Unexpected functions of IncRNAs in gene regulation

Siwen Wang¹ and Elizabeth J Tran¹²*¹

¹Department of Biochemistry; Purdue University; West Lafayette, IN USA; ²Purdue University Center for Cancer Research; Purdue University; West Lafayette, IN USA

Keywords: long non-coding RNAs, kinetics, yeast, transcription, helicase, gene expression, genetics, S. cerevisiae

Long non-coding RNAs (lncRNAs) are a group of molecules that function in gene regulation in yeast, plants and mammals. The precise mechanisms of action for lncRNAs, however, remain largely unclear. The GAL gene cluster has been used as a model system to study the function of these molecules in Saccharomyces cerevisiae, with a historical focus on lncRNA-dependent repression. Strikingly, in characterizing the role of the RNA helicase Dbp2, we discovered that the GAL lncRNAs could also promote transcriptional activation of the targeted GAL protein-coding genes. Interestingly, these lncRNAs help determine how quickly the GAL genes can be induced in response to galactose, without altering final steady-state transcript levels. This unexpected finding suggests that one role for lncRNAs is to promote the timing of gene expression. Herein, we discuss our discoveries in the context of current models of lncRNA functions in eukaryotes, with a key emphasis on future challenges for genomic research.

The mammalian genome is pervasively transcribed, of which a large percentage is non-coding RNAs (ncRNAs).¹² lncRNAs are one group of ncRNAs that are arbitrarily defined as molecules greater than 200 nucleotides in length. lncRNAs regulate the transcription of targeted genes and also function in post-transcriptional regulation of gene transcripts. Based on the few characterized lncRNAs thus far, the proposed functions are largely centered on transcriptional regulation during which lncRNAs can act as molecular signals, decoys, guides and/or scaffolds for transcription factors and epigenetic modifiers. However, IncRNAs have also been shown to encode microRNAs (miRNAs), act as miRNA sponges, and target specific miRNAs for RNA decay.¹ Interestingly, many IncRNAs are associated with inducible genes, suggesting that IncRNAs may aid in gene expression changes in response to extracellular stimuli.¹⁻⁶ However, the precise roles for the vast majority of IncRNAs (> 8,000 in human / ~2,000 in S. cerevisiae) are not known.⁷⁻⁹

The GAL gene cluster in S. cerevisiae encodes the Gal1, Gal10 and Gal7 proteins, which are necessary for metabolism of galactose as a carbon source.¹⁰ Budding yeast preferentially use glucose for production of energy; however, they have the capacity to switch their metabolic pathways in response to the presence of different carbon sources in the media.¹⁰,¹¹ This metabolic reprogramming involves a complex switch to induce the transcription of alternative metabolic genes.¹²⁻¹⁴ The GAL gene cluster is a major component of that switch and has been studied for over 30 y as a model for inducible gene regulation. When budding yeast cells are grown in the presence of glucose, the GAL genes are actively repressed by glucose-dependent repressors Mig1, Nrg1 and associated co-repressors (Cyc8-Tup1).¹⁰,¹⁵,¹⁶ Activation of the GAL genes occurs in the presence of galactose when the genes are activated by the Gal4 transcriptional activator and the co-activators SAGA and Mediator.¹⁴,¹⁵ The GAL genes also exist in a non-induced, non-repressed state when non-fermentable carbon sources such as raffinose are present in the media.¹⁰

In addition to proteinaceous factors, the transcription of the GAL protein-coding genes within the GAL cluster is also regulated by the GAL10 lncRNA, a 4kb transcript that originates from the 3′ end of the GAL10 open reading frame and overlaps the GAL10 and GAL1 genes (Fig. 1).¹⁸⁻²⁰ Initiated adjacent to the GAL10 lncRNA is the shorter GAL10s lncRNA whose function is unknown (Fig. 1). Interestingly, several studies have suggested that the GAL10 lncRNA promotes transcriptional repression.¹⁸⁻²⁰ One of the most striking pieces of evidence in support of a purely repressive mechanism comes from studies of lncRNA decay pathways, which demonstrate that loss of decapping or 5′-3′ decay results in defective GAL gene expression.²⁰ The resulting model from this study is that accumulation of lncRNAs, due to loss of RNA degradation, results in an exacerbated repressive phenotype that reflects the primary function of wild type GAL10 lncRNA.

