

Nucleosome Organization Affects the Sensitivity of Gene Expression to Promoter Mutations

Gil Hornung,¹ Moshe Oren,² and Naama Barkai^{1,*}

¹Department of Molecular Genetics

²Department of Molecular Cell Biology

Weizmann Institute of Science, Rehovot, Israel

*Correspondence: naama.barkai@weizmann.ac.il

DOI 10.1016/j.molcel.2012.02.019

SUMMARY

Gene expression diverges rapidly between related species, playing a key role in the evolution of new phenotypes. The extent of divergence differs greatly between genes and is correlated to promoter nucleosome organization. We hypothesized that this may be partially explained by differential sensitivity of expression to mutations in the promoter region. We measured the sensitivity of 22 yeast promoters with varying nucleosome patterns to random mutations in sequence. Mutation sensitivity differed by up to 10-fold between promoters. This difference could not be explained by the abundance of transcription factor binding sites. Rather, mutation sensitivity positively correlated with the relative occupancy of nucleosomes at the proximal promoter region. Furthermore, mutation sensitivity was reduced upon introduction of a binding site for Reb1, a factor that blocks nucleosome formation, suggesting that nucleosome organization directly regulates mutation sensitivity. Our study suggests an important role for chromatin structure in the evolution of gene expression.

INTRODUCTION

New phenotypes emerge via mutations that change protein function or protein regulation. These two modes of evolution are often complementary, affecting distinct genes and processes (Tirosh and Barkai, 2008a; Soskine and Tawfik, 2010). Most studies of genome evolution focused on mutations in protein sequences that are relatively easy to define using the increasing number of sequenced genomes. However, the findings that most coding regions are tightly conserved between related species, together with the observed large differences in transcriptional programs, emphasize the importance of gene expression divergence between species or strains. Studying the sources of gene expression divergence is thus important to understand the evolutionary process on a molecular level (Tirosh et al., 2009a and references therein).

Notably, the level of interspecies divergence in expression varies between genes, as exemplified most extensively in

studies focusing on budding yeast species (Tirosh and Barkai, 2008a and references therein). Part of the differences in expression divergence between genes is explained by gene function and may therefore be due to differences in purifying or adaptive selection. For example, expression of essential genes diverge significantly less than average, while genes involved in sensing or transporting extracellular nutrients diverge significantly faster. Those functional attributes, however, explain only a small part of the variance between genes (Tirosh et al., 2006; Landry et al., 2007).

The strongest correlate with expression divergence between genes was found to be the pattern of nucleosome organization on gene promoters (Tirosh and Barkai, 2008b; Choi and Kim, 2009; Tirosh et al., 2009a). Thus, the typical yeast promoter contains a well-defined nucleosome-free region (NFR) immediately upstream of the transcription start site. This class of promoters (denoted as DPN for “depleted proximal nucleosome”) is associated with robust expression that changes relatively little between conditions, between individual cells (low expression noise), and also diverges relatively slowly between related species. Further, promoters of the DPN class tend to lack a TATA box and display lower sensitivity to chromatin regulators. On the other end of the spectrum is the group of ~20% of yeast genes that lack a pronounced NFR but bind nucleosomes relatively uniformly along their promoter. This class of promoters (denoted as OPN for “occupied proximal nucleosome”) is associated with flexible expression that changes between conditions, individual cells, and related species. Furthermore, promoters of this class tend to display a TATA box and are highly sensitive to chromatin regulators (Tirosh and Barkai, 2008b; Choi and Kim, 2009).

The rapid expression divergence of OPN promoters can be explained by different, not mutually exclusive processes, including: distinct selection forces, wider spectrum of effective *trans* mutations (Tirosh et al., 2009b), or higher propensity for accumulation of mutations (as was shown for repeat-containing promoters; Vences et al. [2009]). In addition, OPN promoters may be more sensitive to mutations in promoter sequence. Here, we examine this last hypothesis by developing an assay to directly measure the sensitivity of gene expression to mutations in promoters. Although our assay eliminates the effects of selection forces, we still find variation of up to 10-fold in the mutation sensitivity of promoters between the two classes. This difference could not be explained by the abundance of transcription factor binding sites. Rather, mutation sensitivity correlated with the

