

The relative importance of transcription rate, cryptic transcription and mRNA stability on shaping stress responses in yeast

José García-Martínez,¹ Guillermo Ayala,² Vicent Pelechano,³ Sebastián Chávez,⁴ Enrique Herrero⁵ and José E. Pérez-Ortín^{6,*}

¹Departamento de Genética y E.R.I. Biotecmed; Universitat de València; València, Spain; ²Departamento de Estadística e Investigación Operativa; Universitat de València; València, Spain; ³Genome Biology Unit; EMBL; Heidelberg, Germany; ⁴Departamento de Genética; Universidad de Sevilla; Sevilla, Spain; ⁵Departament de Ciències Mèdiques Bàsiques and IRB Lleida; Universitat de Lleida; Lerida, Spain ; ⁶Departamento de Bioquímica y Biología Molecular y E.R.I. Biotecmed; Universitat de València; València, Spain

Key words: transcription rate, cryptic transcription, mRNA stability, stress response, yeast

Abbreviations: TR, transcription rate; ChIP, chromatin immunoprecipitation; GRO, genomic Run-On; SUT, stable unannotated transcripts; CUT, cryptic unstable transcripts; HTS, high throughput sequencing; RP, ribosomal proteins; RNA pol, RNA polymerase; CTD, carboxy terminal domain (of RNAP II); K-S, kolmogorov-smirnov test

It has been recently stated that stress-responding genes in yeast are enriched in cryptic transcripts and that this is the cause of the differences observed between mRNA amount and RNA polymerase occupancy profiles. Other studies have shown that such differences are mainly due to modulation of mRNA stabilities. Here we analyze the relationship between the presence of cryptic transcripts in genes and their stress response profiles. Despite some of the stress-responding gene groups being indeed enriched in specific classes of cryptic transcripts, we found no statistically significant evidence that cryptic transcription is responsible for the differences observed between mRNA and transcription rate profiles.

Introduction

The existence and potential function of nonproductive transcription and of unconventional transcripts is a hot topic of current transcription research.¹ Cryptic transcripts are defined as those that are difficult to detect given their small amount and lack of correspondence with predicted canonical genes. Since functional genomics techniques have been developed, a plethora of different kinds of cryptic transcripts has been detected in all studied eukaryotes, including yeast. The absolute level of yeast cryptic transcription is not known, but estimations for the total amount of antisense transcription go from a rather low, and even controversial, 0.62%² to up to 9.2%.³ Apparently, the more in-depth the cryptic transcription analysis, the larger the proportion of affected genes. A recent high-throughput sequencing (HTS) study found that 60.5% of genes have antisense overlapping transcripts³ but another HTS analysis only detected 310 genes (5.6%) having antisense transcription.⁴

It has been hypothesized that cryptic transcripts play an important role in shaping the stress response in yeast, which brings about differences between mRNA amount and RNA polymerase occupancy profiles.⁵ That study address the potential function of cryptic transcription by comparing the physical presence of RNAP II in genes (“occupancy,” measured by ChIP-chip) and the mature mRNA amount. However, as ChIP cannot

discriminate between active and inactive forms of RNAP II, other methods, such as Genomic Run-On (GRO),⁶ that allow the measurement of the elongation activity should be taken also into account.^{7,8} Especially because the gene-specific differences between mRNA level and RNAP II occupancy could also be due to the gene-specific differences in the relative proportion of active (non-arrested) molecules.⁸ This scenario becomes even more complex when the kinetics of transcription is considered. Part of the discrepancies between changes in mRNA amounts and RNAP II occupancy⁵ (putative transcription rate or TR) could be explained by the kinetic delay that exists between the change in transcription and the corresponding adjustment in mRNA level.⁹ Moreover, several groups working in yeast and in mammalian cells reported variations in the stability of many mRNAs during the stress response by using different techniques.^{4,10-16} This modulation of mRNA half-lives—and not only their initial stability value⁵—should be considered. Nevertheless, the specific contribution of cryptic transcripts, on the one hand, and of mRNA stability regulation, on the other hand, to shaping the stress response has not been systematically evaluated.

