

## The discrepancy between chromatin factor location and effect

Tineke L. Lenstra and Frank C.P. Holstege\*

Molecular Cancer Research; University Medical Centre Utrecht; Utrecht, The Netherlands

The influence of chromatin on many cellular processes is well appreciated. Much has been learned by studying the role of chromatin remodeling and modifying complexes on individual genes. The seemingly straightforward models that inevitably arise from such studies are challenged by genome-wide analyses. Two recent studies in *Saccharomyces cerevisiae* provide unprecedented coverage of both the genome-wide location and the effect on gene expression for the majority of chromatin factors. Comparison of the overlap between location and expression effects reveals a large disconnect, with on average only 2.5% of occupied genes showing changes in expression. It is also interesting that only 24% of all expression effects are associated with chromatin factor occupancy. The large difference between location and effect likely reflects general properties inherent to regulation of gene expression through chromatin in yeast. Explanations for the discrepancy include gene-specific properties that exert a requirement for certain factors only on specific genes, as well as functional redundancy, whereby loss of a particular factor is compensated by others that function in a distinct but nevertheless compensatory manner. Since the majority of chromatin factor perturbations do show significant effects on specific subsets of genes, this implies the presence of different types of gene-specific properties that determine which chromatin factors a particular gene requires for proper expression. Understanding these gene-specific properties should be the focus of future studies aimed at understanding regulation of gene expression through chromatin.

### Introduction

Packaging of DNA into a higher-order structure called chromatin has an enormous impact on eukaryotes. It influences all processes that occur on DNA, such as replication, repair and transcription. Enormous progress has been made in recent years to characterize the factors that interact with chromatin. Chromatin is not a static structure but is dynamically regulated by the chromatin machinery. Chromatin remodelers translocate nucleosomes, thereby controlling accessibility of DNA for binding by other proteins. Chromatin modifiers covalently attach a variety of different modifications, such as methyl, acetyl or phospho-groups, to the histone tails that protrude from nucleosomes. Apart from influencing chromatin condensation, these modifications form binding platforms for other regulatory factors that can in turn affect downstream events.<sup>1–6</sup>

The observation that different factors have different affinities for different (combinations of) chromatin modifications led initially to the elegant histone code theory,<sup>1</sup> now widely recognized as being too simple.<sup>3–6</sup> As a model, the histone code theory has nevertheless been extremely useful, for example in driving efforts to characterize histone modifications, to determine which factors lay down these marks and to discover which factors bind to them. These efforts, in combination with work on nucleosome remodeling complexes, have uncovered a bewildering complexity of different interactions on chromatin. The functional consequences of these interactions have most often been studied in detail on individual genes.

**Keywords:** chromatin, epigenetics, genome-wide, promoter occupancy, transcription

Submitted: 12/23/11

Revised: 01/26/12

Accepted: 01/28/12

<http://dx.doi.org/10.4161/nucl.19513>

\*Correspondence to: Frank C.P. Holstege;  
Email: F.C.P.Holstege@umcutrecht.nl

Extra View to: Lenstra TL, Benschop JJ, Kim T, Schulze JM, Brabers NA, Margaritis T, et al. The specificity and topology of chromatin interaction pathways in yeast. *Mol Cell* 2011; 42:536–49; PMID:21596317; <http://dx.doi.org/10.1016/j.molcel.2011.03.026>

Although this is pivotal for understanding mechanism, the gene-specific nature of such studies confounds systematic testing of general models of gene expression regulation.

The availability of genome-wide technologies, coupled to an increased efficiency in performing such experiments, is now making it possible to systematically investigate the properties of the entire chromatin machinery, for example across the entire genome of the relatively less complex model organism *Saccharomyces cerevisiae*. Here we discuss the outcome of two recent studies, one investigating the location of chromatin factors on DNA,<sup>7</sup> the other investigating the effect on gene expression.<sup>8</sup> The results of both studies are compared. This reveals the enormous disconnect between the globally similar location patterns of chromatin factors across the genome and the specific effects observed on gene expression upon their removal. This disconnect is quite general and several explanations are discussed. The analyses lead to proposals regarding regulation of gene expression through chromatin. These proposals evoke the presence of a diverse set of different gene-specific properties that dictate chromatin factor dependency and therefore specificity. Although speculative, the proposals are supported by the few examples for which the disconnect between location and effect is understood. The ideas presented here therefore provide direction for future studies aimed at understanding how regulation of specific effects on gene expression is nevertheless achieved by globally present chromatin factors.

