A Toggle Involving Cis-Interfering Noncoding RNAs Controls Variegated Gene Expression in Yeast

The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.
Toggle involving cis-interfering noncoding RNAs controls variegated gene expression in yeast

Stacie L. Bumgarner*, Robin D. Dowell*, Paula Grisafi†, David K. Gifford§,∥, and Gerald R. Fink*#

*Whitehead Institute for Biomedical Research, 9 Cambridge Center, Cambridge, MA 02142; and †Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, 32 Vassar Street, Cambridge, MA 02139

Contributed by Gerald R. Fink, August 26, 2009 (sent for review July 24, 2009)

The identification of specific functional roles for the numerous long noncoding (nc)RNAs found in eukaryotic transcriptomes is currently a matter of intense study amid speculation that these ncRNAs have key regulatory roles. We have identified a pair of cis-interfering ncRNAs in yeast that contribute to the control of variegated gene expression at the FLO11 locus by implementing a regulatory circuit that toggles between two stable states. These capped, polyadenylated ncRNAs are transcribed across the large intergenic region upstream of the FLO11 ORF. As with mammalian long intervening (l)ncRNAs, these yeast ncRNAs (ICR1 and PWR1) are themselves regulated by transcription factors (Sfl1 and Flo8) and chromatin remodelers (Rpd3L) that are key elements in phenotypic transitions in yeast. The mechanism that we describe explains the unanticipated role of a histone deacetylase complex in activating gene expression, because Rpd3L mutants force the ncRNA circuit into a state that silences the expression of the adjacent variegating gene.

FLO11 | intergenic transcription | Rpd3L histone deacetylase | transcriptional interference | regulatory RNAs

Recent genome-wide studies of eukaryotic transcriptional landscapes in yeast, mice, and humans have revealed extensive activity in regions previously expected to be transcriptionally inert (1–13). A subset of these noncoding (nc)RNAs are long ncRNAs transcribed across intergenic regions. In mammalian cells, transcription of numerous such long intervening (l)ncRNAs is regulated by the binding of transcription factors critical to mammalian development, including Oct4, Nanog, and Sox2 (5). This observation has engendered speculation that mammalian lncRNAs have key roles in development by regulating expression of protein-coding ORFs via mechanisms distinct from the Dicer-dependent RNAi pathway (5, 14–16). However, experiments that would conclusively test the postulated roles for the vast majority of eukaryotic long ncRNAs have not yet been performed (5). Careful interrogation of specific loci is necessary to distinguish between ncRNAs that represent mere transcriptional “noises” and those that have a bona fide role in regulation and development (17–22). Long intergenic ncRNAs also exist in yeast, and despite the tractability of this model system, most remain uncharacterized.

Recent studies of an intergenic ncRNA that regulates the SER3 gene (23, 24), and other subsequent investigations at specific genes in yeast (25–27), have begun to reveal mechanisms, alternative to the RNAi pathways, via which ncRNAs regulate the expression of protein-coding ORFs. The detection in genome-wide studies of noncoding transcripts within promoter regions and numerous instances of overlapping complementary transcripts points to additional regulatory roles for yeast ncRNAs (2, 3, 8–11).

We have identified a pair of long cis-interfering ncRNAs in yeast that contribute to the control of gene expression at the FLO11 locus via a previously uncharacterized type of regulatory circuit, in which these ncRNAs toggle to control transcription of the downstream protein-coding ORF. Transcription of these yeast ncRNAs is regulated by transcription factors Sfl1 and Flo8, key players in FLO11-dependent developmental transitions that enable this organism to adapt to changing environments (28–35).

Functional characterization of the circuitry involving this pair of ncRNAs helps to explain two puzzling phenomena. First, FLO11 is expressed in a binary or “variegated” fashion in clonal populations of WT cells: FLO11 is transcribed at high levels (“on”) in some cells and is completely transcriptionally silenced (“off”) in others (35). In this report, we present evidence that these ncRNAs contribute to the variegated expression observed at FLO11 by toggling between the transcription of one or the other of these ncRNAs. Second, Rpd3L, a histone deacetylase (HDAC), has an unanticipated net activating effect on FLO11 transcription. Because it was the paradoxical role of Rpd3L as an activator of FLO11 transcription that led us to the discovery of the ncRNAs at the FLO11 locus, our presentation of experimental results begins there.