Table 1. Properties of 22 Genes Used in the Study

Gene Name	Promoter Expression (Fluorescence, A.U.)	Nucleosome Occupancy Ratio	Promoter Type	Conserved TATA Box	Essential Gene	GO SLIM Functions	
HSP150	12,900	2.209	OPN	+		hydrolase activity	structural molecule activity
ERG6	1,800	2.2	OPN			transferase activity	
HXT2	10,500	1.727	OPN	+		transporter activity	
HXT3	1,600	1.719	OPN	+		transporter activity	
GPD1	6,600	1.129	OPN	+		oxidoreductase activity	
RNR2	3,400	0.959	OPN	+	+	oxidoreductase activity	
ERG11	2,200	0.874	OPN	+	+	oxidoreductase activity	
ADH3	1,500	0.642	OPN			oxidoreductase activity	
SAM4	1,700	0.567	OPN			transferase activity	
HSC82	11,300	0.475	Intermediate	+		hydrolase activity	protein binding
NSR1	5,400	0.233	Intermediate			DNA binding	RNA binding
TMA19	7,900	0.006	DPN				
RPB8	1,800	−0.125	DPN		+	transferase activity	nucleotidyltransferase activity
ARO2	1,600	−0.459	DPN			oxidoreductase activity	lyase activity
RPT2	5,000	−0.566	DPN			hydrolase activity	
IPP1	3,100	−0.849	DPN		+	hydrolase activity	protein binding
VMA7	2,100	−0.911	DPN			hydrolase activity	transporter activity
OST1	2,500	−1.004	DPN		+	transferase activity	
DUT1	1,500	−1.074	DPN			hydrolase activity	
TIF6	5,000	−1.136	DPN		+		
PFY1	3,300	−1.532	DPN			protein binding	lipid binding
TOM6	3,100	−1.881	DPN			transporter activity	

relative occupancy of nucleosomes in the proximal region of the promoter. We further examine whether promoter organization directly impacts mutation sensitivity by introducing a binding site for Reb1, a factor that blocks nucleosome formation. Notably, this change led to a significant decrease in mutation sensitivity in three independent cases. Together, our results suggest a key role for promoter nucleosome organization in regulating the capacity of gene expression to evolve via *cis* mutations.

RESULTS

Different Promoters Display Differential Sensitivity to Mutagenesis

We selected 22 promoters that span a range of mid-to-high expression (Table 1). The promoters were selected based on having an OPN or DPN architecture. Eleven promoters belonged to the DPN class, nine promoters belonged to the OPN class, and two promoters were intermediate (Table 1 and examples in Figures 1B and 1C). The promoters represented genes from diverse functions and pathways, as shown by lack of enrichment for GO terms for molecular function or process in the overall

group of genes (at $p < 0.01$). The OPN genes by themselves were enriched for oxidoreductase activity (4/9 genes) and hexose transmembrane transporter activity (2/9 genes), but these enrichments reflect the functional preferences already within all OPN genes in the genome.

Using mutagenic PCR, we generated libraries of at least 300 variants of each promoter and inserted each variant in front of a YFP reporter (Figure 1A). As a control, this procedure was repeated while excluding the mutagenic PCR. Finally, the expression driven by each promoter variant was measured using flow cytometry.

Significant changes in expression were readily observed for some promoters (e.g., RNR2; Figures 1D and 1F), whereas others remained largely unaffected (e.g., OST1; Figures 1E and 1F). Expression tended to decrease upon mutagenesis rather than increase (Figure S1A). We quantified the mutation sensitivity by the fraction of variants that changed expression beyond a certain fold. Mutation sensitivity was correlated with the relative occupancy of the TSS-proximal nucleosome (Figure 1G; $R = 0.54$, $p = 0.01$), a measure that is high for OPN promoters and low for DPN ones (Supplemental Experimental Procedures). This differential sensitivity was consistent for various expression