### Specific Gene Expression Effects Upon Loss of Chromatin Factors

The effect on gene expression upon removal of individual chromatin factor genes has recently been surveyed in *S. cerevisiae*.<sup>8</sup> This yeast has approximately 221 genes that encode chromatin machinery components such as chromatin remodeling and chromatin modifying complexes, histone chaperones, nucleosome assembly factors and transcriptional coregulatory complexes. DNA microarray analyses of gene deletion strains were successfully performed for 165 of these

components. Despite being analyzed under only a single growth condition, 80% of these gene deletion strains show significant changes in gene expression when compared with wild type strains. Besides statistical criteria, significance included a threshold of at least a 1.7-fold change for more than three genes to ensure that only robust differences were further analyzed. For the 80% of strains with expression patterns that differ from wild type, on average 116 genes change, with the largest differential profile observed for deletion of the corepressor Cyc8/Ssn6 (973 differentially expressed genes). The collection of gene expression signatures is useful for a variety of purposes. This includes studying the role of individual factors, structure-function analyses of large protein complexes and analyzing interactions between different chromatin factors. These aspects, including a network topology of all chromatin interactions, are well described in the original study.

An important general outcome is the high degree of specificity of the expression signatures, indicating that the majority of factors are only required for specific sets of genes. The same is true for interactions between complexes. With one exception, identical signatures are only found for subunits of the same protein complexes. Functional interactions between complexes are revealed in the form of partially overlapping signatures. Similar to the specificity of the signatures themselves, this indicates that particular interactions are only important for specific groups of genes. The high degree of specificity observed, both for the factors themselves, as well as for interactions between complexes, is a surprising general property that begs an explanation.

### Chromatin Factor Location

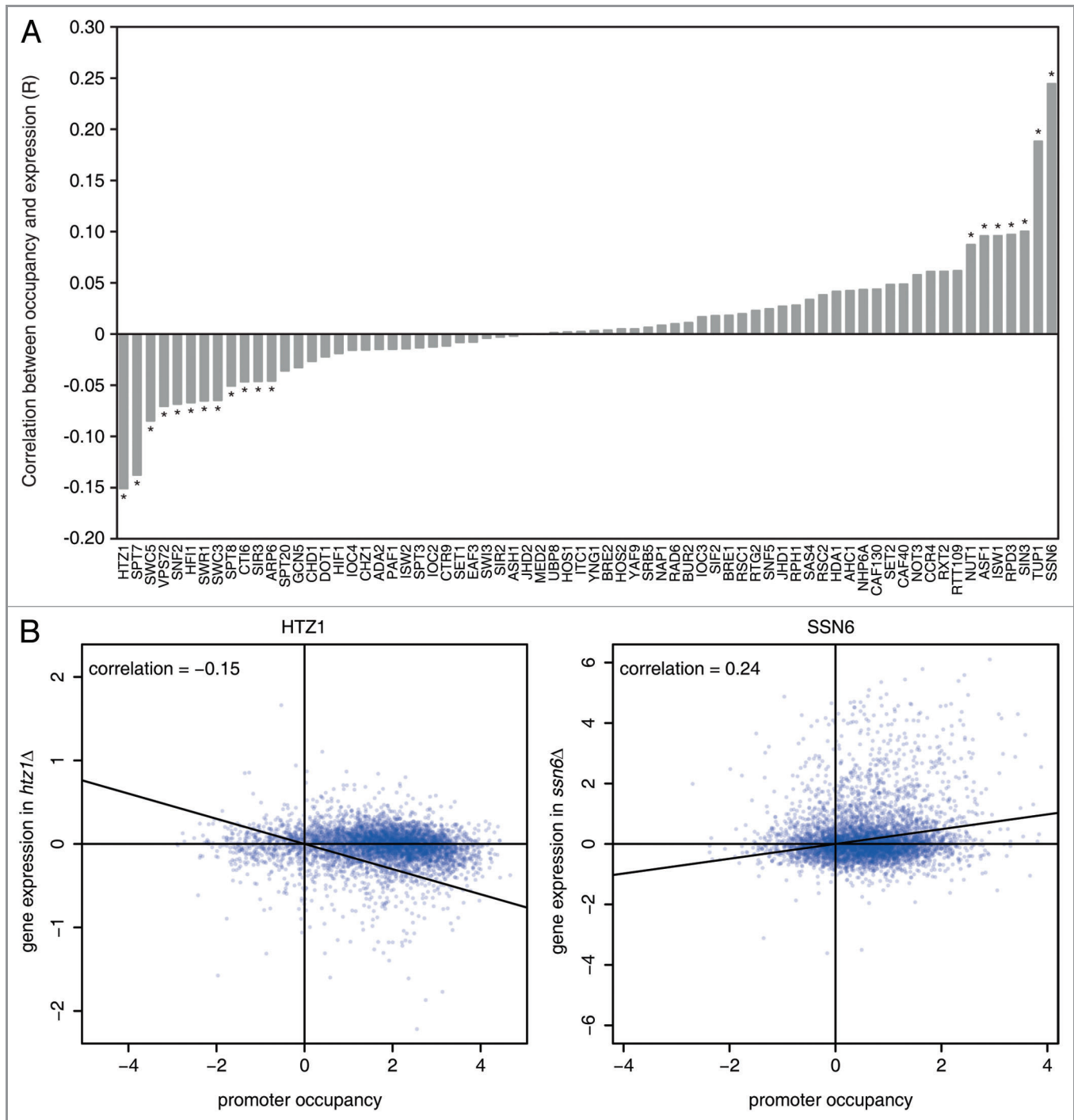
One putative explanation is location. Until recently, the genome-wide location has only been available for a handful of chromatin factors, dispersed over different studies and thereby confounding systematic comparisons. Fortunately, systematic location data for a large subset of the same factors analyzed by expression-profiling has recently become available.<sup>7</sup>

