Concerted evolution of life stage performances signals recent selection on yeast nitrogen use

Sebastian Ibstedt[1]*, Simon Stenberg[2]*, Sara Bagés[1], Arne B. Gjuvsland[2], Francisco Salinas[3], Olga Kourtchenko[1], Jeevan Karloss[2], Anders Blomberg[1], Stig W. Omholt[4], Gianni Liti[3], Gemma Beltran[5], Jonas Warringer[1,2]**

[1] Department of Chemistry and Molecular Biology, University of Gothenburg, PO Box 462, 40530 Gothenburg, Sweden

[2] Centre for Integrative Genetics (CIGENE), Department of Animal and Aquacultural Sciences, Norwegian University of Life Sciences (UMB), PO Box 5003, 1432 Ås, Norway

[3] IRCAN, CNRS UMR 6267, INSERM U998, University of Nice, 06107 Nice, France

[4] NTNU Norwegian University of Science and Technology, Faculty of Natural Sciences and Technology, Department of Biotechnology, N-7491 Trondheim, Norway

[5] Department of Biochemistry and Biotechnology, Universitat Rovira i Virgili, 43007 Tarragona, Spain

* Equal contributions

** Corresponding author: Jonas.Warringer@cmb.gu.se

Keywords: Selection, adaptation, life history, nitrogen, yeast, metabolism

ABSTRACT

Exposing natural selection driving phenotypic and genotypic adaptive differentiation is extraordinary challenging. Given that an organism's life stages are exposed to the same environmental variations, we reasoned that fitness components, such as the lag, rate and efficiency of growth, directly reflecting performance in these life stages should often be selected in concert. We therefore conjectured that correlations between fitness components over natural isolates, in a particular environmental context, would constitute a robust signal of recent selection. Critically, this test for selection requires fitness components to be determined by different genetic loci. To explore our conjecture, we exhaustively evaluated the lag, rate and efficiency of asexual population growth of natural isolates of the model yeast *S. cerevisiae* in a large variety of nitrogen limited environments. Overall, fitness components were well correlated under nitrogen restriction. Yeast isolates were further crossed in all pairwise combinations and co-inheritance of each fitness component and genetic markers were traced. Trait variations tended to map to Quantitative Trait Loci (QTL) that were private to a single fitness component. We further traced QTLs down to single nucleotide resolution and uncovered loss-of-function mutations in *RIM15*, *PUT4, DAL1* and *DAL4* as the genetic basis for nitrogen source use variations. Effects of SNPs were unique for a single fitness component, strongly arguing against pleiotropy between lag, rate and efficiency of reproduction under nitrogen restriction. The strong correlations between life stage performances that cannot be explained by pleiotropy compellingly supports adaptive differentiation of yeast nitrogen source use and suggests a generic approach for detecting selection.

INTRODUCTION

Exposing natural selection and environmental factors driving adaptive differentiation of phenotypes and genotypes is extraordinary challenging. The classical approach aims to unveil co-variations in phenotype, genotype or environmental factors with a reasonable proxy for fitness. Unfortunately, spurious associations emerge due to genetic drift and population structure. Genetic hitchhiking, whereby non-adaptive alleles and their phenotypes piggyback with adaptive variants at nearby loci compounds the problem (Barton 2000), as do pleiotropy, whereby non-adaptive phenotypes hitchhike with adaptive by association to the same gene variant (Stearns 2010). Consequently, only a fraction of environmental factors, genotypes and phenotypes that co-vary with a fitness proxy is directly linked to selection. An elegant approach towards exposing selection on a phenotype was taken by Orr (Orr 1998). If alleles (QTLs) enhancing a trait value are consistently found in one as compared to another lineage of a species, then positive selection is likely to have acted on that trait. Nevertheless, for most traits in most species, too few QTLs are known for the test to constitute a powerful alternative. Grouping traits into aggregates circumvents this lack of power (Fraser, et al. 2010). However, it is rarely clear how a functionally relevant trait grouping is to be achieved.

To be exposed to selection in a particular environment, alleles have to alter net population growth, *i.e.* birth or death, in that environment. All organisms pass through life cycles composed of distinct life stages (Stearns 1992). Success in any of these life stages, either in the form of increased birth relative death or faster progression to a subsequent life stage in which reproduction occurs, directly affects net population growth (Roff 1992). Performance in these life stages is therefore the most immediate components of fitness and lower level phenotypes affect fitness via them. In many organisms, life stages are intimately linked such that they very consistently are subject to the same environmental variations and the same selective pressures. We therefore reasoned that if an environment exposes a population to long term selection, optimization of performance in multiple life stages, would often emerge. In absence of selection however, such fitness components would tend to fluctuate independently of each other, as dictated by the specifics of genetic drift, genetic hitchhiking and pleiotropy. This conjecture directly suggested a test for adaptive differentiation in the form of robust correlations between fitness components. To be valid, this test requires fitness components to be genetically independent, as confounding correlations due pleiotropy and genetic hitchhiking otherwise would emerge.