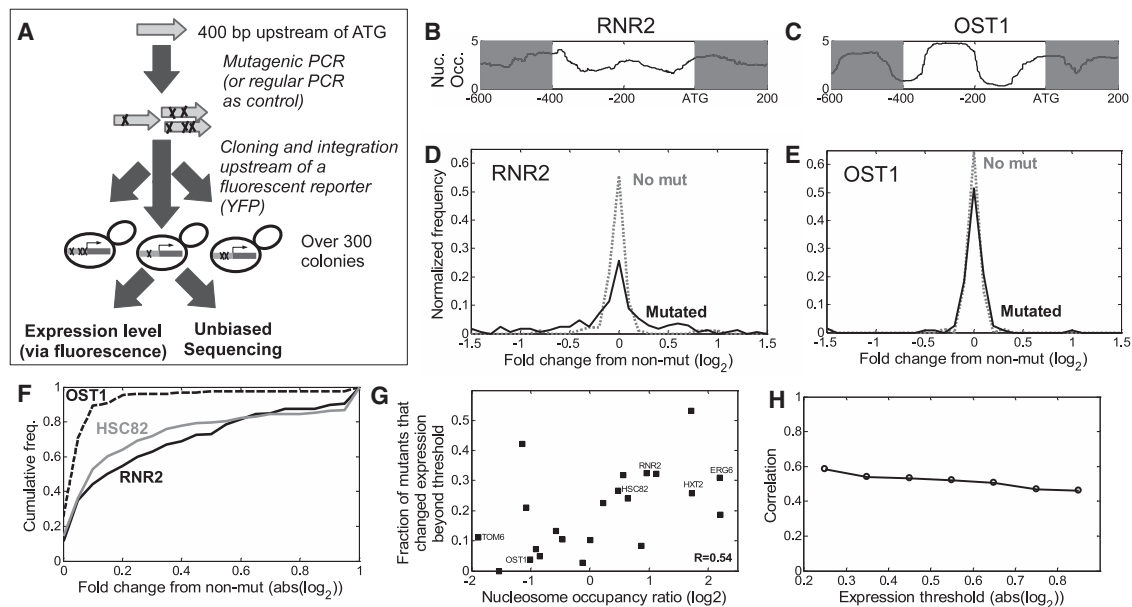


Figure 1. Effect of Random Promoter Mutations on Gene Expression

(A) Experimental design: 400 bp upstream of ATG of 22 different yeast genes were amplified using mutagenic PCR, cloned, and integrated upstream of a yellow fluorescent protein in a neutral genomic locus. Single colonies were picked, and their expression level and sequence were determined.

(B and C) Nucleosome occupancy at the RNR2 and OST1 promoters, respectively (Tsui et al., 2011).

(D and E) Distribution of expression levels in the RNR2 and OST1 libraries, respectively. Control experiments, in which the mutagenic PCR was omitted, are shown in gray broken lines.

(F) Cumulative change in expression: Shown are the cumulative distributions of \log_2 ratio of expression change (in absolute value) for RNR2 (bottom) and OST1 (top, broken line) and an intermediate case of HSC82 (gray). Compare the sharp increase in OST1, where most variants show a similar (wild-type) expression level, with the gradual increase in RNR2, where mutations display a range of expression.

(G) Expression change correlates with occupancy of proximal nucleosome: The fraction of mutant clones that changed expression versus the nucleosome occupancy ratio, defined as the ratio of nucleosome occupancy, close to the transcription start site relative to further upstream (Supplemental Experimental Procedures). A low value for this ratio specifies a DPN architecture, whereas promoters with a high value are OPN. Level of mutagenesis (based on number of mutagenic PCR cycles) is similar. The threshold to define a significant change in expression from the nonmutant expression was $\text{abs}(\log_2)$ of 0.35.

(H) Correlation of expression change with nucleosome occupancy is maintained across expression levels: The Pearson correlation between the fraction of mutant clones that changed expression and the nucleosome occupancy ratio was calculated as in (G). The correlation is shown for different choices of expression change thresholds. Only statistically significant correlations ($p < 0.05$) are presented. See also Figure S1.

thresholds (Figure 1H) and for alternative measures of mutation sensitivity (Figures S1B and S1C).

Mutation Sensitivity Is Correlated to Nucleosome Organization

All promoters in Figure 1G underwent a similar number of mutagenic cycles. Yet, it is possible that the frequency of mutations differed between promoters, leading to a bias in our measure of mutation sensitivity. To control for that, we sought to define a measure for mutation sensitivity that is independent of the number of mutations. A possible drawback for reliable estimation of mutation sensitivity is that for the DPN promoters very few mutants changed expression. We thus repeated the mutagenic process for these promoters with a higher number of PCR cycles, such that the fraction of variants that changed expression was roughly similar for all promoters (Figure S1D). We then sequenced ~75 randomly selected variants from 20 of the 22 libraries (Table S1). For most libraries, the distribution of mutations was roughly even along the promoter sequence (with the exception of high mutation rate in A/T stretches, Supplemental Experimental Procedures). On average, each variant contained

between one to four mutations, depending on the level of mutagenesis used (Table S1). As expected from the number of PCR cycles, libraries of DPN promoters had a higher average number of mutations (~0.8 more mutations, $p = 0.047$).

In each library the fraction of promoter variants that changed expression generally increased with the number of mutations in its sequence (Figure S2A). We defined a measure of mutation sensitivity that accounts for variations in mutation level. The data was fitted to a probabilistic model that assumes an independent contribution of each mutation to the probability of changing expression beyond a certain fold (Experimental Procedures). This fit defines the probability for a single, random, mutation to affect expression (values in Table S3). Once again, this probability correlated with the organization of promoter nucleosomes, increasing monotonically with the occupancy of the TSS-proximal nucleosome (Figure 2A; $R = 0.77$, $p < 0.001$). This correlation persisted for different expression thresholds (Figure 2B). Importantly, in our analysis we ignored the higher mutation rate in A/T-rich stretches that are prevalent in DPN promoters. Taking these mutations into account would reduce the effective mutation sensitivity of DPN promoters even further.

