A Regulatory Role for Repeated Decoy Transcription Factor Binding Sites in Target Gene Expression

Tek-Hyung Lee, Narendra Maheshri

Corresponding author: Narendra Maheshri, MIT

Review timeline:

<table>
<thead>
<tr>
<th>Event</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Submission date</td>
<td>15 February 2011</td>
</tr>
<tr>
<td>Editorial Decision</td>
<td>31 March 2011</td>
</tr>
<tr>
<td>Resubmission received</td>
<td>10 February 2012</td>
</tr>
<tr>
<td>Editorial Decision</td>
<td>02 February 2012</td>
</tr>
<tr>
<td>Revision received</td>
<td>22 February 2012</td>
</tr>
<tr>
<td>Accepted</td>
<td>27 February 2012</td>
</tr>
</tbody>
</table>

Transaction Report:

(Note: With the exception of the correction of typographical or spelling errors that could be a source of ambiguity, letters and reports are not edited. The original formatting of letters and referee reports may not be reflected in this compilation.)

1st Editorial Decision 31 March 2011

Thank you again for submitting your work to Molecular Systems Biology. We have now heard back from the three referees who agreed to evaluate your manuscript. As you will see from the reports below, the referees find the topic of your study of potential interest. They raise, however, substantial concerns on your work, which, I am afraid to say, preclude its publication in its present form.

The reviewers acknowledged that the experiments and modeling provided a potentially interesting investigation of the effects of decoy TF binding sites on gene expression, and noted that the initial experiments raised some surprising observations (e.g. higher predicted TF avidity for the decoy arrays, and diminishing decoy copy number effects for longer arrays). They, however, had clear concerns regarding the conclusiveness of the experiments that attempt to address the mechanisms underlying these observations (in particular, both found the results presented in Fig. 4 unconvincing). The first reviewer indicates that additional experimental evidence directly measuring TF binding avidity for promoter vs decoy tetO7 sites, and for tetO7 sites with different spacer lengths, are needed.

In addition, the reviewers indicated clearly that they felt this manuscript currently lacks important methodological details, suggesting that a more complete summary of the modeling methods should be included in the results and figure legends.

Somewhat related to the previous issue, the editor requests that any revised manuscript include the mathematical model files in a common machine-readable format, as supplementary material. We strongly encourage authors to supply models in SBML when appropriate, and to submit models to a public repository like BioModels or JWS Online.
In addition to our capacity to host datasets in our supplementary information section, we provide a new functionality that allows readers to directly download the 'source data' associated with selected figure panels (e.g. <http://tinyurl.com/365zpej>). This may be particularly appropriate for the experimental data presented in this work. Guidelines have been pasted at the end of this email.

If you feel you can satisfactorily deal with these points and those listed by the referees, you may wish to submit a revised version of your manuscript. Please attach a covering letter giving details of the way in which you have handled each of the points raised by the referees. A revised manuscript will be once again subject to review and you probably understand that we can give you no guarantee at this stage that the eventual outcome will be favorable.

*PLEASE NOTE* As part of the EMBO Publications transparent editorial process initiative (see http://www.nature.com/msb/journal/v6/n1/full/msb201072.html), Molecular Systems Biology now publishes online a Review Process File with each accepted manuscript. Please be aware that in the event of acceptance, your cover letter/point-by-point document will be included as part of this file, which will be available to the scientific community. More information about this initiative is available in our Instructions to Authors. If you have any questions about this initiative, please contact the editorial office (msb@embo.org).

Yours sincerely,

Editor
Molecular Systems Biology

---------------------------------------------------------------------------

Referee comments:

Reviewer #1 (Remarks to the Author):

This report describes use of a yeast version of the synthetic tet-OFF system to study the effect on promoter function of "decoy" transcription factor (TF) binding sites. The TF used here is an E. coli Tet repressor fusion to the herpes virus VP16 transactivator (TA) domain (called tTA). The decoy sites are arrays of Tet operators separated by random sequences of 10 or 30 bp (15, 37, 67, 127 or 240 tetO copies; constructed by Sherratt and colleagues) and the promoter site tested is an array of 7 operators, apparently from Gossen and Bujard (1992), but not described here (i.e. spacing of sites not clear). Test promoters drive expression of either CFP or YFP, and read-outs are made either by FACS or microscopy. Titrating "decoy" arrays are either present on plasmids or integrated into the chromosome.