To explore the power of this test, we considered natural isolates of the model yeast *S. cerevisiae*, which has a well understood life history with fitness components that are straight forward to define, can be precisely measured and that are possible to dissect genetically. Sex in natural *S. cerevisiae* is rare and heavily oriented towards self-fertilization (Ruderfer, et al. 2006; Tsai, et al. 2008). Selection is therefore likely to act primarily on the asexual life cycle (Warringer and Blomberg 2014). Accordingly, the time it takes for an asexual population to pass through the lag phase to the reproductive stage (lag time), its net rate of growth in the reproductive stage (doubling time) and the total population density achieved before the reproductive stage ends (efficiency) are key fitness components. Population growth of chemoheterotrophic microbes, such as yeasts, is limited by access to nutrients, primarily energy, carbon, reducing equivalents and nitrogen. *S. cerevisiae* obtains energy, carbon and reducing equivalents from sugar, while using a wide diversity of nitrogen sources to satisfy its nitrogen needs (Cooper 1982). Decomposing or damaged fruit, flowering plant nectar and tree saps (exudates) are primary microhabitats occupied by yeast (Landry, et al. 2006; Hittinger 2013). These are rich in sugar but poor in nitrogen, with one or a few nitrogen sources dominating each microhabitat (Fig S1A, B) (Gardener and Gillman 2001). Selection for optimal use of single nitrogen sources can therefore be expected to have been strong in natural yeasts and to have driven recent adaptive differentiation. To evaluate the validity of the proposed test for selection, we therefore measured asexual population growth of natural yeasts in a vast array of nitrogen restricted, single nitrogen source environments.

RESULTS

Natural yeasts are highly differentiated with regards to nitrogen source use

To evaluate fitness components of natural yeasts we considered four *S. cerevisiae* strains, each representing one of the non-reproductively isolated yeast populations discovered outside China: the West African DBVPG6044 (WA), the North American YPS128 (NA), the European DBVPG6765 (E), and the Sake Y12 (S). Lineages differed genetically by 0.3-0.7%, corresponding to several million generations of evolution, and encompassed >50% of the known SNP and phenotypic variation outside China (Liti, et al. 2009; Warringer, et al. 2011; Bergstrom, et al. 2014) (Fig S1C). Strains were clonally propagated in 28 nitrogen environments corresponding to all low complexity nitrogen sources used by yeast. These were present at equal and yield limiting nitrogen concentrations (Fig S1D, E). From high resolution

growth curves, it was abundantly clear that natural strains vary greatly in their capacity to use different nitrogen sources (Fig 1A). Quantifying these variations, we resolved the growth supported by the various nitrogen sources into its components and extracted the rate (population doubling time), lag (time to initiate growth) and efficiency of population expansion (total change in population density)(Fig 1B). Variations among natural lineages were pervasive with significant strain differences (False Discovery Rate, FDR, q≤5%) in 45% of the 84 fitness component measures (Fig 1C, S2). The lower the mean capacity to convert nitrogen into growth, the larger was the variation between strains, with the West African standing out as a generally poor performer (Fig 1D-E). Thus, natural *S. cerevisiae* strains are highly differentiated with regards to nitrogen source use and the differentiation is strongest for poor nitrogen sources. The latter is in agreement with recent wine yeast observations (Gutierrez, et al. 2013). Given that entry into stationary phase could occur for reasons other than nitrogen depletion, we investigated whether total population densities achieved truly reflected exhaustion of nitrogen. Surveying the nitrogen remaining in the medium as a function of time in strains WA and NA, we found nitrogen depletion to neatly coincide with exit from the exponential growth phase (Fig 1F). In fact, 10-35% of the total population expansion occurred after this time point, presumably by mobilization of stored intracellular nitrogen. Thus, stationary phase levels reflected depletion of not only external but also internally stored nitrogen.

Concerted selection on lag, rate and efficiency of yeast nitrogen use

To test whether the differentiation of nitrogen source use were caused by adaptation and selection, we compared measures of lag, rate and efficiency of population growth over all nitrogen restricted environments and strains. Correlations between all pairs of fitness components were excellent (Fig 2A). Thus, the efficiency of population growth co-varied extensively (Pearson, $r = 0.85$ and 0.61) with both the rate and lag. This correlation between fitness components has two components: co-variation that is due to genetic variants that are shared between strains and co-variation that is due to genetic variants that differ between strains. The former corresponds to concerted evolution of fitness components in the ancestral lineage. The latter corresponds to concerted evolution of fitness components in lineage(s) after their separation. As >98% of nucleotides are shared between strains, we expect much of the co-variation between fitness components to arise from these invariant positions. We can

