Gene duplication and the evolution of ribosomal protein gene regulation in yeast

Ilan Wapinski[a,b,1], Jenna Pfiffner[b], Courtney French[b], Amanda Socha[b], Dawn Anne Thompson[b], and Aviv Regev[b,c,1]

[a]Department of Systems Biology, Harvard Medical School, Boston, MA 02115; [b]Broad Institute of MIT and Harvard and [c]Howard Hughes Medical Institute and Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02142

Edited* by Eric S. Lander, Broad Institute of MIT and Harvard, Cambridge, MA, and approved February 1, 2010 (received for review October 15, 2009)

Coexpression of genes within a functional module can be conserved at great evolutionary distances, whereas the associated regulatory mechanisms can substantially diverge. For example, ribosomal protein (RP) genes are tightly coexpressed in *Saccharomyces cerevisiae*, but the *cis* and *trans* factors associated with them are surprisingly diverged across *Ascomycota* fungi. Little is known, however, about the functional impact of such changes on actual expression levels or about the selective pressures that affect them. Here, we address this question in the context of the evolution of the regulation of RP gene expression by using a comparative genomics approach together with cross-species functional assays. We show that an activator (Ifh1) and a repressor (Crf1) that control RP gene regulation in normal and stress conditions in *S. cerevisiae* are derived from the duplication and subsequent specialization of a single ancestral protein. We provide evidence that this regulatory innovation coincides with the duplication of RP genes in a whole-genome duplication (WGD) event and may have been important for tighter control of higher levels of RP transcripts. We find that subsequent loss of the derived repressor led to the loss of a stress-dependent repression of RPs in the fungal pathogen *Candida glabrata*. Our comparative computational and experimental approach shows how gene duplication can constrain and drive regulatory evolution and provides a general strategy for reconstructing the evolutionary trajectory of gene regulation across species.

stress response | comparative functional genomics | regulatory modules | expression profiling

The coordinated expression of modules of functionally related genes, such as ribosomal proteins or oxidative phosphorylation enzymes, is often conserved at great evolutionary distances (1). This is consistent with a selective pressure to conserve coordinated transcript levels to maintain functional cellular modules. Recent studies have shown that the regulatory mechanisms controlling conserved modules can diverge, most notably by switching from one regulatory system to another while preserving coregulation (1, 2). However, because previous studies have typically relied on functional expression data from only one or two species, it is unknown whether these changes in regulatory mechanisms affect the expression levels of a module's genes at all or whether both coexpression and expression levels are conserved.

A prominent example of a conserved regulatory module is the ribosomal protein (RP) module. Genes encoding RPs are tightly coexpressed in organisms from bacteria to humans (3, 4), consistent with a selective pressure to conserve coordinated transcript levels to maintain a stoichoimetric balance in ribosome assembly. The transcription factors controlling RP gene expression have changed several times since the last common ancestor of the *Ascomycota* fungi, which span *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe* (4–6). These dramatic changes include the loss of the ancestral regulator Tbf1, the emergence of Rap1 as a key activator among the *Hemiascomycota* (4, 5), as well as the addition of *Mcm1* as a regulator in *Kluyveromyces lactis* (7). This phenomenon of regulatory substitutions, first identified in the RP module (4, 7), has since been recognized as a general feature of module evolution in *Ascomycota*, suggesting that the RP module can serve as a general model for regulatory evolution.

It remains unclear whether this regulatory overhaul has affected RP gene expression. Under normal growth conditions, RP mRNA transcripts constitute a large fraction (>30%) of the total mRNA in *S. cerevisiae* cells (8, 9) (Fig. S1). These genes have been shown to be strongly corepressed in *S. cerevisiae* under environmental stress or nutrient limitation conditions (10, 11). In contrast, a previous study in *Candida albicans* failed to show similar repression under environmental stress (12), suggesting that RP expression levels may have diverged between species. However, all of the RP regulatory circuits characterized to date have a similar functional organization with some components that are constitutively localized on the RP promoter (Tbf1 in *C. albicans* and Rap1 in *S. cerevisiae*; both are also associated with telomere maintenance), and some that are regulated by nutrient- and environmental-response pathways (e.g., Ifh1 and Sfp1) (6, 13–16).

Here, we use a combined computational and experimental strategy to trace the evolution of gene regulation within the RP module. We find that an activator and a repressor that are known to control RP gene regulation in normal and stress conditions in *S. cerevisiae* are derived from the duplication of a single ancestral protein, followed by divergence of the repressor. This regulatory innovation coincides with the duplication of RP genes in a whole-genome duplication (WGD) event and may have been important for tighter control of higher levels of RP transcripts. To test this hypothesis, we used comparative expression profiling across six yeast species and found evidence for functional specialization between these regulators. We also show that subsequent loss of the derived repressor coincides with the loss of a stress-dependent repression of RPs in the fungal pathogen *Candida glabrata* and with the loss of duplicate RP genes in this species. Our study is an early example of a systematic, hypothesis-driven, functional phylogenomic study. This approach can be adopted for the study of gene regulation in a wide range of organisms and regulatory modules.