The experiments reported here largely consist of measuring promoter output as a function of decoy site copy number. The results are presented in terms of a model for the doxycycline-tTA interaction, developed by the authors previously and expanded here in Supplemental Material, which predicts that repeats will, in a number-dependent fashion, convert the dose-response curve from a graded to an "ultrasensitive" one. To fit their dose-response data to this simplified model they had to estimate the ratio of tTA to repeat number and to affinity of tetO sites in the repeats. Their model fits the data reasonably well for low tTA concentrations but systematically overestimates the competitive effect of the repeats at high tTA levels. Taken on face value the results are potentially quite interesting and may have important implications for evolution of transcription factor network function, since repeat number variation is a very frequent form of spontaneous genetic alteration. However, there are several unexplained features of the data that raise significant concerns regarding the authors' interpretation of their results.

Critique:

1. The authors' model predicts an "effective binding avidity" for the decoy sites that is several orders
of magnitude higher than that for the tetO7 array present at the promoter. The reason for this difference is not understood, but the authors present evidence (Fig. 6) that it persists even if a tetO7 site is placed adjacent to a decoy array. They also describe evidence arguing that this difference is not due to either tTA turnover through the ubiquitin system or to locus-specific interactions with the transcriptional machinery. In any event, tTA association with both decoy and promoter (tetO7) sites should be measured by quantitative chromatin immunoprecipitation (ChIP) experiments that should give a very reliable estimate of the relative in vivo occupancy of the different sites. The authors' attempts to quantify tTA binding to arrays by fluorescence microscopy (Fig. 4) are completely unconvincing and these data should be eliminated from the manuscript.

2. The concern raised above begs the question of the effect of tetO7 site arrays as decoys, where these arrays should (?) display the same binding avidity as they do in the context of the promoter (a single tetO7 site).

3. Another curious and completely unexplained observation is that the effect of decoy copy number diminishes significantly beyond 37 sites, though much less so if the arrays are present on separate plasmids. Again, analysis of this phenomenon by qPCR ChIP of tTA on the different site arrays would provide important insights. As stated on page 6: It seems like 67 or more repeats have nearly identical effects (on repression and ultrasensitivity). What is the minimal size of repeats needed to see these effects? Or similarly: What is the minimal spacer length between single tetO sites to see an additive effect?

Minor points:

1. Fig. 2 B,C&E: Use same color for similar tetO arrays.

2. In the strain list should read "ura3" not "urs3". W303 (the strain used in these experiments) is from Rodney Rothstein's laboratory, and the reference should be: Thomas BJ, Rothstein R (1989) Elevated recombination rates in transcriptionally active DNA. Cell 56: 619-630

Reviewer #2 (Remarks to the Author):

Review: Lee and Maheshri

The manuscript deals with an interesting question, the effects of having multiple "decoy" transcription-factor binding sites in the genome on gene expression. The authors combine quantitative experiments with simple theoretical modeling to address that question. Thus, both the subject matter and the approach are suitable for MSB.

I found the manuscript non-uniform in quality. Important things were not explained well; critical details seemed to be missing while not-so-interesting findings were given too much space. Below I detail my main points of criticism:

1) The most interesting part of the story (for me) was the excellent agreement between theory and experiment in predicting the effect of "decoy" sites on gene expression (pages 4-5 and Figure 2). However, I was unable to learn much from this agreement and to assess the actual quality of the agreement. This is because the details of the mathematical models are never given in the main text. The very long description in the Supplementary Material is not a solution. Describing the model needs to be an integrated part of the narrative (remember that MSB has a pretty quantitative audience). Similarly, when experimental results are presented and compared to predictions of the model, the process of comparing theory and experiment needs to be explained in detail: Did any parameters need to be fitted? If so, how many? Which ones? Typically in such studies, it is exactly the identity and values of fitting parameters that allows us to understand the system better through the model.

2) Other quantitative aspects are also quite lacking (or are well hidden in the Supplementary Material). For example: How was the amount of tTA quantified in Figures 2B,C and elsewhere?
What are the error estimates on the data points in Figure 2E? How is "ultrasensitive response" defined and quantified?

3) I found the later parts of the work (pages 6-8 and Figure 3-6) to be less impressive. The results are mostly qualitative in nature, and are either of a negative nature ("Factors that affect tTA stability and binding do not appear to alter the ultrasensitive response"), ambiguous (Figure 4) or at least not surprising (the creation of bimodal behavior). Unless more quantitative "meat" can be put on these results, my view is that this part can be shortened, with details going into the Supplementary Material, in favour of a more rigorous treatment in the first part of the paper, as discussed above.

