Published as: *FEBS Lett*. 2009 December 17; 583(24): 3959.

Fungal regulatory evolution: *cis* and *trans* in the balance

Dawn Anne Thompson[1] and **Aviv Regev**[1,2,†]

[1] Broad Institute of MIT and Harvard, 7 Cambridge Center, Cambridge, MA 01242

[2] Howard Hughes Medical Institute, Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 01240

Abstract

Regulatory divergence is likely a major driving force in evolution. Comparative genomics is being increasingly used to infer the evolution of gene regulation. *Ascomycota* fungi are uniquely suited among eukaryotes for regulatory evolution studies, due to broad phylogenetic scope, many sequenced genomes, and tractability of genomic analysis. Here we review recent advances in the identification of the contribution of *cis* and *trans* factors to expression divergence. Whereas current strategies have led to the discovery of surprising signatures and mechanisms, we still understand very little about the adaptive role of regulatory evolution. Empirical studies including experimental evolution, comparative functional genomics and hybrid and engineered strains are showing early promise toward deciphering the contribution of regulatory divergence to adaptation.

Keywords

Fungal; regulatory; evolution; expression; divergence

Introduction

Divergence in the regulatory mechanisms that control gene expression has been repeatedly postulated to play a major role in evolution. Examples of regulatory differences between species are known in a wide range of species including bacteria [1], fungi [2], flies [3,4], and mammals [5]. However, the mechanisms through which regulatory systems evolve are still poorly understood, and in most cases the adaptive importance of regulatory changes is unknown.

Comparative genomics approaches based on whole-genome sequences of diverse organisms are being increasingly used to infer the evolution of gene regulation. These studies rely on two main strategies: (**1**) computational comparison of *cis*-regulatory organization between promoters of orthologous genes, and (**2**) comparative functional analysis of mRNA profiles and TF-promoter interactions measured across different organisms. The latter empirical approach, while less prevalent, is gaining increasing attention and beginning to shed light on the relation between sequence evolution, changes in gene expression and adaptation.

†Correspondence should be addressed to AR (aregev@broadinstitute.org).

The *Ascomycota* fungi are a particularly suitable group for studies of regulatory evolution. A large number of eukaryotic species with characterized life styles belong to this monophyletic group, which spans at least 300 million years of evolution (Figure 1A). These include two extensively studied model organisms, *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*, as well as important human pathogens, such as *Candida albicans*. Many organisms in this group are easy to grow in the lab, and are amenable to genetic manipulation and environmental perturbations, allowing us to effectively delineate the molecular mechanisms underlying biological responses. *Ascomycota* genomes are small and compact enough to be computationally tractable, while still having many of the hallmarks of a eukaryotic system, thus providing an excellent model.

An unprecedented amount of genomics knowledge has been accrued on *Ascomycota*. On the one hand, the molecular systems in the model organism *S. cerevisiae* have been studied using a wide range of genomics tools, from extensive transcription profiling studies (over 2000 profiles available, [6]), through single cell proteomics [7], and large-scale screens of protein and genetic interactions [8,9]. On the other hand, sequencing and extensive analysis of over 100 genomes has delineated functional elements in specific genomes as well as global phylogenetic trends. In particular, a whole genome duplication (WGD) event has occurred in the phylogeny [10,11], and sequenced genomes are available from before and after this important event.

Finally, strong evidence suggests that regulatory changes were associated with divergence in major physiological responses among *Ascomycota*. For example, although central carbon metabolism follows the same general outline in all yeasts, important biochemical, genetic and regulatory variations exist. Some species, including *S. cerevisiae* and close relatives, follow a respiro-fermentative growth during aerobic growth on glucose (characterized by high glucose uptake, high ethanol secretion rate and low biomass yield); whereas other species (*e.g. Kluyveromyces*) favor respiratory growth in the same conditions (low glucose uptake and high biomass yield). A shift to a respiro-fermentative lifestyle has occurred more than once in the phylogeny, most notably following the whole genome duplication event [12,13] and independently in *Schizosaccharomyces* [14]. These metabolic differences are also accompanied by divergence of gene regulation, including the introduction of a host of glucose-dependent repressive mechanisms on respiratory metabolism, the differential transcriptional regulation of isozymes [13], and the repression of mitochondrial biogenesis genes in log phase growth [15,16].

In this review, we focus on recent advances made in understanding the evolution of gene regulation in *Ascomycota* from short evolutionary timescales (hundreds of generations to 5 millions years) that are typical to intra-species variation to long timescales spanning tens of millions of years involving extensive adaptive radiation and speciation. We examine conservation and divergence at three levels of study. **First**, we wish to characterize and quantify the *key evolutionary signatures* that are observed in these species. **Second**, we wish to understand the *molecular mechanisms*, in *cis* and in *trans*, underlying these signatures. **Finally**, we wish to understand the relative role of neutral changes and selection in shaping conservation and divergence of regulatory systems. As we show, whereas current strategies have led to the discovery of surprising signatures and mechanisms, we still understand very little about the adaptive role of regulatory evolution. Empirical studies including experimental evolution, comparative functional genomics and hybrid and engineered strains are showing early promise toward deciphering the contribution of regulatory divergence to adaptation.