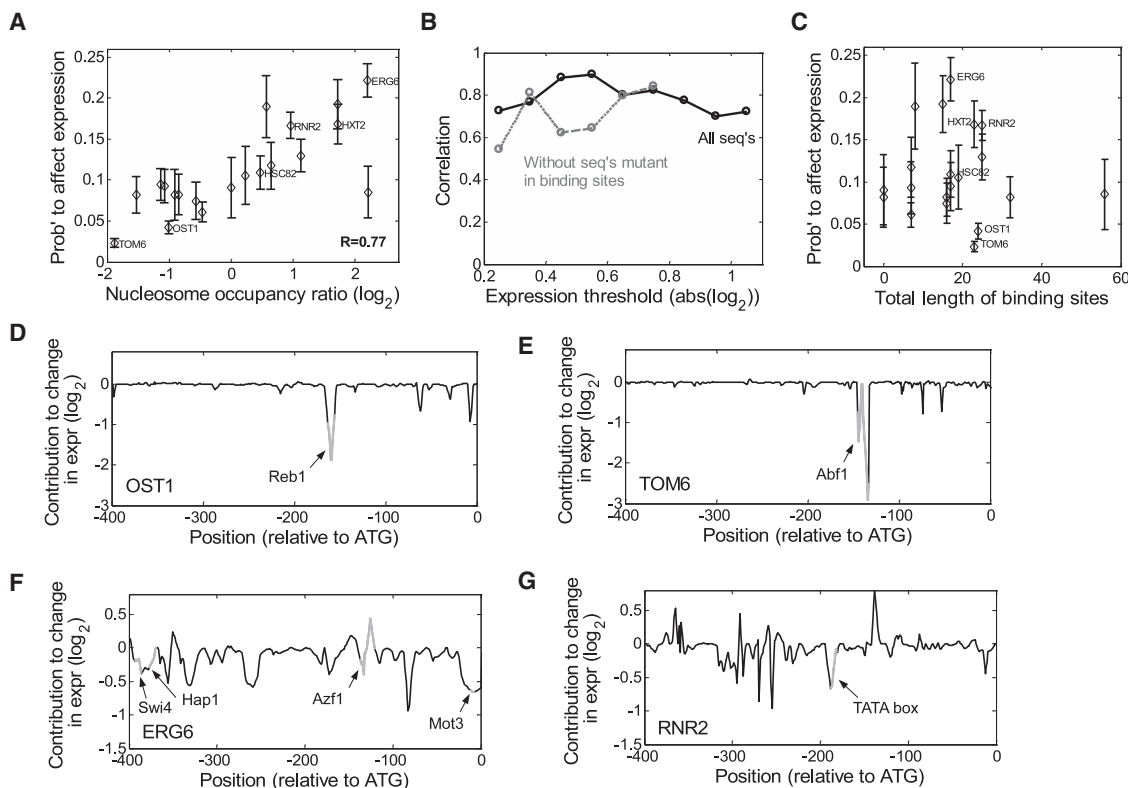


Figure 2. Sensitivity to Random Mutations Correlates with the Relative Proximal Nucleosome Occupancy

(A and B) Mutation sensitivity correlates with nucleosome occupancy: The probability that a random mutation will affect expression (mutation sensitivity) was estimated for all promoters based on the sequences and expression levels of each mutant (Experimental Procedures). This probability is plotted as a function of the proximal nucleosome occupancy ratio (Supplemental Experimental Procedures; results were not affected by choice of nucleosome occupancy dataset; Figures S4C and S4D). The $\text{abs}(\log_2)$ expression change threshold in (A) is 0.35, and the correlation value for different thresholds is shown in (B). Correlation values obtained after removing clones mutated at known binding sites (MacIsaac et al., 2006) are shown in gray in (B). Only correlations with high statistical confidence are shown. Error bars are based on bootstrapping (Experimental Procedures).

(C) Mutation sensitivity does not correlate with the total length of transcription factor binding sites: The probability (as in A) plotted against total length of binding sites within the promoter, predicted by PSSM (Badis et al., 2008).

(D–G) Spatial distribution of sensitive regions along the promoter: The effect of a mutation at specific promoter regions on the fold expression change (\log_2 basis) is shown for OST1, TOM6, ERG6, and RNR2, respectively. Method of calculation is detailed in Supplemental Experimental Procedures. Note the change of scale between (D) and (E) and between (F) and (G). See also Figure S2.

Together, we find that OPN promoters in our data are significantly more sensitive to promoter mutations than DPN ones. We have examined the correlation of mutation sensitivity with other properties of nucleosome organization: Mutation sensitivity was positively correlated with the “fuzziness” of nucleosomes on the promoter (a measure of how delocalized these nucleosomes are, which is also characteristic of OPN promoters (Figures S2B and S2C; Tirosh and Barkai, 2008b; Zaugg and Luscombe, 2012), but sensitivity was not significantly correlated with the length of nucleosome-free DNA (quantified by number of base pairs having a low nucleosome occupancy signal; Figures S2D and S2E). Furthermore, there was an inverse correlation between total nucleosome occupancy on the promoter and mutation sensitivity, albeit only at a narrow range of expression thresholds (Figures S2F–S2G). Notably, mutation sensitivity was independent of the pre-mutation expression level (Figure S2H).

