Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Genetics & Development

The molecular basis of phenotypic variation in yeast

Justin C Fay

The power of yeast genetics has now been extensively applied to phenotypic variation among strains of *Saccharomyces cerevisiae*. As a result, over 100 genes and numerous sequence variants have been identified, providing us with a general characterization of mutations underlying quantitative trait variation. Most quantitative trait alleles exert considerable phenotypic effects and alter conserved amino acid positions within protein coding sequences. When examined, quantitative trait alleles influence the expression of numerous genes, most of which are unrelated to an allele's phenotypic effect. The profile of quantitative trait alleles has proven useful to reverse quantitative genetics approaches and supports the use of systems genetics approaches to synthesize the molecular basis of trait variation across multiple strains.

Addresses
Department of Genetics and Center for Genome Sciences and Systems Biology, Washington University, St. Louis, MO, United States

Corresponding author: Fay, Justin C (jfay@genetics.wustl.edu)

Current Opinion in Genetics & Development 2013, **23**:672–677

This review comes from a themed issue on **Genetics of system biology**

Edited by **Shamil Sunyaev** and **Fritz Roth**

For a complete overview see the Issue and the Editorial

Available online 21st November 2013

0959-437X/$ – see front matter,

http://dx.doi.org/10.1016/j.gde.2013.10.005

Introduction

Most variable phenotypes have a complex genetic basis. While it has become relatively easy to coarsely map the genetic basis of such traits, identifying causative genes and genetic changes within those genes remains challenging. As such, the power of yeast genetics has made *Saccharomyces cerevisiae* an attractive model for dissecting complex traits and revealing the molecular bases of a number of traits in exquisite detail [1].

Perhaps the greatest challenge to quantitative trait mapping is being able to map any type of change, for example, single nulceotide polymorhism (SNP), insertion/deletion polymorphism (InDel), or change in chromosome structure, and to determine how it affects a trait, for example, changes in protein structure or gene expression. In many organisms, meeting this challenge is difficult due to technological limitations. For example, noncoding changes are more difficult to identify and their effects may not always be recapitulated outside of their native context. *S. cerevisiae* has proven adept at meeting these challenges and revealing the types variants in a population that contribute to quantitative trait variation.

Currently, over 100 quantitative trait genes (QTGs) and half as many quantitative trait nucleotides (QTNs) have been identified in yeast. Here, I review insights that have emerged from the consensus of these studies and highlight remaining challenges that need to be addressed. While the QTN program has been justly criticized as not being reflective of evolutionary change [2], the high resolution genetic analysis conducted in yeast is relevant to our understanding of how trait variation is generated and maintained in a population.

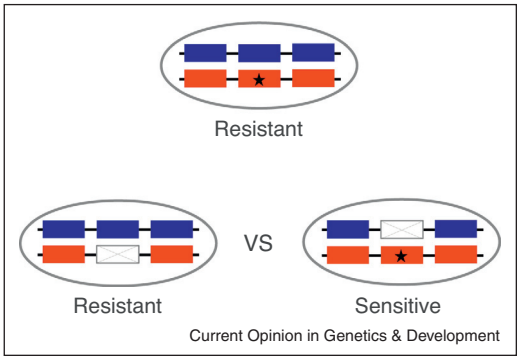
The case for QTN: linked QTGs and multiple QTNs

Genomics has made quantitative trait locus (QTL) mapping a tenable means of identifying the genetic basis of phenotypic variation. As a result, the number of QTL and their effect sizes have been documented for many traits. However, there has always been a concern that single QTL of large effect might be caused by multiple linked QTGs of smaller effect [3]. The first example of such in yeast involved a major effect high temperature growth QTL that was elegantly shown to be caused by three out of 15 genes in a 32 kb region [4]. The genes were identified using reciprocal hemizygosity analysis (Figure 1), which has since become the standard for efficient and robust identification of QTGs in yeast. More recent studies have documented other examples of linked QTGs [5–7], which intriguingly also occur in the same region on chromosome XIV but involve different genes. Currently, this 75 kb region harbors 10 known QTGs linked to a variety of traits and represents a hotspot of quantitative trait variation (Figure 2). In a similar vein, multiple QTNs have been shown to occur within a single QTG. Such genes include *FLO11* [8], *HO* [9], *IME1* [10], *MLH1* [11], *PCA1* [12], *PMS1* [11], *RAD5* [13], *AQY1* and *AQY2* [14], highlighting the importance of carefully dissecting causal variants within a QTG. As summarized in Figure 3, out of 110 quantitative trait alleles that have been identified in yeast, half (54) have been delineated to specific nucleotide changes.