Results

Two Key Regulators of RP Gene Expression Are Paralogs Derived from a WGD Event. We hypothesized that changes in *trans* factors could conserve RP coexpression while diverging module expression. We therefore examined the *Ascomycota* gene orthologs (17) of all transcription factors previously implicated in RP gene regulation in *S. cerevisiae* (13–16, 18, 19) (Fig. 1*A* and Fig. S2). We discovered that two of these regulators, Ifh1 and Crf1 (Fig. 1*B*), are in fact paralogs that date to a WGD that occurred ~150 million years ago (20, 21). These genes encode two transcriptional cofactors that affect RP

Author contributions: I.W., D.A.T., and A.R. designed research; I.W., J.P., C.F., A.S., and D.A.T. performed research; I.W. and A.R. analyzed data; and I.W. and A.R. wrote the paper.

The authors declare no conflict of interest.

*This Direct Submission article had a prearranged editor.

Freely available online through the PNAS open access option.

Data deposition: Gene expression data is available at http://www.broadinstitute.org/~ilan/PNAS2010

[1]To whom correspondence may be addressed. E-mail: ilan_wapinski@hms.harvard.edu or aregev@broad.mit.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0911905107/DCSupplemental.

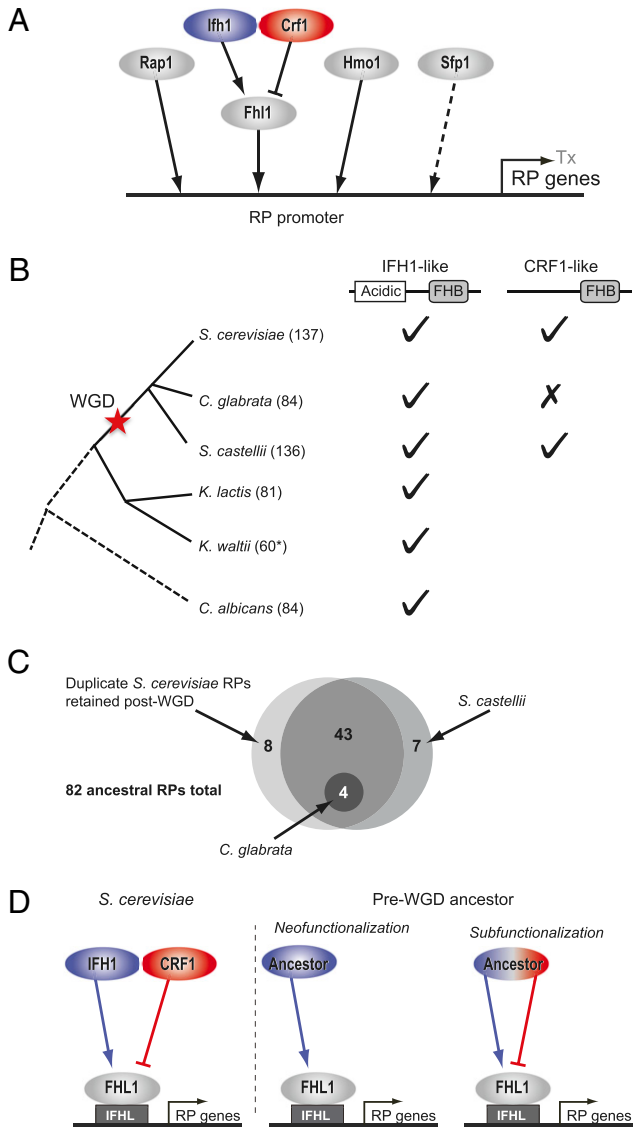

Fig. 1. The evolutionary history of RP genes and their regulators *IFH1* and *CRF1*. (*A*) Shown are the key regulators previously associated with RP gene promoters in *S. cerevisiae*. Regulators are shown with no specific orientation along the promoter. Rap1 directly binds to RP promoters (4) whereas Ifh1 and Crf1 interact with RP promoters via Fhl1 (13–15). It is still unclear whether Sfp1 and Hmo1 affect RP gene expression through direct binding (18, 19) or indirectly (6, 33). (*B*) *IFH1* and *CRF1* are paralogs originating from an ancestral WGD (20, 21, 23). (*Left*) A phylogenetic tree of *Hemiascomycota* fungi that diverged before and after a WGD event (red star), the number of RP genes found in each genome (in parentheses), and the presence (check) or absence (X mark) of an *IFH*-like or *CRF*-like gene in these species (*Right*). Although *IFH1* was retained in all lineages after the WGD (red star), *CRF1* was lost in *C. glabrata*, consistent with the pattern of paralogous RP retention. *IFH1* retains sequence features similar to those of its non post-WGD orthologs, and *CRF1* has lost an acidic N terminus region (white box) but has retained a conserved FHB domain (13) (gray box). (*C*) A Venn diagram of the RP genes retained in duplicate in each of the three post-WGD species. Nearly all paralogous RP copies were lost in *C. glabrata*, whereas *S. cerevisiae* and *S. castellii* have retained a significant portion of them. (*D*) (*Left*) In *S. cerevisiae*, *IFH1* (blue) induces RP gene expression whereas *CRF1* (red) is a stress-responsive repressor (13), and both interact with *FHL1* (gray). (*Right*) Hypothetical roles of the *IFH1*/*CRF1* ancestor as solely an activator (blue, suggesting neofunctionalization of *CRF1*) or as both an activator and a repressor (blue/red, consistent with subfunctionalization).

gene expression in *S. cerevisiae* by condition-dependent binding to the *Fhl1* transcription factor at RP gene promoters (13) (Fig. 1*D*, *Left*). Previous studies have shown that *Ifh1* binds *Fhl1* under rich growth conditions and induces RP expression (14–16), whereas stress-dependent binding by *Crf1* represses RP expression within some *S. cerevisiae* strains but not others (13, 16). (It is unknown which strain's Crf1-deletion phenotype is ancestral in *S. cerevisiae* and which is derived.)