Results and Discussion

In order to test whether cryptic transcription is important in shaping the stress response in yeast, we performed a comparative

*Correspondence to: José E. Pérez Ortín; Email: jose.e.perez@uv.es
Submitted: 11/21/11; Revised: 12/26/11; Accepted: 01/18/12
<http://dx.doi.org/10.4161/trns.3.1.19416>

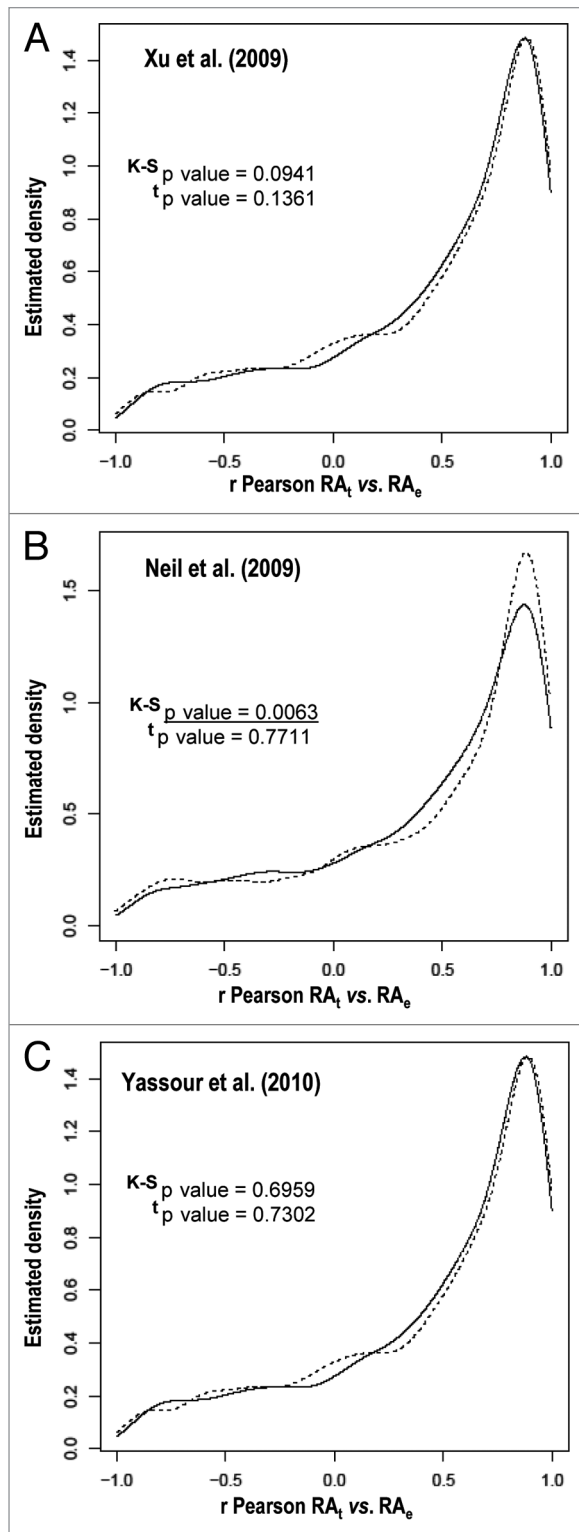


Figure 1. Cryptic transcription does not correlate with the differences between the transcription rate and mRNA profiles in heat stress response. We display the kernel density estimators for the r Pearson correlations between the theoretical and experimental mRNA levels in the heat stress response. The estimated densities correspond to genes with cryptic transcripts (dotted line) and to genes without cryptic transcripts (solid line), as determined by Xu et al.¹⁷ (A), Neil et al. (B) and Yassour et al.² (C) Note that the same bandwidth has been used for both densities. We have tested the normality hypothesis using the Shapiro-Wilk test and the normal null hypothesis is rejected for all data sets. However, because we have large groups of genes to be compared the means can be compared using a two-sample t test (the Welch correction was used). Additionally, the null hypothesis of a common distribution for both groups will be tested using the two-sample Kolmogorov-Smirnov test (K-S). All the statistical tests can be found in reference 19. The p values for the K-S test and t tests are indicated in each panel. Statistically significant values are underlined. Theoretical mRNA (RA_t) levels were calculated using the equation described in reference 12 and assuming a constant degradation constant (k_d) equal to the steady-state value observed before stress (time 0). From that equation, the theoretical mRNA values at different times were calculated from an initial mRNA amount and the experimental TR values.¹² Experimental mRNA (RA_e) levels were those determined in reference 12. r Pearson correlations were calculated for individual gene values of the predicted theoretical mRNA levels plotted vs. the experimental mRNA data. See reference 10 and 12 for further details of the data sets used.