Another potential source for this increased sensitivity could be the larger number of transcription factor binding sites encoded

by OPN promoters (Field et al., 2008; Tirosh and Barkai, 2008b). To examine that, we removed sequences mutated in known binding sites (based on ChIP-chip data from MacIsaac et al. [2006]) from our analysis, yet the mutation sensitivity remained highly correlated with the OPN measure (Figure 2B). Moreover, mutation sensitivity was not correlated with the total length of binding sites, as predicted by sequence-based TF-recognition motifs or transcription factor ChIP data (Figures 2C and S2I, respectively) nor with the number of binding sites (Figure S2J). Thus, nucleosome arrangement correlates with the mutation sensitivity independently of the content of transcription factor binding sites.

Previous studies have found that expression from OPN promoters tends to diverge more between species (Tirosh and Barkai, 2008b; Choi and Kim, 2009). We thus compared our estimate of mutation sensitivity to the measured divergence of gene expression across related yeast species (Figure S2K). Surprisingly, we did not find a correlation between the two parameters,

which can be explained if divergence in expression in the tested species is driven mostly by *trans*-related factors, rather than mutations in *cis* (Tirosh et al., 2009b). Another property of OPN promoters is their higher expression noise, which is captured in the measure of burst size (population variance over the mean expression) (Field et al., 2008; Tirosh and Barkai, 2008b; Choi and Kim, 2009). We estimated the burst size of our promoters (G.H., R. Bar-Ziv, D. Rosin, N. Tokuriki, D. Tawfik, M.O., and N.B., unpublished data) and found that this parameter has no significant correlation with mutation sensitivity (Figure S2L). A possible cause for the lack of correlation is that promoters with the highest burst size are OPN promoters with a TATA box, whereas the OPN promoters with the highest mutation sensitivity do not necessarily have a TATA box.

The Distribution of Regions that Affect Expression Differs between OPN and DPN Promoters

To define the promoter regions affecting expression, we focused on two high-sensitivity OPN promoters (ERG6 and RNR2) and two low-sensitivity DPN ones (OST1 and TOM6) and sequenced over ~250 variants from each library (Table S1). We approximated the sensitivity of each promoter region assuming that mutation effects are additive and that neighboring base pairs exert similar effects (Supplemental Experimental Procedures). Notably, in the two DPN promoters, regions affecting expression were clustered in just a few hotspots (Figures 2D and 2E). In contrast, sensitive regions were widely distributed across the entire OPN promoters (Figures 2F and 2G). Moreover, hotspots in the insensitive promoters were associated with known regulatory elements, whereas regions affecting expression in the highly sensitive promoters were only loosely linked to known regulatory sites (based on ChIP).

Mutation Sensitivity Can Be Modulated through Binding Sites of a Nucleosome Remodeler

Next we attempted to modulate the mutation sensitivity by minimum perturbations to promoter sequence. Reb1 is a DNA-binding factor that can antagonize nucleosome formation, leading to establishment of an NFR (Angermayr et al., 2003; Raisner et al., 2005; Hartley and Madhani, 2009; Bai et al., 2011). By mutating one to three base pairs in the highly sensitive ERG6, RNR2, and HXT2 OPN promoters (lacking NFR), we engineered a consensus Reb1 binding site at the proximal promoter region (Table S4). Expression was only slightly changed by these mutations (with the exception of one case where a TATA box was also eliminated; Table S4). Next, we compared the mutation sensitivity of the engineered promoters by repeating the analysis above (Tables S2 and S6). Indeed, mutation sensitivity was significantly reduced (Figures 3A and S3A–S3F), strongly suggesting that nucleosome organization is causal for high mutation sensitivity. Finally, we repeated these experiments with promoters in which we mutated the TATA box, an additional element correlated with expression divergence between species (Tirosh et al., 2006, 2009a; Landry et al., 2007; Tirosh and Barkai, 2008b). Promoters bearing these mutations, as well as other unrelated mutations as control (Table S4), did not display reduced mutation sensitivity (HXT2 in Figures 3B and S3G; GPD1 and HSP150 in Figures S3H and S3I).