We next compared the protein sequences of these two paralogs to that of their pre- and post-WGD orthologs and found that the pre-WGD orthologs are highly similar to the activator Ifh1. In contrast, the *S. cerevisiae* repressor Crf1 and its post-WGD orthologs all lack an ancestral acidic N-terminal domain (13) that is important for *trans*-activation (Fig. S3*C*). Furthermore, we found an elevated (~4.5-fold) amino acid substitution rate (21) in Crf1 compared to Ifh1 (Fig. S3 *A* and *B*). Taken together, these findings suggest that Ifh1's role as an inducer is ancestral, whereas Crf1's function as a repressor is derived following the WGD (Fig. 1*B*), which is likely associated with the loss of the acidic domain.

Evolutionary History of Ifh1/Crf1 Traces That of RP Genes. We found that Crf1 orthologs are present in most post-WGD species, such as the other sensu stricto *Saccharomyces* and *Saccharomyces castellii*, but that it was lost from the genome of *C. glabrata*, another post-WGD species (unrelated to the pre-WGD species *C. albicans*) (22) (Fig. 1*B* and Fig. S2*B*). Remarkably, duplicate copies of the RP genes themselves are present in the same species as the repressor Crf1 (17, 23); a large number (55, or 69%) of RP genes remain in duplicate copies in *S. cerevisiae* and *S. castellii*, but very few (4, or 5%) duplicate RPs were retained in *C. glabrata* (17, 21–23) (Fig. 1*C*). When present, paralogous RP genes are highly conserved in function, protein-coding sequence (21), and regulatory program (Table S1).

Taken together, our analysis is consistent with a potential divergence in the regulation of RP gene expression post-WGD. In this model, the retained paralogous RPs in post-WGD lineages result in higher total RP transcript content, which requires a more complex control to coordinate RP repression under stress because of their increased transcriptional burden (8). Crf1 can fulfill this role by competing with Ifh1 for Fhl1 binding, thus rapidly substituting a transcriptional activator with a repressor. We further hypothesized that Crf1's role as a repressor of RP genes in stress arose after the WGD but before the divergence between the *S. cerevisiae* and *S. castellii* lineages. If this hypothesis is correct, then we expect that, in stress conditions, *S. cerevisiae* and *S. castellii* RPs will be coordinately repressed, but *C. glabrata*'s RPs will not.

Comparative Expression Profiling in Three Post-WGD Species Shows That Loss of Crf1 Is Associated with Lack of RP Repression in *C. glabrata*. To test our hypothesis, we next compared the regulation of RP genes in stress conditions in three species from three post-WGD clades: *S. cerevisiae*, *C. glabrata*, and *S. castellii*. In each species, we used species-specific microarrays (*Materials and Methods*) to measure genome-wide mRNA expression responses under comparable growth and stress conditions (Fig. 2 and Figs. S4 and S5). We verified the presence of a stress response in each organism by three criteria: (1) change in growth rate upon treatment; (2) induction of known induced environmental stress response genes (Fig. S4*A*); and (3) repression of ribosome biogenesis (RiBi) genes (8, 11) (Fig. 2 and Fig. S4*B*). We found that heat shock resulted in the most consistent and prominent stress response across all species; for *C. glabrata*, we tested shock from 30 °C to both 37 °C and 42 °C, since it is a commensal human pathogen adapted to 37° C. In normal growth conditions, total RP gene transcripts (*Materials and Methods*) contributed a higher fraction (~35–40%) in *S. cerevisiae* and *S. castellii*, than in *C. glabrata* (~25–30%), as expected given the presence of paralogous RP genes in these species (Fig. S1).