An evolutionary and functional dichotomy of expression conservation and divergence in gene orthologs

Expression profiles collected across organisms allow us to determine the extent of conservation or divergence in the mRNA levels and regulation across orthologous genes. Within *Ascomycota*, large compendia of mRNA profiles exist for the model organisms *S. cerevisiae*, *S. pombe* and *C. albicans*, whereas smaller datasets are available for other Ascomycota (e.g. other Saccharomyces [17], C. glabrata [18], K. lactis [19], and some *Euascomycota* [20,21]) as well as different *S. cerevisiae* strains [22–26]. Using such profiles, and a good mapping of groups of orthologous genes [6] we can determine the degree of expression divergence (ED) – a quantitative measure of the differences in the expression of a pair of orthologs between two species [17].

Interestingly, divergence in the expression of gene orthologs follows a broad functional and evolutionary dichotomy [25–28]: genes with conserved expression typically encode proteins involved in growth control and general metabolism ('growth branch'), whereas those with divergent expression are often subtelomeric, responsive to external and internal signals (*e.g.* stress response) and are nonessential. This dichotomy in variation is preserved at multiple levels: from variation in isogenic cells in a population [29], through genetic variants of *S. cerevisiae* [24,26,30], to different species in the *sensu stricto* clade [17]. Furthermore, it is also reflected by concomitant constraints on copy number variation at great phylogenetic distances [6]: genes from the low-ED 'growth' branch have few duplication and loss events, whereas those in the high-ED 'stress and metabolism' end are volatile, and experience substantial variation in copy number between species. For low-ED genes, this suggests a strong selective pressure and functional constraint on the specific amount of gene products in the cell. For high-ED genes, it is tempting to conversely suggest a pressure for flexibility in gene regulation. However, this conclusion must be interpreted with care as we discuss below.

The impact of *cis* regulatory elements and promoter organization on expression divergence

Both *cis* and *trans* regulatory mutations/polymorphisms can contribute to expression divergence. A genetic change can affect expression directly in *cis*, by altering transcription factor binding sites in the promoter region, changing chromatin organization or affecting mRNA stability, or indirectly, by modifying the activity of the gene product and causing expression changes through feed-back control. Alternatively, a polymorphism in one gene can affect the expression of other genes in *trans* (Figure 1B).

Many *cis*-regulatory elements are conserved in closely related species. In some cases, the specific site and its location in the promoter is conserved, a feature exploited for motif identification using alignments of orthologous regulatory regions [31,32]. In other cases, gain and loss of *cis*-regulatory motifs, and the potential for corresponding changes in transcription factor binding, occur on relatively short time scales (on the order of 5 – 20 my), both within and between species [28,33–36]. Doniger *et al.* [33,34] estimated that, of the lineage-specific binding-site losses within *sensu stricto Saccharomyces*, over half correspond to newly emerged binding sites in the same regulatory regions. Turnover of one binding site in a promoter for a functionally equivalent one can explain how gene expression can be maintained despite change in regulatory sequences.

In other cases the apparent loss of a binding site corresponds to loss of TF control and a change in gene expression pattern. For example, most species have enriched Rapid Growth Elements (RGE, AATTTT) upstream of all ribosomal proteins (RP), but post-WGD species that can

decouple fermentation from respiration have lost the RGE sites upstream of mitochondrial RP genes [15], consistent with the loss of coregulation of mitochondrial function and cell growth.

More generally, several promoter components can affect the expression plasticity of a gene, including point mutations in binding sites [35], sequence features affecting its chromatin structure [16,17,27], and the presence of unstable tandem repeats [37]. For example, the dichotomy between high- and low-ED genes discussed above also corresponds to distinct chromatin organization and transcriptional mechanisms. The promoters of genes with conserved expression (low-ED) have well-positioned nucleosomes, and most of their regulatory elements are situated within a substantial nucleosome free region (NFR). Their transcription is TATA-independent and they are less susceptible to chromatin remodeling. Conversely, high-ED genes are associated with promoters with more distributed nucleosomes, their transcription is TATA-dependent, and is more sensitive to chromatin remodeling. Furthermore, the sensitivity of gene's expression level to mutation increases in the presence of a TATA box. Interestingly, recent studies have shown that promoters of high-ED genes are also associated with the presence of unstable tandem repeats [37], and that changes in such repeats may drive changes in nucleosome organization and gene expression. The promoters of these genes are enriched for TF binding sites resulting in more potential for combinatorial interactions proposed to enhance evolutionary divergence.

Notably, these promoter features are associated with expression variability both between isogenic cells and between genetically distinct strains and species, suggesting that both short-term 'responsiveness' of gene expression and long term evolvability may be inter-twined through promoter organization. However, it is unclear if the latter is the result of direct selection or is simply a by-product of the type of regulation required to respond to environmental stimuli.

Divergence in gene expression through changes in *trans* factors

Changes in *trans*-factors contribute to expression divergence through either a change in the factor's responsiveness to upstream signals, binding to newly emerging sites upstream of new targets, or through the factor's ability to bind different 'non-canonical' sites. There are several known cases where changes in a TF's binding preferences co-evolved with changes in the regulatory sequences upstream of orthologous targets. For example, *in vitro* binding studies showed that the ancestral Rpn4 TF bound a wider set of sequences than the modern-day *S. cerevisiae* protein. The change in sequence specificity corresponds to changes in motif usage upstream of the target proteasome genes – *S. cerevisiae* targets no longer contain sites that the TF cannot bind, even though these are prominently upstream of *C. albicans* proteasome genes [2]. It is unclear whether the co-evolution of Rpn4 specificity and its targets' upstream motifs was driven by selection or emerged simply through neutral drift in one followed by co-evolution of the other.