Results

HDAC Rpd3L Is a Net Activator of FLO Gene Expression. Null mutations (Rpd3L−) in components of Rpd3L, including Cti6, Rxt2, and Pho23, result in increased silencing of the FLO11 and FLO10 promoters, indicating that Rpd3L is a net transcriptional activator of these genes. This role for Rpd3L is demonstrated in three ways. First, promoter activity was assayed in strains in which the endogenous promoter is fused to a reporter gene (P_FLO11-URA3, Fig. 1A; P_FLO11-GFP or P_FLO10-GFP, Fig. 1A), which precisely replaces the FLO ORF (Table S1). Detection of ura− (5-FOA resistant) or gfp− cells in WT vs. Rpd3L− strains indicates that FLO promoter silencing is elevated in Rpd3L− cell populations. Second, Northern blot analysis (Fig. 1B) shows that FLO11 mRNA is reduced in Rpd3L− (cit4) compared with WT. Third, disruption of Rpd3L function results in loss of FLO11-dependent phenotypes. Homozygous Rpd3L− diploids fail to form pseudohyphae (Fig. 1C) and Rpd3L− haploids do not adhere to YPD agar (Fig. 1D). These phenotypes are observed in rpd3Δ deletion mutants, indicating that the catalytic component of the Rpd3L HDAC is required for net activation of FLO11. Rpd3L− strains transformed with a P_TEF-FLO11 plasmid...
recover both diploid filamentation and haploid adhesion (Fig. 1 C and D), demonstrating that the phenotypic defects observed in Rpd3L\textsuperscript{−} mutant strains are a direct consequence of their loss of FLO11 expression. These phenotypes are specific to the Rpd3L and not the Rpd3S complex. FLO11 promoter activity is not affected in a strain lacking Eaf3 (Fig. 1 A), unique to Rpd3S (36, 43). Experiments described below use null alleles of Ctf6, unique to Rpd3L, to assay the effects of disrupting Rpd3L function.

Epistasis data suggest that (i) Rpd3L works upstream of FLO-specific transcription factors Sfl1 (repressor) (31, 33, 44) and Flo8 (activator) (28, 29, 33); and (ii) net activation of FLO11 by Rpd3L depends on Sfl1 function. The phenotype of the Rpd3L\textsuperscript{−} sfl1 double mutant is indistinguishable from that of the sfl1 mutant, in which FLO11 promoter silencing is lost in all cells (Fig. 1 A). As in the flo8 strain, all cells have a silenced FLO11 promoter in the Rpd3L\textsuperscript{−} flo8 double mutant (Fig. 1 A). SFL1 mRNA levels in Rpd3L\textsuperscript{−} mutant strains do not differ significantly from WT (Fig. S1 B), indicating that the role of Rpd3L in activating FLO11 expression is not via an indirect mechanism involving transcriptional repression of SFL1.

Rpd3L Localization to the FLO11 Promoter Alters Transcription Factor Binding and Chromatin Remodeling. Genome-wide ChIP-chip detects Rpd3 localization at two regions within the upstream intergenic region of FLO11: \(~1,250\) and \(~2,850\) bp upstream of the ATG of FLO11 (Fig. 2 A). Gene-specific ChIP shows that enrichment of Rpd3 upstream of FLO11 is at least 4-fold higher than at unbound regions, and exceeds enrichment at the INO1 promoter (Fig. 2 A and Fig. S2 A), where Rpd3 localization is reported (45). Compared with WT, localization of the transcriptional activator Flo8 to the FLO11 promoter is significantly decreased in the Rpd3L\textsuperscript{−} mutant and, as previously reported (33), is increased in sfl1 (Fig. 2 B). In the Rpd3L\textsuperscript{−} sfl1 double mutant, Flo8 binding is restored, but not to the levels observed in sfl1. Thus, Flo8 binding remains impaired in the Rpd3L\textsuperscript{−} mutant even in the absence of Sfl1. Yeast TATA box-binding protein (TBP) localization to the FLO11 TATA box (\(~92\) bp) is absent in Rpd3L\textsuperscript{−} and is elevated above WT levels in sfl1 (Fig. 2 C). Histone H4 localization shows that nucleosome eviction fails to occur at the FLO11 core promoter in Rpd3L\textsuperscript{−} cells compared with sfl1 cells (Fig. 2 D). Differential enrichment of TBP and H4 is not merely an artifact of differential overall signal on the arrays, because signal is similar at control regions (Fig. S2 C and D).