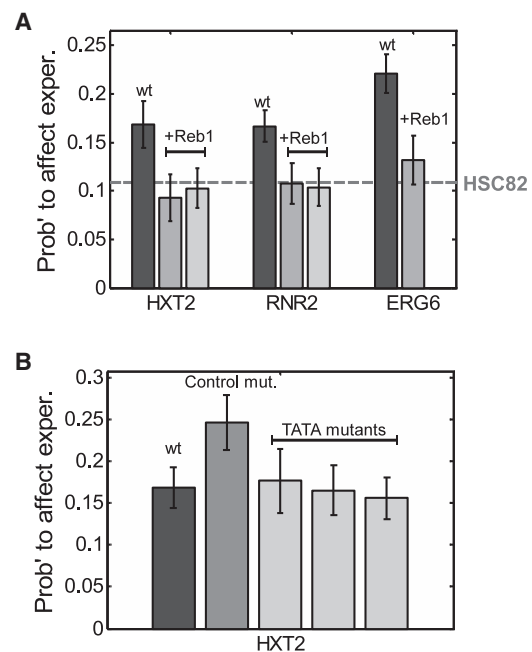


Figure 3. Reb1 Binding Site Reduces Sensitivity to Random Mutations

(A) Mutation sensitivity of original and engineered promoters: Probability of a mutation to change expression beyond an $\text{abs}(\log_2)$ expression threshold of 0.35, for original promoters and their modified version with Reb1. Results were consistent over a range of thresholds (Figure S3). Position of Reb1 binding site from left to right: HXT2: –250, –141; RNR2: –193, –247; ERG6: –167 (for sequences, see Table S4). Dashed gray line is the mutation sensitivity of HSC82, a promoter with intermediate nucleosome occupancy levels (Figures 1F, 2A, and Table 1). Error bars are based on bootstrapping (Experimental Procedures).

(B) Mutations in TATA do not reduce sensitivity to random mutations: Probability of a mutation to change expression beyond an $\text{abs}(\log_2)$ expression threshold of 0.35, for the original HXT2 promoter and versions with a mutation in the TATA box. Also shown is a control mutant with random mutations that lowered the expression (Table S4). TATA mutants (from left to right): TATAAAcA, TATgAAAA, TATAgAAA (see also Table S4). Data is tabulated in Table S5. See also Figure S3.

DISCUSSION

Gene expression evolution was studied extensively by comparing the transcription profiles of related species (Tirosh et al., 2009a). Changes observed in such comparisons are the result of long-term evolutionary dynamics, consisting of drift and selection, and depend, among other factors, on the frequency by which mutations arise, the effect of mutations on gene expression, and the different selection forces that may act on gene expression. In this study we described an alternative approach that singles out the effect of (random) mutations in gene promoter on gene expression while eliminating selection forces. Notably, we find that the probability that a random mutation will change expression differs greatly between promoters. In fact, mutation sensitivity changed by up to 10-fold within our dataset of 22 promoters.

Differences in mutation sensitivity were not explained by the abundance of binding sites. Rather, they were strongly

correlated with the organization of nucleosomes along the promoter. In fact, we observe a linear correlation between mutation sensitivity and the occupancy of nucleosomes at the proximal promoter region. Most notably, we were able to modulate this sensitivity—significantly reducing it—by mutating only one to three base pairs in the promoter sequence, thereby introducing a Reb1 binding site in the proximal region of the promoter. Reb1 was shown to block nucleosome binding and is therefore likely to reduce mutation rate by changing the organization of the promoter nucleosomes. While our results are based on only 22 promoters, the strong correlation (and practically the only correlation) that we observe between mutation sensitivity and nucleosome occupancy ratio, together with our ability to modify mutation sensitivity through nucleosome-disturbing sites, suggests a role for nucleosome organization in determining the effect of mutations.

The mechanistic basis for increased mutation sensitivity of OPN promoters is not clear. A revealing observation from our work was that the distribution across the promoter of mutations with an effect on expression varied between DPN and OPN promoters, being tightly localized to specific regulatory positions in the former, but widely distributed along the full promoter region in the latter. This suggests that expression from DPN-type promoters is regulated by well-defined transcription factor binding sites, whereas transcription from OPN promoters requires concerted action from many promoter regions. We can hypothesize several explanations for this phenomenon: (1) In OPN promoters the nucleosome is much more delocalized than in DPN promoters. It is possible that in such a case even few mutations in many possible sites can have a large impact on the stability of the nucleosome, thereby shifting the competition between transcription factor and the nucleosome binding. (2) It is possible that activation of OPN promoters relies on weak cooperative binding of many transcription factors and chromatin regulators (Tanay, 2006). Thus, the binding sites that have the highest affinity (as predicted from PSSM or ChIP) may not have the strongest effect on transcription, yet there are many small effects from low-affinity binding sites all throughout the promoter. (3) Another possibility is that in OPN promoters, which have a more bendable structure, seemingly distant regions of the promoter actually act together due to formation of loops, and thus mutations in distant regions are coupled in their effect. Additional experiments (such as DNase footprinting or ChIP of nucleosomes and relevant transcription factors) need to be performed in order to corroborate one or more of these models.