5

estimate this ancestral co-evolution by considering trait averages, over the four strains, as a proxy for ancestral traits. Comparing trait averages for different fitness components over the different environments, we find these to be strongly correlated (Fig S3A). Thus, fitness components have largely co-evolved with regards to nitrogen source use in the ancestor of current yeasts. Fitness component co-evolution after the separation of lineages would be reflected in a correlation over the four strains. Given the low sample size (four), which adds substantial randomness that hides any existing co-variation, we cannot provide an accurate estimate of recent fitness component co-evolution. The average correlation (r=0.30, considering each environment and pair of fitness component independently) is likely to be a substantial underestimate and should be regarded as lower bound for co-evolution. Under an assumption of variations in fitness components being due to different alleles, these correlations are challenging to explain without invoking selection. However, if underlying alleles are pleiotropic and consistently affect multiple fitness components in the same direction, selection need not necessarily be involved. Rejecting pleiotropy is not straight forward for alleles that are shared between strains. However, it can be done for alleles that differ between strains because following mating between two strains, we expect variations in two fitness components to co-segregate in offspring recombinants only if they are explained by the same, or genetically linked, alleles. We therefore mated the four natural lineages in all six pairwise combinations. Resulting hybrids were sporulated to obtain 92 F1 meiotic progeny from each cross. F1 segregants were genotyped at nucleotide sites for which parents were known from genome sequencing to be polymorphic (Cubillos, et al. 2011). We then precisely followed the net population growth of these 552 recombinants in all nitrogen environments, extracting >42,000 fitness component measures. Finally, we evaluated the co-segregation of each SNP marker and each life history trait by QTL linkage mapping, for each cross and nitrogen environment separately. Overall, we detected 230 robust (permutation test, α=0.05) QTL. 87.4% of all QTL were unique for a single fitness component (Fig 2B), rejecting the null hypothesis of genetic variants with pleiotropic effects on fitness components. For example, QTLs private to lag, rate or efficiency determined population growth in the West African x Wine/European cross in citrulline (Fig 2C), leucine and isoleucine (Fig S3B). In agreement with this fitness component specificity of QTLs, only weak correlation (r=0.15) between fitness components remained in the average environment and cross. Together with the strong correlations between natural isolates, this supports that natural variation in yeast nitrogen source use indeed reflects adaptive differentiation.

QTLs were largely unique to a single nitrogen source (Fig 2D, S4A), meaning that different alleles tend to control variation in the use of different nitrogen sources and are acted upon independently by selection. Nevertheless, a few extremely pleiotropic QTL, affecting the same fitness component in a large number of nitrogen restricted microhabitats, were observed. For example, a single QTL on chromosome VI accounted for poor growth efficiency of the Wine/European strain in almost all nitrogen restricted niches, suggesting a general relaxation of selection for efficient nitrogen use in this strain. Predominantly, QTL also emerged only in a single of the three crosses in which a particular genetic background was represented (Fig S4B). Therefore, the penetrance of the underlying alleles was highly dependent on genetic context, suggesting widespread epistasis. Although the results reported above provide strong indications, the limitations of QTL data should be acknowledged. Due to lack of power, detected QTLs do not explain all of the heritable variation in traits (Bloom, et al. 2013). Furthermore, the break-up of parental allele structures during meiosis and the emergence of novel allele combinations can both disrupt and promote epistasis, affecting trait values. Finally, QTL represent the combined effect of all alleles in a region. These effects call for some caution and means that stringent conclusions can only be reached by identifying and measuring the effects of individual mutations.

Alleles private to a single fitness component control natural variation in yeast nitrogen source use

To address the shortcoming of the QTL analysis, we traced the genetic basis of three QTLs down to single alleles. All non-essential genes within these three QTL regions were identified. We then crossed reference strain (BY4741) gene knockouts for each of these genes to both well and poor performing natural strains, resulting in pairs of diploid hemizygotes. A trait difference within such hemizygotic pairs was taken to imply that the two alleles at the hemizygotic locus affect the trait differently. To exclude confounding effects from haploinsufficiency, we next mated a poor and well performing natural strain. In two of the resulting diploid hybrids, we reciprocally deleted either of the two alleles. A trait difference between these was considered final confirmation of the causative loci. We first focused on the chromosome VI QTL controlling variation exclusively in population growth efficiency in all crosses involving the Wine/European strain, and in almost all nitrogen environments (Fig 3A, Fig S5A). Hemizygotic diploids only containing the WE allele of *RIM15* mimicked the poor

growth efficiency of the WE parent (Fig 3B). Furthermore, the WE *RIM15* consistently imposed poor efficiency on hemizygotes from crosses with other natural yeasts (Fig 3C, S5B). In all cases, *RIM15* fully explained the defects of the Wine/European parent. The WE *RIM15* contains an early two base pair insertion, *rim15c.459_460insCA*, shifting the reading frame to cause an early stop codon (Fig S5C). Thus, WE *RIM15* is a null allele. No rate or lag QTL were found in the *RIM15* region in the relevant crosses and nitrogen sources and diploids only containing the WE *RIM15* allele showed no impairments of the rate or lag (Fig S5D-F). *RIM15* encodes a poorly understood protein kinase known to control stationary phase entry (Wanke, et al. 2005) and sporulation efficiency (Bergstrom, et al. 2014). As sporulation is induced by nitrogen depletion, it is tempting to speculate that *rim15* sporulation defects arise as a consequence of nitrogen use impairment.

We next considered the strong chromosome XV QTL affecting population growth rate on proline in all crosses involving the West African DBVPG6044 (Fig 3D). Hemizygotic diploids from crosses with reference strain gene knockouts identified the West African *PUT4*, encoding a high affinity proline permease (Andreasson, et al. 2004), as the only allele in the region contributing to the trait variation (Fig 3E). Reciprocal *PUT4* hemizygotes from crosses between the West African and other natural strains confirmed that the West African *PUT4* impairs proline growth rate, accounting for 54 to 96%, depending on cross, of the West African defect (Fig 3F). No efficiency or lag proline QTL were found in the *PUT4* region and no diploids hemizygotic for the WA *PUT4* allele showed any impairments of population growth efficiency or lag (Fig S6A-C). Thus, the *PUT4* defect is private to growth rate. WA *PUT4* allele harbors no non-synonymous but several synonymous and promoter mutations, implying that expression differences cause the proline growth variation (Fig S7A). WA *PUT4* is inherited by the SK1 and Y55 lab strains, but the proline defect is completely buffered in Y55, emphasizing that its exposure to selection depends on genetic context (Fig S7B, C).