In other cases, while the specificity of the *trans*-factor remains unchanged, it facilitates the acquisition of new targets under its control. For example, Borneman *et al*. [35] used ChIP-chip to examine binding of two TFs, Ste12p and Tec1p, in *S. cerevisiae*, *S. mikatae*, and *S. bayanus*, estimated to have diverged 20 mya. TF binding events were conserved across all three species in only 20% of promoters, suggesting substantial gain- or loss of individual targets. In many cases the loss of TF binding correlated with loss of the binding site. However, in a number of cases TF binding occurred despite absence of an identifiable underlying DNA motif, confirming that TFs can bind non-canonical sites [38]. Such 'promiscuous' binding may be important for acquisition of new targets, since weak TF binding to a non-canonical sequence followed by selection for optimal binding could support the emergence of a recognizable TF element [36]. Similar divergence (15% between *S. cerevisiae*, *K. lactis*, and *C. albicans*) and

promiscuity was observed for the direct targets of the Mcm1 transcription factors, suggesting a general trend [39].

Duplication and divergence of *trans* factors can have a large impact on expression divergence. For example, the Yeast –specific AP-1 (YAP1) bZIP family of TFs that are conserved from yeast to human are a clear example of the special role of TF duplication in *trans*-divergence [40]. Changes is specificity of the eight paralogous Yap transcription factors of *S. cerevisiae* is attributable to both differences in DNA binding motifs and variation in the regulatory domains that mediate response to a variety of stresses. Other factors that could contribute to changes in specificity include cooperative binding with other TFs, TF-homo- or heterodomerization, or different DNA binding kinetics.

Consistent with the major impact of promoter chromatin organization on expression divergence, several recent studies have shown that chromatin remodeling factors can have a major impact on expression divergence. For example, the changes in expression accompanying the perturbation (mutation or deletion) of various chromatin regulators revealed that many high-ED genes are markedly regulated at the chromatin level [41]. Furthermore, much of the divergence in expression level between wild and lab strains of *S. cerevisiae* can be explained by *trans* differences in chromatin modifiers [42], as we discuss below.

It can be challenging to reconcile this substantial evolutionary diversity in *cis*- and *trans*-factors controlling the expression of individual genes with the functional organization of regulatory networks into co-regulated modules (‘regulons’) of functionally related genes [43–45]. Comparative studies from bacteria to yeast to human have established that modules of co-expressed genes can be highly conserved. However, how are multiple evolutionary events coordinated across dozens and hundreds of genes to sustain regulons? In some cases, changes occur in *trans*, thus conserving *co*-expression while diverging the mRNA levels of all transcripts in a module, while *cis* changes may primarily serve to tune membership in modules. In addition, as we discuss in a separate review [46] multiple forms of functional redundancy also allow for more complex divergence of regulatory mechanism while maintaining module identity.

Quantifying *cis*- and *trans*- contributions to expression divergence

While the examples above are instructive, they are insufficient to assess the relative importance of distinct mechanisms to expression divergence. Two types of studies have distinguished the magnitude of *cis*- and *trans*- effects on expression divergence. Within-species studies rely on expression quantitative trait loci (eQTL) analysis using segregants from a cross between distinct strains [23,42,47,48] and monitoring allele-specific expression in intraspecific hybrids [30,49]; between-species studies distinguish *cis*- and *trans*-effects by comparing interspecific hybrids to the individual species. Since these strategies rely on crosses or hybrids they are limited to the < 20 Mya scale [28].

The most extensive mapping of *cis*- and *trans*- effects has been conducted with a cross between a lab (BY) and wine (RM) strain of *S. cerevisiae*, that have substantial differences in gene expression likely due to adaptation to different niches [22,23]. Extensive eQTL analysis has shown that a large fraction (70%) of the variation in gene expression among segregants in this cross can be attributed to *trans*-effects. Interestingly, many of these *trans*-effects may involve variation in chromatin modi ers, consistent with their mechanistic role in affecting the expression of high-ED genes, as discussed above. Notably, analysis of allele-specific expression in intraspecific hybrids of two strains can provide more mechanistic information. For example, regulatory variation that acts directly in *cis* would result in an allele-specific expression pattern. Indeed, in the majority of cases in which the expression level of a gene is linked to it’s own locus based on the segregant analysis, its expression pattern was allele-

specific in the BY and RM hybrid, indicating direct *cis* action due to alterations in the promoter sequence [49].

The effect of genetic variation on gene expression phenotypes often depends on environmental conditions. A recent study estimated the effects of gene-environment interaction (GEI) on transcript abundance by profiling expression in the segregants of the BY and RM cross in both glucose and ethanol [50]. Numerous loci demonstrated GEI as defined by having opposite effects in glucose and ethanol. These corresponded to polymorphisms that influence *trans*-factors. Furthermore, genes affected by GEI were nearly twice as numerous as those with genetic-only effects [26] and were enriched for loci exhibiting the hallmarks of high ED genes. However, an important factor that was not considered in these studies, is transient changes in gene expression as cells transition between environments, a common ecological scenario. Indeed, a recent study in *S. cerevisiae* strains of different genetic backgrounds responding to heat shock found that half of the transcripts only showed GEI effects during the transition between environments but not in acclimated cells [51]. Transcripts with persistent GEI were enriched for classic high ED genes as in previous studies, whereas those displaying transient GEI were enriched for essential genes [51].