The other cryptic transcripts data sets used long-SAGE HTS to find sense and antisense CUTs,¹⁸ or employed HTS, but only studied antisense SUTs.²

To test if cryptic transcription accounts for the discrepancies between RNAP II occupancy and mRNA level, we compared the experimentally determined mRNA levels throughout the heat stress response time course (reviewed in ref. 10 and 12 for a detailed description of the stress response analysis protocol) to the theoretical mRNA level value expected if mRNA stability did not change during the heat stress response (calculated using the kinetic equation described in refs. 9 and 10) and the TR experimental data determined by GRO.¹² The plot for each gene was then analyzed and the Pearson's correlation coefficient (r) was calculated. An r value of 1 means that both experimental and theoretical mRNA levels are identical and, thus, there is no other influence on mRNA level but that of TR changes. Lower r values imply that the TR is not the only determinant of mRNA level and that other effects, such as the changes in either mRNA stability or in cryptic transcription, also play a role during the stress response. We used two kinds of representations: First, in **Figure 1** we display the density functions estimated using a kernel density estimator.¹⁹ The same plot contains the estimated density functions for two different groups of genes (those with and those without cryptic transcripts). For all cryptic transcripts data sets,^{2,17,18} no significant difference is seen between genes with or without cryptic transcripts when using a t test (i.e., between the mean of the correlations). In the case of the Neil et al. data set, only a residual significance is observed using a Kolmogorov-Smirnov test (K-S).

Another way to analyze the data are by plotting the correlation coefficient against the richness in cryptic transcripts (represented as a 100-genes sliding window). This reveals (**Fig. S1A**) that the mere existence of cryptic transcripts from the Xu et al.

analysis of data recently obtained by our group during response to heat stress experiments.¹² While we used several different cryptic transcription data sets,^{2,5,17,18} previous studies in reference 5, are restricted to just one of them.¹⁷ This last data set contains a genome-wide distribution of stable unannotated transcripts (SUT) from wild-type cells and of cryptic unstable transcripts (CUT) present only in nuclear exosome-deficient cells (*rrp6*).