As with the expression data, the collection of location data constitutes a rich resource, suitable for different types of analyses. An important outcome of the location study is the general discrimination of genes and factors corresponding to regulation through the coregulators TFIIID or SAGA.<sup>9</sup> Interestingly, the binding patterns of the different factors are a lot less specific than the expression signatures of their deletion. As is presented below, direct comparison of the two data sets reveals striking disconnects between location and effect upon removal, demonstrating that specific location can generally be ruled out as the major explanation for the specific effects observed in gene expression upon chromatin factor removal.

### Disconnect Between Location and Effect on Expression

A total of 70 of the 165 deletions analyzed in the expression data set were also directly present in the location study. Because almost all chromatin factors are large, multi-subunit complexes, the actual overlap in factors analyzed by both studies is very high. Correlation analysis reveals a statistically significant correlation ( $p < 0.01$ ) between expression and occupancy for 19 of the 70 directly shared factors. As is often the case with genome-wide comparisons, statistically significant results can be misleading when taken on their own and usually require additional inspection to determine biological significance. Even for the statistically significant overlaps, the actual correlation values are very low, ranging from -0.15 to 0.24 (Fig. 1A). Three subunits of SAGA show a negative correlation between occupancy and gene expression, confirming their activating role on transcription.<sup>10</sup> Similar negative correlations are detected for the histone variant Htz1 (Fig. 1B) and for four subunits of its remodeling complex SWR1.<sup>11-13</sup> Positive correlations are observed for two subunits of the Rpd3L histone deacetylase complex and for the general corepressor complex Tup1 and Cyc8/Ssn6 (Fig. 1B).<sup>14</sup>

The correlations may be low because some of the expression effects are indirect, thereby negatively influencing the correspondence. Determining the overlap

