Wu and Lai *BMC Systems Biology* 2015, **9**(Suppl 6):S2
http://www.biomedcentral.com/1752-0509/9/S6/S2

BMC Systems Biology

RESEARCH Open Access

Functional redundancy of transcription factors explains why most binding targets of a transcription factor are not affected when the transcription factor is knocked out

Wei-Sheng Wu*, Fu-Jou Lai

From Joint 26th Genome Informatics Workshop and Asia Pacific Bioinformatics Network (APBioNet) 14th International Conference on Bioinformatics (GIW/InCoB2015)
Tokyo, Japan. 9-11 September 2015

Abstract

Background: Biologists are puzzled by the extremely low percentage (3%) of the binding targets of a yeast transcription factor (TF) affected when the TF is knocked out, a phenomenon observed by comparing the TF binding dataset and TF knockout effect dataset.

Results: This study gives a plausible biological explanation of this counterintuitive phenomenon. Our analyses find that TFs with high functional redundancy show significantly lower percentage than do TFs with low functional redundancy. This suggests that functional redundancy may lead to one TF compensating for another, thus masking the TF knockout effect on the binding targets of the knocked-out TF. In addition, we show that seven classes of genes (lowly expressed genes, TATA box-less genes, genes containing a nucleosome-free region immediately upstream of the transcriptional start site (TSS), genes with low transcriptional plasticity, genes with a low number of bound TFs, genes with a low number of TFBSs, and genes with a short average distance of TFBSs to the TSS) are insensitive to the knockout of their promoter-binding TFs, providing clues for finding other biological explanations of the surprisingly low percentage of the binding targets of a TF affected when the TF is knocked out.

Conclusions: This study shows that one property of TFs (functional redundancy) and seven properties of genes (expression level, TATA box, nucleosome, transcriptional plasticity, the number of bound TFs, the number of TFBSs, and the average distance of TFBSs to the TSS) may be useful for explaining a counterintuitive phenomenon: most binding targets of a yeast transcription factor are not affected when the transcription factor is knocked out.

Background

The binding of transcription factors (TFs) to the promoters of their target genes is one of the most important mechanisms for transcriptional regulation of gene expression. Therefore, knowing the binding targets of TFs is helpful for understanding how cells respond to stimuli by regulating the gene expression repertoire. In 2004, Harbison et al. [1] performed ChIP-chip experiments to determine the binding targets of 203 yeast TFs in the rich media condition. Since then, many computational methods have used this TF binding dataset to reconstruct yeast transcriptional regulatory networks [2-5]. These methods are all based on one assumption: most, if not all, binding targets of a TF are regulated by this TF. This assumption is supported by three computational studies which estimate about 60% of the binding targets of a TF are indeed the regulatory targets of this TF [6-8].

However, this assumption was challenged by an experimental study conducted by Hu et al. in 2007 [9]. They performed microarray experiments to identify the

* Correspondence: wessonwu@mail.ncku.edu.tw
Department of Electrical Engineering, National Cheng Kung University, Tainan, Taiwan

differentially expressed genes in each of 263 TF knockout strains in the rich media condition. Then they compared the set of genes bound by a TF (retrieved from Harbison et al.'s study [1]) with the set of genes differentially expressed when this TF is knocked out (retrieved from their own study [9]). Surprisingly, they found that only 3% of the binding targets of a TF are affected by the knockout of this TF. That is, only 3% of the binding targets of a TF are indeed regulated by this TF.

Biologists are puzzled by this extremely low percentage and researchers have tried to explain this counter-intuitive phenomenon. Several computational studies showed that by cleaning the noises in the TF binding dataset and applying advanced statistical analysis tools for the identification of differentially expressed genes in the TF knockout effect dataset, the percentage can only be improved to 6%, indicating that data analysis issue is not the main reason that causes this extremely low percentage [9-14]. Therefore, researchers started to find biological explanations for the low percentage. Two computational studies have shown that co-expression, protein sequence homology and shared protein-protein interactions may lead to one TF compensating for another, thus masking the TF knockout effect on the binding targets of the knocked-out TF [10,15]. In this study, our goal is to find out other plausible biological explanations for the surprisingly low percentage. Our analyses suggest that one TF property (functional redundancy) may lead to one TF compensating for another, thus masking the TF knockout effect on the binding targets of the knocked-out TF. In addition, we show that seven gene properties (low expression level, lacking a TATA box, containing a nucleosome-free region immediately upstream of the transcriptional start site (TSS), low transcriptional plasticity, a low number of bound TFs, a low number of TFBSs, and a short average distance of TFBSs to the TSS) are associated with a gene being insensitive to the knockout of its promoter-binding TFs.