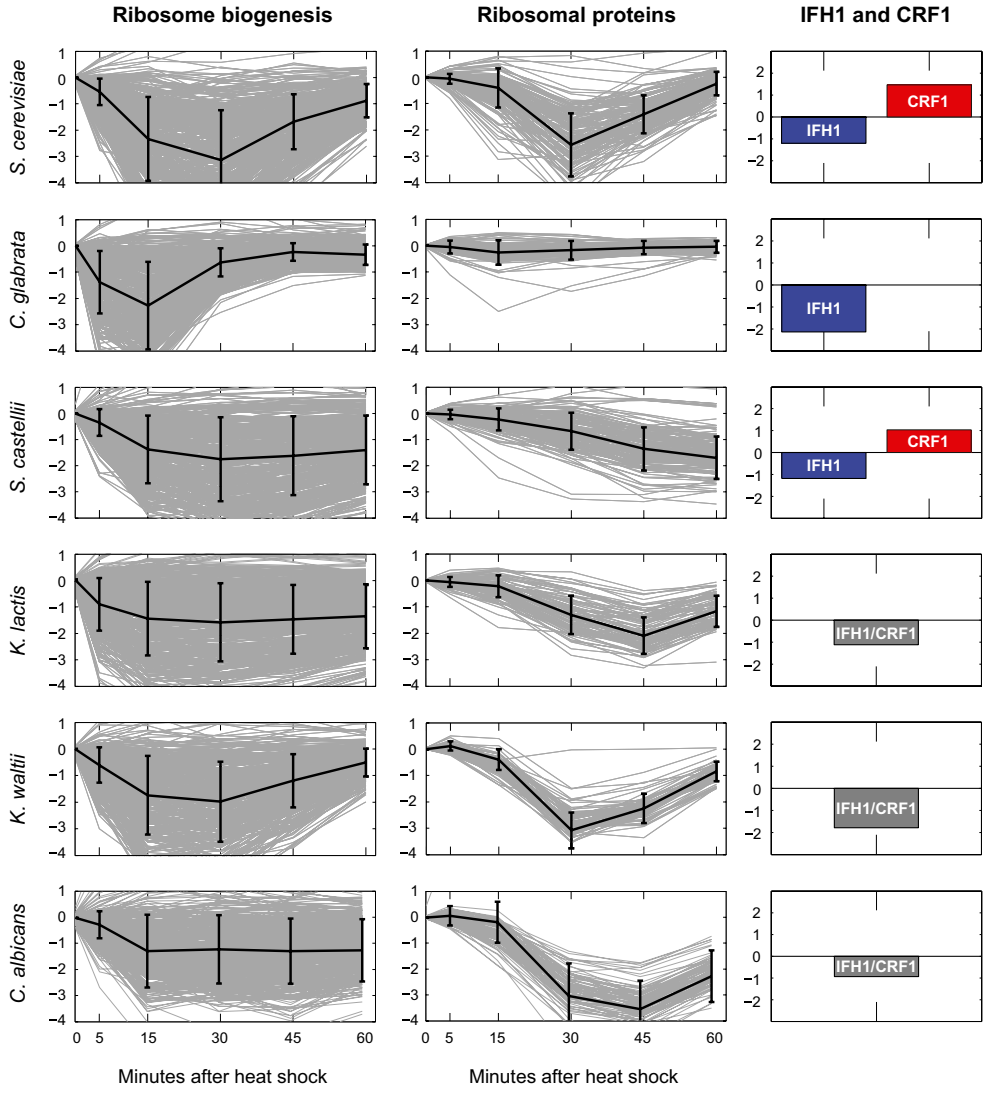

Fig. 2. RP and *IFH1/CRF1* expression in all species. Shown are the $\log_2$ fold-changes in expression levels of ribosomal protein (RP, *Center*) and ribosome biogenesis (RiBi, *Left*) genes in each of the species at each time point, relative to time point 0. Gray, individual genes; black, mean and standard deviation of the whole set. (*Right*) The change in expression levels in each species' copy of *IFH1* (blue) and *CRF1* (red) or their preduplication orthologs (gray) during the peak of each species' response to stress (peak times for species are from top to bottom: 30, 15, 30, 45, 30, and 30 min).

We found that RP mRNA levels are markedly repressed upon stress in both *S. cerevisiae* and *S. castellii*, but are unaffected by any stress condition tested in *C. glabrata* (Fig. 2 and Fig. S5). The lack of RP repression in *C. glabrata* is not merely a consequence of low basal RP expression levels under normal growth conditions (Fig. S1*B*). Furthermore, under stress, both *S. cerevisiae* and *S. castellii* repress *IFH1* expression and induce *CRF1* expression (Fig. 2). These findings are consistent with the model that Crf1 became a repressor of RP gene expression following the WGD but before the divergence between *S. cerevisiae* and *S. castellii*. Loss of this repressor in *C. glabrata* resulted in its inability to repress RP transcription levels in response to stress.

Expression Profiling of Pre-WGD Species Suggests a Model for Ifh1/Crf1 Evolution. What was the evolutionary trajectory of these changes? One possibility is *neofunctionalization*: the preduplication ancestor of Ifh1/Crf1 performed the role of an activator, but, following duplication, Crf1 lost its *trans*-activating domain and function, resulting in a repressor (Fig. 1*D*). An alternative is *subfunctionalization* either via complementary degenerative mutations or by enabling an escape from adaptive conflict within a multifunctional protein (24–27). In this scenario, the preduplication ancestor had both activator and repressor functions, which were separately assumed by the Ifh1 and Crf1 paralogs after their duplication and before the divergence of the post-WGD clades (Fig. 3*B*). This could have occurred either neutrally or due to selective pressure, e.g., in response to the duplication of RP genes. The fact that the pre-WGD species *C. albicans* was previously reported to lack substantial RP repression under stress (12) led us to hypothesize that the ancestral protein was only an activator.