The emerging field of population genomics represents further fertile ground for distinguishing the role of *cis* and *trans* variation within a species. Two recent studies determined the whole genome sequence of *S. cerevisiae* and *S. paradoxus* strains from a large variety of sources and locations [52,53]. This repository of natural variation represents a powerful tool to dissect the genetic basis of regulatory variation underlying natural phenotypic diversity. For example, a recent study [54] has shown that variation in sporulation efficiency between a strain isolated from an oak tree and a vineyard strain is due to allelic variation in the genes encoding the transcription factors ImeI, Rme1 and Rsf1. In this case, the interactions between alleles (epistasis) affecting transcription and hence sporulation efficiency were non-additive and complex.

Hybrids between closely related species offer a complementary approach to quantify the relative contributions of *cis*- and *trans*-factors to expression divergence. Such studies compare the expression of orthologs in each species alone to that measured for each ortholog (‘allele’) when they share a common cell in a hybrid. A recent study using an inter-species hybrid between *S. cerevisiae* and *S. paradoxus* found that *cis* effects dominate variation in gene expression [28]. This is consistent with previous reports in flies [4] and mammals [5], and is in contrast to the larger contribution of *trans*-factors to intra-species variation [23,30,47]. In contrast, *trans* effects were condition-specific, primarily attributable to differential responses to sensory signals and not to variation in direct transcriptional regulators. This observation is consistent with the prominence of *trans* effects in GEI studies in *S. cerevisiae* strains [26,50, 51].

What may explain the prevalence of *cis* variation between species and the high levels of intra-specific *trans* variation? One clue comes from a recent study [55] showing that *trans* variation is more subject to dominance effects than *cis* variation. Thus, variation due to *trans*-regulatory alleles is biased toward greater deviation from an additive contribution when affecting a complex phenotype than is *cis* variation. On shorter timescales, the pervasive pleiotropic effects and the much higher rate by which *trans*-variation is produced can account for the gene expression variation within populations. Over longer timescales, purifying selection could purge *trans*-regulatory variation. Conversely, although *cis*-variation is produced at a slower rate, positive selection may act more efficiently to fix *cis* changes due to their higher additivity and weaker pleiotropic effects. Notably, population genetic modeling of the evolutionary forces affecting the pattern of variation for the *cis*-regulatory QTL in the RM and BY cross [56] concluded that purifying selection against mildly deleterious alleles is the dominant force

governing *cis*-regulatory evolution and found evidence that positive selection has played a role in the evolution of major *trans* acting QTLs.

The adaptive significance of expression divergence

What is the adaptive significance of regulatory divergence? How to distinguish between adaptive changes and regulatory neutrality and drift? In some cases, regulatory changes are clearly coupled to other adaptive changes in lifestyle. For example, studies comparing regulatory modules between *C. albicans* and *S. cerevisiae* showed how a specific loss of an ancestral *cis*-regulatory element from the promoters of genes in the mitochondrial ribosomal proteins has changed their chromatin organization and decoupled their regulation from cell growth in respiro-fermentative species that no longer relied on respiration for growth in high glucose [15].

In many other cases, it is unclear whether the regulatory change is adaptive or neutral. For example, Tsong *et al*. compared the mating transcriptional modules in *C. albicans*, *K. lactis* and *S. cerevisiae*, and reconstructed a series of *cis*- and *trans*- regulatory changes that have resulted in a transition from an activator-based control to a repressor-based regulation of the mating response [57]. Since the overall regulatory logic was unchanged by this transition, one possibility is that it is a result of neutral “regulatory drift” rather than an adaptive change [36].

Experimental evolution and the adaptive role of expression divergence

Experimental evolutionary approaches have the potential to disentangle adaptive changes and from cases where alternate mechanisms have evolved to perform the same function. Most evolutionary studies infer the trajectory of evolution from sampling variation in extant populations and thus are limited in their ability to address evolutionary dynamics. In contrast, experimental evolution studies [58–60] can observe adaptation in real time and under known selective pressures at short time scales. Recent advances in genomic technologies, in particular the advent of rapid and cheap re-sequencing [61] allow us to efficiently identify all the genetic changes that have occurred in evolved lines subjected to different selective pressures. By quantifying the effects of one or more genetic changes on growth and other measures of fitness and function, we can distinguish adaptive versus neutral regulatory architectures. Finally, by analyzing multiple lines evolved in parallel under the same selective pressure we can assess the range of possible evolutionary trajectories. Notably, the exceptional genetic tractability of *S. cerevisiae* has rendered it an excellent model for experimental evolution studies [60,62].

This power has been recently demonstrated in a study of chemostat cultures subjected to either glucose, sulfate or phosphate limitation [63] for ~200 generations. The genetic variation in each of the evolved strains was assessed using tiling microarrays. In addition to point mutations, the spectrum of genetic alterations included frequent genomic amplifications and rearrangements as well as retrotransposition events. Retrospective analysis of the observed frequencies of mutations over the course of evolution in the chemostat suggested that these mutations originated in the batch phase growth of the cultures prior to chemostat inoculation. When comparing multiple strains evolved under each of the selective pressures, Gresham *et al*. observed several distinct genotypic and phenotypic evolutionary trajectories in the glucose- or phosphate-limited environments, whereas a single trajectory dominated in the sulfate-limited populations. In all cases, adaptation to nutrient-limitation results in massive remodeling of global gene expression. Importantly, even distinct genetic changes often led to convergent mRNA profiles. For example, in several populations that were independently evolved in glucose-limited conditions *HXT* genes encoding high affinity glucose transporters were amplified, while in another population a retrotranposition event within the *MTH1* gene (a negative regulator of glucose sensing) is the likely cause of the observed increase in the

expression of several *HXT* genes. Furthermore, a number of other mutated loci in clones evolved in glucose-limited conditions have known roles as key regulators in carbon metabolism, suggesting a major role for *trans*-regulation. This is consistent with the observation that *trans* effects are predominant in intra-specific expression divergence suggesting this may be a major evolutionary strategy for adaptation on shorter timescales.