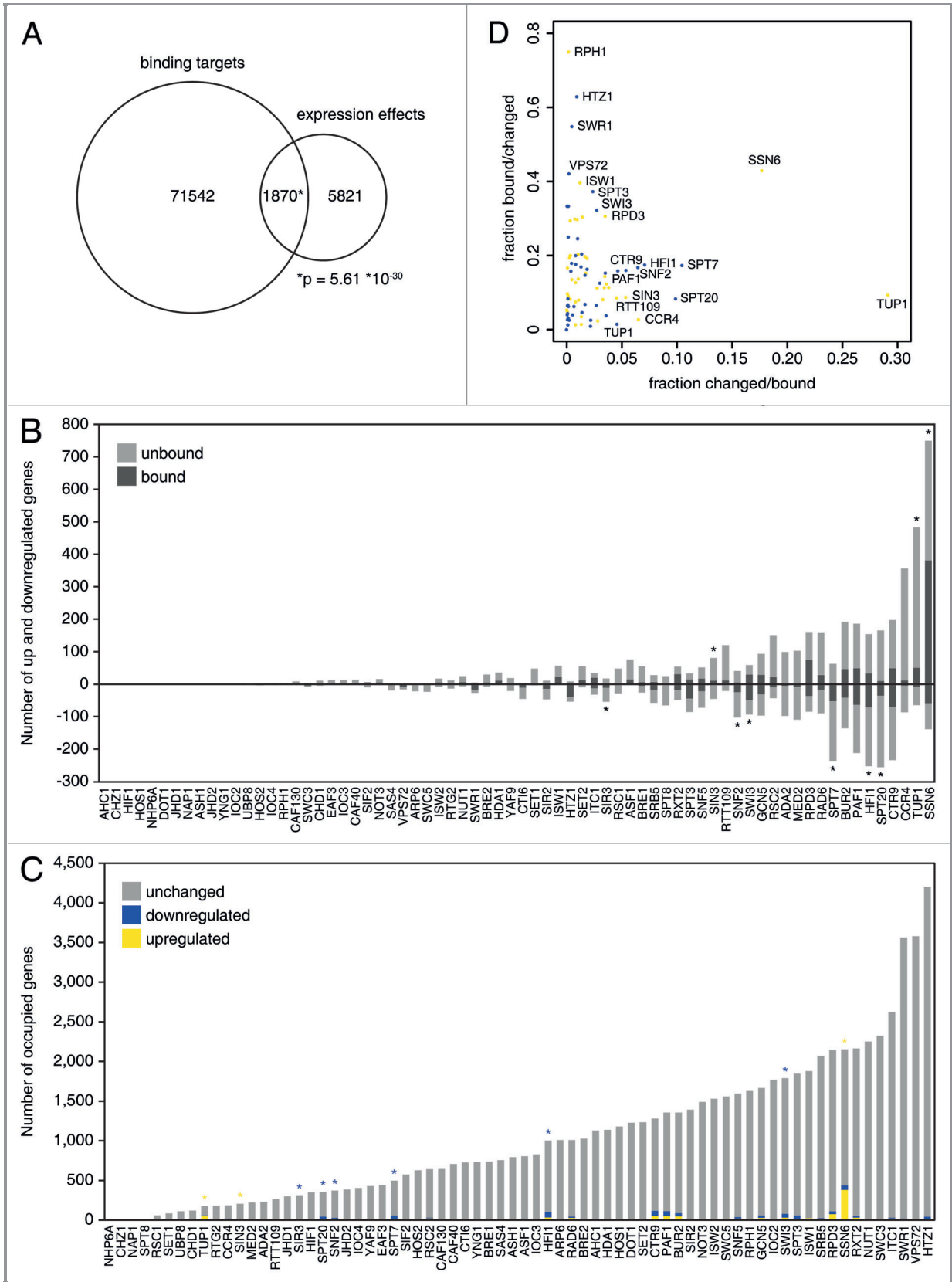


**Figure 1.** (A) Correlation analysis between occupancy and expression values for 70 chromatin regulators. For the 70 chromatin regulators that overlap between the location and expression data sets, the uncentered cosine correlation (R) is shown between occupancy values at the promoter (maximum value of UAS and TSS probe)<sup>7</sup> and mRNA expression changes upon deletion.<sup>8</sup> The correlation is based on the 5,482 genes shared between the 2 data sets. Correlation values marked by a star (\*) indicate significant correlations ( $p < 0.01$ , after Bonferroni multiple testing correction). (B) Correlation plots of promoter occupancy (x-axis) and expression (y-axis) for the highest positive (Ssn6) and highest negative (Htz1) correlation.

between expression changes in the deletion mutants (expression effects) and ChIP binding in promoter regions (binding targets) may help separate direct from indirect transcriptional changes. For the

70 chromatin regulators that overlap between the two data sets, a total of 73,412 promoter binding events are reported (5% FDR). For deletion mutants of the same factors, there are 7691

significant expression changes ( $FC > 1.7$ ,  $p < 0.05$ ). In total there are 1,870 cases of occupancy coupled to changes in expression (Fig. 2A,  $p = 5.61 \times 10^{-30}$ ). Overall, 24% of the expression effects in



**Figure 2.** For figure legend, see page 217.

**Figure 2 (See opposite page).** Overlap between occupancy and expression. (A) Venn diagram of overlap between binding targets (5% FDR) and expression effects (fold change > 1.7,  $p < 0.05$ ) for all overlapping factors. (B) Number of significant up and downregulated genes for deletion mutants of the factors shown on the x-axis. Genes that are also bound by the same factor are colored dark gray. Bars marked by a star (\*) show significant overlap between binding and up or downregulated genes (hypergeometric test,  $p < 0.01$ ). (C) Number of genes occupied by the factors shown on the x-axis. Genes that also show expression effects in the corresponding deletion mutants are colored yellow (upregulation) or blue (downregulation). Yellow and blue stars indicate significant overlap between occupancy and up or downregulated genes, respectively (hypergeometric test,  $p < 0.01$ ). (D) The fractions bound/changed and changed/bound are plotted for the upregulated genes (yellow dots) and downregulated genes (blue dots) separately.

the deletion mutants can be linked to promoter binding. Inversely, only 2.5% of the occupancies show a change in expression in the knockouts. Most chromatin factors thus occupy the promoters of many genes, but only a few of these show expression effects upon deletion.