Methods

TF binding dataset and TF knockout effect dataset

The TF binding dataset was downloaded from Harbison et al.'s study [1]. They performed ChIP-chip experiments to determine the significantly (determined by the p-value threshold) bound genes of 203 yeast TFs in the rich media condition. The TF knockout effect dataset was downloaded from Hu et al.'s study [9]. They performed microarray experiments to identify the significantly (determined by the p-value threshold) differentially expressed genes in each of 263 TF knockout strains grown in the rich media condition, the same growth condition used in the ChIP-chip experiments conducted by Harbison et al. [1]. A previous study showed that using the p-value threshold of 0.005 yields the highest overlap between the TF binding dataset and TF knockout effect dataset [10], so we adopted 0.005 as the p-value threshold in this study. Of the 203 TFs in the TF binding dataset, 173 were also in the TF knockout effect dataset. Therefore, the binding and knockout effect data of these 173 TFs were used in this study, which contained 11374 TF-gene binding relationships and 11986 TF-gene knockout effects among 173 TFs and 4065 genes. Then we compared the set of genes bound by a TF with the set of genes differentially expressed when this TF is knocked out. Similar to previous studies [9-14], on average only 4% (453/11374) of the TF-gene binding relationships had the TF-gene knockout effects. That is, only 4% of the TF binding dataset was overlapped with the TF knockout effect dataset.

Calculation of the functional redundancy of each TF

The procedure of calculating the functional redundancy of TF t is as follows. First, calculate the functional similarity (FS) between TF t and TF q using the Jaccard similarity coefficient

$$FS(t,q) = \frac{|A_t \cap A_q|}{|A_t \cup A_q|},$$

where A_t(or A_q) is the set of functional annotation terms assigned to TF t (or TF q) according Gene Ontology database [16] and MIPS functional catalogue database [17] and $|A_t \cap A_q|$ is the number of common functional annotation terms of TF t and TF q. Note that $0 \le FS(t, q) \le 1$. $FS(t, q) = 1$ when TF t and TF q have completely the same set of functional annotation terms and $FS(t, q) = 0$ when TF t and TF q have completely different sets of functional annotation terms. Then the functional redundancy (FR) of TF t is defined as

$$FR(t) = \max_q FS(t,q). \quad (1)$$

Note that $0 \le FR(t) \le 1$. TF t would have high functional redundancy if there exists another TF q whose functions are highly similar to the functions of TF t.

Overlap percentage (OP) calculation

Following Gitter et al.'s approach [10], the percentage of the TF binding dataset that is overlapped with the TF knockout effect dataset for M (a set of TFs with some property, e.g. high functional redundancy) is calculated as

$$OP = \frac{\sum_{t \in M} |G_B(t) \cap G_K(t)|}{\sum_{t \in M} |G_B(t)|} \quad (2)$$

where G_B(t) is the set of genes significantly bound by TF t, $G_K(t)$ is the set of genes significantly affected by

the knockout of TF *t*, and $|G_B(t)|$ is the number of genes significantly bound by TF *t*.

Similarly, the percentage of the TF binding dataset that is overlapped with the TF knockout effect dataset for *N* (a set of genes with some property, e.g. high expression levels) is calculated as

$$OP = \frac{\sum_{g \in N} |T_B(g) \cap T_K(g)|}{\sum_{g \in N} |T_B(g)|} \quad (3)$$

where $T_B(g)$ is the set of TFs which significantly bind to gene g and $T_K(g)$ the set of TFs which significantly affect the expression of gene g when they are knocked out.

Results and discussion

The overlap percentage varies among different TFs and different genes

Although on average only 4% of the TF binding dataset is overlapped with the TF knockout effect dataset, the percentage actually varies among different TFs and different genes. As shown in Figure 1, the percentage for different TFs varies between 0% and 36% and the percentage for different genes varies between 0% and 100% (see Additional file 1 for details). Identifying biological features that are associated with the overlap percentage may lead to biological explanations of the surprisingly low percentage of the binding targets of a TF affected when this TF is knocked out.

Functional redundancy of TFs explains why most binding targets of a transcription factor are not affected when the transcription factor is knocked out

In order to test whether functional redundancy may lead to one TF compensating for another, thus masking the TF knockout effect on the binding targets of the knocked-out TF, let us define two sets of TFs. The first is the set of TFs with high functional redundancy, which is defined as those TFs whose functional redundancy calculated using Equation (1) are among the top X% (X = 10, 20, 30, 40 or 50) of the 173 TFs under study. The other is the set of TFs with low functional redundancy, which is defined as those TFs whose functional redundancy are among the bottom X% of the 173 TFs under study. As shown in Figure 2, TFs with high functional redundancy show significantly lower overlap percentage (calculated using Equation (2)) than do TFs with low functional redundancy, suggesting that functional redundancy may explain why most binding targets of a TF are not affected when the TF is knocked out. Note that our result is robust against different choices (10, 20, 30, 40 or 50) of X and different sources (MIPS or GO) of functional annotation terms being used.

Lowly expressed genes have lower overlap percentage

Since both ChIP-chip and TF knockout experiments were performed on the yeast cells grown in the rich media condition, we speculate that lowly expressed genes in the rich media condition have lower percentage of the TF binding dataset overlapped with the TF knockout effect dataset than do highly expressed genes. To test our speculation, let us define two sets of genes. The first is the set of lowly expressed genes, which is defined as those genes whose expression levels are among the bottom X% (X = 10, 20, 30, 40 or 50) of the 4065 genes under study. The other is the set of highly expressed genes, which is defined as those genes whose expression levels are among the top X% of the 4065 genes under study (see Additional file 2 for details). The gene expression data in the rich media condition was downloaded from Holstege et al.'s study [18] and