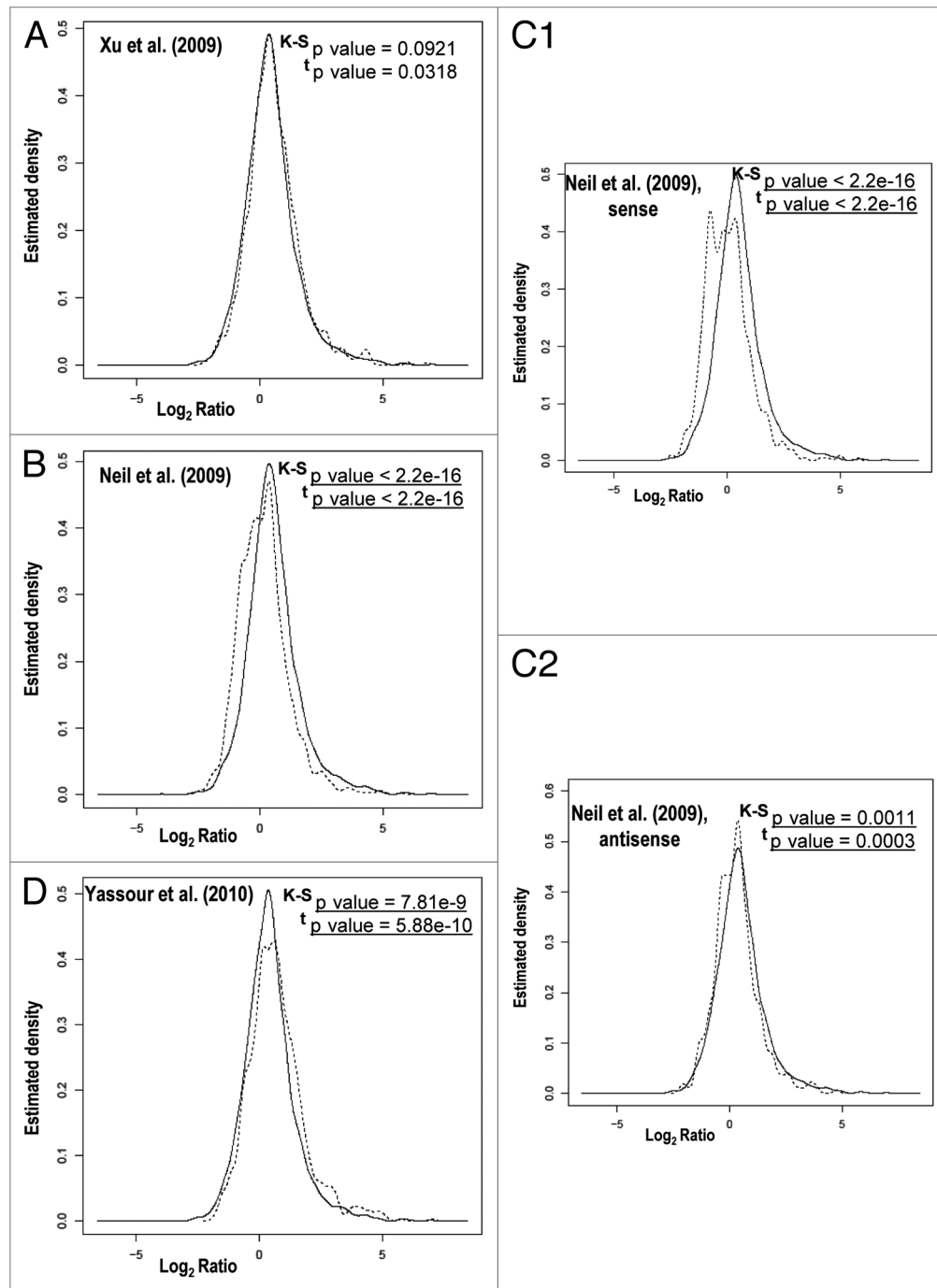


Figure 2. Different correlation patterns in comparisons between cryptic transcription data sets and change ratios after heat stress. We show a representation of the average ratio change (in \log_2 scale) of each time point after stress in relation to the zero time (x axis) against the density (y axis) of the distribution function for genes with cryptic transcripts (dotted line) and for genes without cryptic transcripts (solid line), as determined by Xu et al.¹⁷ (A), Neil et al.² (B) and Yassour et al.² (D). The data set of Neil et al.² is split into sense (C1) and antisense (C2). As in Figure 1, the distributions and the means of correlation coefficients were tested by using the K-S and t tests, respectively. The p values for the K-S and t tests are indicated in each part. Statistically significant values are underlined.

data set¹⁷ in a gene has very little effect (a very low slope, 0.0005) on the discrepancy between the mRNA level and the TR at which it is synthesized (r value). Moreover, when using the other two cryptic transcripts data sets,^{2,18} we found an even lesser effect of cryptic transcription on the discrepancy between theoretical

and experimental mRNA levels, with a flat or slightly negative slope (Fig. S1B and C). Even when considering CUTs and SUTs separately,¹⁷ or sense and antisense separately,¹⁸ we obtained no significant tendency (not shown). Similar results were also found when we used the gene expression data from a different study