We unexpectedly uncovered an additional mechanism by which the GAL10 lncRNA could affect gene regulation in our studies of the DEAD-box protein Dpb2 in S. cerevisiae. DEAD-box proteins are non-processive RNA helicases that modulate the conformation of RNA or RNA-protein complexes.²¹ Our lab found that loss of Dpb2 results in accumulation of a 3′ extended GAL10s lncRNA.²² In line with the previously proposed role of the GAL10 lncRNA, we anticipated that the GAL genes would be transcriptionally repressed in the absence of Dpb2. However, in testing this hypothesis, we found that loss of Dpb2 results in rapid transcriptional induction of all...
3 GAL cluster genes from repressive (+glucose) to activated (+galactose) conditions. This rapid induction is dependent on the GAL IncRNAs, suggesting that GAL IncRNAs promote transcriptional activation in dbp2Δ cells. Strikingly, this was also the case for a yeast strain lacking XRN1, a 5′ to 3′ RNA exonuclease specifically required for the degradation of GAL IncRNAs. Rapid induction of GAL gene transcription correlates with reduced association of the Cyc8 co-repressor at the GAL gene promoters and faster recruitment of RNAPII upon glucose galactose switch. Moreover, the GAL lncRNAs also enhanced gene promoters and faster recruitment of RNAPII upon glucose GAL lncRNAs. Rapid induction of exonuclease specifically required for the degradation of GAL also the case for a yeast strain lacking xrn1Δ, a 5′ to 3′ RNA exonuclease specifically required for the degradation of GAL IncRNAs. Rapid induction of GAL gene transcription correlates with reduced association of the Cyc8 co-repressor at the GAL gene promoters and faster recruitment of RNAPII. These rapid inductions are consistent with previous studies.20,24 Further analysis, however, showed defective transcription induction in protein-coding genes.27 Alternatively, it is possible that IncRNA-dependent repression and activation occur through a common mechanism, whereby the GAL IncRNAs generally interfere with the binding of transcription factors at targeted promoters. Consistent with the reduced association of Cyc8, we predict that the GAL IncRNAs interfere with glucose-dependent repressors in the presence of glucose in wild type, xrn1Δ and dbp2Δ cells. In the absence of DCP2, however, the GAL IncRNAs accumulate to very high levels.20 This may cause interference with other DNA-binding proteins, such as the transcriptional activator Gal4 and/or co-activator SAGA.17 This model is consistent with the reduction of H3K18 histone acetylation seen in the absence of DCP2, a histone modification facilitated by the SAGA complex.20,28 Thus, the GAL IncRNAs may not be generally repressive or activating but may cause differential effects on gene transcription depending on the growth conditions and/or mutant strain backgrounds. This may explain why the repressive role of the GAL10 IncRNA was only detectable in wild type strains grown in mixed sugar sources, whereby both activators and repressors are simultaneously present.

Although the mechanism for IncRNA-dependent induction is not known, several models are plausible. One possibility is that the GAL IncRNAs form RNA-DNA hybrids with the GAL protein-coding gene promoters, thereby altering chromatin structure (Fig. 2, left). A role for RNA-DNA hybrids in IncRNA-dependent gene regulation is supported by in vitro assays between the mammalian DHFR DNA and the associated DHFR IncRNA.28 However, in order for the GAL IncRNAs to stimulate activation, these hybrids would have to be selective for repressors rather than activators. Another option is that the GAL IncRNAs function through transcriptional interference, whereby the synthesis of the IncRNA, rather than the IncRNA molecule itself, alters protein-coding gene expression (Fig. 2, middle and 1,30). This mechanism is utilized by the mammalian Airn IncRNA to repress the insulin-like growth factor type 2 receptor (Igf2r) gene.31 If the GAL IncRNAs function through transcriptional interference, Xrn1 and Dcp2 may function co-transcriptionally to degrade nascent RNA from the 5′-3′ direction. This would be consistent with the association of these decay factors with chromatin and the proposed role of Dhp2 in co-transcriptional assembly of RNA-binding proteins on nascent RNA.32,33 Finally, work from several labs has shown that the GAL genes become associated with the nuclear face of the nuclear pore complex (NPC) upon the transcription induction.33,34 Recent work from the Stutz laboratory has now shown that this

![Figure 1. Schematic diagram of the GAL gene cluster with the GAL IncRNAs. The GAL gene cluster contains GAL1, GAL10, and GAL7 protein-coding genes (thick arrows with the arrow indicating the transcription direction) as well as the GAL10 and GAL10s IncRNAs (dashed arrows). The GAL10 IncRNA is initiated from the 3′ end of GAL10 open reading frame and is transcribed antisense to GAL10, across the GAL10 and GAL1 protein-coding genes. Synthesis of the GAL10 IncRNA is dependent on the Rebl transcription factor and is mutually exclusive with induction of GAL gene transcription.10,26 The GAL10s IncRNA is initiated adjacent to the GAL10 IncRNA and transcribed to the sense direction of the GAL10 protein-coding gene.9 The GAL cluster schematic is adapted from 25.](image-url)
gene-to-NPC association promotes deSUMOylation of the Cyc8 (Ssn6) co-repressor by the NPC-associated SUMO protease Ulp1, thereby facilitating rapid induction.35 Thus, it is possible that the GAL lncRNAs facilitate recruitment of the GAL genes to the NPC through an as-of-yet uncharacterized mechanism (Fig. 2, right). The fact that several RNA-binding proteins that require Dbp2 to associate mRNA, such as Nab2, Yra1 and Mex67, associate directly with NPC components supports this model.32,36-38