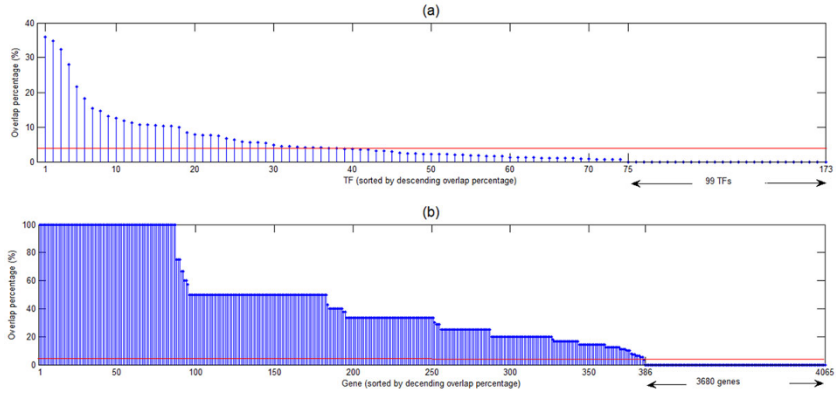

Figure 1 The overlap percentage varies among different TFs and genes. Although the overlap percentage for all 173 TFs (or all 4065 genes) under study is 4% (shown as the red line), the percentage actually varies among different TFs and different genes. (a) The overlap percentage for different TFs varies between 0% and 36%. (b) The overlap percentage for different genes varies between 0% and 100%.

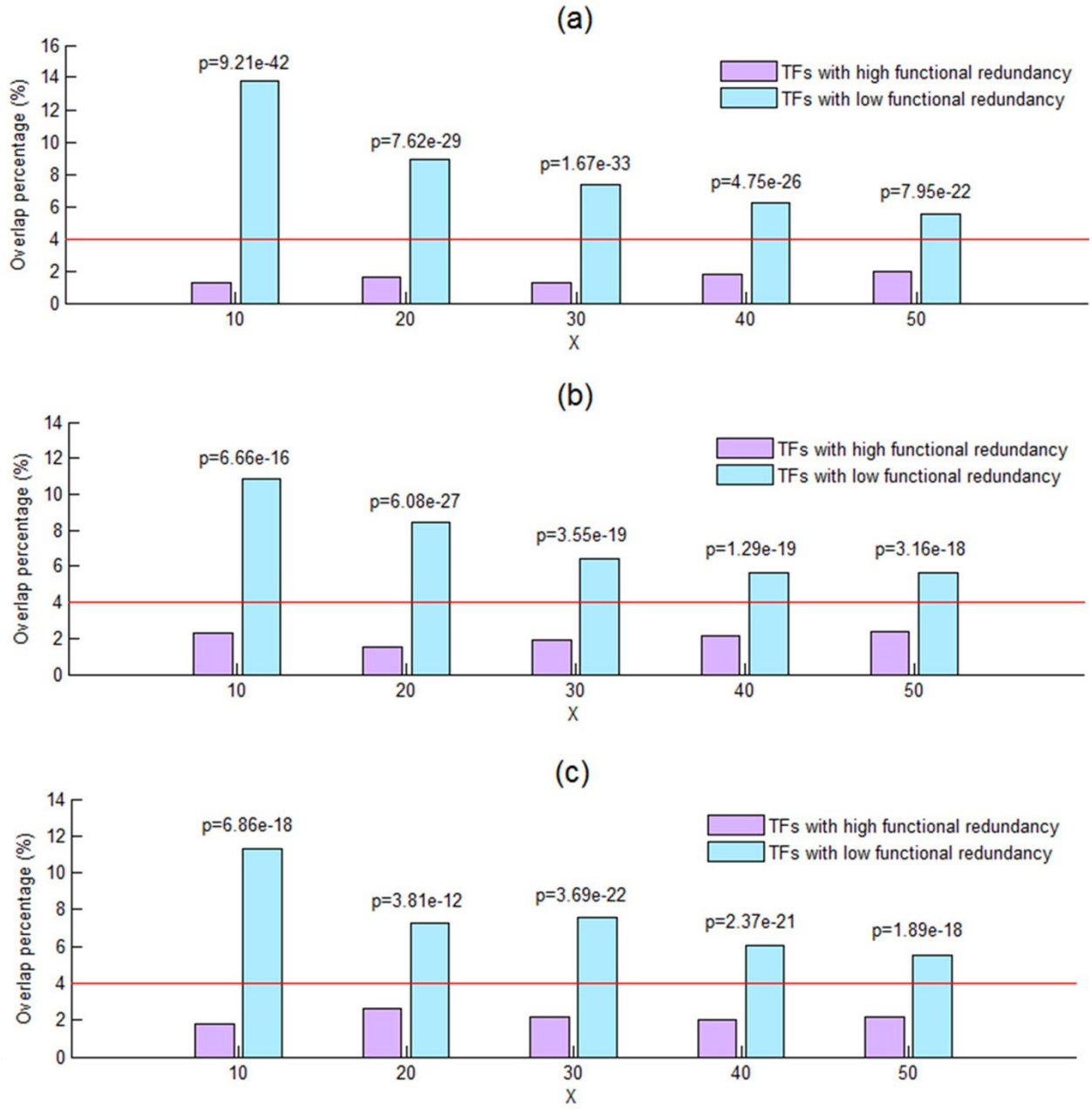

Figure 2 Functional redundancy explains the low overlap percentage. The set of TFs with high/low functional redundancy is defined as those TFs whose functional redundancy are among the top/bottom X% of the 173 TFs under study. By using the one-sided two-sample proportion test [28], we find that TFs with high functional redundancy show significantly (p-value << 0.001) lower overlap percentage than do TFs with low functional redundancy, suggesting that functional redundancy may explain why most binding targets of a TF are not affected when the TF is knocked out. Note that our result is robust against different choices (10, 20, 30, 40 or 50) of X and different sources ((a)MIPS [17], (b) GO:BP [16] or (c) GO:MF [16]) of functional annotation terms being used. The red line indicates the overlap percentage (4%) for all 173 TFs under study.