Minor comments:

4) Regarding the observation that "Similar amounts of tTA appear to be sequestered by the 67x, 113x, and 240x contiguous repeats" (page 6), how does this observation stand with regards to previous studies? For example, the work of Y.M. Wang et al (PNAS 2005).

5) There are a few typos that need correcting. For example, I think that the title contains one: what does the word "on" refer to? If it refers to "Role of...", then "on" should be replaced with "in". But perhaps I am misunderstanding the title...

6) In reviewing the prevalence of "tandem repeats", which the authors do a good job of (page 3), could they say something about the situation in bacteria?

7) On the same page, the authors state that "Over 45% of the human genome contains repeated sequences". What does that mean? That 45% of the genome ARE repeated sequences? I could not understand.

Reviewer #3 (Remarks to the Author):

The goal of this manuscript is to investigate the effect of the number and position of "decoy sites" on gene expression regulated by a transcription factor that binds the same type of sites in a target gene's promoter. This is an interesting problem, addressed elegantly through the introduction of various numbers of chromosomal or plasmid-born TetO sites into yeast cells carrying a fluorescent reporter controlled by the tTA activator. Increasing the level of the activator capable of promoter binding by either adding inducer doxycycline to the growth medium or by increasing the strength of the promoter driving tTA expression, an ultrasensitive response is observed when the number of decoy sites becomes large enough. Moreover, when combined with positive feedback, the ultrasensitive response caused by decoy sites leads to bistability.

The question is interesting and the experiments are elegant. However, many details are unclear or missing, making the paper very difficult to understand and evaluate. The paper will be highly suitable for publication once the following concerns are addressed.

Major comments:

(1) Three different methods are used to introduce tandem repeats and perhaps reporter genes into yeast cells: on high copy plasmids, on single copy centromeric plasmids, and by chromosomal integration. For each subsection and each figure panel, it should be meticulously and carefully described and specified whether: (1) the gene encoding tTA; (2) the tandem repeats; and (3) the promoter+reporter constructs are (1) on high copy plasmids; (2) on single copy centromeric plasmids; or (3) chromosomally integrated. If there are differences between the behaviors of these setups, they need to be carefully discussed and interpreted.

(2) In all the figures where the data and the mathematical model are co-plotted, the model is overemphasized by the use of thick lines, and as a result the experimental data are obscured. The emphasis should be reversed on all plots, bringing the data to the forefront while plotting the modeling results with thin lines and dim colors.
(3) In all the figures where the active tTA concentration was titrated using doxycycline, the x-axis is showing tTA concentrations predicted by modeling. There should be corresponding figures, at least as Supplemental Data showing the reporter as a function of the inducer (actual experimental data on both axes). What would the model predict the reporter concentrations to be as a function of doxycycline?

(3) Cell population-level averaging combined with noise may make the dose-responses appear more graded than they actually are at the single cell level. Is this apparent in Figure 4, or can it be studied?

(4) On page 5, paragraph 2, a key statement is "This implies the degradation rate of tTA bound to the repeats is not significantly different than unbound tTA; otherwise, adding decoy sites would not decrease expression." While the authors refer to Burger, Walczak & Wolynes, it should be better explained in this manuscript and then computationally explored how the ultrasensitivity would be affected if the degradation of decoy-bound tTA is much larger/smaller than that of free tTA.

(6) A similar statement is reiterated in the Discussion: 1) decoy-bound tTA is degraded at the same rate as unbound tTA, 2) the effective binding avidity of tTA to the decoys is several orders of magnitude higher than to the promoter, and 3) the sequestration effect is limited by the number of available decoy sites. What happens to the ultrasensitivity and the overall dose-response if these conditions are not satisfied? This should be explored computationally.

Minor comments:

(1) It should be stated whether the individual TetO site sequences in the promoters were completely identical with the sequences in the repeats.

(2) Page 3 and onwards: "TR’s" is a possessive form, which is grammatically incorrect here.

(3) Page 7: "1xtetO promoter response in the absence of feedback is linear": "linear" is probably an overstatement.

(4) Figure 6: Four illustrations are needed in panel A to correspond to the four lines in panel B.

(5) Could it be discussed why the model breaks down at high tTA?

(6) There is something wrong with the caption of Figure 2E.