Our studies suggest that lncRNAs promote the timing of gene expression. This is a key insight, as the functions of the vast majority of the 8,000 lncRNA in humans (~2,000 in budding yeast) remain unknown.7-9 Interestingly, a substantial fraction of lncRNAs in mammalian cells is associated with transcription factor genes that are tightly regulated at the temporal level during cellular differentiation and development.39-41 This suggests that a kinetic role for lncRNAs may be conserved between fungi and humans. Interestingly, a recent large scale analysis of lncRNAs in mouse stem cells found that knockdown of these molecules only altered the final mRNA levels of less than 10% of their neighboring protein-coding genes.41 Either a large portion of these lncRNAs examined are non-functional or that these molecules function in a different process, like transcriptional poising. We suggest that the discovery of temporal roles for lncRNAs offers a new and exciting challenge to the future of genomic research by adding a temporal level of control to the eukaryotic transcriptome.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

We thank all the members from Tran lab for advice and discussions. Special thanks to Barb Golden for critical reading of this manuscript. We apologize to colleagues whose work could not be cited due to space constraints. This work was supported by NIH grant R01GM097332 to EJT.

References

1. Ulitsky I, Bartel DP. lncRNAs: genomics, evolution, and mechanisms. Cell 2013; 154:26-46; PMID:23827673; http://dx.doi.org/10.1016/j.cell.2013.06.020
2. Djebali S, Davis CA, Merkel A, Dobin A, Lassmann T, Mortazavi A, Tanzer A, Lagarde J, Lin W, Schlesinger F, et al. Landscape of transcription in human cells. Nature 2012; 489:101-8; PMID:22955628; http://dx.doi.org/10.1038/nature11233
3. Castelnuovo M, Rahman S, Guffanti E, Infantino V, Stutz F, Zenklusen D. Bimodal expression of PHO84 is modulated by early termination of antisense transcription. Nat Struct Mol Biol 2013; 20:851-8; PMID:23770821; http://dx.doi.org/10.1038/nstmh.2598
4. Camblong J, Iglesias N, Fieckenscher C, Diepois G, Stutz F. Antisense RNA stabilization induces transcriptional gene silencing via histone deacetylation in S. cerevisiae. Cell 2007; 131:706-17; PMID:18022356; http://dx.doi.org/10.1016/j.cell.2007.09.014
5. Martens JA, Laprade L, Winston F. Intergenic transcription is required to repress the Saccharomyces cerevisiae SER3 gene. Nature 2004; 429:571-4; PMID:15175754; http://dx.doi.org/10.1038/nature02538
6. Hung T, Wang Y, Lin MF, Koegel AK, Korake Y, Grant GD, Hoflings HM, Shah N, Umbricht C, Wang P, et al. Extensive and coordinated transcription of noncoding RNAs within cell-cycle promoters. Nat Genet 2011; 43:621-9; PMID:21642992; http://dx.doi.org/10.1038/ng.848
7. Neel H, Malabat C, d'Aubenton-Carafa Y, Xu Z, Steinmetz LM, Jacquier A. Widespread bidirectional promoters are the major source of cryptic transcripts in yeast. Nature 2009; 457:1038-42; PMID:19169244; http://dx.doi.org/10.1038/nature07747

Figure 2. GAL lncRNAs promote the kinetics of transcriptional induction. Under repressive conditions, the GAL lncRNAs displace glucose-dependent repressors from chromatin, generating a transiently derepressed state. The RNA helicase Dbp2 and RNA decay factors Xm1 and Dcp2 maintain the fidelity of this transcriptional switch by facilitating removal and subsequent decay of the GAL lncRNAs, respectively. The possible mechanisms for GAL lncRNA function include the formation of lncRNA-DNA hybrids, transcriptional interference and/or tethering of the GAL genes to the NPC. Both lncRNA-DNA hybrid formation and transcriptional interference would be predicted to interfere with chromatin architecture and association of transcription factors. NPC tethering would facilitate deSUMOylation of the Cyc8 co-repressor.35 These models are not necessarily mutually exclusive and future experiments are necessary to determine the mechanism and order of events for lncRNA-dependent poising.
8. van Dijk EL, Chen CL, d’Aubenton-Carafa Y, Gourvenec S, Kwapisz M, Roche V, Bertrand C, Silvain M, Legois-Né P, Roellet S, et al. XUTs are a class of Xn1-sensitive antiens regulator non-coding RNA in yeast. Nature 2011; 475:114-7; PMID:21697827; http://dx.doi.org/10.1038/nature10118