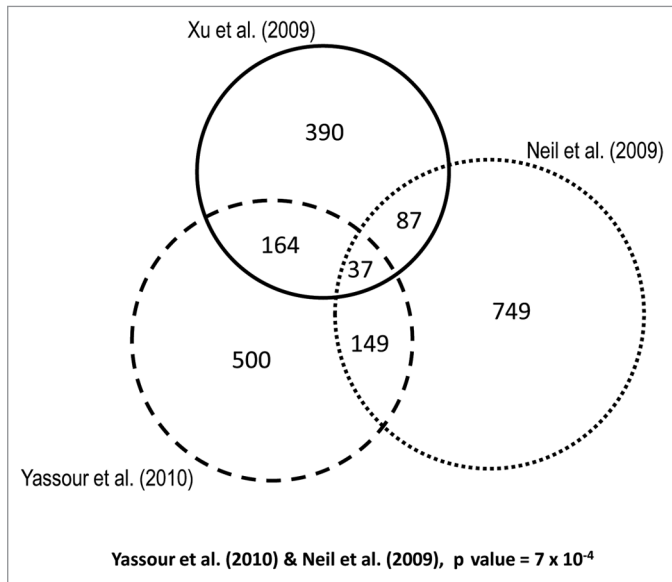


Figure 3. Different cryptic transcription data sets do not significantly overlap. The lists of genes with cryptic transcripts from Xu et al.¹⁷ Neil et al.² and Yassour et al.² mostly contain different genes. We used the original data sets but selected only those genes having overlapping cryptic transcripts that were also present in our previous heat stress-response study.¹² All the overlappings are not statistically significant when using a hypergeometric test, except for the overlap of Neil et al. and that of Yassour et al.² which are significantly larger than expected with a p value of 7×10^{-4} .

of oxidative stress response (not shown).¹⁰ Since we previously experimentally demonstrated that mRNA stabilities vary according to predictions during stress responses,^{10,12} here we argue that changes in the mRNA half-life are the main cause of the discrepancies between mRNA level and the TR.

Other authors have postulated that genes involved in stress responses are enriched in long antisense transcripts and have found that there is a negative correlation in their expression relative to the expression of regular genes.² We analyzed the differences between genes with or without overlapping cryptic transcripts in relation to their mRNA level change during the stress response (Fig. 2). In most cases, a significant difference between the two gene classes was observed. Only the cryptic transcripts data set from Xu et al.¹⁷ shows no significant difference. For this data set, we also performed a sliding window analysis of the effect of the presence of any kind of cryptic transcription and the strength of the gene response during heat stress.¹² Figure S2 illustrates that there is no clear tendency (upper graph) when using the data set of Xu et al.¹⁷ unlike the very clear effect noted for the data set of Neil et al. Downregulated genes tend to be enriched in cryptic transcription, which is mainly due to sense cryptic transcription with a poor antisense cryptic transcription effect (Figs. 2C1 and S2C1). The genes in that part of the graph are enriched in the translation-related categories [e.g., ribosomal proteins genes (RP)—p value < 10^{-49}]. This result may correspond to an abundance of truncated transcripts in the RP genes provoked by a large proportion of blocked RNAPs seen in their 5' part, as we previously showed in reference 7. However, in the data

set of Yassour et al.² which corresponds exclusively to antisense cryptic transcription, the tendency is the opposite, with upregulated genes being enriched in cryptic transcription (Figs. 2D and S2D). This matches the results described by the authors, who found an over-representation of the stress-induced genes among their data set.² Why such different results? As described above, all three cryptic transcription data sets differ considerably and, in fact, the overlapping among them is quite small and mostly non significant (Fig. 3). This can be due to the fact that each method uses a different protocol, which can be biased or affected by particular technical artifacts, therefore detecting different types of cryptic transcripts.^{2,17,18} Hence, the molecular mechanisms underlying the appearance of the different cryptic RNA types and their functions can be quite different. Thus, it is not surprising that they all provide different results in our analyses. In any case, our study suggests that genes responding to stress tend to be enriched in sense and antisense cryptic transcription by either decreasing or increasing their TR, respectively. This result does not necessarily mean, however, that cryptic transcription is the cause of the particular mRNA level stress response profile of those genes. In fact, we found that those genes with more marked changes in their RNA level during the stress response are not those with a lower correlation in theoretical and experimental mRNA profiles; there is even a weak tendency to a better correlation in mRNA profiles (r Pearson calculated as in Fig. 1) with the absolute ratio of change during heat stress (Fig. 4A). When analyzing the r coefficient for the mRNA profiles of those genes with both strong responses to stress and the presence of cryptic transcription (Fig. 4B), it is evident that they are not especially biased toward a low r Pearson coefficient (as defined in Fig. 1). In fact, as Figure 4B illustrates, those RP genes with a strong presence of sense cryptic transcripts (to the left end of the x axis in Graph S2C1) show a higher average r (0.896) than the total gene population (0.45). This is also true (average r = 0.793) for those genes induced by heat shock with a strong presence of antisense transcripts (to the right end in Graph D from Fig. S2). Finally, when the general profile is analyzed (dotted curve in Fig. 4B), it is clear that most genes have a fairly good r coefficient. All these results lead to the conclusion that a change in the TR is the main cause of a change in mRNA level during the stress response, and that this is even clearer for those genes with any kind of cryptic transcription.