This disconnect is also observed if the overlap between location and expression is assessed for factors individually. Assuming that most chromatin regulators are either activators or repressors, it would be expected that either the upregulated or the downregulated genes are enriched for binding. This tendency is observed for most of those regulators that do show a significant overlap, such as Ssn6 and Spt20 (Fig. 2B). In contrast, the majority of factors do not have a significant overlap and what overlap there is, is equally distributed between up and downregulated genes (Fig. 2B). In general, the fraction of occupied genes that also show expression effects is very small (Fig. 2C), because the binding study identified many more targets than the expression signatures. Deletion of either Ssn6 or Tup1 induces transcriptional changes in 18% and 29% of the occupied genes, respectively (Fig. 2D), but for the remaining regulators, less than 10% of the binding target genes are affected upon removal (Fig. 2D). Simultaneous analysis of both fractions (bound/changed and changed/bound, Figure 2D) reveals that most factors only have a (relatively) high value for one of the two fractions, with the exception of Ssn6. For example, subunits of the SWR1 complex and the histone variant Htz1 are positioned at many sites in the genome. The majority of genes with expression effects in these mutants are therefore associated with these factors, but the genes that show effects are only a small fraction of all bound genes (Fig. 2D). In conclusion, although there is some overlap between location and expression, most chromatin factors occupy the promoters of

many genes, but only a few of those genes change expression upon removal of the factor. The specificity observed in the expression signatures cannot be explained by specific binding patterns of the chromatin factors.

### Technical Differences Between Data Sets do not Explain the Disconnect

There are several differences in the setup of the two data sets, including differences in growth temperature (25°C for location data set vs. 30°C for expression data set), media (YPD vs. SC, respectively) and mating type. Removing genes whose expression is specifically affected by media and mating type (together 142 genes) only increases the correlation value significantly for the Rpd3L subunit Rxt2. The correlation values of all other factors change only slightly, and in either direction (data not shown). The microarrays used for the binding study contain three probes per gene; at the upstream activating sequence (UAS); at the transcription start site (TSS); and at the 3' end of the genes. Since for most factors it is unknown where a factor binds within a promoter, the original study used the highest value from either the UAS or the TSS probe to determine promoter occupancy. For the calculations of the correlations described above, the same data are used, and these correlations do not improve if either the UAS or the TSS binding scores are used (data not shown).

It should be noted that measurements of expression changes are based on the average mRNA levels of millions of cells. Small effects in specific cell cycle phases may be diluted too much by population averaging and are thus missed. Moreover, several regulators may act to fine tune the transcriptional response. Deleting these regulators may affect the transcriptional noise, i.e., the variation between cells, but

not the mean expression level.<sup>15-18</sup> Genes that are occupied by these regulators may be bona fide targets, but with no changes observed in the mean expression levels upon deletion. Even though these technical problems may have reduced the overlap, the scale of the disconnect (only 2.5% overlap) suggests that it is unlikely caused by technical limitations, but rather that this is a genuine biological phenomenon. Location of a chromatin factor is thus not very predictive of a transcriptional effect upon its removal.

### Possible Explanations for the Disconnect Between Location and Effect

The above comparison highlights the difference between location and effect for chromatin regulators. A similar disconnect has been noted before for histone modifications, such as H3K4, H3K36 and H3K79 methylation. Removal of the modifications results in highly specific changes in gene expression,<sup>8</sup> even though the modifications are present on almost every gene.<sup>19</sup> The systematic nature of the current studies suggests that the previously observed disconnect is not an exception but a general feature of chromatin regulation.<sup>7,8</sup> Although this comparison is limited to yeast, studies in mammalian cells have reported similar disconnects for a few individual factors.<sup>20-22</sup> It therefore remains to be seen whether the disconnect observed in yeast is general.

A single promoter is occupied by at least 75 proteins,<sup>7</sup> raising the question of what all these proteins are doing at those genes. Generally, there are two possible explanations why a certain factor would bind at the promoter of a gene, but show no effect on gene expression: either its function is simply not required for proper expression of that gene under the condition assayed, or lack of its function can be compensated by another factor upon removal. The most

well understood compensatory mechanism is redundancy.<sup>23,24</sup> In its strictest definition, this entails a take-over of function by a gene that performs an identical biochemical role. Most chromatin modifying enzymes do not have biochemically redundant partners in yeast. If compensatory mechanisms do play a role, these would therefore have to entail a form of compensation that has the same functional outcome rather than an identical biochemical mechanism. Although speculative, this explanation is supported by the many negative synthetic genetic interactions between complexes that perform biochemically distinct processes.<sup>25</sup> The alternative idea, that many chromatin factors or marks occupy many genes without an important role in gene expression, may seem unattractive in terms of efficiency or “design” but does need to be considered as a possibility when viewed in light of tinkering theories of evolution.<sup>26</sup>

Regardless of this issue of compensation vs. lack of function, the majority of chromatin factor deletions do yield some specific effects on gene expression, leaving the question of what makes those genes sensitive. The specificity of the expression profiles implies that there are certain gene-specific properties that explain why specific genes are sensitive and other genes insensitive to certain perturbations. The important question is what these properties are. In order to illustrate such properties, the few cases for which the distinguishing gene-specific properties are known are discussed.