Interestingly, although the OPN architecture and the existence of a TATA box almost exclusively appear together (Field et al., 2008; Tirosch and Barkai, 2008b), our results suggest that their effect on the promoter is different. Despite the large effects of nucleosome occupancy on mutation sensitivity, we find that disrupting the TATA box impacts expression level without significant changes to mutation sensitivity. This is consistent with results from a study on synthetic promoters (Mogno et al., 2010) suggesting that the effect of the TATA box on expression is independent of upstream binding sites.

A clearer picture on the relationship between chromatin, transcription factor binding, and expression is yet to emerge. Yet

regardless of the mechanism, the fact that promoters differ in their mutation sensitivity according to nucleosome architecture may be beneficial for maintaining the expression of some genes robust against mutations while allowing others to rapidly evolve.

EXPERIMENTAL PROCEDURES

Generation of Mutant Libraries

Promoter fragments (400 bp upstream to ATG) were cloned from the yeast genome and subjected to random mutagenesis using the GeneMorph II Random Mutagenesis kit (Stratagene). PCR fragments were cloned into a specifically designed plasmid. The plasmid contained a selection marker and sequences that are homologous to genomic regions in a yeast strain with an integrated YFP. Plasmid was linearized and transformed into the yeast. Single colonies were collected and kept for further analysis. Complete details of plasmids, yeast strains, PCR, and cloning procedure are found in [Supplemental Experimental Procedures](#).

Measurements of Expression Using Flow Cytometry

Full details of measurement procedure are given in the [Supplemental Experimental Procedures](#). Briefly, yeast libraries were replicated into 150 μ l YPD in U-bottom 96-well plates and grown overnight at 30°C. Saturated cultures were resuspended and diluted \sim 1:150 into prewarmed 96-well plates with 130 to 150 μ l SC –His media. Diluted cells were grown for \sim 5.5 to 7 hr at 30°C with vigorous shaking, and fluorescence was measured by flow cytometry on the BD LSR II system (BD Biosciences) with a High Throughput Sampler extension (HTS). Excitation wavelength was 488 nm and emission was collected via 525/50 filter. Complete details of analysis and normalization procedures are provided in the [Supplemental Experimental Procedures](#).

Calculating Mutation Sensitivity

We used the combined data of sequence and expression to calculate the probability that a single random mutation will affect gene expression. This was done by fitting the data to a probabilistic model that depends on two parameters: (1) the mutation-independent background probability for change in expression, P_{back} ; (2) the mutation sensitivity P_{aff} , defining the probability that a single mutation will affect expression. The model assumes that mutation effects are independent. Hence, the probability P_m that a sequence containing m mutations will differ in expression level is given by:

$$P_m = 1 - (1 - P_{\text{back}})(1 - P_{\text{aff}})^m. \quad (1)$$

If we have N sequence variants with m mutations, the probability that n of them will change expression is given by the binomial distribution:

$$P_{\text{tot}}(m) = \binom{n}{N} P_m^n (1 - P_m)^{N-n}. \quad (2)$$

From our data, we find n and N for each value of m . The values for P_{back} and P_{aff} are estimated by maximizing the sum $\sum_m \log(P_{\text{tot}})$. Note that the value of n depends on the threshold that defines the significant change in expression. The analysis was therefore repeated for different threshold values showing consistent results.

In order to avoid potential sequencing artifacts in counting mutations, we only considered mutations with high sequencing quality scores ([Supplemental Experimental Procedures](#)). For the group with $m = 0$ mutations we only took into account sequences with high quality scores for all bases. As is shown in [Figures S4A and S4B](#), the fit of the data to the model was generally very good. For high expression thresholds the low number of variants that affect expression reduced the quality of the fit.

Error bars were estimated by a bootstrapping method. Bootstrapped sets of n and N for each number of mutations were drawn randomly with replacement from our original dataset. Then, P_{aff} was re-estimated for each bootstrapped set using the maximum likelihood procedure above, while maintaining P_{back} constant. The procedure was repeated 200 times, and standard deviation was calculated from the bootstrapped P_{back} values.

SUPPLEMENTAL INFORMATION

Supplemental Information includes four figures, seven tables, Supplemental Experimental Procedures, and Supplemental References and can be found with this article online at [doi:10.1016/j.molcel.2012.02.019](https://doi.org/10.1016/j.molcel.2012.02.019).

ACKNOWLEDGMENTS

We thank the lab of Dan Tawfik, and in particular Nobuhiko Tokuriki, for assistance with the cloning procedure. We thank Einat Dayan for technical assistance. We thank the lab of Eran Segal for AW1 yeast strain and protocols. Furthermore, we thank Itay Tirosh, Ilya Soifer, and all our lab members for helpful discussions. G.H. would like to express his gratitude to N. Hornung Broudo for support and understanding. This work was supported by NIH (P50GM068763), by the ERC, The Israel Science Foundation, and by the Hellen and Martin Kimmel award for innovative investigations. N.B. is the incumbent of the Lorena Greenberg Scherzer Professorial chair.