To distinguish between the neofunctionalization and subfunctionalization models, we measured genome-wide mRNA expression profiles during stress in three pre-WGD species: *K. lactis*, *Kluyveromyces waltii*, and *C. albicans*. We found a coordinated stress-dependent repression of RP gene expression in all three species (Fig. 2). In *C. albicans*, this repression was not observed at 37 °C heat shock [consistent with a previous study (12)], but was very robust at 42 °C, consistent with the behavior in other species. Overall, these data strongly suggest that there is a conserved transcriptional pro-

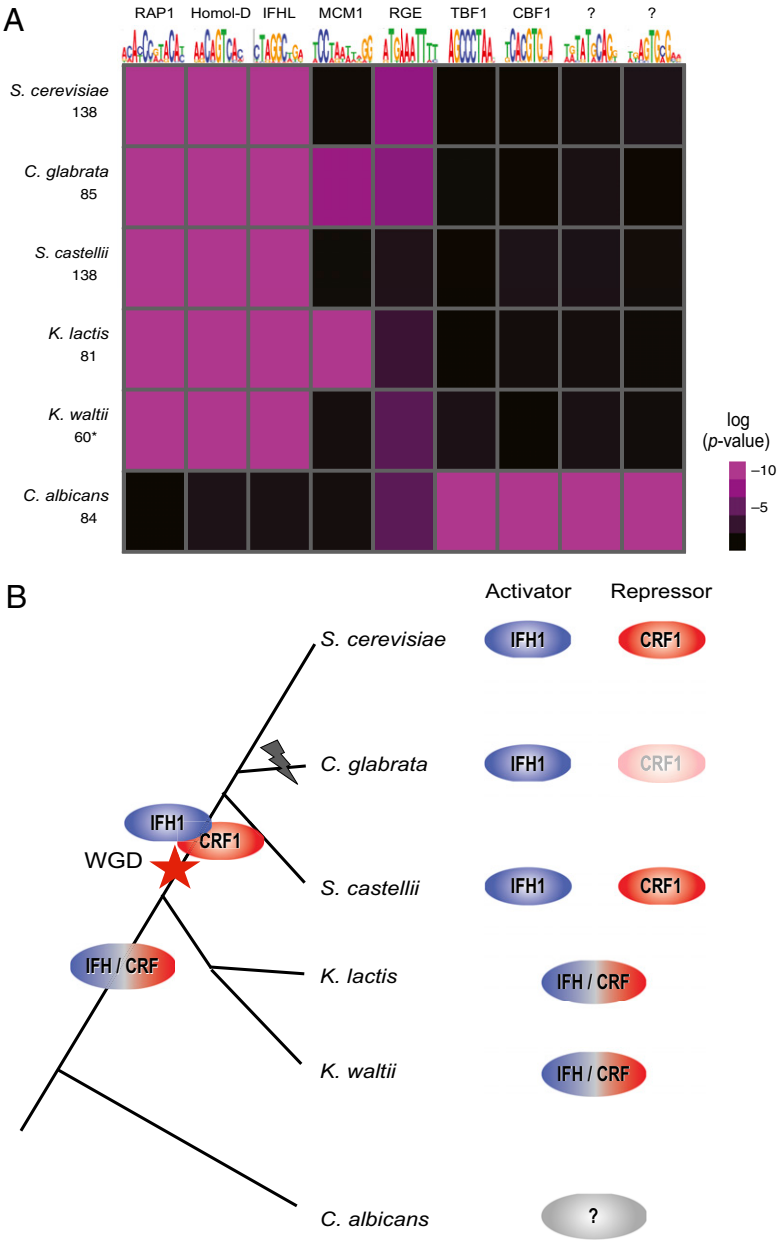

Fig. 3. *cis*- and *trans*-regulatory evolution of RP genes across pre- and post-WGD species. (*A*) Representative *cis*-regulatory motifs (columns) found to be enriched in RP promoters in at least one species (rows) and the enrichments of RP promoters associated with the motif in each species (dark purple, $P < 10^{-10}$; black, $P = 1$). The total number of annotated RPs in each species is denoted next to the species name. (*B*) A model of an evolutionary trajectory of *IFH1/CRF1* functional roles. The pre-WGD ortholog of *IFH1/CRF1* performed dual roles, both inducing and repressing RP expression levels in a condition-dependent manner. After duplication, each paralog specialized to act as a separate activator (*IFH1*) and repressor (*CRF1*). The repressor was subsequently lost in the lineage leading to *C. glabrata*.

gram for RP repression that predates the WGD and supports a subfunctionalization model for the evolution of Crf1.

Cross-Species Analysis of *cis*-Regulatory Programs Supports Our Model of Evolution Through Changes in *Trans* Regulation. To further test our model of evolution through changes in *trans* factors and to explore the role of additional factors, we next examined the organization of *cis*-regulatory elements in RP gene promoters in each species. We first searched for overrepresented sequence motifs within the promoters of RP genes of these six species (28) (*Materials and Methods*). Our analysis shows that IFHL sites (4, 5), known to be bound by Fhl-Ifh1/Crf1 in *S. cerevisiae*, are significantly enriched in the RP gene promoters of both pre- and post-WGD species, including *K. lactis*, *K. waltii*, *S. castellii*, and *C. glabrata* (Fig. 3*A* and Fig. S6). Furthermore, these are colocated with the RAP1 binding site, suggestive of cooperative interactions [as previously reported (4)]. This supports the role of Ifh1 in regulating RP gene expression in each of these species. As previously shown, the outgroup species *C. albicans* has a different *cis*-regulatory organization, dominated by binding sites for Tbf1 and Cbf1 and lacking directly discernible IFHL sites (5). Because the Ifh1/Fhl1 complex is physically associated with RP gene promoters in *C. albicans* (6), there are two possibilities for this discrepancy: (1) Fhl1 and Ifh1 bind indirectly through the Tbf1 protein (6), or (2) the Fhl1 protein recognizes a variant site, similar to the Tbf1 consensus (1, 4). Both cases are consistent with Ifh1's role as a regulator of RP gene expression in *C. albicans*.