Future prospects: the role of empirical studies

While novel genomic technologies will continue to fuel the great advances that have been made to the study of the evolution of gene regulation, the relative roles of selective forces driving divergence versus neutral drift remain largely theoretical. Purifying selection can be effectively invoked for the conservation of *cis*-regulatory elements in closely related species, and for the low-ED of genes involved in general growth processes. However, it is unclear how much of the increased divergence of highED genes is due to direct positive selection, and how much is a by-product of regulatory mechanisms required for environmental responsiveness. Understanding how changes in regulatory control relate to upstream changes in signal sensing and processing may shed light on this question, and recent studies on the evolution of signal transduction in *Ascomycota* [64,65] will be instrumental in this endeavor.

Most studies to date focused on the effect of *trans* and *cis* changes on transcription initiation, but divergence in mRNA levels can also be affected by changes in *cis*- and *trans*- factors that impact transcription elongation or termination and mRNA processing and stability. Furthermore, a recent comparative study has discovered that RNA interference (an RNA-silencing pathway), while absent in *S. cerevisiae*, is present in many other budding yeasts such as *S. castellii*, *C. albicans* and *K. polysporus* [66]. Finally, gene expression is influenced by processes that are downstream of transcription such as nuclear export, translation initiation, elongation and termination, and protein degradation. For example, a recent study found that a genetic change in Mkt1, a protein that affects P-body sequestration of mRNAs encoding mitochondrial proteins, is responsible for a major change in expression between a wild and a lab strain of *S. cerevisiae* [42]. Furthermore, a proteomics study in the segregants from the BY X RM cross showed that loci influencing protein abundance differed from those that affected transcripts levels highlighting the importance of direct analysis of the proteome [67]. Evolutionary studies that address these additional levels of regulation are scarce [64]. However, new technologies (*e.g.*, Ingolia *et. al.* [68]) are emerging that allow genome-wide investigation of translational control and proteomic profiling hold promise for understanding the contribution of post-transcriptional regulation to evolutionary divergence in gene expression.

The foremost advantage of *Ascomycota* for studies of evolution of gene expression is in the facility of experiments in both model- and non-model organisms. We discern three major trends towards empirical studies of regulatory evolution. Comparative functional genomics follows in the footsteps of sequencing studies by measuring the transcriptional responses and molecular mechanisms across a set of extant species in a phylogeny, and uses these measurements and phenotypic differences to infer the history of gene regulation. Forward evolution studies focus on the immediate impact of selection, by following traces of strains collected along an experiment, and use sequencing, genetics and molecular profiles to infer regulatory evolution in ‘real time’. Finally, engineered strains and hybrids allow us to test evolutionary hypotheses and quantify the contribution of distinct factors to regulatory changes. These range in increasing evolutionary distance from eQTL mapping in segregants from crosses of distinct strains of the same species [23,42,47,48], to hybrids between species [28], to engineered strains swapping molecular elements from distant species for their endogenous orthologs or introducing ‘random’ engineered variation [69]. Together with elaborate phenotyping these should allow us to decipher the functional and adaptive implications of regulatory variation.

Acknowledgments

We thank Audrey P. Gasch, Dana J. Wohlbach, the anonymous reviewers and members of the Regev lab for fruitful discussions. AR was supported by the Howard Hughes Medical Institute, the Human Frontiers Science Program, an NIH PIONEER award, a Career Award at the Scientific Interface from the Burroughs Wellcome Fund, and the Sloan Foundation.