Received: September 19, 2011

Revised: January 14, 2012

Accepted: February 28, 2012

Published online: March 29, 2012

REFERENCES

- Angermayr, M., Oechsner, U., and Bandlow, W. (2003). Reb1p-dependent DNA bending effects nucleosome positioning and constitutive transcription at the yeast profilin promoter. *J. Biol. Chem.* 278, 17918–17926.
- Badis, G., Chan, E.T., van Bakel, H., Pena-Castillo, L., Tillo, D., Tsui, K., Carlson, C.D., Gossett, A.J., Hasinoff, M.J., Warren, C.L., et al. (2008). A library of yeast transcription factor motifs reveals a widespread function for Rsc3 in targeting nucleosome exclusion at promoters. *Mol. Cell* 32, 878–887.
- Bai, L., Ondracka, A., and Cross, F.R. (2011). Multiple sequence-specific factors generate the nucleosome-depleted region on CLN2 promoter. *Mol. Cell* 42, 465–476.
- Choi, J.K., and Kim, Y.-J. (2009). Intrinsic variability of gene expression encoded in nucleosome positioning sequences. *Nat. Genet.* 41, 498–503.
- Field, Y., Kaplan, N., Fondufe-Mittendorf, Y., Moore, I.K., Sharon, E., Lubling, Y., Widom, J., and Segal, E. (2008). Distinct modes of regulation by chromatin encoded through nucleosome positioning signals. *PLoS Comput. Biol.* 4, e1000216.
- Hartley, P.D., and Madhani, H.D. (2009). Mechanisms that specify promoter nucleosome location and identity. *Cell* 137, 445–458.
- Landry, C.R., Lemos, B., Rifkin, S.A., Dickinson, W.J., and Hartl, D.L. (2007). Genetic properties influencing the evolvability of gene expression. *Science* 317, 118–121.
- MacIsaac, K.D., Wang, T., Gordon, D.B., Gifford, D.K., Stormo, G.D., and Fraenkel, E. (2006). An improved map of conserved regulatory sites for *Saccharomyces cerevisiae*. *BMC Bioinformatics* 7, 113.
- Mogno, I., Vallania, F., Mitra, R.D., and Cohen, B.A. (2010). TATA is a modular component of synthetic promoters. *Genome Res.* 20, 1391–1397.
- Raisner, R.M., Hartley, P.D., Meneghini, M.D., Bao, M.Z., Liu, C.L., Schreiber, S.L., Rando, O.J., and Madhani, H.D. (2005). Histone variant H2A.Z marks the 5' ends of both active and inactive genes in euchromatin. *Cell* 123, 233–248.
- Soskine, M., and Tawfik, D.S. (2010). Mutational effects and the evolution of new protein functions. *Nat. Rev. Genet.* 11, 572–582.
- Tanay, A. (2006). Extensive low-affinity transcriptional interactions in the yeast genome. *Genome Res.* 16, 962–972.
- Tirosh, I., and Barkai, N. (2008a). Evolution of gene sequence and gene expression are not correlated in yeast. *Trends Genet.* 24, 109–113.
- Tirosh, I., and Barkai, N. (2008b). Two strategies for gene regulation by promoter nucleosomes. *Genome Res.* 18, 1084–1091.
- Tirosh, I., Weinberger, A., Carmi, M., and Barkai, N. (2006). A genetic signature of interspecies variations in gene expression. *Nat. Genet.* 38, 830–834.
- Tirosh, I., Barkai, N., and Verstrepen, K.J. (2009a). Promoter architecture and the evolvability of gene expression. *J. Biol.* 8, 95.
- Tirosh, I., Reikhav, S., Levy, A.A., and Barkai, N. (2009b). A yeast hybrid provides insight into the evolution of gene expression regulation. *Science* 324, 659–662.
- Tsui, K., Dubuis, S., Gebbia, M., Morse, R.H., Barkai, N., Tirosh, I., and Nislow, C. (2011). Evolution of nucleosome occupancy: conservation of global properties and divergence of gene-specific patterns. *Mol. Cell. Biol.* 31, 4348–4355.
- Vinces, M.D., Legendre, M., Caldara, M., Hagihara, M., and Verstrepen, K.J. (2009). Unstable tandem repeats in promoters confer transcriptional evolvability. *Science* 324, 1213–1216.
- Zaugg, J.B., and Luscombe, N.M. (2012). A genomic model of condition-specific nucleosome behavior explains transcriptional activity in yeast. *Genome Res.* 22, 84–94.