Rpd3L, Sfl1, and Flo8 Control a Pair of cis-Acting ncRNAs, Implementing a Toggle That Contributes to FLO11 Regulation. The findings that Rpd3L localizes to the FLO11 promoter and activates FLO11 expression presented a paradox, because HDACs normally function as repressors of transcription by condensing chromatin (46). This paradox could be resolved if Rpd3L repressed the transcription of a cis-acting ncRNA, itself responsible for repression of FLO11 transcription via a promoter occlusion mechanism (23, 24). To test this possibility, we assayed for polyadenylated transcripts deriving from the \(~3.6\)-kb region upstream of FLO11. Strand-specific microarrays provided an initial view of transcription surrounding the FLO11 locus. These arrays detected Crick- and Watson-strand transcription (no ORFs \(>3.0\) bp; Fig. S3 A) across several kilobases of the upstream intergenic region of FLO11 (Fig. 3 A). An analogous result was observed at the variating FLO10 locus (Fig. S3 B) (35). To quantify and determine the size of the ncRNAs upstream of FLO11, Northern blot analysis was performed on oligo(dT)-selected RNAs with strand-specific RNA probes (Fig. 3 B–D). Probes for Crick-strand transcription detect a \(~3.2\)-kb ncRNA, designated ICRI (interfering Crick RNA), transcribed across much of the upstream intergenic region of FLO11 (Fig. 3 C). Low levels of an \(~8\)-kb Crick-strand transcript, which may represent a species transcribed across the FLO11 promoter and ORF, are also detected in some mutants (Fig. 3 C and D). A probe specific for Watson-strand transcription at a region far upstream of the FLO11 ORF detects another ncRNA: \(~1.2\) kb in length and designated PWR1 (promoting Watson RNA) (Fig. 3 C).

Cap-dependent RACE was used to map the \(5^{\prime}\) and \(3^{\prime}\) ends of ICRI and PWR1. The \(5^{\prime}\) RACE identified start sites for ICRI over a 250-bp range. 3,445–3,197 bp upstream of FLO11 (Fig. 3 B and Table S2). The \(3^{\prime}\) RACE identified a strong stop site for ICRI 209 bp upstream of FLO11 and other stops closer to (6, 4, and 2 bp upstream) and within (10 + 24 bp) the FLO11 ORF itself (Fig. 3 B and Table S2). The \(5^{\prime}\) RACE for PWR1 identified start sites over a 160-bp range, 2,190 to 2,339 bp upstream of FLO11 (Fig. 3 B and Table S3). PWR1 is complementary to \(~1.2\) kb of the \(5^{\prime}\) end of ICRI and terminates in the region where ICRI initiates, between 3246 and 3409 bp upstream of FLO11 (Fig. 3 B and Table S3). This configuration suggests...
possible regulatory roles for ICR1 and PWR1 (23–25): ICR1 could repress FLO11 transcription by occluding its promoter, whereas PWR1 could promote FLO11 transcription by interfering with ICR1 binding and chromatin remodeling. (A) ChIP-chip experiments were performed using a functional Myc-tagged allele of Rpd3 in a WT haploid (Fig. S2A). The plot shows fold enrichment of Rpd3-Myc in chromatin immunoprecipitated (IP) with an anti-Myc antibody normalized to the whole-cell extract (WCE). (Inset) Quantitative PCR was performed on IP and WCE using primers specific for the FLO11 promoter (~1,400 bp), for positive binding control Pمول، and for unbound regions APL2 and ARG2. Data were normalized to unbound region ARK1 and are expressed as fold enrichment ± SD. (B) Localization of Flo8 using a Myc-tagged allele in WT and mutant haploids was assayed by qPCR with primers specific for the FLO11 promoter on IP (anti-Myc) and WCE. Data were normalized to unbound region ACT1 and are expressed as fold enrichment ± SEM. (C) Localization of TBP was assayed by ChIP-chip in haploid WT, cti6, and sfl1 cells. The plot shows fold enrichment of TBP at the FLO11 promoter in IP (anti-TBP) normalized to WCE. (D) Localization of histone H4 was assayed by ChIP-chip in haploid WT, cti6, and sfl1 cells. The plot shows fold enrichment of H4 at the FLO11 promoter in IP (anti-H4) normalized to WCE.
transcription. The analogous pair of ncRNAs transcribed upstream of FLO11 (Fig. S3) adds support to this model.