Finally, we considered the exceedingly slow growth of the West African when utilizing allantoin (Fig 3G), the nitrogen secretion product of mammals other than apes (Young, et al. 1944). Allantoin lag time is unaffected, but the enormous growth rate effect precluded growth efficiency estimation. The defect affected close to 50% of WA offspring, regardless of cross, suggesting a monogenic effect (Fig S8A). Arguing against monogenicity however, results from the lab strain cross implied that neither *DAL4* nor *DAL1 WA* alleles were able to support allantoin growth (Fig 3H). *DAL4* and *DAL1* encode enzymes catalyzing the first two steps in allantoin uptake and degradation and are arranged back to back in an allantoin use gene

cluster (Fig S8B, C). This arrangement fuelled concern that deletion of one *DAL* gene impairs expression of its neighbor, as suggested (Naseeb and Delneri 2012). Inspecting WA *DAL1* and *DAL4,* we uncovered a single nucleotide frameshifting insertion in *DAL4* and an early proline to serine non-synonymous mutation, predicted to be strongly detrimental, in *DAL1* (Fig S8D). Repairing either of these mutations with S288C variants failed to restore allantoin growth (Fig 3I), implying that they independently disrupt allantoin utilization. To confirm this, we repaired either of the two mutations while also introducing a centromeric plasmid with the functional S288C version of the other allele. Perfect allantoin growth was restored in both constructs. Thus, the *DAL1* and *DAL4* are null alleles arising from *dal4c.1201delA* and *dal1c.415C>T* mutations respectively. Each mutation disrupts growth to a degree that is comparable to the effect of both mutations together, reflecting positive epistasis between loss-of-function mutations in components of a linear pathway (Lehner 2011). Curiously the lab strain Y55, which has inherited both WA null alleles, achieved reasonably fast allantoin growth suggesting the existence of unknown routes for allantoin use (Fig S8E, F).

DISCUSSION

From a genetics perspective it is not surprising that distinct alleles control variation in lag, rate and efficiency of population expansion. Effects on any single one of these direct components of fitness, both by environmental factors (Warringer, et al. 2008) and gene deletion (Warringer, et al. 2003), is common. Biochemically, a positive coupling between the rate and efficiency of population expansion is controversial because the rate of metabolic reactions is negatively correlated with the energy remaining after reactions (Westerhoff, et al. 1983; Heinrich, et al. 1997; Pfeiffer, et al. 2001). Under energy restriction, this is expected to force a trade-off between population growth rate and efficiency and there is ample experimental support for this (Postma, et al. 1989; Spor, et al. 2008; Spor, et al. 2009). However, under nitrogen restriction, accelerated burning of energy by fast metabolic reactions does not necessarily reduce biomass yields and increase in biomass yield does not require slow metabolism. Energy is available in excess. Consequently, the observed correlation between rate and efficiency does not violate thermodynamics. Assimilated nitrogen can either be stored or used inside cells or channeled into reproduction. Reflecting this distinction, the nitrogen content per cell can vary over orders of magnitude. Enhanced population growth efficiency under nitrogen restriction represents the prioritized channeling of nitrogen to the

next generation. This is important, because it suggests an evolutionary mechanism whereby efficiency enhancing mutations may be selected. In absence of internal storages, efficiency enhancing mutations cannot increase frequency in well mixed populations because non-mutants enjoy equal access to the resource and benefit equally from efficiency enhancements (Hardin 1968; MacLean 2008). Privatization of nitrogen by internal accumulation means that only genomes carrying efficiency enhancing mutations benefit from them, enabling selection on population growth efficiency.

The fundamental assumption of this paper is that a strong correlation between fitness components that reflect performance in different life stages, but are non-pleiotropic, only can emerge through concerted selection. Concerted selection follows from the intrinsic link between life-stages underlying these fitness components, which are exposed to the same environmental variations and selective pressures. However, the conjecture does not necessarily predict such a strong correlation as here observed. It is tempting to speculate that diminishing return of consecutive mutations on a trait (Chou, et al. 2011; Khan, et al. 2011; Kryazhimskiy, et al. 2014) may be part of the explanation for the strength of detected correlations. First, it provides a powerful incentive for parallel accumulation of mutations enhancing different fitness components. Two mutations enhancing different fitness components will simply tend to have a larger aggregate fitness contribution than two mutations enhancing the same fitness component. Second, the diminishing return of consecutive mutations means that even if different numbers of beneficial mutations have been accumulated in different fitness components, their aggregate trait contributions will tend to be rather similar.