9. Cabili MN, Trapnell C, Goff L, Kozol M, Tazon-Vega B, Regov A, Rinn JL. Integrative annotation of large human intergenic noncoding RNAs reveals global properties and specific subclasses. Genes Dev 2011; 25:1915-27; PMID:21890647; http://dx.doi.org/10.1101/gad.1744611

10. Sellick CA, Campbell RN, Reece RJ. Galactose metabolism in yeast-structure and regulation of the leucine pathway enzymes and the genes encoding them. Int Rev Cell Mol Biol 2008; 269:111-50; PMID:18779058; http://dx.doi.org/10.1016/S1373-6488(08)10003-4

11. Gancedo JM. Yeast carbon catabolite repression. Microbiol Mol Biol Rev 1998; 62:334-61; PMID:9618445

12. Platt A, Reece RJ. The yeast galactose genetic switch is mediated by the formation of a Gal4p-Gal80p-Gal3p complex. EMBOJ 1998; 17:4086-91; PMID:9760023; http://dx.doi.org/10.1093/emboj/17.14.4086

13. Yano K, Fukasawa T. Galactose-dependent reversible transcription attenuates promoter activation. EMBO J 2004; 23:5302-15; PMID:12897126; http://dx.doi.org/10.1012/EMBOJ.23.16.5302-5515.2003

14. Houseley J, Rubbi L, Grunstein M, Tollervey D. The Snf1 kinase controls glucose repression in Saccharomyces cerevisiae. Proc Natl Acad Sci U S A 1997; 94:1721-6; PMID:9050845; http://dx.doi.org/10.1073/pnas.94.5.1721

15. Zhou H, Winston F. NRG1 is required for glucose energin HB, Silver PA. Genome-wide localization of the nuclear transport machinery couples transcriptional status and nuclear organization. Cell 2004; 117:427-59; PMID:15137957; http://dx.doi.org/10.1016/S0092-8674(04)01448-9

16. Pluvier CA, Campbell RN, Reece RJ. Galactose metabolism in yeast-structure and regulation of the leucine pathway enzymes and the genes encoding them. Int Rev Cell Mol Biol 2008; 269:111-50; PMID:18779058; http://dx.doi.org/10.1016/S1373-6488(08)10003-4

17. Gancedo JM. Yeast carbon catabolite repression. Microbiol Mol Biol Rev 1998; 62:334-61; PMID:9618445

18. Platt A, Reece RJ. The yeast galactose genetic switch is mediated by the formation of a Gal4p-Gal80p-Gal3p complex. EMBOJ 1998; 17:4086-91; PMID:9760023; http://dx.doi.org/10.1093/emboj/17.14.4086

19. Yano K, Fukasawa T. Galactose-dependent reversible interaction of Gal3p with Gal80p in the induction pathway of Gal4p-activated genes of Saccharomyces cerevisiae. Proc Natl Acad Sci U S A 1997; 94:1721-6; PMID:9050845; http://dx.doi.org/10.1073/pnas.94.5.1721

20. Traven A, Jelicic B, Sopita M. Yeast Gal4: a transcriptional paradigm revisited. EMBO Rep 2006; 7:496-9; PMID:16670683; http://dx.doi.org/10.1038/sj.emboj.7400679

21. Zhou H, Winston F. NRG1 is required for glucose repression of the SCU2 and GAL genes of Saccharomyces cerevisiae. BMC Genet 2001; 2:5; PMID:11281938; http://dx.doi.org/10.1186/1471-2165-2-5

22. Papamichos-Chronakis M, Gligoris T, Tzamarias D. The Snf1 kinase controls glucose repression in yeast by modulating interactions between the Mig1 repressor and the Cyc8-Tup1 co-repressor. EMBO Rep 2004; 5:368-72; PMID:15031717; http://dx.doi.org/10.1038/sj.emboj.7400120

23. Wreake VM, Workman JL. Inducible gene expression: diverse regulatory mechanisms. Nat Rev Genet 2010; 11:426-37; PMID:20421872; http://dx.doi.org/10.1038/nrg2781

24. Houseley J, Rabbi L, Grunstein M, Tollervey D, Vogelauer M. An mRNA modulates histone modification and mRNA induction in the yeast GAL gene cluster. Mol Cell 2008; 32:685-95; PMID:19066643; http://dx.doi.org/10.1016/j.molcel.2008.09.027

25. Pinkaya M, Gourvenec S, Morillon A. H3 lysine 4 di – and tri-methylation deposited by cryptic transcription attenuates promoter activation. EMBO J 2009; 28:1697-707; PMID:19407817; http://dx.doi.org/10.1038/emoj.2009.108