RP promoters in all species (except the outgroup *C. albicans*) were also enriched for three additional sites: RAP1, HomolD, and RGE (4, 5) (Fig. 3*A* and Fig. S6). It is not known which proteins, if any, bind to the HomolD element (4). The MCM1 element (7) was enriched in the *C. glabrata* and *K. lactis* promoters. Because *K. lactis* RPs are strongly repressed in stress, the MCM1 element alone cannot account for the lack of RP repression in *C. glabrata*. This overall conserved organization in species that have diverged before and after the WGD suggests that it is unlikely that the lack of RP repression in *C. glabrata* results from *cis*-regulatory changes, rather highlighting the role of *trans* changes in this evolutionary event.

Finally, we examined the possible role of Sfp1, another transcription factor that was suggested to impact both RP (18, 19) and RiBi (18, 29) gene expression under nutrient limitation in *S. cerevisiae*. A recent in vitro study showed that *S. cerevisiae* Sfp1 binds a motif highly similar to the RGE element (30). Sfp1 orthologs are present in all pre- and post-WGD species (with WGD paralogs retained in *C. glabrata* and *S. castellii*). Furthermore, we found that the RGE motif is enriched in RP gene promoters across all species in our study, including *C. glabrata* (Fig. 3*A* and Fig. S6). The presence of the RGE site and the Sfp1 proteins may explain why *C. glabrata* RPs are substantially repressed following glucose depletion in the diauxic shift (Fig. S7).

Discussion

In this study, we used a combined computational and experimental approach to study the evolution of RP gene regulation following a WGD event. Our results support a model of *trans* specialization (through subfunctionalization) in RP gene regulation in the post-WGD lineages. In this model (Fig. 3*B*), the pre-WGD *IFH1*/*CRF1* ancestor was both an activator of RP expression under rich growth conditions and a repressor under stress. Following the WGD, the paralogous genes may have specialized, resulting in a separate activator (Ifh1) and repressor (Crf1). The loss of the Crf1 ortholog eliminated the stress-induced repressor function in *C. glabrata*, thus accounting for the lack of RP repression under stress treatments in this species. Ifh1 is still functional in all species, as indicated by the enrichment of IFHL sites in all species' RP promoters.

What were the evolutionary factors affecting this process? One possibility is that, after the duplication of the *IFH1*/*CRF1* ancestor, each copy was more receptive to mutations that were buffered by the presence of a paralog (24, 27). Such neutral changes eventually led to specialized inducers and repressors in the same regulatory program. Alternatively, the emergence of specialized regulators may have been more favorable for managing the increased dosage of RP mRNA following the WGD, which may require a more refined regulatory program with specialized repression under stress. When this pressure was relieved by the loss of RP duplicates in *C. glabrata*, the repressor was lost, resulting in the lack of RP repression in stress. This scenario would be consistent with patterns of regulatory shifts observed across bacterial evolution, where repressors tend to be lost in a genome only after their target genes (31).

The fact that *C. glabrata* adapts the most quickly to environmental stresses in our experiments may also explain why it is able to maintain basal RP expression levels under transient environmental perturbations (although not under nutrient limitation conditions). It is surprising, however, that *C. glabrata* does repress its RiBi genes under stress. One mechanism to balance this discrepancy could be that *C. glabrata* regulates RP levels only posttranscriptionally, as is the case in other organisms (8).

Our work provides a comprehensive strategy for testing hypotheses regarding the evolution of a key molecular pathway by using a comparative functional approach. Applying comparative functional genomics together with an analysis of coding and regulatory sequences revealed how the *cis* and *trans* inputs and expression outputs of an essential gene module have evolved over hundreds of millions of years. Our approach can be widely applied to understand how molecular networks have evolved as organisms have adapted to a variety of habitats and environmental conditions.

Materials and Methods

Gene Orthologies and Phylogenetic Profiles. All gene orthologies and gene trees were calculated using the Synergy algorithm, as previously described (17). Orthologies are available from http://www.broadinstitute.org/regev/orthogroups/. Notably, many RP genes are missing from the *K. waltii* genome annotations but are in fact present in the genome sequences.

Strains and Growth Conditions. We used the following strains for each species: *S. cerevisiae* Bb32 (3), *C. glabrata* CBS 138, *S. castellii* CLIB 592, *K. lactis* CIIB 209, *K. waltii* NCYC 2644, and *C. albicans* SC 5314. Cultures were grown in the following rich medium: yeast extract (1.5%), peptone (1%), dextrose (2%), SC amino acid mix (Sunrise Science) 2 g/L, adenine 100 mg/L, trptophan 100 mg/L, uracil 100 mg/L, at 200 rpm in a New Brunswick Scientific Edison, New Jersey air shaker model I26R and water bath model C76. The medium was chosen to minimize cross-species variation in growth. Following the experimental treatments described below, stressed and mock-treated cultures were transferred to shaking water baths.