The first example is the expression profile of *set2Δ*. *Set2* methylates histone 3 on lysine 36 in the body of actively

transcribed genes. Recognition of this modification by the Rpd3S complex results in deacetylation of the coding regions, thereby preventing cryptic initiation of transcription within genes.<sup>27-29</sup> Loss of *Set2* or Rpd3S components results in increased acetylation and inappropriate binding of the pre-initiation complex to cryptic promoters. Although this effect is genome-wide, only specific genes show increased expression in the deletion mutants. These expression profiles are enriched for long genes, simply because longer genes contain more cryptic promoters. In addition, cryptic transcripts are inherently easier to detect in genes with lower transcription frequency. This explains why specific RNAs are upregulated, even though the action of *Set2* and Rpd3S complex is independent of gene length or transcription frequency.<sup>30,31</sup>

The second example involves the histone variant H2A.Z, which is encoded by the *HTZ1* gene. This histone variant is present at the +1 position of every gene. At the subtelomers, H2A.Z prevents spreading of the Sir complex into the chromosome.<sup>32</sup> Deletion of *Htz1* does not affect all genes, but a specific gene group that is enriched for subtelomerically located genes, which reflects its specific role near the telomeres. In this example, the proximity to the chromosome ends, and not the binding of H2A.Z itself determines whether a particular gene will be affected by loss of H2A.Z.

These two examples demonstrate that various characteristics, such as the presence of cryptic promoters, itself dictated by sequence and gene length, as well as chromosomal position, can explain why

the specificity of the expression profiles differ from the global binding patterns. The specific effects observed in the other expression profiles are likely also dictated by other, similarly gene-specific properties. These properties may include different characteristics, either alone or in combination: promoter structure,<sup>33</sup> length of the nucleosome free region,<sup>34</sup> presence of a TATA box,<sup>35</sup> type of gene-specific transcription factor,<sup>36</sup> nucleosome dynamics,<sup>37</sup> the presence of introns,<sup>38</sup> regulation by antisense transcripts,<sup>39</sup> etc. The combination of these and other variables may determine which factors a particular gene needs for proper expression, similarly to what is observed for *Set2/Rpd3S*. Considering the intense investigation of chromatin factors over the last decade, it is striking that the specific expression effects of only two chromatin factors can be explained. In order to improve our understanding of transcription regulation through chromatin, it will be important to determine which gene-specific properties contribute to the specificity of the expression effects for many more chromatin factors. Future work should thus be focused on comprehending the identity of these properties and how they influence regulation of gene expression through chromatin.

#### Acknowledgments

We thank E. O Duibhir for critical reading and K. Sameith and T. Margaritis for useful discussions. Supported by the Netherlands Bioinformatics Centre (NBIC) and the Netherlands Organization of Scientific Research (NWO), grants 016108607, 81702015, 05071057, 91106009, 021002035 (T.L.L.).

#### References

1. Strahl BD, Allis CD. The language of covalent histone modifications. *Nature* 2000; 403:41-5; PMID: 10638745; <http://dx.doi.org/10.1038/47412>
2. Narlikar GJ, Fan HY, Kingston RE. Cooperation between complexes that regulate chromatin structure and transcription. *Cell* 2002; 108:475-87; PMID: 11909519; [http://dx.doi.org/10.1016/S0092-8674\(02\)00654-2](http://dx.doi.org/10.1016/S0092-8674(02)00654-2)
3. Berger SL. The complex language of chromatin regulation during transcription. *Nature* 2007; 447: 407-12; PMID:17522673; <http://dx.doi.org/10.1038/nature05915>
4. Lee J-S, Smith E, Shilatifard A. The language of histone crosstalk. *Cell* 2010; 142:682-5; PMID:20813257; <http://dx.doi.org/10.1016/j.cell.2010.08.011>
5. Campos EI, Reinberg D. Histones: annotating chromatin. *Annu Rev Genet* 2009; 43:559-99; PMID: 19886812; <http://dx.doi.org/10.1146/annurev.genet.032608.103928>
6. Rando OJ, Chang HY. Genome-wide views of chromatin structure. *Annu Rev Biochem* 2009; 78:245-71; PMID:19317649; <http://dx.doi.org/10.1146/annurev.biochem.78.071107.134639>
7. Venters BJ, Wachi S, Mavrich TN, Andersen BE, Jena P, Sinnamon AJ, et al. A comprehensive genomic binding map of gene and chromatin regulatory proteins in *Saccharomyces*. *Mol Cell* 2011; 41:480-92; PMID: 21329885; <http://dx.doi.org/10.1016/j.molcel.2011.01.015>
8. Lenstra TL, Benschop JJ, Kim T, Schulze JM, Brabers NACH, Margaritis T, et al. The specificity and topology of chromatin interaction pathways in yeast. *Mol Cell* 2011; 42:536-49; PMID:21596317; <http://dx.doi.org/10.1016/j.molcel.2011.03.026>
9. Huisinga KL, Pugh BF. A genome-wide housekeeping role for TFIID and a highly regulated stress-related role for SAGA in *Saccharomyces cerevisiae*. *Mol Cell* 2004; 13:573-85; PMID:14992726; [http://dx.doi.org/10.1016/S1097-2765\(04\)00087-5](http://dx.doi.org/10.1016/S1097-2765(04)00087-5)