In summary, we consider that understanding the functional role of cryptic transcription is becoming a very interesting research focal point. However, we also consider that changes in the TR are the main cause of changes in mRNA levels during stress responses.^{4,12,20} As we have shown, most of the discrepancy between RNAP occupancy (or TR) and mRNA levels during transcriptional responses can be explained by modulation in mRNA stabilities.^{4,6,10-16} Although the effect can be more marked for particular genes with a large proportion of cryptic transcripts, this cannot account for most of the discrepancies detected between mRNA changes and RNAP II changes. Therefore, it is more experimentally supported to argue that most discrepancies are due to both the delay required by kinetic laws and the changes occurring in mRNA stability.^{10-14,20-22}

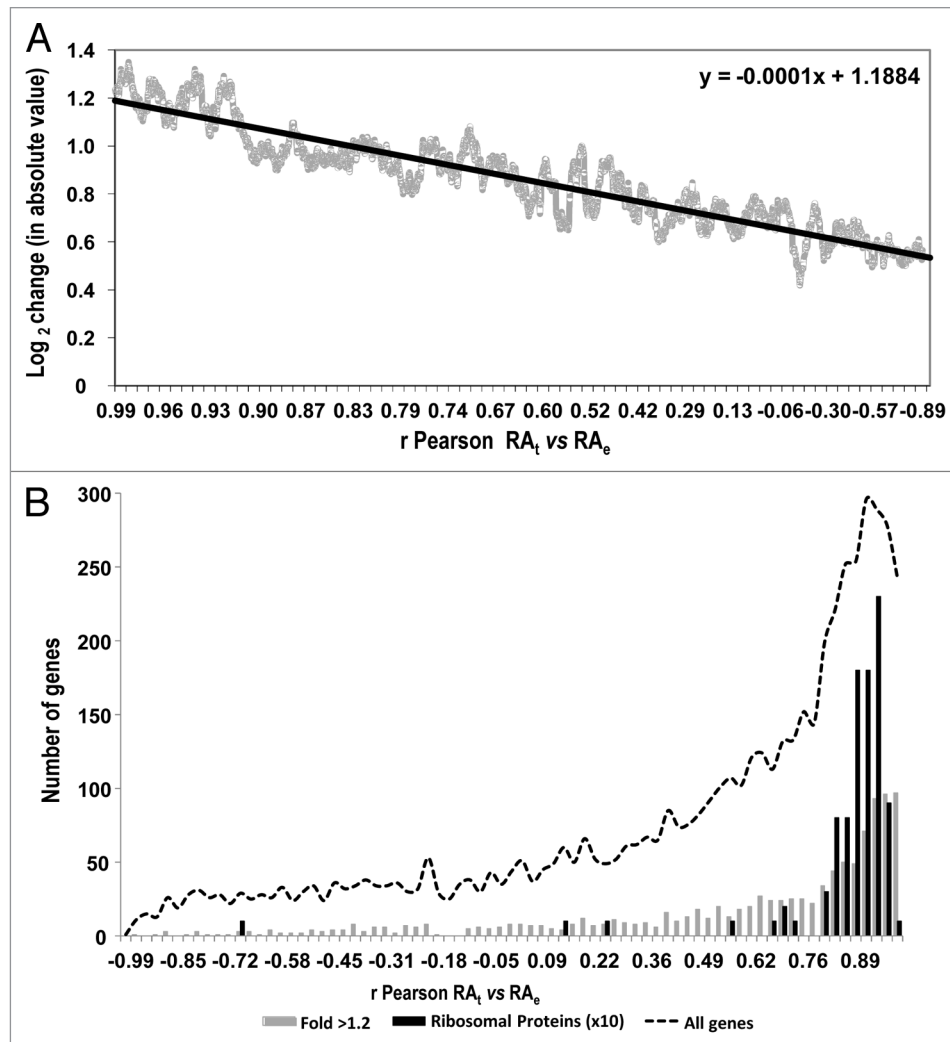


Figure 4. Analyses of up- and downregulated genes after heat stress show a main dependence on the transcription rate change. (A) There is a slight positive dependence of the ratio of change (in \log_2 with absolute values, either up or downregulated) after stress in relation to the r correlation coefficient, calculated as in Figure 1. (B) Histogram of the correlation coefficients for the whole set of genes (dotted line) and for two particular groups [96 ribosomal protein (RP) genes (black bars) showing a considerable downregulation and a large proportion of sense cryptic transcription (shown in Fig. S2), and 1,051 upregulated genes (>1.2 on the \log_2 scale, see Fig. S2; gray bars)]. Note that the scale for the RP genes is magnified ten times.

Acknowledgments

This work has been supported by grants from the Spanish MICINN: BFU2010-17656/BMC to E.H., BFU2010-21975-CO3-01/BMC to J.E.P.O. and BFU2010-21975-CO3-02/BMC to S.C., from the Generalitat Valenciana Prometeo 2011/088 to J.E.P.O. and from Junta de Andalucía (grants P07-CVI02623 and P08-CVI-03508) to S.C. MICINN and Junta de

Andalucía grants involve FEDER resources from the European Union. V.P. is supported by an EMBO fellowship.

Note

Supplemental material can be found at:
www.landesbioscience.com/journals/trans/article/19416

Note added in proof:

In a recent paper Xu et al. (2011) made an exhaustive study analysis of antisense transcription in yeast and concluded that, although it affects the regulation of the gene expression the main driving force of canonical (sense) expression is not its associated antisense expression.

Xu Z, Wei W, Gagneur J, Clauder-Münster S, Smolik M, Huber W, Steinmetz L. Antisense expression increases gene expression variability and locus interdependency. *Mol. Syst. Biol.* 7:468.