There is an inverse correlation observed between ICR1 and PWR1 transcription. ICR1, but not PWR1, is transcribed at the highest levels detected in this study in mutants (Rpd3L−, flo8, Rpd3L−/flo8, Rpd3L−/flo8 sfl1; Fig. 3 C and D) where transcription of FLO11 is largely silenced. These data implicate Flo8 and Rpd3L as repressors of ICR1. ICR1 is barely detectable in sfl1 mutants in which Rpd3L function is still intact, indicating that Sfl1 function normally promotes ICR1 transcription. PWR1 is detected only in the strains in which FLO11 is also transcribed at high levels (Figs. 1 A and B and 3 C and D). PWR1 transcription requires Flo8 and is promoted by Rpd3L activity, but is repressed by Sfl1 function (Fig. 3 C and D). Both PWR1 and ICR1 are detected in the mixed population of FLO11 on and off cells in the variegating WT strain (Fig. 3C). Quantitative qPCR assays support the presence of the ICR1 transcript, the quantitative differences in its transcription observed by Northern blot analysis, and an inverse correlation between ICR1 and FLO11 transcription (Fig. 3E and Fig. S4, A).

If ICR1 transcription across the FLO11 promoter has a causal role in repressing FLO11, then termination of the ICR1 transcript should block this inhibition and restore FLO11 expression. This prediction was tested with strains in which ICR1 is terminated by constructs (T1–T3) containing a transcriptional terminator (Fig. 4A). The control construct (C) contains an ORF sequence with no terminator (Fig. 4A). Insertion of T1, T2, or T3 at a site 3,041 bp upstream of FLO11 (~350 bp downstream of ICR1 initiation) restores FLO11-dependent adhesion in Rpd3L− mutants (Fig. 4D). The extent of rescue correlates directly with the strength of the terminator (Fig. 4C) and the resulting increase in FLO11 expression (Fig. 4B). Control construct C inserted at the same site does not terminate ICR1 and does not restore adhesion to the Rpd3L− mutant (Fig. 4B–D).

ICR1 and PWR1 show evidence of reciprocal transcriptional interference. This interference is suggested by the inverse correlation in their transcription and by Northern blot bands indicative of a range of transcript sizes that could result from interference (Fig. 3C and D). A genomic comparison of four yeasts closely related to Saccharomyces cerevisiae (49, 50) shows that the region of overlap between PWR1 and ICR1 represents the least conserved DNA sequence in this region, suggesting that transcription per se, rather than specific DNA sequence, is important there (Fig. S4B). A URA3 gene inserted as a surrogate-initiated similarly to ICR1 revealed PWR1-imposed interference on URA3 expression (Fig. S5). Last, termination of ICR1 increases PWR1 levels in the Rpd3L− background (Fig. 4C). The fact that low level ICR1 is detected even in sfl1 mutants (Fig. 3 C and D) suggests that ICR1 may be constitutive, supporting a model in which its levels are tuned by PWR1 transcription.

The insertion of a terminator into just one copy of the FLO11 promoter in Rpd3L−/H11002 diploids up-regulates expression of the downstream ORF only in cis (Fig. 4E). Overexpression of ICR1 or PWR1 in trans has no effect on FLO11 promoter activity in WT, sfl1, flo8, or Rpd3L− strains (Fig. S6). These results show that ICR1 and PWR1 function in cis to regulate FLO11 transcription.

Together, these data support a mutual interference between PWR1 and ICR1, and suggest a model for transcriptional variation at the FLO11 locus involving a toggle between these ncRNAs (Fig. 5).