Nagalakshmi et al.'s study [19]. As shown in Figures 3a and 3b, lowly expressed genes show significantly lower overlap percentage (calculated using Equation (3)) compared with highly expressed genes, suggesting that low expression level is associated with a gene being insensitive to the knockout of its promoter-binding TFs. Note that our result is robust against different choices (10, 20, 30, 40 or 50) of X and different sources (Holstege et al.'s study or Nagalakshmi et al.'s study) of gene expression data being used.

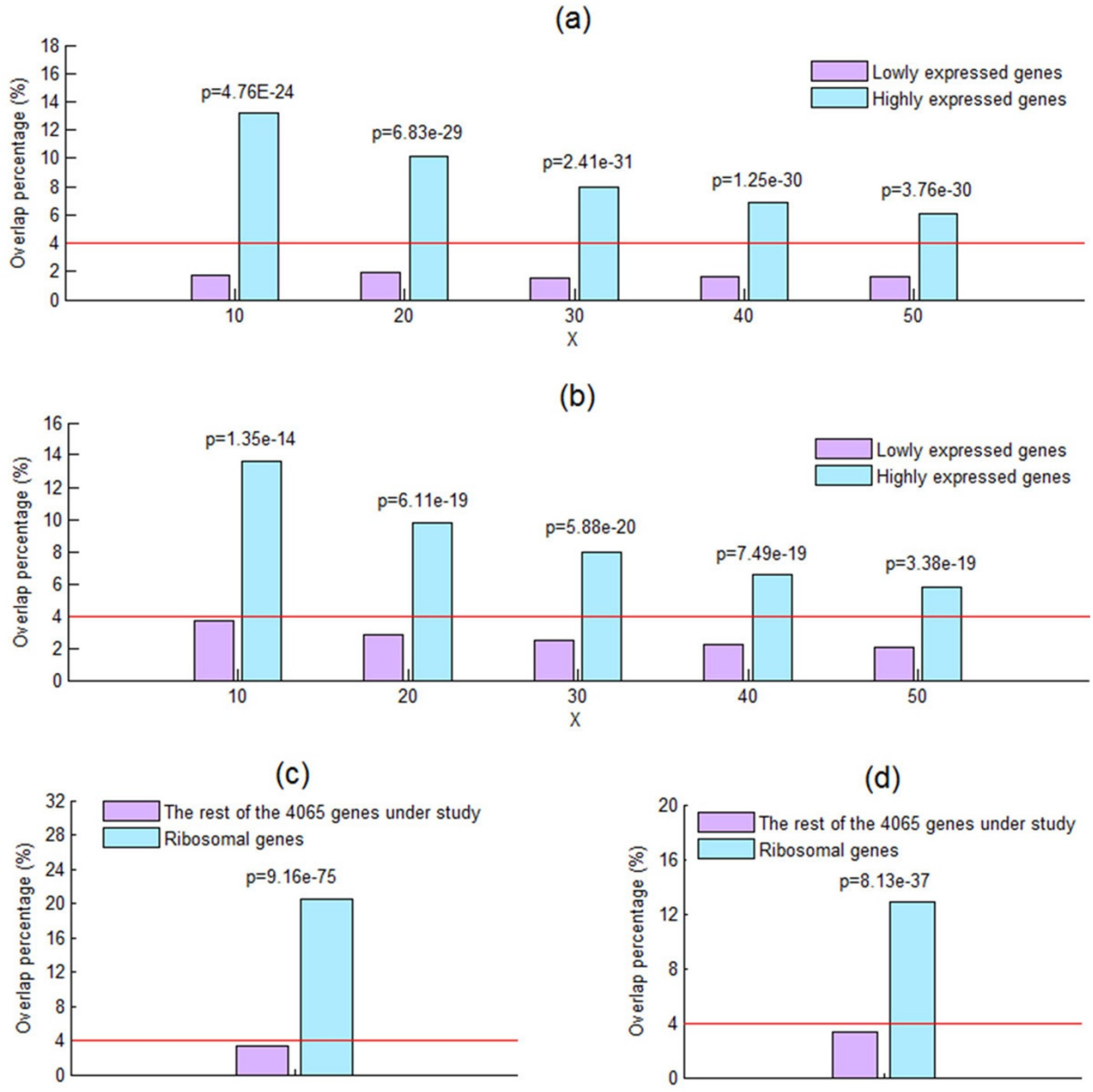

Figure 3 High expression level and ribosomal genes are associated with high overlap percentage. The set of lowly/highly expressed genes is defined as those genes whose expression levels are among the bottom/top X% of the 4065 genes under study. By using the one-sided two-sample proportion test [28], we find that lowly expressed genes show significantly (p-value << 0.001) lower overlap percentage compared with highly expressed genes, suggesting that low expression level is associated with a gene being insensitive to the knockout of its promoter-binding TFs. Note that our result is robust against different choices (10, 20, 30, 40 or 50) of X and different sources ((a) Holstege et al.'s study [18] or (b) Nagalakshmi et al.'s study [19]) of gene expression data being used. The red line indicates the overlap percentage (4%) for all 4065 genes under study. In addition, ribosomal genes show significantly (using the one-sided two-sample proportion test) higher overlap percentage compared with the rest of 4065 genes under study. This result further supports our finding that highly expressed genes show significantly higher overlap percentage. Note that our result is robust against different sources ((c) KEGG [20] or (d) MIPS [17]) of the list of ribosomal genes being used.