References

- van Bakel H, Nislow C, Blencowe BJ, Hughes TR. Most “dark matter” transcripts are associated with known genes. *PLoS Biol* 2010; 8:1000371; PMID:20502517; <http://dx.doi.org/10.1371/journal.pbio.1000371>.
- Yassour M, Pfiffner J, Levin JZ, Adiconis X, Gnirke A, Nusbbaum C, et al. Strand-specific RNA sequencing reveals extensive regulated long antisense transcripts that are conserved across yeast species. *Genome Biol* 2010; 11:87; PMID:20796282; <http://dx.doi.org/10.1186/gb-2010-11-8-r87>.
- Ozsolak F, Kapranov P, Foissac S, Kim SW, Fishilevich E, Monaghan AP, et al. Comprehensive polyadenylation site maps in yeast and human reveal pervasive alternative polyadenylation. *Cell* 2010; 143:1018-29; PMID:21145465; <http://dx.doi.org/10.1016/j.cell.2010.11.020>.
- McKinlay A, Araya CL, Fields S. Genome-wide analysis of nascent transcription in *Saccharomyces cerevisiae*. *Genes Genomes Genet* 2011; 1:549-58.
- Kim TS, Liu CL, Yassour M, Holik J, Friedman N, Buratowski S, et al. RNA polymerase mapping during stress responses reveals widespread nonproductive transcription in yeast. *Genome Biol* 2010; 11:75; PMID:20637075; <http://dx.doi.org/10.1186/gb-2010-11-7-r75>.
- García-Martínez J, Aranda A, Pérez-Ortín JE. Genomic run-on evaluates transcription rates for all yeast genes and identifies gene regulatory mechanisms. *Mol Cell* 2004; 15:303-13; PMID:15260981; <http://dx.doi.org/10.1016/j.molcel.2004.06.004>.
- Pelechano V, Jimeno-González S, Rodríguez-Gil A, García-Martínez J, Pérez-Ortín JE, Chávez S. Regulon-specific control of transcription elongation across the yeast genome. *PLoS Genet* 2009; 5:1000614; PMID:19696888; <http://dx.doi.org/10.1371/journal.pgen.1000614>.
- Rodríguez-Gil A, García-Martínez J, Pelechano V, Muñoz-Centeno Mde L, Geli V, Pérez-Ortín JE, et al. The distribution of active RNA polymerase II along the transcribed region is gene-specific and controlled by elongation factors. *Nucleic Acids Res* 2010; 38:4651-64; PMID:20385590; <http://dx.doi.org/10.1093/nar/gkq215>.
- Pérez-Ortín JE, Alepuz PM, Moreno J. Genomics and gene transcription kinetics in yeast. *Trends Genet* 2007; 23:250-7; PMID:17379352; <http://dx.doi.org/10.1016/j.tig.2007.03.006>.
- Molina-Navarro MM, Castells-Roca L, Bellí G, García-Martínez J, Marín-Navarro J, Moreno J, et al. Comprehensive transcriptional analysis of the oxidative response in yeast. *J Biol Chem* 2008; 283:17908-18; PMID:18424442; <http://dx.doi.org/10.1074/jbc.M800295200>.