Although there are few intrinsic limitations to the applicability of the proposed test for selection, caveats need to be recognized. First, it is prudent to point out that our evidence for co-evolution of fitness components is strongest with regards to co-evolution in the ancestral lineage. In contrast, our evidence for non-pleiotropy between fitness components relates to variants that differ between strains, i.e. more recent evolutionary events. It is hard to envision why fixed variants would be largely pleiotropic, whereas polymorphisms would be largely non-pleiotropic with regards to effects on fitness components. Nevertheless, we can currently not completely rule out this possibility. Second, the test is straight forward to apply only to traits that directly reflect performance in different life stages. Great care should be taken before extending the test to organisms with hard to define, hard to measure or hard to genetically dissect fitness components. Third, absence of correlation between fitness

components is not a sufficient ground for rejecting selection. Over shorter time periods, optimization of a single fitness component may occur by sampling error, even if multiple fitness components are under selection, as few mutations are accumulated in each genome. This has been made abundantly clear in artificial laboratory selections (MacLean and Gudelj 2006; Novak, et al. 2006; Bachmann, et al. 2013). Optimization of single fitness components may also occur in organisms where life stages are less intimately linked and occur in different environmental contexts, as selection pressures then may be radically different. Fourth, the test is unable to distinguish between positive and recently relaxed selection. Here, the *RIM15*, *PUT4*, *DAL1* and *DAL4* alleles were lineage specific loss-of-function mutations, implying recent relaxation of selection in that lineage (Zorgo, et al. 2012). Nevertheless, extension-of-function mutations in *RIM15*, *DAL4* and *PUT4* emerge rapidly in nitrogen restricted artificial laboratory selections (Hong and Gresham 2014). Hence, it cannot be excluded that loss-of-function mutations are ancestral and that independent repair recently have occurred in well-functioning lineages. Fifth, all fitness components should ideally be accounted for. We could not quantify effects on sexual recombination, which although exceedingly rare may be under selection, or on diploid growth traits, which often differs from haploid growth traits (Zorgo, et al. 2013). Furthermore, we could not measure performance in the yeast spore state, which may be under selection in nature (Neiman 2011), or in the net death phase, when growth has ceased. From a general perspective, not accounting for all possible phenotypic and genetic relationships between fitness components leaves a margin for error. Assigning selection may still be valid, but the interpretation of which fitness components that are under selection may be confounded due to undetected pleiotropy between them.

MATERIALS AND METHODS

Strains: *S. cerevisiae* natural isolates were collected and haploidized as described (Cubillos, et al. 2009; Liti, et al. 2009). Haploids were crossed in all pairwise combinations. 23 tetrads from each cross were sporulated to obtain 92 F1 recombinants. These were genotyped at 164-180 polymorphic positions (Cubillos, et al. 2011). Diploid hemizygotes were obtained by mating haploids to single gene deletion BY4741 gene deletions using robotics. *RIM15* and *PUT4* reciprocal hemizygotes of crosses between natural isolates were obtained by one-step PCR and manual crossing. WA *dal4c.1201delA* and *dal1c.415C>T* mutations were replaced by the S288C variants using site specific *in vivo* mutagenesis (Stuckey, et al. 2011). WT and

versions with either the *dal4c.1201delA* or *dal1c.415C>T* repaired were transformed with centromeric p5472 MoBY plasmids (Ho, et al. 2009). A haploid S288C derivative, YSBN10, was used as control.

Medium and cultivation: Micro-cultivation was performed in Synthetic Defined (SD) medium as described (Warringer and Blomberg 2003; Warringer, et al. 2003), using a single nitrogen source present at 30mg N/L. Experiments were run over 72-144h until all isolates had entered stationary phase (Warringer and Blomberg 2003). Where a stationary phase had not been reached at the end of the experiment, experiments were discarded to avoid confounding effects. Population growth lag, rate and efficiency were extracted as described (Warringer, et al. 2008). Fitness components were $\log_2$ transformed and normalized to those of at least four controls. Means of repears (n=2) were used for linkage analysis. Other experiments were performed with larger numbers of repeats, as indicated.

Nitrogen uptake: Micro-cultivation of yeast cells was performed in Synthetic Defined (SD) medium as described above. At each sampling time point, 20 biological replicates of each micro-cultivated sample were pooled. Supernatants were collected and analyzed by DEEMM derivatization and HPLC analysis as described (Gomez-Alonso, et al. 2007). Concentrations of the relevant nitrogen source were calculated using internal and external standards.

QTL analysis: QTL analysis was performed using a non-parametric model in r/QTL at 2cM density (Broman, et al. 2003). Significance was estimated by permutation tests (α=0.05). QTL positions were counted as identical if positions were within one average marker distance.

Sequence and SNP analysis: Sequence data was taken from (Bergstrom, et al. 2014). Negative SNP consequences were predicted by Sorting Intolerant From Tolerant (SIFT) analysis (Kumar, et al. 2009).

Statistics: Two group comparisons were performed using a homoscedastic two-tailed Student's t-test and False Discovery Rates (FDR) at $q \leq 5\%$. Significance of pleiotropy was tested using randomized permutations and a Mann-Whitney-Wilcoxon test.

SUPPLEMENTARY INFORMATION

Supplementary Materials and Methods

Supplementary figures S1-S8

REFERENCES

Andreasson C, Neve EP, Ljungdahl PO. 2004. Four permeases import proline and the toxic proline analogue azetidine-2-carboxylate into yeast. Yeast 21:193-199.

Bachmann H, Fischlechner M, Rabbers I, Barfa N, Branco Dos Santos F, Molenaar D, Teusink B. 2013. Availability of public goods shapes the evolution of competing metabolic strategies. Proc Natl Acad Sci U S A 110:14302-14307.

Barton NH. 2000. Genetic hitchhiking. Philos Trans R Soc Lond B Biol Sci 355:1553-1562.

Bergstrom A, Simpson JT, Salinas F, Barre B, Parts L, Zia A, Nguyen Ba AN, Moses AM, Louis EJ, Mustonen V, et al. 2014. A High-Definition View of Functional Genetic Variation from Natural Yeast Genomes. Mol Biol Evol.