(7) Page 11, last paragraph: the following are further interesting examples of molecular titration resulting in ultrasensitivity that the Authors may consider citing: ATc titration by cytoplasmic TetR: PNAS 106(13), 5123-8 (2009) and Nucl. Acids Res. 38(8), 2712-26 (2010); sigma factor titration by anti-sigma factor: Phys Biol. 7(3):036005 (2010).
Reviewer #1 (Remarks to the Author):

This report describes use of a yeast version of the synthetic tet-OFF system to study the effect on promoter function of "decoy" transcription factor (TF) binding sites. The TF used here is an E. coli Tet repressor fusion to the herpes virus VP16 transactivator (TA) domain (called tTA). The decoy sites are arrays of Tet operators separated by random sequences of 10 or 30 bp (15, 37, 67, 127 or 240 tetO copies; constructed by Sherratt and colleagues) and the promoter site tested is an array of 7 operators, apparently from Gossen and Bujard (1992), but not described here (i.e. spacing of sites not clear). Test promoters drive expression of either CFP or YFP, and read-outs are made either by FACS or microscopy. Titrating "decoy" arrays are either present on plasmids or integrated into the chromosome.

The experiments reported here largely consist of measuring promoter output as a function of decoy site copy number. The results are presented in terms of a model for the doxycycline-tTA interaction, developed by the authors previously and expanded here in Supplemental Material, which predicts that repeats will, in a number-dependent fashion, convert the dose-response curve from a graded to an "ultrasensitive" one. To fit their dose-response data to this simplified model they had to estimate the ratio of tTA to repeat number and to affinity of tetO sites in the repeats. Their model fits the data reasonably well for low tTA concentrations but systematically overestimates the competitive effect of the repeats at high tTA levels. Taken on face value the results are potentially quite interesting and may have important implications for evolution of transcription factor network function, since repeat number variation is a very frequent form of spontaneous genetic alteration. However, there are several unexplained features of the data that raise significant concerns regarding the authors' interpretation of their results.

Critique:

1. The authors' model predicts an "effective binding avidity" for the decoy sites that is several orders of magnitude higher than that for the tetO7 array present at the promoter. The reason for this difference is not understood, but the authors present evidence (Fig. 6) that it persists even if a tetO7 site is placed adjacent to a decoy array. They also describe evidence arguing that this difference is not due to either tTA turnover through the ubiquitin system or to locus-specific interactions with the transcriptional machinery. In any event, tTA association with both decoy and promoter (tetO7) sites should be measured by quantitative chromatin immunoprecipitation (ChIP) experiments that should give a very reliable estimate of the relative in vivo occupancy of the different sites. The authors' attempts to quantify tTA binding to arrays by fluorescence microscopy (Fig. 4) are completely unconvincing and these data should be eliminated from the manuscript.

We thank this reviewer for the suggestion to perform ChIP experiments – we were also contemplating the same. We have now performed these experiments (Figure 4 and detailed in a new section “The effects of repeats on expression can be explained in terms of tTA binding”), as the results reflect the notion that binding to the promoter at low levels of tTA is weaker than binding to the decoy sites. Because of the non-equilibrium nature of the cross-linking step in ChIP, the ChIP signal may not correspond in a linear manner to tTA occupancy over the entire range, so there is a danger in using it as a complete quantitative readout to directly estimate the difference in occupancy. Nevertheless, qualitatively it is clear that there is a large difference.
2. The concern raised above begs the question of the effect of tetO7 site arrays as decoys, where these arrays should (?) display the same binding avidity as they do in the context of the promoter (a single tetO7 site).

We found this to be a terrific observation and helped us better interpret our ChIP experiments. We found by ChIP that (1) the difference in binding to the 7tetO promoter versus the various repeated decoy sites was not as significant as we would have expected (Figure 4B). However, we also found by the fluorescent reporter assays that multiple 7tetO sites were fully capable of reducing target gene expression and generating a more sigmoidal-like response, regardless of if they were within or without the promoter context (the downstream minimal CYC1 promoter which the 7tetO sites are normally adjacent to) (Supplementary Figure 6). Therefore, at least a subset of these 7tetO sites are serving as strong or high avidity binding sites – but our model implies that these binding events cannot be important for gene expression. To test this idea, we repeated the ChIP experiment with a 1xtetO promoter (Figure 4C) and found a much clearer low tTA region where significant binding to decoy sites occurred with little binding to the single tetO site at this promoter (which presumably has to be important for gene expression).