- Romero-Santacreu L, Moreno J, Pérez-Ortín JE, Alepuz P. Specific and global regulation of mRNA stability during osmotic stress in *Saccharomyces cerevisiae*. *RNA* 2009; 15:1110-20; PMID:19369426; <http://dx.doi.org/10.1261/rna.1435709>.
- Castells-Roca L, García-Martínez J, Moreno J, Herrero E, Bellí G, Pérez-Ortín JE. Heat shock response in yeast involves changes in both transcription rates and mRNA stabilities. *PLoS One* 2011; 6:17272; PMID:21364882; <http://dx.doi.org/10.1371/journal.pone.0017272>.
- Mata J, Marguerat S, Bähler J. Post-transcriptional control of gene expression: a genome-wide perspective. *Trends Biochem Sci* 2005; 30:506-14; PMID:16054366; <http://dx.doi.org/10.1016/j.tibs.2005.07.005>.
- Keene JD. The global dynamics of RNA stability orchestrates responses to cellular activation. *BMC Biol* 2010; 8:95; PMID:20619007; <http://dx.doi.org/10.1186/1741-7007-8-95>.
- Rabani M, Levin JZ, Fan L, Adiconis X, Raychowdhury R, Garber M, et al. Metabolic labeling of RNA uncovers principles of RNA production and degradation dynamics in mammalian cells. *Nat Biotechnol* 2011; 29:436-42; PMID:21516085; <http://dx.doi.org/10.1038/nbt.1861>.
- Schwanhäusser B, Busse D, Li N, Dittmar G, Schuchhardt J, Wolf J, et al. Global quantification of mammalian gene expression control. *Nature* 2011; 473:337-42; PMID:21593866; <http://dx.doi.org/10.1038/nature10098>.
- Xu Z, Wei W, Gagneur J, Perocchi F, Clauder-Münster S, Camblong J, et al. Bidirectional promoters generate pervasive transcription in yeast. *Nature* 2009; 457:1033-7; PMID:19169243; <http://dx.doi.org/10.1038/nature07728>.
- Neil 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>.
- Rosner B. *Fundamentals of Biostatistics*. Brooks/Cole Cengage Learning 2010.
- Pérez-Ortín JE. Genomics of mRNA turnover. *Brief Funct Genomic Proteomic* 2007; 6:282-91; PMID:18216027; <http://dx.doi.org/10.1093/bfpg/elm029>.
- Grigull J, Mnaimneh S, Pootoolal J, Robinson MD, Hughes TR. Genome-wide analysis of mRNA stability using transcription inhibitors and microarrays reveals posttranscriptional control of ribosome biogenesis factors. *Mol Cell Biol* 2004; 24:5534-47; PMID:15169913; <http://dx.doi.org/10.1128/MCB.24.12.5534-47.2004>.
- Wang Y, Liu CL, Storey JD, Tibshirani RJ, Herschlag D, Brown PO. Precision and functional specificity in mRNA decay. *Proc Natl Acad Sci USA* 2002; 99:5860-5; PMID:11972065; <http://dx.doi.org/10.1073/pnas.092538799>.