Ribosomal genes are known to be highly transcribed in the rich media condition. If our finding is biologically meaningful, we expect that ribosomal genes have higher overlap percentage compared with the rest of the 4065 genes under study. To test this assertion, we downloaded two lists of ribosomal genes from KEGG ribosome pathway: sce03010 [20] and MIPS functional category: 12.01.01 ribosomal proteins [17]. As expected,

ribosomal genes show significantly higher overlap percentage (calculated using Equation (3)) compared with the rest of the 4065 genes under study (see Figures 3c and 3d), thus further strengthen our finding. Note that our result is robust against different sources (KEGG or MIPS) of the list of ribosomal genes being used.

TATA box-less genes have lower overlap percentage

It is known that TATA box-less and TATA box-containing genes are distinctly regulated [21]. TATA box-less genes tend to be housekeeping genes, have a sharply peaked TF binding site (TFBS) distribution and are constitutively expressed, while TATA box-containing genes are usually associated with environmental stress responses, dispersed TFBS distribution and variably expressed under different conditions [21-25]. It is interesting to know whether these two classes of genes differ in their overlap percentage. The lists of TATA box-less genes and TATA box-containing genes were downloaded from Basehoar et al.'s study [21]. Depending on how stringent the criterion for defining a TATA box is, three possible lists of TATA box-containing genes were defined by Basehoar et al. [21]. As shown in Figure 4, TATA box-less genes show significantly lower overlap percentage (calculated using Equation (3)) compared with TATA box-containing genes, suggesting that lacking a TATA box is associated with a gene being insensitive to the knockout of its promoter-binding TFs. Note that our result is robust against different criteria of defining TATA box-containing genes.

Genes containing a nucleosome-free region (NFR) have lower overlap percentage

In yeast, the capacity to modulate gene expression upon changing conditions (i.e., transcriptional plasticity) correlates with the organization of their promoter nucleosomes [26]. Genes containing an NFR immediately upstream of the transcriptional start site (TSS) are characterized by low transcriptional plasticity, while genes lacking an NFR immediately upstream of the TSS are characterized by high transcriptional plasticity. It is interesting to know whether these two classes of genes differ in their overlap percentage. The lists of genes containing and lacking an NFR were both downloaded from Tirosh and Baikai's study [26]. As shown in Figure 5a, genes containing an NFR show significantly lower overlap percentage (calculated using Equation (3)) compared with genes lacking an NFR, suggesting that containing an NFR immediately upstream of the TSS is associated with a gene being insensitive to the knockout of its promoter-binding TFs.

It is known that genes lacking an NFR are subjected to greater regulation by specific chromatin remodelling factors than are genes containing an NFR [26]. If our finding is biologically meaningful, we expect that TFs involved in chromatin remodelling have higher overlap percentage compared with the rest of the 173 TFs under study. To test this assertion, we downloaded the list of TFs involved in chromatin remodelling from Ozonov and van Nimwegen's study [27]. As expected, TFs involved in chromatin remodelling show significantly higher overlap percentage (calculated using Equation (2)) compared with the rest of the 173 TFs under study (see Figure 5b), thus further strengthen our finding.

Genes with low transcriptional plasticity have lower overlap percentage

We have shown that two classes of genes (TATA box-less genes and genes containing an NFR) have lower overlap percentage. Since both classes of genes are known to have low transcriptional plasticity [21,26], this prompts us to speculate that genes with low transcriptional plasticity have lower percentage of the TF binding dataset overlapped with the TF knockout effect dataset

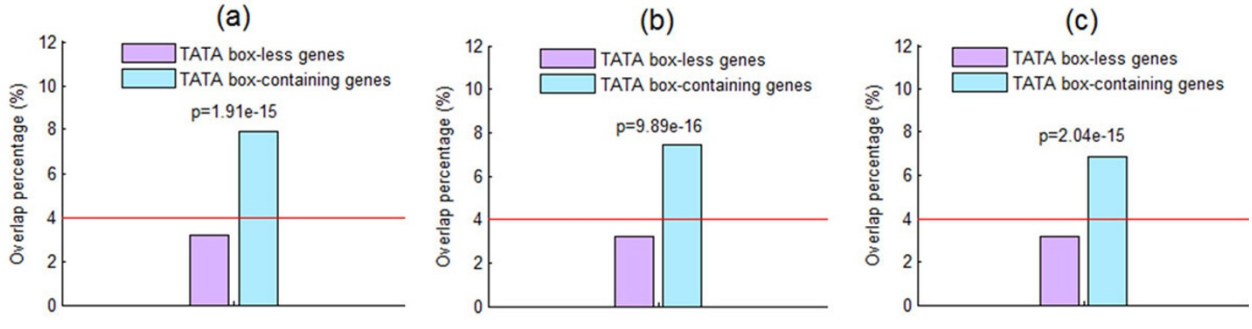

Figure 4 Lacking a TATA box is associated with a gene being insensitive to the knockout of its promoter-binding TFs. By using the one-sided two-sample proportion test [28], we find that TATA box-less genes show significantly (p-value << 0.001) lower overlap percentage compared with TATA box-containing genes, suggesting that lacking a TATA box is associated with a gene being insensitive to the knockout of its promoter-binding TFs. Note that our result is robust against three different criteria ((a), (b), and (c)) for defining TATA box-containing genes being used (see Basehoar et al.'s study [21] for more details). The red line indicates the overlap percentage (4%) for all 4065 genes under study.