10. Grant PA, Schieltz D, Pray-Grant MG, Steger DJ, Reese JC, Yates JR, 3rd, et al. A subset of TAF(II)s are integral components of the SAGA complex required for nucleosome acetylation and transcriptional stimulation. *Cell* 1998; 94:45-53; PMID:9674426; [http://dx.doi.org/10.1016/S0092-8674\(00\)81220-9](http://dx.doi.org/10.1016/S0092-8674(00)81220-9)
11. Kobor MS, Venkatasubrahmanyam S, Meneghini MD, Gin JW, Jennings JL, Link AJ, et al. A protein complex containing the conserved Swi2/Snf2-related ATPase Swr1p deposits histone variant H2A.Z into euchromatin. *PLoS Biol* 2004; 2:E131; PMID:15045029; <http://dx.doi.org/10.1371/journal.pbio.0020131>
12. Krogan NJ, Keogh M-C, Datta N, Sawa C, Ryan OW, Ding H, et al. A Snf2 family ATPase complex required for recruitment of the histone H2A variant Htz1. *Mol Cell* 2003; 12:1565-76; PMID:14690608; [http://dx.doi.org/10.1016/S1097-2765\(03\)00497-0](http://dx.doi.org/10.1016/S1097-2765(03)00497-0)
13. Mizuguchi G, Shen X, Landry J, Wu W-H, Sen S, Wu C. ATP-driven exchange of histone H2AZ variant catalyzed by SWR1 chromatin remodeling complex. *Science* 2004; 303:343-8; PMID:14645854; <http://dx.doi.org/10.1126/science.1090701>
14. Smith RL, Johnson AD. Turning genes off by Ssn6-Tup1: a conserved system of transcriptional repression in eukaryotes. *Trends Biochem Sci* 2000; 25:325-30; PMID:10871883; [http://dx.doi.org/10.1016/S0968-0004\(00\)01592-9](http://dx.doi.org/10.1016/S0968-0004(00)01592-9)
15. Newman JRS, Ghaemmaghami S, Ihmels J, Breslow DK, Noble M, DeRisi JL, et al. Single-cell proteomic analysis of *S. cerevisiae* reveals the architecture of biological noise. *Nature* 2006; 441:840-6; PMID:16699522; <http://dx.doi.org/10.1038/nature04785>
16. Raj A, van Oudenaarden A. Nature, nurture, or chance: stochastic gene expression and its consequences. *Cell* 2008; 135:216-26; PMID:18957198; <http://dx.doi.org/10.1016/j.cell.2008.09.050>
17. Bar-Even A, Paulsson J, Maheshri N, Carmi M, O'Shea E, Pilpel Y, et al. Noise in protein expression scales with natural protein abundance. *Nat Genet* 2006; 38:636-43; PMID:16715097; <http://dx.doi.org/10.1038/ng1807>
18. Maheshri N, O'Shea EK. Living with noisy genes: how cells function reliably with inherent variability in gene expression. *Annu Rev Biophys Biomol Struct* 2007; 36:413-34; PMID:17477840; <http://dx.doi.org/10.1146/annurev.biophys.36.040306.132705>
19. Pokholok DK, Harbison CT, Levine S, Cole M, Hannett NM, Lee TI, et al. Genome-wide map of nucleosome acetylation and methylation in yeast. *Cell* 2005; 122:517-27; PMID:16122420; <http://dx.doi.org/10.1016/j.cell.2005.06.026>
20. Fan Y, Nikitina T, Zhao J, Fleury TJ, Bhattacharyya R, Bouhassira EE, et al. Histone H1 depletion in mammals alters global chromatin structure but causes specific changes in gene regulation. *Cell* 2005; 123:1199-212; PMID:16377562; <http://dx.doi.org/10.1016/j.cell.2005.10.028>
21. Jiang H, Shukla A, Wang X, Chen WY, Bernstein BE, Roeder RG. Role for Dpy-30 in ES cell-fate specification by regulation of H3K4 methylation within bivalent domains. *Cell* 2011; 144:513-25; PMID:21335234; <http://dx.doi.org/10.1016/j.cell.2011.01.020>
22. Yildirim O, Li R, Hung J-H, Chen PB, Dong X, Ec L-S, et al. Mbd3/NURD complex regulates expression of 5-hydroxymethylcytosine marked genes in embryonic stem cells. *Cell* 2011; 147:1498-510; PMID:22196727; <http://dx.doi.org/10.1016/j.cell.2011.11.054>
23. Prince VE, Pickett FB. Splitting pairs: the diverging fates of duplicated genes. *Nat Rev Genet* 2002; 3:827-37; PMID:12415313; <http://dx.doi.org/10.1038/nrg928>