Discussion

We report the discovery of two long intergenic ncRNAs, ICR1 and PWR1, that have key roles in regulating transcription of the nearby protein-coding ORF FLO11. The ~3.2-kb ICR1 ncRNA is initiated far upstream (~3.4 kb) from the FLO11 ORF and is transcribed across much of the large promoter of FLO11 (53), repressing FLO11 transcription in cis. Our data support a "promoter occlusion" model (23, 24), in which transcription of ICR1 blocks access to general transcription factors and to chromatin remodelers required for nucleosome ejection. The ~1.2-kb PWR1 ncRNA is transcribed from the strand complementary to that encoding ICR1, and promotes FLO11 transcription by interfering with ICR1.

In our model (Fig. 5), the competitive binding of Sfl1 or Flo8 at their respective binding domains (33) initiates events that
contribute to either (i) stabilization of the silent state (Sfl1 binding) or (ii) stabilization of the competent state (Flo8 binding) in each cell. Competition between Sfl1 and Flo8 determines which of two mutually exclusive ncRNA transcription programs occurs. Lack of PWR1 transcription in the absence of Flo8 allows transcription of ICR1 to occlude downstream sequences that recruit other trans-activators of FLO11 expression. Reciprocally, the absence of Sfl1 binding allows Flo8 binding and PWR1 transcription that interferes with ICR1, preventing occlusion of downstream sites and promoting FLO11 transcription. This interplay between Flo8 and Sfl1 is reminiscent of exclusive toggle switches in prokaryotes (54–56) where binding domains for transcription factors that control two adjacent operons overlap, such that only one factor can bind at a time. In such systems, binding competition determines which operon is exclusively transcribed. Competition between Sfl1 and Flo8 generates opposing outputs from the same regulatory region and forms the basis of an analogous switch between two transcriptional states. A similar configuration detected at the variegating FLO10 locus (Fig. S3B) (35) supports this model.

This ncRNA circuitry helps to solve a puzzle concerning Rpd3L. This HDAC is a net activator of FLO11, a surprising role given its function in silencing other target genes (36–38). We propose that Rpd3L activates FLO11 expression by having its anticipated role in condensing chromatin at a site far upstream of the FLO11 core promoter. Localized chromatin condensation by Rpd3L at an upstream site could (i) hinder the access of Sfl1 to its binding site, but promote Flo8 binding (Fig. 5), and/or (ii) directly repress ICR1 transcription, in either case toggling the FLO11 promoter toward a state competent for transcription of the protein-coding ORF. Three results support a proposed role for Rpd3L. First, Rpd3L localizes to the same region (Fig. 2A) where Sfl1 and Flo8 bind (33, 57). Second, double mutant analysis suggests that Rpd3L’s net-activation of FLO11 occurs via inhibition of Sfl1-mediated repression: When Sfl1 function is lost, Rpd3L activity is not needed for transcription of PWR1 (Fig. 3D) or FLO11 (Fig. 1A). Third, if Rpd3L function normally hinders Sfl1 and/or promotes Flo8 binding, then Sfl1 would be expected to occupy its site more often and Flo8 less often (Fig. 2B) in an Rpd3L– mutant. When both Rpd3L and Sfl1 functions are lost, Flo8 could access its binding site more readily (Fig. 2B). However, we cannot exclude the possibility that up-regulation of PWR1 itself in sfl1 dominates over the repressive effects of increased ICR1 when Rpd3L function is also lost, a consequence of the coupled regulation in this toggle switch. Weak Rpd3 localization detected −2,850 bp upstream of FLO11 (Fig. 2A) and the observation that ICR1 transcription is lower in the sfl1 flo8 double mutant compared with the Rpd3L–sfl1 flo8 triple mutant (Fig. 3D) point to the possibility of some Sfl1-independent role for Rpd3L in repressing ICR1 transcription. ICR1 and PWR1 are implicated in controlling an epigenetic phenomenon in yeast (35) that involves the reversible transition from a chromatin state that is competent for transcription of a protein-coding ORF to one that is silenced for its transcription. The roles proposed for these ncRNAs, which share features with mammalian lincRNAs (5), may have general significance for epigenetic regulation in other eukaryotes. There is evidence that epigenetic phenomena, such as imprinting and X-inactivation in mammals, involve ncRNAs (58). Our discovery of a circuitry involving two ncRNAs at the yeast FLO11 locus suggests that regulation of other epigenetic phenomena that involve a progression from an unstable or bistable condition to a stable transcriptional state (either on or off) may, like the FLO11 gene, be controlled by underlying ncRNA regulatory networks.