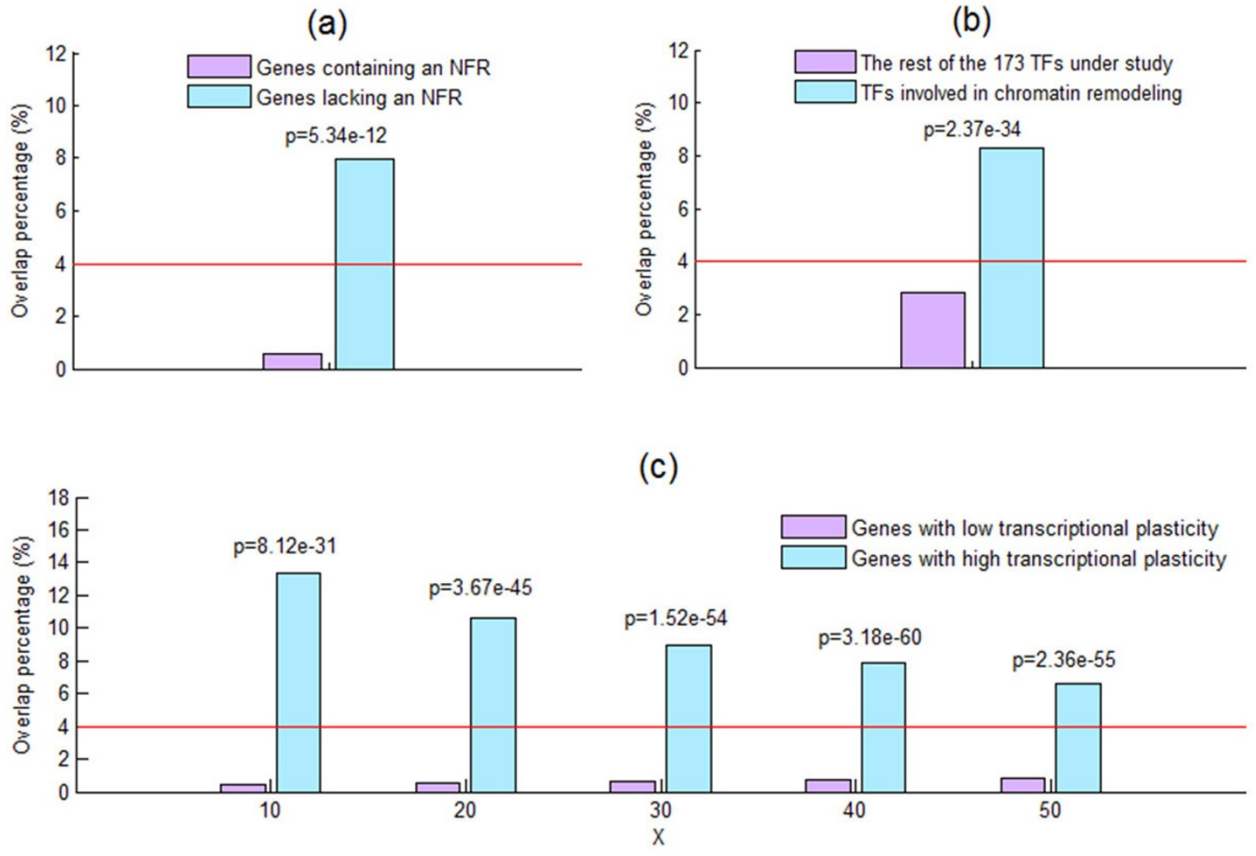

Figure 5 NFR, chromatin remodelling TFs and transcriptional plasticity are associated with the overlap percentage. (a) By using the one-sided two-sample proportion test [28], we find that genes containing an nucleosome free region (NFR) show significantly (p-value << 0.001) lower overlap percentage compared with genes lacking an NFR, suggesting that containing an NFR immediately upstream of the TSS is associated with a gene being insensitive to the knockout of its promoter-binding TFs. The red line indicates the overlap percentage (4%) for all 4065 genes under study. (b) TFs involved in chromatin remodelling show significantly (using the one-sided two-sample proportion test) higher overlap percentage compared with the rest of the 173 TFs under study. This result further supports our finding that genes lacking an NFR show significantly higher overlap percentage. (c) Low transcriptional plasticity is associated with a gene being insensitive to the knockout of its promoter-binding TFs. The set of genes with low/high transcriptional plasticity is defined as those genes whose transcriptional plasticity are among the bottom/top X% of the 4065 genes under study. By using the one-sided two-sample proportion test, we find that genes with low transcriptional plasticity show significantly (p-value << 0.001) lower overlap percentage than do genes with high transcriptional plasticity, suggesting that low transcriptional plasticity is associated with a gene being insensitive to the knockout of its promoter-binding TFs. Note that our result is robust against different choices (10, 20, 30, 40 or 50) of X being used.