24. Hartman JL, 4th, Garvik B, Hartwell L. Principles for the buffering of genetic variation. *Science* 2001; 291:1001-4; PMID:11232561; <http://dx.doi.org/10.1126/science.291.5506.1001>
25. Collins SR, Miller KM, Maas NL, Roguev A, Fillingham J, Chu CS, et al. Functional dissection of protein complexes involved in yeast chromosome biology using a genetic interaction map. *Nature* 2007; 446:806-10; PMID:17314980; <http://dx.doi.org/10.1038/nature05649>
26. Jacob F. Evolution and tinkering. *Science* 1977; 196:1161-6; PMID:860134; <http://dx.doi.org/10.1126/science.860134>
27. Carrozza MJ, Li B, Florens L, Suganuma T, Swanson SK, Lee KK, et al. Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. *Cell* 2005; 123:581-92; PMID:16286007; <http://dx.doi.org/10.1016/j.cell.2005.10.023>
28. Joshi AA, Struhl K. Eaf3 chromodomain interaction with methylated H3-K36 links histone deacetylation to Pol II elongation. *Mol Cell* 2005; 20:971-8; PMID:16364921; <http://dx.doi.org/10.1016/j.molcel.2005.11.021>
29. Keogh M-C, Kurdistani SK, Morris SA, Ahn SH, Podolny V, Collins SR, et al. Cotranscriptional set2 methylation of histone H3 lysine 36 recruits a repressive Rpd3 complex. *Cell* 2005; 123:593-605; PMID:16286008; <http://dx.doi.org/10.1016/j.cell.2005.10.025>
30. Li B, Gogol M, Carey M, Pattenden SG, Seidel C, Workman JL. Infrequently transcribed long genes depend on the Set2/Rpd3S pathway for accurate transcription. *Genes Dev* 2007; 21:1422-30; PMID:17545470; <http://dx.doi.org/10.1101/gad.1539307>
31. Lickwar CR, Rao B, Shabalin AA, Nobel AB, Strahl BD, Lieb JD. The Set2/Rpd3S pathway suppresses cryptic transcription without regard to gene length or transcription frequency. *PLoS One* 2009; 4:e4886; PMID:19295910; <http://dx.doi.org/10.1371/journal.pone.0004886>
32. Meneghini MD, Wu M, Madhani HD. Conserved histone variant H2A.Z protects euchromatin from the ectopic spread of silent heterochromatin. *Cell* 2003; 112:725-36; PMID:12628191; [http://dx.doi.org/10.1016/S0092-8674\(03\)00123-5](http://dx.doi.org/10.1016/S0092-8674(03)00123-5)
33. Filion GJ, van Bommel JG, Braunschweig U, Talhout W, Kind J, Ward LD, et al. Systematic protein location mapping reveals five principal chromatin types in *Drosophila* cells. *Cell* 2010; 143:212-24; PMID:20888037; <http://dx.doi.org/10.1016/j.cell.2010.09.009>
34. Cairns BR. The logic of chromatin architecture and remodelling at promoters. *Nature* 2009; 461:193-8; PMID:19741699; <http://dx.doi.org/10.1038/nature08450>
35. Basehoar AD, Zanton SJ, Pugh BF. Identification and distinct regulation of yeast TATA box-containing genes. *Cell* 2004; 116:699-709; PMID:15006352; [http://dx.doi.org/10.1016/S0092-8674\(04\)00205-3](http://dx.doi.org/10.1016/S0092-8674(04)00205-3)
36. Harbison CT, Gordon DB, Lee TI, Rinaldi NJ, Macisaac KD, Danford TW, et al. Transcriptional regulatory code of a eukaryotic genome. *Nature* 2004; 431:99-104; PMID:15343339; <http://dx.doi.org/10.1038/nature02800>
37. Hager GL, McNally JG, Misteli T. Transcription dynamics. *Mol Cell* 2009; 35:741-53; PMID:19782025; <http://dx.doi.org/10.1016/j.molcel.2009.09.005>
38. Nagalakshmi U, Wang Z, Waern K, Shou C, Raha D, Gerstein M, et al. The transcriptional landscape of the yeast genome defined by RNA sequencing. *Science* 2008; 320:1344-9; PMID:18451266; <http://dx.doi.org/10.1126/science.1158441>
39. Tisseur M, Kwapisz M, Morillon A. Pervasive transcription—Lessons from yeast. *Biochimie* 2011; 93:1889-96; PMID:21771634; <http://dx.doi.org/10.1016/j.biochi.2011.07.001>