Bloom JS, Ehrenreich IM, Loo WT, Lite TL, Kruglyak L. 2013. Finding the sources of missing heritability in a yeast cross. Nature 494:234-237.

Broman KW, Wu H, Sen S, Churchill GA. 2003. R/qtl: QTL mapping in experimental crosses. Bioinformatics 19:889-890.

Chou HH, Chiu HC, Delaney NF, Segre D, Marx CJ. 2011. Diminishing returns epistasis among beneficial mutations decelerates adaptation. Science 332:1190-1192.

Cooper TG. 1982. Nitrogen metabolism in Saccharomyces cerevisiae. In: Strathern JN, Jones EW, Broach JR, editors. The molecular biology of the yeast Saccharomyces cerevisiae: metabolism and gene expression. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press p. 39-99.

Cubillos FA, Billi E, Zorgo E, Parts L, Fargier P, Omholt S, Blomberg A, Warringer J, Louis EJ, Liti G. 2011. Assessing the complex architecture of polygenic traits in diverged yeast populations. Mol Ecol 20:1401-1413.

Cubillos FA, Louis EJ, Liti G. 2009. Generation of a large set of genetically tractable haploid and diploid Saccharomyces strains. FEMS Yeast Res 9:1217-1225.

Fraser HB, Moses AM, Schadt EE. 2010. Evidence for widespread adaptive evolution of gene expression in budding yeast. Proc Natl Acad Sci U S A 107:2977-2982.

Gardener MC, Gillman MP. 2001. Analyzing variability in nectar amino acids: composition is less variable than concentration. J Chem Ecol 27:2545-2558.

Gomez-Alonso S, Hermosin-Gutierrez I, Garcia-Romero E. 2007. Simultaneous HPLC analysis of biogenic amines, amino acids, and ammonium ion as aminoenone derivatives in wine and beer samples. J Agric Food Chem 55:608-613.

Gutierrez A, Beltran G, Warringer J, Guillamon JM. 2013. Genetic basis of variations in nitrogen source utilization in four wine commercial yeast strains. PLoS One 8:e67166.

Hardin G. 1968. The tragedy of the commons. The population problem has no technical solution; it requires a fundamental extension in morality. Science 162:1243-1248.

Heinrich R, Montero F, Klipp E, Waddell TG, Melendez-Hevia E. 1997. Theoretical approaches to the evolutionary optimization of glycolysis: thermodynamic and kinetic constraints. Eur J Biochem 243:191-201.

Hittinger CT. 2013. Saccharomyces diversity and evolution: a budding model genus. Trends Genet 29:309-317.

Ho CH, Magtanong L, Barker SL, Gresham D, Nishimura S, Natarajan P, Koh JL, Porter J, Gray CA, Andersen RJ, et al. 2009. A molecular barcoded yeast ORF library enables mode-of-action analysis of bioactive compounds. Nat Biotechnol 27:369-377.

Hong J, Gresham D. 2014. Molecular specificity, convergence and constraint shape adaptive evolution in nutrient-poor environments. PLoS Genet 10:e1004041.

Khan AI, Dinh DM, Schneider D, Lenski RE, Cooper TF. 2011. Negative epistasis between beneficial mutations in an evolving bacterial population. Science 332:1193-1196.

Kryazhimskiy S, Rice DP, Jerison ER, Desai MM. 2014. Microbial evolution. Global epistasis makes adaptation predictable despite sequence-level stochasticity. Science 344:1519-1522.

Kumar P, Henikoff S, Ng PC. 2009. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat Protoc 4:1073-1081.

Landry CR, Townsend JP, Hartl DL, Cavalieri D. 2006. Ecological and evolutionary genomics of Saccharomyces cerevisiae. Mol Ecol 15:575-591.

Lehner B. 2011. Molecular mechanisms of epistasis within and between genes. Trends Genet 27:323-331.

Liti G, Carter DM, Moses AM, Warringer J, Parts L, James SA, Davey RP, Roberts IN, Burt A, Koufopanou V, et al. 2009. Population genomics of domestic and wild yeasts. Nature 458:337-341.

MacLean RC. 2008. The tragedy of the commons in microbial populations: insights from theoretical, comparative and experimental studies. Heredity (Edinb) 100:471-477.

MacLean RC, Gudelj I. 2006. Resource competition and social conflict in experimental populations of yeast. Nature 441:498-501.

Naseeb S, Delneri D. 2012. Impact of chromosomal inversions on the yeast DAL cluster. PLoS One 7:e42022.

Neiman AM. 2011. Sporulation in the budding yeast Saccharomyces cerevisiae. Genetics 189:737-765.

Novak M, Pfeiffer T, Lenski RE, Sauer U, Bonhoeffer S. 2006. Experimental tests for an evolutionary trade-off between growth rate and yield in E. coli. Am Nat 168:242-251.

Orr HA. 1998. Testing natural selection vs. genetic drift in phenotypic evolution using quantitative trait locus data. Genetics 149:2099-2104.

Pfeiffer T, Schuster S, Bonhoeffer S. 2001. Cooperation and competition in the evolution of ATP-producing pathways. Science 292:504-507.