3. Another curious and completely unexplained observation is that the effect of decoy copy number diminishes significantly beyond 37 sites, though much less so if the arrays are present on separate plasmids. Again, analysis of this phenomenon by qPCR ChIP of tTA on the different site arrays would provide important insights. As stated on page 6: It seems like 67 or more repeats have nearly identical effects (on repression and ultrasensitivity). What is the minimal size of repeats needed to see these effects? Or similarly: What is the minimal spacer length between single tetO sites to see an additive effect?

The ChIP experiments performed on different size arrays, yet probing the exact same region, confirmed that binding in the probed region (spanning 6 tetO sites) was significantly less with a 240x array versus a 67x array. This difference in the “per site occupancy” must be the case if we believe that the 67x and 240x array sequester equal numbers of tTA molecules (since 240x array has more sites). The difference was not exactly 4-fold (closer to 2-fold – Supplementary Figure 5) as one might expect, but it is close given the caveat of how quantitative the ChIP signal is.

As to the other point about the minimal spacer length between single tetO sites for an additive effect – we did not directly address this question. However, as we now make clear in Figure 2B, the 240x tetO array consists of a 113x tetO array and a 127x tetO array separated by 775 bp of sequence (that encodes a prokaryotic gentamycin resistance gene). Therefore, we can say that the minimal spacer length must be > 775 bp. Our suspicion is that the long-range anticooperative effect observed on these decoy arrays at higher TTA occupancy probably spans many kilobases. We explicitly discuss this in p.14 in the discussion section. While we agree that this is an important question to address, we submit that this would be the first step in mechanistic understanding of the anticooperativity phenomena, and is the subject of another study. Furthermore, there is some precedence for this effect in vitro that Reviewer 3 pointed out, and we point this out in the Discussion section.

Minor points:

1. Fig. 2 B,C&E: Use same color for similar tetO arrays.
Now in Figure 2 and throughout the manuscript we use the same color for similar tetO number arrays.

2. In the strain list should read "ura3" not "urs3". W303 (the strain used in these experiments) is from Rodney Rothstein's laboratory, and the reference should be: Thomas BJ, Rothstein R (1989) Elevated recombination rates in transcriptionally active DNA. Cell 56: 619-630

We have made the correction in the strain list table and added this reference in the Materials and Method section.
The manuscript deals with an interesting question, the effects of having multiple "decoy" transcription-factor binding sites in the genome on gene expression. The authors combine quantitative experiments with simple theoretical modeling to address that question. Thus, both the subject matter and the approach are suitable for MSB.

We agree with the reviewer and thank them for this kind comment.

I found the manuscript non-uniform in quality. Important things were not explained well; critical details seemed to be missing while not-so-interesting findings were given too much space. Below I detail my main points of criticism:

1) The most interesting part of the story (for me) was the excellent agreement between theory and experiment in predicting the effect of "decoy" sites on gene expression (pages 4-5 and Figure 2). However, I was unable to learn much from this agreement and to assess the actual quality of the agreement. This is because the details of the mathematical models are never given in the main text. The very long description in the Supplementary Material is not a solution. Describing the model needs to be an integrated part of the narrative (remember that MSB has a pretty quantitative audience). Similarly, when experimental results are presented and compared to predictions of the model, the process of comparing theory and experiment needs to be explained in detail: Did any parameters need to be fitted? If so, how many? Which ones? Typically in such studies, it is exactly the identity and values of fitting parameters that allow us to understand the system better through the model.

We have made significant changes to address this comment. First, we have introduced a “MODEL” section where we describe in detail the model in the main text. We show all the reactions considered as well as the expressions used for the fit in the main text. We only leave formulation of the differential equations behind the model to the supplementary information.

In addition, we now clearly state in the first part of the RESULTS section (p.7-10) how we did the fitting and what parameters were fit, etc. Moreover, in Supplementary Figure 3 we show a type of sensitivity analysis of the model fit to key parameters – primarily the ratio of decoy binding site to promoter binding site affinity (Kd/Kp) – by explicitly showing fit results in graphical format for alternate values of Kd/Kp, as well as the associated residuals. Finally, we now explicitly state in the main text that the usefulness of the model is to identify that Kd/Kp must be significantly less than 1 (a main finding of our study) and are careful to acknowledge that the relation of these parameters to chemical properties is complicated by the simplicity of our model.