Abbreviations

TF	Transcription factor
RP	Ribosomal proteins
WGD	Whole genome duplication

References

1. McAdams HH, Srinivasan B, Arkin AP. The evolution of genetic regulatory systems in bacteria. Nat Rev Genet 2004;5:169–178. [PubMed: 14970819]
2. Gasch AP, Moses AM, Chiang DY, Fraser HB, Berardini M, Eisen MB. Conservation and evolution of cis-regulatory systems in ascomycete fungi. PLoS Biol 2004;2:e398. [PubMed: 15534694]
3. Prud'homme B, Gompel N, Carroll SB. Emerging principles of regulatory evolution. Proc Natl Acad Sci USA 104 Suppl 2007;1:8605–12.
4. Wittkopp P, Haerum B, Clark A. Regulatory changes underlying expression differences within and between *Drosophila* species. Nat Genet 2008;40:346–350. [PubMed: 18278046]
5. Khaitovich P, Enard W, Lachmann M, Pääbo S. Evolution of primate gene expression. Nat Rev Genet 2006;7:693–702. [PubMed: 16921347]
6. Wapinski I, Pfeffer A, Friedman N, Regev A. Natural history and evolutionary principles of gene duplication in fungi. Nature 2007;449:54–61. [PubMed: 17805289]
7. Newman JR, Ghaemmaghami S, Ihmels J, Breslow DK, Noble M, DeRisi JL, Weissman JS. Single-cell proteomic analysis of S. cerevisiae reveals the architecture of biological noise. Nature 2006:840–6. [PubMed: 16699522]
8. Tong AH, et al. Global mapping of the yeast genetic interaction network. Science 2004:808–13. [PubMed: 14764870]
9. Ito T, Chiba T, Ozawa R, Yoshida M, Hattori M, Sakaki Y. A comprehensive two-hybrid analysis to explore the yeast protein interactome. Proc Natl Acad Sci U S A 2001;98:4569–74. [PubMed: 11283351]
10. Kellis M, Birren BW, Lander ES. Proof and evolutionary analysis of ancient genome duplication in the yeast *Saccharomyces cerevisiae*. Nature 2004;428:617–24. [PubMed: 15004568]
11. Wolfe KH, Shields DC. Molecular evidence for an ancient duplication of the entire yeast genome. Nature 1997;387:708–13. [PubMed: 9192896]
12. Piskur J, Rozpedowska E, Polakova S, Merico A, Compagno C. How did Saccharomyces evolve to become a good brewer? Trends in Genetics 2006:183–186. [PubMed: 16499989]
13. Conant G, Wolfe K. Increased glycolytic flux as an outcome of whole-genome duplication in yeast. Mol Syst Biol 2007;3:12.
14. Kurtzman, CP. The Yeasts a Taxonomic Study. Elsevier; 2000.
15. Ihmels J, Bergmann S, Gerami-Nejad M, Yanai I, McClellan M, Berman J, Barkai N. Rewiring of the yeast transcriptional network through the evolution of motif usage. Science 2005;309:938–40. [PubMed: 16081737]
16. Field, Y., et al. Nat Genet. 2009. Gene expression divergence in yeast is coupled to evolution of DNA-encoded nucleosome organization; p. 438-45.
17. Tirosh I, Weinberger A, Carmi M, Barkai N. A genetic signature of interspecies variations in gene expression. Nat Genet 2006;38:830–4. [PubMed: 16783381]

18. Lelandais G, Tanty V, Geneix C, Etchebest C, Jacq C, Devaux F. Genome adaptation to chemical stress: clues from comparative transcriptomics in *Saccharomyces cerevisiae* and *Candida glabrata*. Genome Biol 2008;9:R164. [PubMed: 19025642]
19. Suleau A, Gourdon P, Reitz-Ausseur J, Casaregola S. Transcriptomic analysis of extensive changes in metabolic regulation in *Kluyveromyces lactis* strains. Eukaryot Cell 2006;5:1360–1370. [PubMed: 16896219]
20. Andersen MR, Vongsangnak W, Panagiotou G, Salazar MP, Lehmann L, Nielsen J. A trispecies *Aspergillus* microarray: comparative transcriptomics of three *Aspergillus* species. Proc Natl Acad Sci USA 2008;105:4387–92. [PubMed: 18332432]
21. Tian C, Kasuga T, Sachs MS, Glass NL. Transcriptional profiling of cross pathway control in *Neurospora crassa* and comparative analysis of the Gcn4 and CPC1 regulons. Eukaryot Cell 2007;6:1018–29. [PubMed: 17449655]
22. Townsend J. Population genetic variation in genome-wide gene expression. Mol Biol Evol 2003;20:955–963. [PubMed: 12716989]
23. Brem RB, Yvert G, Clinton R, Kruglyak L. Genetic dissection of transcriptional regulation in budding yeast. Science 2002;296:752–5. [PubMed: 11923494]
24. Kvitek DJ, Will JL, Gasch AP. Variations in stress sensitivity and genomic expression in diverse *S. cerevisiae* isolates. PLoS Genet 2008;4:e1000223. [PubMed: 18927628]
25. Landry CR, Lemos B, Rifkin SA, Dickinson WJ, Hartl DL. Genetic properties influencing the evolvability of gene expression. Science 2007;317:118–21. [PubMed: 17525304]
26. Landry CR, Oh J, Hartl DL, Cavalieri D. Genome-wide scan reveals that genetic variation for transcriptional plasticity in yeast is biased towards multi-copy and dispensable genes. Gene 2006;366:343–51. [PubMed: 16427747]
27. Tirosh I, Barkai N. Two strategies for gene regulation by promoter nucleosomes. Genome Res 2008;18:1084–1091. [PubMed: 18448704]
28. Tirosh I, Reikhav S, Levy AA, Barkai N. A yeast hybrid provides insight into the evolution of gene expression regulation. Science 2009;324:659–62. [PubMed: 19407207]
29. Raser J. Control of sochasticity in eukaryotic gene expression. Science 2004;304:1811–1814. [PubMed: 15166317]
30. Wang D, et al. Expression evolution in yeast genes of single-input modules is mainly due to changes in trans-acting factors. Genome Res 2007;17:1161–9. [PubMed: 17615293]
31. Kellis M, Patterson N, Endrizzi M, Birren B, Lander ES. Sequencing and comparison of yeast species to identify genes and regulatory elements. Nature 2003;423:241–54. [PubMed: 12748633]
32. Cliften P. Finding functional features in *Saccharomyces* genomes by phylogenetic footprinting. Science 2003;301:71–76. [PubMed: 12775844]
33. Doniger SW, Kim HS, Swain D, Corcuera D, Williams M, Yang SP, Fay JC. A catalog of neutral and deleterious polymorphism in yeast. PLoS Genet 2008;4:e1000183. [PubMed: 18769710]
34. Doniger SW, Fay JC. Frequent gain and loss of functional transcription factor binding sites. PLoS Comput Biol 2007;3:e99. [PubMed: 17530920]
35. Borneman AR, et al. Divergence of transcription factor binding sites across related yeast species. Science 2007;317:815–9. [PubMed: 17690298]
36. Tuch BB, Li H, Johnson AD. Evolution of eukaryotic transcription circuits. Science 2008;319:1797–9. [PubMed: 18369141]
37. Vinces M, Legendre M, Caldara M, Hagihara M, Verstrepen K. Unstable tandem repeats in promoters confer transcriptional evolvability. Science 2009;324:1213–1216. [PubMed: 19478187]
38. Tanay A. Extensive low-affinity transcriptional interactions in the yeast genome. Genome Res 2006;16:962–72. [PubMed: 16809671]
39. Tuch BB, Galgoczy DJ, Hernday AD, Li H, Johnson AD. The evolution of combinatorial gene regulation in fungi. PLoS Biol 2008;6:e38. [PubMed: 18303948]
40. Tan K, Feizi H, Luo C, Fan SH, Ravasi T, Ideker TG. A systems approach to delineate functions of paralogous transcription factors: role of the Yap family in the DNA damage response. Proc Natl Acad Sci USA 2008;105:2934–9. [PubMed: 18287073]