Types of changes

Mirroring the relative abundance of different types of DNA polymorphism [15], most mapped variants are single nucleotide changes (Figure 2 and Figure 3). However, chromosome rearrangements, copy number variants (CNVs) and InDels have also been identified,

Figure 1



Reciprocal hemizygosity test. The test compares the phenotypes (e.g. resistant/sensitive) of two heterozygous diploid strains that are the same except for being hemizygous for a gene of interest. One strain (left) carries a deletion (x) of one allele, the other strain (right) carries a deletion of the other allele. The test can identify a recessive or partially recessive allele (star) of a gene, regardless of whether the causal mutation occurs in a coding or noncoding sequence.

potentially more often than what one might expect based on the genome abundance of these types of polymorphisms relative to SNPs [15]. In the case of the tandemly duplicated *ENA* locus, sodium/lithium tolerance is associated with a recent introgression of an *S. paradoxus* allele into some but not all strains of *S. cerevisiae* [15,16], a novel source of variation that is governed by reproductive barriers rather than mutation rate.

Coding versus noncoding changes

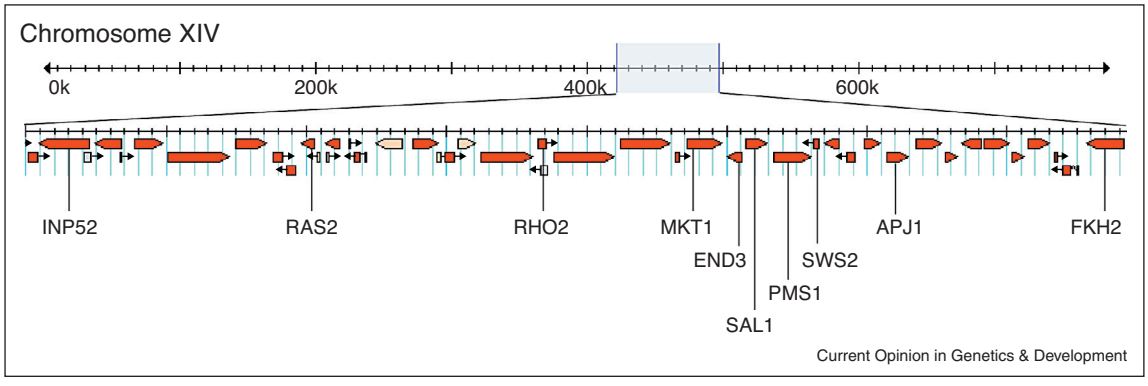
Similar to other organisms [17], the majority (69%) of QTNs lie within protein coding sequences (Figure 3). However, contrary to other organisms there is arguably little bias towards successfully mapping coding relative to noncoding QTN in yeast. Nevertheless, the relative abundance of protein coding changes is nearly identical to the 72% of the yeast genome that encodes proteins [18], which is substantially higher than most plant and animal genomes.

Small versus large effects

The vast majority of alleles that have been identified generate moderate to large phenotypic effects. In the context of QTL mapping, there is undoubtedly a bias towards identifying alleles of large effect; they are the first to be pursued and also the easiest to resolve to single genes or genetic changes. Even so, alleles of moderate to large effect can explain most variation in a cross. In two studies of sporulation efficiency, 88% of variation in a cross [10] and 92% of the parental difference [19] was explained by QTN in three genes, only one of which, *RME1*, was shared between the two studies. More generally, a modest number of QTL was found to explain an average of 88% of additive genetic variation across 46 traits [20]. Nevertheless, quite a few of the QTL identified in this latter study can be considered small effect loci as they would not have been found with even a moderate number of 100 segregants. Thus, large effect alleles, while commonly found, don't preclude the existence of numerous alleles of small effect.

Recent studies have successfully targeted and identified alleles of small effect. In these studies, small effect alleles were either linked to large effect alleles or masked by interactions with them such that they were only identified by first fixing the alleles of large effect, either through backcrossing [21,22] or allele replacement [23•,24]. The observation that QTL were found at or nearby three previously discovered large effect QTGs adds further evidence of multiple linked QTGs or QTNs [23•]. However, targeting alleles of small effect is not easy; their effects are more difficult to distinguish from subtle differences in genetic background that can arise between nearly identical strains. For example, a non-complementation screen using the yeast deletion collection faithfully

Figure 2



Quantitative trait gene hotspot. Ten QTGs in a 70 kb region on chromosome XIV are labeled below a graphic view of genes indicated by red pointed boxes.