Postma E, Verduyn C, Scheffers WA, Van Dijken JP. 1989. Enzymic analysis of the crabtree effect in glucose-limited chemostat cultures of Saccharomyces cerevisiae. Applied and Environmental Microbiology 55:468-477.

Roff DA. 1992. The Evolution of Life Histories: Theory and Analysis. New York: Chapman and Hall.

Ruderfer DM, Pratt SC, Seidel HS, Kruglyak L. 2006. Population genomic analysis of outcrossing and recombination in yeast. Nat Genet 38:1077-1081.

Spor A, Nidelet T, Simon J, Bourgais A, de Vienne D, Sicard D. 2009. Niche-driven evolution of metabolic and life-history strategies in natural and domesticated populations of Saccharomyces cerevisiae. BMC Evol Biol 9:296.

Spor A, Wang S, Dillmann C, de Vienne D, Sicard D. 2008. "Ant" and "grasshopper" life-history strategies in Saccharomyces cerevisiae. PLoS One 3:e1579.

Stearns FW. 2010. One hundred years of pleiotropy: a retrospective. Genetics 186:767-773.

Stearns SC. 1992. The evolution of life histories. Oxford: Oxford university press.

Stuckey S, Mukherjee K, Storici F. 2011. In vivo site-specific mutagenesis and gene collage using the delitto perfetto system in yeast Saccharomyces cerevisiae. Methods Mol Biol 745:173-191.

Tsai IJ, Bensasson D, Burt A, Koufopanou V. 2008. Population genomics of the wild yeast Saccharomyces paradoxus: Quantifying the life cycle. Proc Natl Acad Sci U S A 105:4957-4962.

Wanke V, Pedruzzi I, Cameroni E, Dubouloz F, De Virgilio C. 2005. Regulation of G0 entry by the Pho80-Pho85 cyclin-CDK complex. EMBO J 24:4271-4278.

Warringer J, Anevski D, Liu B, Blomberg A. 2008. Chemogenetic fingerprinting by analysis of cellular growth dynamics. BMC Chem Biol 8:3.

Warringer J, Blomberg A. 2003. Automated screening in environmental arrays allows analysis of quantitative phenotypic profiles in Saccharomyces cerevisiae. Yeast 20:53-67.

Warringer J, Blomberg A. 2014. Yeast Phenomics-Large scale mapping of the genetic basis for organismal traits. In: Hancock JM, editor. Phenomics. Boca Raton: CRC Press. p. 172-207.

Warringer J, Ericson E, Fernandez L, Nerman O, Blomberg A. 2003. High-resolution yeast phenomics resolves different physiological features in the saline response. Proc Natl Acad Sci U S A 100:15724-15729.

Warringer J, Zorgo E, Cubillos FA, Zia A, Gjuvsland A, Simpson JT, Forsmark A, Durbin R, Omholt SW, Louis EJ, et al. 2011. Trait variation in yeast is defined by population history. PLoS Genet 7:e1002111.

Westerhoff HV, Hellingwerf KJ, Van Dam K. 1983. Thermodynamic efficiency of microbial growth is low but optimal for maximal growth rate. Proc Natl Acad Sci U S A 80:305-309.

Young EG, Wentworth HP, Hawkins WW. 1944. The absorption and excretion of allantoin in mammals. J. Pharmacol. Experi. Therapeutics 81:1-9.

Zorgo E, Chwialkowska K, Gjuvsland AB, Garre E, Sunnerhagen P, Liti G, Blomberg A, Omholt SW, Warringer J. 2013. Ancient Evolutionary Trade-Offs between Yeast Ploidy States. PLoS Genet 9:e1003388.

Zorgo E, Gjuvsland A, Cubillos FA, Louis EJ, Liti G, Blomberg A, Omholt SW, Warringer J. 2012. Life history shapes trait heredity by accumulation of loss-of-function alleles in yeast. Mol Biol Evol 29:1781-1789.

FIGURE LEGENDS

Figure 1 Natural yeasts are highly differentiated for nitrogen source use

A) Population density as a function of time for Y12 (S), YPS128 (NA) and the S288C derivative YSBN in sample single nitrogen source environments (n=2). B) Extraction of lag (time to initiate proliferation), rate (population doubling time) and efficiency (total change in population density) of population growth from high density growth curves. C) Natural yeasts vary in capacity to convert different nitrogen sources into population growth. Nitrogen sources significantly (FDR, α=5%) better (red) or worse (green) for one natural yeast lineage (n=6) than others (n=18) are shown. D) Variation in nitrogen source utilization between natural yeast lineages is largest for poor nitrogen sources. Variance over the four natural strains is plotted as a function of mean. Population growth efficiency is displayed. Line = linear regression. E) The North American is superior and the West African inferior at utilizing single nitrogen sources. Fraction of fitness component measures in which each strain (n=6) is significantly (FDR, α=5%) better (blue) or worse (red) than other natural yeasts (n=18) is shown. F) Yeast variation in total population density reached reflects variations in nitrogen use efficiency. Population size (left y-axis, continuous bold lines) and nitrogen remaining in the medium (right y-axis, broken bold lines) was quantified as a function of time in West African and North American strains. Broken lines = time of external nitrogen depletion.

.