41. Choi JK, Kim YJ. Epigenetic regulation and the variability of gene expression. Nat Genet 2008;40:141–7. [PubMed: 18227874]
42. Lee SI, Pe'er D, Dudley AM, Church GM, Koller D. Identifying regulatory mechanisms using individual variation reveals key role for chromatin modification. Proc Natl Acad Sci U S A 2006;103:14062–7. [PubMed: 16968785]
43. Ihmels J, Friedlander G, Bergmann S, Sarig O, Ziv Y, Barkai N. Revealing modular organization in the yeast transcriptional network. Nat Genet 2002;31:370–7. [PubMed: 12134151]
44. Segal E, Shapira M, Regev A, Pe'er D, Botstein D, Koller D, Friedman N. Module networks: Identifying regulatory modules and their condition-specific regulators from gene expression data. Nat Genet 2003;34:166–76. [PubMed: 12740579]
45. Tanay A, Sharan R, Kupiec M, Shamir R. Revealing modularity and organization in the yeast molecular network by integrated analysis of highly heterogeneous genomewide data. Proc Natl Acad Sci USA 2004;101:2981–6. [PubMed: 14973197]
46. Wolbach DJ, Thompson DA, Gasch AP, Regev A. From elements to modules: regulatory evolution in Ascomycota fungi. Current Opinion in Genetics & Development. 2009 In Press.
47. Yvert G, Brem RB, Whittle J, Akey JM, Foss E, Smith EN, Mackelprang R, Kruglyak L. Trans-acting regulatory variation in *Saccharomyces cerevisiae* and the role of transcription factors. Nat Genet 2003;35:57–64. [PubMed: 12897782]
48. Brem RB, Storey JD, Whittle J, Kruglyak L. Genetic interactions between polymorphisms that affect gene expression in yeast. Nature 2005:701–703. [PubMed: 16079846]
49. Ronald J, Brem RB, Whittle J, Kruglyak L. Local regulatory variation in *Saccharomyces cerevisiae*. PLoS Genet 2005;1:e25. [PubMed: 16121257]
50. Smith EN, Kruglyak L. Gene-environment interaction in yeast gene expression. PLoS Biol 2008;6:e83. [PubMed: 18416601]
51. Eng KH, Kvitek DJ, Keles S, Gasch AP. Transient genotype-environment interactions following environmental shock provide a source of expression variation for essential genes. 2009 submitted.
52. Liti, G., et al. Nature. 2009. Population genomics of domestic and wild yeasts; p. 337-341.
53. Schacherer J, Shapiro J, Ruderfer D, Kruglyak L. Comprehensive polymorphism survey elucidates population structure of Saccharomyces cerevisiae. Nature 2009:342–345. [PubMed: 19212320]
54. Gerke J, Lorenz K, Cohen B. Genetic interactions between transcription factors cause natural variation in yeast. Science 2009;323:498–501. [PubMed: 19164747]
55. Lemos B, Araripe LO, Fontanillas P, Hartl DL. Dominance and the evolutionary accumulation of cis- and trans-effects on gene expression. Proc Natl Acad Sci U S A 2008;105:14471–6. [PubMed: 18791071]
56. Ronald J, Akey JM. The evolution of gene expression QTL in Saccharomyces cerevisiae. PLoS One. 2007
57. Tsong AE, Tuch BB, Li H, Johnson AD. Evolution of alternative transcriptional circuits with identical logic. Nature 2006;443:415–20. [PubMed: 17006507]
58. Holder KK, Bull JJ. Profiles of adaptation in two similar viruses. Genetics 2001;159:1393–404. [PubMed: 11779783]
59. Herring CD, et al. Comparative genome sequencing of Escherichia coli allows observation of bacterial evolution on a laboratory timescale. Nat Genet 2006;38:1406–12. [PubMed: 17086184]
60. Replansky T, Koufopanou V, Greig D, Bell G. Saccharomyces sensu stricto as a model system for evolution and ecology. Trends in Ecology & Evolution 2008:494–501. [PubMed: 18656281]
61. Medini D, Serruto D, Parkhill J, Relman DA, Donati C, Moxon R, Falkow S, Rappuoli R. Microbiology in the post-genomic era. Nat Rev Microbiol 2008;6:419–30. [PubMed: 18475305]
62. Kao K, Sherlock G. Molecular characterization of clonal interference during adaptive evolution in asexual populations of *Saccharomyces cerevisiae*. Nat Genet 2008;40:1499–1504. [PubMed: 19029899]
63. Gresham D, et al. The repertoire and dynamics of evolutionary adaptations to controlled nutrient-limited environments in yeast. PLoS Genet 2008;4:e1000303. [PubMed: 19079573]