Materials and Methods

Strains, Media, Microbiological Techniques, and Growth Conditions. Yeast strains used in this study (Table S1) are derived from Σ1278b (28). Standard yeast media were prepared and genetic manipulation techniques were carried out as described (59). For experiments with Pnor-URA3 strains, YPD liquid media were grown overnight at 15°C, and diluted 1:50, and grown at 30°C. Cultures densities were adjusted to equivalence, serially diluted 4-fold, and spotted onto synthetic complete (SC), SC-Ura, and SC + 5-FOA (0.1%) agar plates (60). Haploid adhesion tests were performed as described (30). To induce pseudohyphal growth, strains were grown on SLAD media (61). For Northern blot analysis, qPCR, ChIP, RACE, and microarray expression analysis, cells were grown overnight in YPD liquid, diluted 1:50, and grown to O.D500 0.8–1.2 for use in experiments. Plasmids are listed in Table S1.

Northern Blot Analysis. For the Northern blot analysis in Fig. 1B, total RNA isolated by standard acid phenol extraction was used. For all other blots, total RNA was oligo(dt)-selected to enrich for polyadenylated transcripts. RNAs were separated on formaldehyde-agarose denaturing gels and blotted as described (62). Hybond membranes were hybridized with 32P-labeled DNA probes (Fig. 1C and load controls) or 32P-labeled RNA probes generated with the Ambion T7 Maxiscript Kit (all other hybridizations).

ChIPs. Protocols have been described (63). Briefly, IPs were performed with Dynal Protein G magnetic beads preincubated with antibodies against Myc-epitope (Covance 9E-11 MMS-164P), yeast TBP (Santa Cruz SC-33736), or histone H4 (Upstate Millipore 05-885). For gene-specific ChIP, Cy-5 labeled IP and Cy-3 WCE were hybridized to Σ1278b custom genomic microarrays (Agilent, strand-specific probes – every 50 bp). Data were normalized as follows: Cross-talk normalization provided coefficients for Cy3–Cy5 and Cy5–Cy3 to correct intensities in each channel. Resulting values were median normalized. The data were transformed under the assumption that Cy3 = Cy5 is a good fit. JBD algorithm identified binding events (64).

qPCR. Total RNA obtained by standard acid phenol extraction was reversed transcribed (Qiagen QuantiTect Kit); cDNAs were analyzed with primers specific to targets, SYBR Green reagents (Applied Biosystems), and the ABI 7500 qPCR system.

Genome-Wide Transcription Profiling. Cy3- or Cy5-labeled cDNAs were generated using SuperScript II Reverse Transcriptase on Poly(A) RNA, hybridized to Σ1278b custom genomic microarrays (Agilent, strand-specific probes – every 25 bp), and scanned (Agilent). Data were normalized as follows: Cross-talk normalization provided coefficients for Cy5–Cy3 and Cy3–Cy5 to correct intensities in each channel. Resulting values were median normalized. The data were trans-
formed under the assumption that Cy3 – Cy5 is a good fit. Differential expression between samples on the same array was determined as the difference in median intensity of the set of probes associated with a given transcript.

**RACE.** Mapping of 5’ and 3’ ends of capped, polyadenylated RNA was carried out with specific primers and the Invitrogen GeneRacer Kit. RACE products were cloned (pCR4-TOPO) and sequenced.


**ACKNOWLEDGMENTS.** We thank R. Young (Whitehead Institute) for reagents to generate Myc-tagged Rpd3; A. Rolfe and T. Danford for their contributions to design and analysis of the S.1278B microarrays; O. Ryan and C. Boone (University of Toronto, Toronto) for their p23A strains; C. Wu (Whitehead Institute) for sharing strains yCW91 and yCW180 and primers CW352-CW355; and F. Lewitter for sharing bioinformatics expertise. This work was supported by National Institutes of Health Grants GM035010, R01GM069676, and DK076284 (to R.D.D.). G.R.F. is an American Cancer Society Professor.