Figure 2 Concerted selection on lag, rate and efficiency of yeast nitrogen use

A) The efficiency, rate and lag of population growth are strongly correlated in natural yeast strains. The $\log_2$ of population growth efficiency, rate and lag were pairwise compared over all natural yeasts and nitrogen environments. Means ($n = 6$), Pearson correlations (r) and linear regressions are shown. B) Efficiency, rate and lag of nitrogen source use are genetically

independent. QTL (permutation test, α = 0.05) private to a fitness component or shared between components were summed over six yeast crosses and 28 nitrogen environments. C) QTL private to a single fitness component control variation in the efficiency, rate and lag of citrulline use in the WE x WA cross. LOD scores represent co-segregation of growth and genetic markers in 92 F1 recombinants. Bands represent chromosome boundaries. Dotted lines show threshold for significance (permutation test, α = 0.05). D) Nitrogen use QTL are private to a single nitrogen source. Fraction of total QTL detected for one or more than one nitrogen source is displayed. Each cross and fitness component was considered separately. There is weak support for more environment pleiotropy than expected by chance (Mann-Whitney-Wilcoxon, $p = 0.016$)

Figure 3 Alleles private to single fitness components control variation in yeast nitrogen source use

A-C) *RIM15* accounts for poor population growth efficiency of the Wine/European lineage during nitrogen limitation. A) LOD score plots of the co-segregation of growth efficiency and genetic markers among 92 F1 recombinants from the WE x S cross in proline. Bands represent chromosome boundaries and dotted line shows threshold for significance (permutation test, α = 0.05). B) Proline growth efficiency of diploid hemizygotes retaining only the WE (blue) or S (red) allele in crosses between WE or S and single BY4741 gene deletions. Gene deletions correspond to non-essential genes in the chromosome VI QTL region. Mean (n=2) log_2 values were normalized to those of the YSBN control ($n = 4$). Error bars = SEM. C) Proline growth efficiency of reciprocal diploid hemizygotes retaining only the WE or only the alternate *RIM15* allele in crosses between WE and other natural lineages. Mean (n = 8) log_2 values were normalized to those of the YSBN control (n = 5). Error bars = SEM, p-values = Student's test. D-F) *PUT4* allele accounts for poor proline growth rate in the West African lineage. D) LOD score plot of co-segregation of proline growth rate and genetic markers among 92 F1 recombinants from the WA x WE cross. Bands represent chromosome boundaries. Dotted line indicates threshold for significance (permutation test, α=0.05). E) Proline growth rates of diploid hemizygotes retaining only the WE (blue) or WA (red) allele in crosses between WE or WA and single BY4741 gene deletions. Gene deletions correspond to non-essential genes in the chromosome XV QTL region. Mean (n=2) log_2 values were normalized to those of the YSBN control (n=4). Error bars = SEM, p-values = Student's test.

F) Proline growth rates of reciprocal diploid hemizygotes retaining only the WA or only the alternate *PUT4* allele in crosses between WE and other natural lineages. Mean (n=8) $\log_2$ values were normalized to those of the YSBN control (n=5). Error bars = SEM (n=2), p-values = Student's test. *PUT4* explained 97±6% of the WA-NA, 67±4% of the WA-WE and 54±3% of the WA-S variation. G-I) A frameshift mutation in *DAL4* and a non-synonymous mutation in *DAL1* both impairs WA growth rate on allantoin. G) LOD score plot of co-segregation of allantoin growth rate and each genetic marker among 92 F1 recombinants from the WA x WE cross. Bands represent chromosome boundaries. Dotted line shows threshold for significance (permutation test, α=0.05). H) Allantoin growth rates of diploid hemizygotes retaining only the WE (blue) or WA (red) allele in crosses between WE or WA and single BY4741 gene deletions. Gene deletions correspond to non-essential genes in the chromosome IX QTL region. Mean (n=8) $\log_2$ values were normalized to those of the YSBN control (n=5). Error bars = SEM. K) Addition of the S288C *DAL1* or *DAL4* to a haploid WA via a centromeric plasmid and replacement of the candidate SNP in the other *DAL* gene with the S288C nucleotide variant identified *dal4c.1201delA* and *dal1c.415C>T* as independently impairing the allantoin growth rate of WA. Mean (n=6-12) $\log_2$ population growth rates are displayed. Error bars = SEM.

A

Log₂ density (OD)

Valine
Glycine

YPS128
Y12
YSBN

Time (h)

B

Log₂ density (OD₄₅₀₋₅₈₀ₙₘ)

$\Delta OD/\Delta t$
ΔOD
Δt

Time (h)

C

Performance
Poor
Average
Good

DBVPG6044 (WA) — Efficiency, Rate, Lag
Y12 (S) — Efficiency, Rate, Lag
DBVPG6765 (WE) — Efficiency, Rate, Lag
YPS128 (NA) — Efficiency, Rate, Lag

Adenine
Allantoin
Ammonium
Arginine
Asparagine
Citrulline
Cytosine
GABA
Glutamic acid
Glutamine
Glutathione
Glycine
Isoleucine
Leucine
Methionine
Ornithine
Phenylalanine
Proline
Serine
Threonine
Tryptophan
Tyrosine
Urea
Valine

D

$r = -0.55$

Variance

Log₂ density change (OD)

E

Poor performance
Good performance

YPS128 (NA)
Y12 (S)
DBVPG6765 (WE)
DBVPG6044 (WA)

Frequency of traits

F

Log₂ population density (OD)

Residual nitrogen mg N/L

NA
WA
Methionine

NA
WA
Phenyl alanine

NA
WA
Tyrosine