than do genes with high transcriptional plasticity. To test our speculation, let us define two sets of genes. The first is the set of genes with low transcriptional plasticity, which is defined as those genes whose transcriptional plasticity are among the bottom X% of the 4065 genes under study. The other is the set of genes with high transcriptional plasticity, which is defined as those genes whose transcriptional plasticity are among the top X% of the 4065 genes under study (see Additional file 2 for details). The transcriptional plasticity each gene in the yeast genome was downloaded from Lin et al.'s study [22]. As shown in Figure 5c, genes with low transcriptional plasticity show significantly lower overlap percentage (calculated using Equation (3)) than do genes with high transcriptional plasticity, suggesting that low transcriptional plasticity is associated with a gene being insensitive to the knockout of its promoter-binding TFs. Note that our result is robust against different choices (10, 20, 30, 40 or 50) of X being used.

Several gene properties are not associated with the overlap percentage

In the previous sections, we show that four gene properties (expression level, TATA box, nucleosome, and transcriptional plasticity) are associated with the overlap percentage. Actually, five other gene properties are also tested but do not have statistically significant association with the overlap percentage. These five gene properties

include the 5'UTR length, 3'UTR length, gene essentiality, number of physical interaction partners and number of genetic interaction partners.

More analyses motivated by Cusanovich et al's study

In Cusanovich et al.'s paper [29], they reported that functional TF binding is enriched in the regulatory regions with a larger number of bound TFs and more binding sites. Moreover, functional TF binding tends to occur further from the TSS (i.e. in the enhancer regions). Motivated by their findings, we perform extra analyses and have the following three observations: (i) a low number of bound TFs in a gene, (ii) a low number of TFBSs in a gene, and (iii) a short average distance of TFBSs to the TSS in a gene are all associated with a gene being insensitive to the knockout of its promoter-binding TFs (see Additional file 3 for details).

Conclusions

This study gives a plausible biological explanation of a counterintuitive phenomenon: most binding targets of a yeast transcription factor are not affected when the transcription factor is knocked out. Our analyses find that TFs with high functional redundancy show significantly lower percentage than do TFs with low functional redundancy. This suggests that functional redundancy may lead to one TF compensating for another, thus masking the TF knockout effect on the binding targets of the knocked-out TF. In addition, identifying biological features that are associated with the overlap percentage may provide clues for finding other biological explanations. We show that seven gene properties (low expression level, lacking a TATA box, containing a nucleosome-free region immediately upstream of the transcriptional start site, low transcriptional plasticity, a low number of bound TFs, a low number of TFBSs, and a short average distance of TFBSs to the TSS) are associated with a gene being insensitive to the knockout of its promoter-binding TFs.

Additional material

Additional file 1: The details of the 4065 genes under study. For each of the 4065 genes under study, its $T_B(g)$, $T_K(g)$ and $T_B(g) \cap T_K(g)$ are provided.

Additional file 2: The details of the top/bottom X% of the expression level and transcriptional plasticity. For each top/bottom X % case, their gene names, $T_B(g)$, $T_K(g)$, and $T_B(g) \cap T_K(g)$ are provided.

Additional file 3: The details of the analyses motivated by Cusanovich et al's study.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

WSW conceived the research topic and provided essential guidance. WSW and FJL developed the method and wrote the manuscript. FJL did all the simulations. Both authors read, edited and approved the final manuscript.

Acknowledgements

This study was supported by National Cheng Kung University and Ministry of Science and Technology of Taiwan MOST-103-2221-E-006 -174 -MY2.

Declarations

The publication of this paper was funded by National Cheng Kung University and Ministry of Science and Technology of Taiwan MOST-103-2221-E-006 -174 -MY2.

This article has been published as part of *BMC Systems Biology* Volume 9 Supplement 6, 2015: Joint 26th Genome Informatics Workshop and 14th International Conference on Bioinformatics: Systems biology. The full contents of the supplement are available online at http://www.biomedcentral.com/bmcsystbiol/supplements/9/S6.