Heat Shock. Overnight cultures for each species were grown in 650 ml of media at 22 °C to between 3×10^7 and 1×10^8 cell/mL (OD_{600} = 1.0 for *S. cerevisiae*, *S. castellii*, and *K. lactis*; 0.7 for *C. glabrata* and 0.85 for *C. albicans*). The shift to the heat-shock temperatures was carried out as follows: First the overnight culture was split into two 300-ml cultures and cells from each were collected by removing the media via vacuum filtration (Nanopore). The cell-containing filters were resuspended in pre-warmed media to either control (22 °C) or heat-shock temperatures (37 °C or 42 °C). Density measurements were taken approximately 1 min after cells were resuspended to ensure that concentrations did not change during the transfer from overnight media. A total of 12 ml of culture was harvested 5, 15, 30, 45, and 60 min after resuspension by quenching them in liquid methanol at −40 °C, which was later removed by centrifugation at −9 °C and stored overnight at −80 °C. Cell density measurements were repeatedly taken every 5–15 min for the first 2 hr after treatment. Harvested cells were later washed in RNase-free water and archived in RNAlater (Ambion) for future preparations. Cells were also harvested from cultures just before treatment for use as controls.

Salt. Overnight cultures for each species were grown in 600 ml of media at 30 °C until reaching a final concentration of 3×10^7 and 1×10^8 cell/mL The culture was split into two parallel cultures of 250 ml and sodium chloride was added to one culture for a final concentration of 0.3 M NaCl. Cells were harvested by vacuum filtration 5, 15, 30, and 60 min after the addition of sodium chloride and from cultures immediately before the addition of sodium chloride for use as controls (time 0 min). Filters were placed in liquid nitrogen and stored at −80 °C and were later archived in RNAlater for future use.

Hydrogen Peroxide. Cultures were grown exactly as for salt stress, except that hydrogen peroxide (H_2O_2) was added to a final concentration of 5 mM.

RNA Preparation, Probe Preparation, and Microarray Hybridization. Total RNA was isolated using the RNeasy midi or mini kits (Qiagen) according to the provided instructions for mechanical lysis. Samples were quality controlled with the RNA 6000 Nano II kit of the Bioanalyzer 2100 (Agilent). Total RNA samples were labeled with either Cy3 or Cy5 using a modification of the protocol developed by Joe DeRisi (University of California at San Francisco) and Rosetta Inpharmatics that can be obtained at http://www.microarrays.org.

Microarray Data Analysis. Between two and four biological replicates for each time point were hybridized against the mock T = 0 control on two-color Agilent 55- or 60-mer oligo-arrays in the 4×44 K format for the *S. cerevisiae* strain (commercial array; four to five probes per target gene) or the custom 8×15 K format for all other species (two probes per target gene). After hybridization and washing per the manufacturer's instructions, arrays were scanned using an Agilent scanner and analyzed with Agilent's Feature Extraction software (release 10.5.1.1). The median relative intensities across probes were used to estimate the expression values for each gene, and these median values across replicates were used to estimate the overall expression response per gene per time point.

Estimation of Absolute Transcript Abundance. To assess the absolute levels of each gene's mRNA transcript, we summed each gene probe's raw processed signal from the control channel of the microarray. We then confirmed that this procedure renders consistent and accurate estimations by comparing its values across multiple biological replicates and by checking its correlation to the transcript levels from recent mRNA sequencing data (9). The values were highly consistent across replicates and correlated well (R^2 = 0.75–0.85) with RNA-seq data (9).

Promoter Sequence Analysis. RP genes were identified for each non-*S. cerevisiae* species by orthologous projection using orthologs available from http://www.broadinstitute.org/regev/orthogroups (17). Promoter sequences for each RP gene were defined as 600 bases upstream of ATG and truncated when neighboring ORFs overlapped with this region. *Cis*-regulatory motifs were discovered using the Amadeus software package (28), searching for up to 5 motifs of lengths 8–12 that are significantly enriched as compared with the background set of promoters. Motif targets were identified via the TestMOTIF software program (32) using a three-order Markov background model estimated from the entire set of promoters per genome. Motifs were then clustered according to their targets, and nonredundant motif sets were determined according to maximal coverage of the RP gene set.

ACKNOWLEDGMENTS. We thank Oliver Rando and Audrey Gasch for helpful discussions and comments on previous versions of this manuscript. We thank Leslie Gaffney for assistance with preparing figure graphics. I.W. is the Howard Hughes Medical Institute Fellow of the Damon Runyon Cancer Research Foundation and was supported by a Lawrence Summers Fellowship. This work was supported by the Human Frontiers Science Program, the Howard Hughes Medical Institute, a Career Award at the Scientific Interface from the Burroughs Wellcome Fund, a National Institutes of Health PIONEER award, the Broad Institute, and a Sloan Fellowship (A.R.).

