All progeny were fully genotyped for all segregating SNPs, gene-cluster size alterations, and mating type, all of which segregated ~1:1 in tetrads (Tables S6 and S7). As before, we tested each pairwise combination of loci for interaction by means of ANOVA (Table S8).

Among the offspring of the S2 × M8 cross in the low-glucose environment at 24 hr (Figure 4A; Table S6), we found only one marginal p value of 0.015 for a PMA1e-MKT1e negative fitness interaction (in the presence of other segregating alleles). Because the initial value was marginal, we tested this preliminary evidence for an interaction in two additional independent experiments.

In the first, we measured the fitness of 24 genotyped offspring of two crosses (three tetrads from each cross) that segregated at only the two SNP sites in PMA1 and MKT1 (no other evolved alleles were present in the cross). Here we found that the fitness of offspring carrying both evolved alleles was depressed over the entire growth cycle in low glucose (Figure 4B), most prominently at the 21 and 24 hr time points (the same time point as in Figure 4A). At 24 hr, an overall ANOVA of additive variation over the four genotypes was statistically significant (p < 0.016, one test only), and
In this study, we used whole-genome sequencing of progenitor and evolved strains, along with genotyping, fitness assays, and mRNA profiling, to identify and characterize the genetic and molecular basis of early events associated with divergent selection in experimental yeast populations. We found six key determinants, each of which contributes to ecological isolation in which genotypically mixed hybrids are not as well matched to either environment as the pure evolved strains. The DM interaction between \textit{PMA1} and \textit{MKT1} is the first reported between evolved alleles of known genes in experimental populations derived from a common ancestor. Although it is tempting to speculate on how such an incompatibility might affect natural yeast populations, our study was limited to haploid effects. One possibility is that a DM incompatibility like that reported here would quickly be eliminated under the low rate of outcrossing in \textit{S. cerevisiae} \cite{25}; in such a case, the incompatibility would persist in hybrid populations. These possibilities remain to be investigated.

No consistent functional theme has yet emerged among the known “speciation genes” implicated in DM interactions among species in nature \cite{5–10}. Here we show that the adaptive mechanisms evolved in response to strong directional selection in two environments have substantial effects on gene regulation and phenotype and that at least two of the adaptive determinants produce an intrinsic clash resulting in a fitness reduction characteristic of a DM interaction. In extant species examined to date, the majority of DM incompatibilities occur in genes unrelated to ecological adaptation \cite{6}. Our experiments with the same strains at 24 hr, finding a significant interaction between the \textit{PMA1} and \textit{MKT1} alleles, when fitting a linear mixed model treating strain as a random effect and tested against a null model of no interaction between \textit{PMA1} and \textit{MKT1} (\textit{PMA1aMKT1a:} \textit{0.69} ± \textit{0.02}, \textit{PMA1aMKT1e:} \textit{0.70} ± \textit{0.02}, \textit{PMA1eMKT1a:} \textit{0.66} ± \textit{0.01}, \textit{PMA1eMKT1e:} \textit{0.46} ± \textit{0.03}, \textit{p} < \textit{10}^{-4})). This interaction is also significant on log scale (\textit{p} < \textit{4} × \textit{10}^{-4})). This fulfills the criterion for a DM interaction \cite{2}. Similar assays with offspring segregating for \textit{MDS3} and \textit{PMA1} showed no such negative interaction (Figure 4C).

We independently confirmed the negative interaction between the \textit{PMA1} and \textit{MKT1} genotypes in competition experiments in the low-glucose environment at an early time point (17 hr under conditions matching those in Figure 4B), showing a negative reduction in the number of doublings in the \textit{PMA1eMKT1e} genotype strain (\textit{MKT1e:} \textit{0.87} ± \textit{0.01 SE} \textit{[n = 3]; PMA1e,} \textit{0.89} ± \textit{0.02 SE} \textit{[n = 3]; PMA1eMKT1e,} \textit{0.7} ± \textit{0.07 SE} \textit{[n = 3]}, all relative to the doublings by the progenitor). As a control, we confirmed the expected differential effect of \textit{MDS3e} in the competition assay (1.28 ± 0.01 \textit{SE} \textit{[n = 2]}). The difference in fitness among the genotypes fell just short of being significant (\textit{p} < \textit{0.061}, one-way ANOVA, linear scale), likely reflecting the smaller sample size and the earlier (17 hr) time point. Nevertheless, each of these three experiments supported the conclusion of negative interaction between the evolved alleles of \textit{PMA1} and \textit{MKT1}, most notably at 24 hr. In contrast, there was no evidence for a DM interaction in the \textit{S2 × M8} and \textit{S6 × M8} offspring in high salt and the \textit{S6 × M8} offspring in low glucose (Tables S7 and S8), where all adaptive determinants had effects similar to those in crosses of the evolved strains and the progenitor (Figures S1A and S1B).
study, in which we experimentally set the conditions thought to foster incipient speciation, documents a counterexample in which divergent adaptive changes themselves confer a DM incompatibility. It is possible that newly evolved adaptive mechanisms under other conditions will have similarly far-reaching consequences, with potential for DM incompatibility. We propose that the potential pool of speciation genes includes genes conveying adaptation under strong selection in the earliest stages of speciation—that functional diversity in speciation genes could reflect the diversity of adaptive mechanisms.

Supplemental Information

Supplemental Information includes Supplemental Experimental Procedures, two figures, and eight tables and can be found with this article online at doi:10.1016/j.cub.2010.06.022.

Acknowledgments

J.B.A. and L.M.K. were supported by Discovery Grants from the Natural Sciences and Engineering Research Council of Canada. We thank the Broad Sequencing Platform, and in particular H. Spurling, for sequencing work, as well as I. Gat-Viks, M. Chan, J. Konieczka, D. Mohammad, C. Ye, M. Gutman, and other members of the Regev laboratory for helpful discussions and comments. J.F. was supported by the National Science Foundation Graduate Research Fellowship, D.A.T. was supported by the Human Frontiers Science Program, and A.R. was supported by the Howard Hughes Medical Institute, a Career Award at the Scientific Interface from the Burroughs Wellcome Fund, a National Institutes of Health PIONEER award, and a Sloan Fellowship.

Received: January 27, 2010
Revised: June 8, 2010
Accepted: June 9, 2010
Published online: July 15, 2010

References