64. Beltrao P, Trinidad JC, Fiedler D, Roguev A, Lim WA, Shokat KM, Burlingame AL, Krogan NJ. Evolution of phosphoregulation: comparison of phosphorylation patterns across yeast species. PLoS Biol 2009;7:e1000134. [PubMed: 19547744]

65. Mody A, Weiner J, Ramanathan S. Modularity of MAP kinases allows deformation of their signalling pathways. Nat Cell Biol 2009;11:484–91. [PubMed: 19295513]

66. Drinnenberg IA, Weinberg DE, Xie KT. RNAi in Budding Yeast. Science. 2009

67. Foss EJ, Radulovic D, Shaffer SA. Genetic basis of proteome variation in yeast. Nat Genet. 2007

68. Ingolia NT, Ghaemmaghami S, Newman JRS. Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome …. Science. 2009

69. Wang HH, Isaacs FJ, Carr PA, Sun ZZ, Xu G, Forest CR, Church GM. Programming cells by multiplex genome engineering and accelerated evolution. Nature. 2009

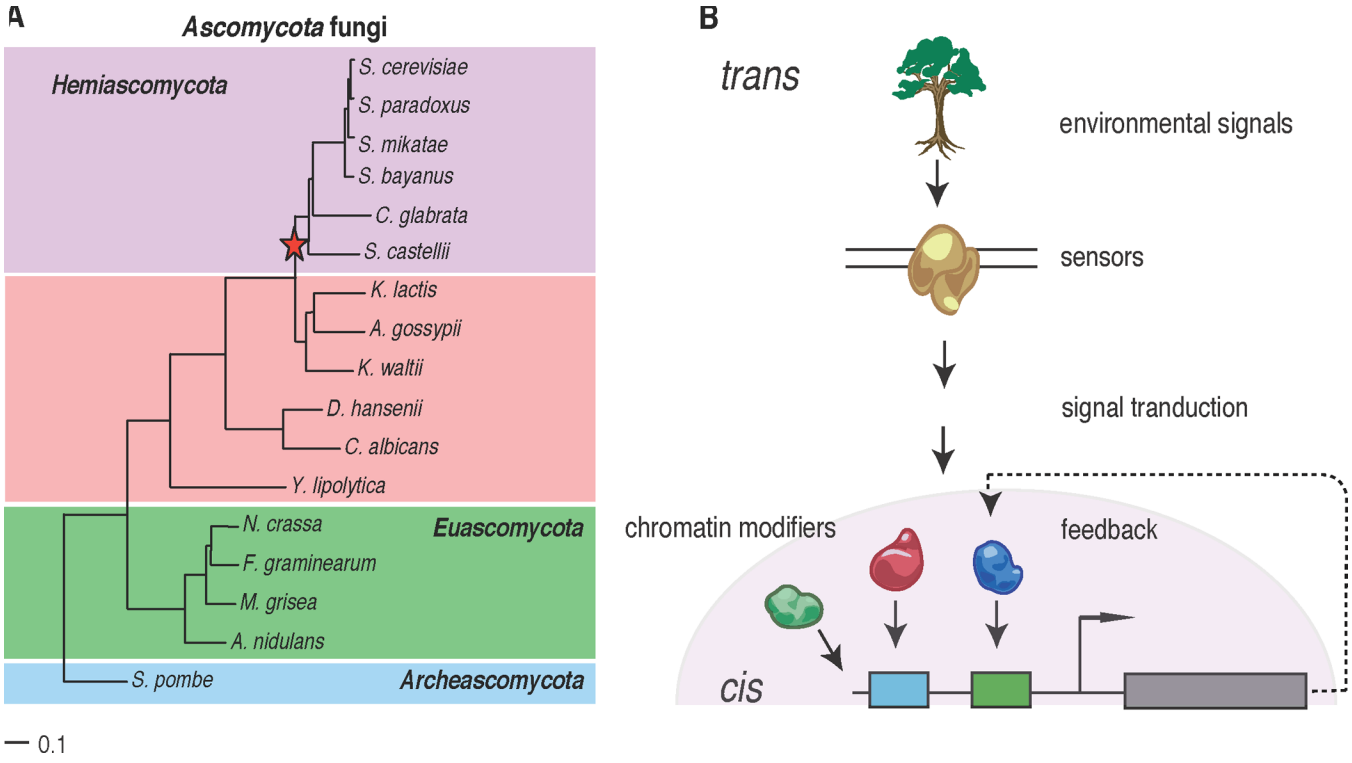

Figure 1.

A. *Ascomycota* fungi. A phylogenetic tree showing sequenced species from the major clades of the *Ascomycota* fungi. Red star – WGD. Tree is drawn to scale and adapted from [6]. **B. Factors affecting regulatory evolution**. *trans* factors, including differential interpretation of environmental signals by sensory and signaling proteins, chromatin modifiers (green ovoid), transcription factors (red and blue ovoids), as well as *cis* regulatory elements (boxes) affect gene expression and drive regulatory evolution. Figure 1B adapted from [46].