1. Wohlbach DJ, Thompson DA, Gasch AP, Regev A (2009) From elements to modules: Regulatory evolution in Ascomycota fungi. *Curr Opin Genet Dev* 19:571–578.
2. Tuch BB, Li H, Johnson AD (2008) Evolution of eukaryotic transcription circuits. *Science* 319:1797–1799.
3. Ihmels J, Bergmann S, Berman J, Barkai N (2005) Comparative gene expression analysis by differential clustering approach: Application to the Candida albicans transcription program. *PLoS Genet* 1:e39.
4. Tanay A, Regev A, Shamir R (2005) Conservation and evolvability in regulatory networks: The evolution of ribosomal regulation in yeast. *Proc Natl Acad Sci USA* 102: 7203–7208.
5. Hogues H, et al. (2008) Transcription factor substitution during the evolution of fungal ribosome regulation. *Mol Cell* 29:552–562.
6. Lavoie H, Hogues H, Whiteway M (2009) Rearrangements of the transcriptional regulatory networks of metabolic pathways in fungi. *Curr Opin Microbiol* 12:655–663.
7. Tuch BB, Galgoczy DJ, Hernday AD, Li H, Johnson AD (2008) The evolution of combinatorial gene regulation in fungi. *PLoS Biol* 6:e38.
8. Warner JR (1999) The economics of ribosome biosynthesis in yeast. *Trends Biochem Sci* 24:437–440.
9. Yassour M, et al. (2009) Ab initio construction of a eukaryotic transcriptome by massively parallel mRNA sequencing. *Proc Natl Acad Sci USA* 106:3264–3269.
10. Causton HC, et al. (2001) Remodeling of yeast genome expression in response to environmental changes. *Mol Biol Cell* 12:323–337.
11. Gasch AP, et al. (2000) Genomic expression programs in the response of yeast cells to environmental changes. *Mol Biol Cell* 11:4241–4257.
12. Enjalbert B, Nantel A, Whiteway M (2003) Stress-induced gene expression in Candida albicans: absence of a general stress response. *Mol Biol Cell* 14:1460–1467.
13. Martin DE, Soulard A, Hall MN (2004) TOR regulates ribosomal protein gene expression via PKA and the Forkhead transcription factor FHL1. *Cell* 119:969–979.
14. Schawalder SB, et al. (2004) Growth-regulated recruitment of the essential yeast ribosomal protein gene activator Ifh1. *Nature* 432:1058–1061.
15. Wade JT, Hall DB, Struhl K (2004) The transcription factor Ifh1 is a key regulator of yeast ribosomal protein genes. *Nature* 432:1054–1058.
16. Zhao Y, et al. (2006) Fine-structure analysis of ribosomal protein gene transcription. *Mol Cell Biol* 26:4853–4862.
17. Wapinski I, Pfeffer A, Friedman N, Regev A (2007) Natural history and evolutionary principles of gene duplication in fungi. *Nature* 449:54–61.
18. Jorgensen P, et al. (2004) A dynamic transcriptional network communicates growth potential to ribosome synthesis and critical cell size. *Genes Dev* 18:2491–2505.
19. Marion RM, et al. (2004) Sfp1 is a stress- and nutrient-sensitive regulator of ribosomal protein gene expression. *Proc Natl Acad Sci USA* 101:14315–14322.
20. Dietrich FS, et al. (2004) The Ashbya gossypii genome as a tool for mapping the ancient Saccharomyces cerevisiae genome. *Science* 304:304–307.
21. Kellis M, Birren BW, Lander ES (2004) Proof and evolutionary analysis of ancient genome duplication in the yeast Saccharomyces cerevisiae. *Nature* 428:617–624.
22. Dujon B, et al. (2004) Genome evolution in yeasts. *Nature* 430:35–44.
23. Byrne KP, Wolfe KH (2005) The Yeast Gene Order Browser: Combining curated homology and syntenic context reveals gene fate in polyploid species. *Genome Res* 15:1456–1461.
24. Conant GC, Wolfe KH (2008) Turning a hobby into a job: How duplicated genes find new functions. *Nat Rev Genet* 9:938–950.
25. Force A, et al. (1999) Preservation of duplicate genes by complementary, degenerative mutations. *Genetics* 151:1531–1545.
26. Hughes AL (2005) Gene duplication and the origin of novel proteins. *Proc Natl Acad Sci USA* 102:8791–8792.
27. Lynch M, Conery JS (2000) The evolutionary fate and consequences of duplicate genes. *Science* 290:1151–1155.
28. Linhart C, Halperin Y, Shamir R (2008) Transcription factor and microRNA motif discovery: The Amadeus platform and a compendium of metazoan target sets. *Genome Res* 18:1180–1189.
29. Jorgensen P, Nishikawa JL, Breitkreutz BJ, Tyers M (2002) Systematic identification of pathways that couple cell growth and division in yeast. *Science* 297:395–400.
30. Zhu C, et al. (2009) High-resolution DNA-binding specificity analysis of yeast transcription factors. *Genome Res* 19:556–566.
31. Hershberg R, Margalit H (2006) Co-evolution of transcription factors and their targets depends on mode of regulation. *Genome Biol* 7:R62.
32. Barash Y, Elidan G, Kaplan T, Friedman N (2005) CIS: Compound importance sampling method for protein-DNA binding site p-value estimation. *Bioinformatics* 21:596–600.
33. Hall DB, Wade JT, Struhl K (2006) An HMG protein, Hmo1, associates with promoters of many ribosomal protein genes and throughout the rRNA gene locus in Saccharomyces cerevisiae. *Mol Cell Biol* 26:3672–3679.