Published: 9 December 2015

References

1. Harbison CT, Gordon DB, Lee TI, Rinaldi NJ, Macisaac KD, Danford TW, *et al*: **Transcriptional regulatory code of a eukaryotic genome.** *Nature* 2004, **431(7004)**:99-104.
2. Wu WS, Li WH, Chen BS: **Computational reconstruction of transcriptional regulatory modules of the yeast cell cycle.** *BMC Bioinformatics* 2006, **7**:421.
3. Lemmens K, Dhollander T, De Bie T, Monsieurs P, Engelen K, Smets B, Winderickx J, De Moor B, Marchal K: **Inferring transcriptional modules from ChIP-chip, motif and microarray data.** *Genome Biology* 2006, **7(5)**:R37.
4. Ernst J, Vainas O, Harbison CT, Simon I, Bar-Joseph Z: **Reconstructing dynamic regulatory maps.** *Mol Syst Biol* 2007, **3**:74.
5. Wu WS, Li WH: **Identifying gene regulatory modules of heat shock response in yeast.** *BMC Genomics* 2008, **9**:439.
6. Gao F, Foat BC, Bussemaker HJ: **Defining transcriptional networks through integrative modeling of mRNA expression and transcription factor binding data.** *BMC Bioinformatics* 2004, **5**:31.
7. Wu WS, Li WH, Chen BS: **Identifying regulatory targets of cell cycle transcription factors using gene expression and ChIP-chip data.** *BMC Bioinformatics* 2007, **8**:188.
8. Lai FJ, Chang JS, Wu WS: **Identifying a transcription factor's regulatory targets from its binding targets.** *Gene Regul Syst Bio* 2010, **4**:125-133.
9. Hu Z, Killion PJ, Iyer VR: **Genetic reconstruction of a functional transcriptional regulatory network.** *Nat Genet* 2007, **39(5)**:683-687.
10. Gitter A, Siegfried Z, Klutstein M, Fornes O, Oliva B, Simon I, Bar-Joseph Z: **Backup in gene regulatory networks explains differences between binding and knockout results.** *Mol Syst Biol* 2009, **5**:276.
11. Reimand J, Vaquerizas JM, Todd AE, Vilo J, Luscombe NM: **Comprehensive reanalysis of transcription factor knockout expression data in Saccharomyces cerevisiae reveals many new targets.** *Nucleic Acids Res* 2010, **38(14)**:4768-4777.
12. Yang TH, Wu WS: **Identifying biologically interpretable transcription factor knockout targets by jointly analyzing the transcription factor knockout microarray and the ChIP-chip data.** *BMC Systems Biology* 2012, **6**:102.
13. Yang TH, Wang CC, Wang YC, Wu WS: **YTRP: a repository for yeast transcriptional regulatory pathways.** *Database (Oxford)* 2014, bau014.
14. Yang TH, Wu WS: **Inferring functional transcription factor-gene binding pairs by integrating transcription factor binding data with transcription factor knockout data.** *BMC Systems Biology* 2013, **7(Suppl 6)**:S13.
15. Dai Z, Dai X, Xiang Q, Feng J: **Robustness of transcriptional regulatory program influences gene expression variability.** *BMC Genomics* 2009, **10**:573.
16. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, *et al*: **Gene ontology: tool for the unification of biology. The Gene Ontology Consortium.** *Nat Genet* 2000, **25(1)**:25-29.
17. Ruepp A, Zollner A, Maier D, Albermann K, Hani J, Mokrejs M, *et al*: **The FunCat, a functional annotation scheme for systematic classification of proteins from whole genomes.** *Nucleic Acids Res* 2004, **32(18)**:5539-5545.

18. Holstege FC, Jennings EG, Wyrick JJ, Lee TI, Hengartner CJ, Green MR, *et al*: **Dissecting the regulatory circuitry of a eukaryotic genome.** *Cell* 1998, **95(5)**:717-728.
19. Nagalakshmi U, Wang Z, Waern K, Shou C, Raha D, Gerstein M, Snyder M: **The transcriptional landscape of the yeast genome defined by RNA sequencing.** *Science* 2008, **320(5881)**:1344-1349.
20. Kanehisa M, Goto S, Sato Y, Kawashima M, Furumichi M, Tanabe M: **Data, information, knowledge and principle: back to metabolism in KEGG.** *Nucleic Acids Res* 2014, **42(Database issue)**:D199-D205.
21. Basehoar AD, Zanton SJ, Pugh BF: **Identification and distinct regulation of yeast TATA box-containing genes.** *Cell* 2004, **116(5)**:699-709.
22. Lin Z, Wu WS, Liang H, Woo Y, Li WH: **The spatial distribution of cis regulatory elements in yeast promoters and its implications for transcriptional regulation.** *BMC Genomics* 2010, **11**:581.
23. Tirosh I, Weinberger A, Carmi M, Barkai N: **A genetic signature of interspecies variations in gene expression.** *Nat Genet* 2006, **38(7)**:830-834.
24. Landry CR, Lemos B, Rifkin SA, Dickinson WJ, Hartl DL: **Genetic properties influencing the evolvability of gene expression.** *Science* 2007, **317(5834)**:118-121.
25. Wu WS: **Different functional gene clusters in yeast have different spatial distributions of the transcription factor binding sites.** *Bioinform Biol Insights* 2011, **5**:1-11.
26. Tirosh I, Barkai N: **Two strategies for gene regulation by promoter nucleosomes.** *Genome Res* 2008, **18(7)**:1084-1091.
27. Ozonov EA, van Nimwegen E: **Nucleosome free regions in yeast promoters result from competitive binding of transcription factors that interact with chromatin remodelling modifiers.** *PLoS Comput Biol* 2013, **9(8)**:e1003181.
28. Glantz SA: *Primer of Biostatistics*. 7 edition. McGraw-Hill Medical, New York; 2011.
29. Cusanovich DA, Pavlovic B, Pritchard JK, Gilad Y: **The functional consequences of variation in transcription factor binding.** *PLoS Genet* 2014, **10(3)**:e1004226.

doi:10.1186/1752-0509-9-S6-S2
Cite this article as: Wu and Lai: **Functional redundancy of transcription factors explains why most binding targets of a transcription factor are not affected when the transcription factor is knocked out.** *BMC Systems Biology* 2015 **9**(Suppl 6):S2.