2) Other quantitative aspects are also quite lacking (or are well hidden in the Supplementary Material). For example: How was the amount of tTA quantified in Figures 2B,C and elsewhere? What are the error estimates on the data points in Figure 2E? How is "ultrasensitive response" defined and quantified?

Our major overhaul of aspects of model development and fitting will have hopefully addressed these questions. We are now explicit about the process by which we convert doxycycline levels to relative tTA levels in p 7-9 of the RESULTS section, and we give examples of the data plotted versus both doxycycline and relative tTA (Figure 2, Supplementary Figure 4, Supplementary Figure 9).

We have eliminated the data in Figure 2E. Aggregate data corresponding the methionine-titrated tTA is included in Supplementary Figure 2, along with error estimates.

Because we found that the term “ultrasensitive” meant different things to different readers, we have eliminated the term from the manuscript. We describe the altered dose-response as a transition from a linear, graded response to a
sharper, more sigmoidal-like threshold response. These terms are all descriptive. The sharpness of this response is quantified in the fit parameter Kn/Kp which describes the relative binding affinity of tTA to decoy sites versus promoter binding sites. The result of the fit is provided in Table 1, and the effect of varying Kn/Kp on the shape of the curve is described in general in Figure 1, and specifically in the context of real data and the doxycycline model in Supplementary Figure 3.

Alternatively, we could quantify the steepness of the response, for example, by looking at the slope at the inflection point, but am not sure what this descriptive measure adds.

3) I found the later parts of the work (pages 6-8 and Figure 3-6) to be less impressive. The results are mostly qualitative in nature, and are either of a negative nature ("Factors that affect tTA stability and binding do not appear to alter the ultrasensitive response"), ambiguous (Figure 4) or at least not surprising (the creation of bimodal behavior). Unless more quantitative "meat" can be put on these results, my view is that this part can be shortened, with details going into the Supplementary Material, in favour of a more rigorous treatment in the first part of the paper, as discussed above.

We partially agree with this comment and have removed Figure 4 and 6, and also removed discussion of various negative results. However, we still believe that the confirmation of the bimodal behavior is an important qualitative result because it is consistent with our main findings that the dose response has become sharper. This IS obvious if our dose-response measurements are qualitatively correct, and so it is an independent confirmation. This we believe is the most surprising result of the paper, because one would not expect binding affinities of the promoter and decoy sites to be different apriori.

Again, we have eliminated the term “ultrasensitive” in favor of a describing a “sharper, steeper curve” with ‘increased concavity” or a more “sigmoidal curve”.

Minor comments:

4) Regarding the observation that "Similar amounts of tTA appear to be sequestered by the 67x, 113x, and 240x contiguous repeats" (page 6), how does this observation stand with regards to previous studies? For example, the work of Y.M. Wang et al (PNAS 2005).

We thank the reviewer for this insightful comment – we now discuss how are results may indeed be similar to the above mentioned work in the discussion section.

5) There are a few typos that need correcting. For example, I think that the title contains one: what does the word "on" refer to? If it refers to "Role of...", then "on" should be replaced with "in". But perhaps I am misunderstanding the title...

We thank the reviewer for noting this – we have changed the title to: A regulatory role for decoy transcription factor binding sites in target gene expression.

6) In reviewing the prevalence of "tandem repeats", which the authors do a good job of (page 3), could they say something about the situation in bacteria?

Tandem repeats are present in many prokaryotes, although not as prevalent. One important place they lie is in contingency loci, where there growth or expansion leads to frameshifts and production of a functional or nonsensical protein product. We have included this in the main text.

7) On the same page, the authors state that "Over 45% of the human genome contains repeated sequences". What does that mean? That 45% of the genome ARE repeated sequences? I could not understand.

We thank the reviewer for this comment and have endeavored to be more careful in our language. Clearly, if a 100 kb region CONTAINS repeated sequence, that does not necessarily mean it is completely comprised of repeated
sequences. But that was what we intended to mean. We now state “That 45% of the genome is repeated sequence”, as suggested by the reviewer.

Reviewer #3 (Remarks to the Author):

The goal of this manuscript is to investigate the effect of the number and position of "decoy sites" on gene expression regulated by a transcription factor that binds the same type of sites in a target gene's promoter. This is an interesting problem, addressed elegantly through the introduction of various numbers of chromosomal or plasmid-born TetO sites into yeast cells carrying a fluorescent reporter controlled by the tTA activator. Increasing the level of the activator capable of promoter binding by either adding inducer doxycycline to the growth medium or by increasing the strength of the promoter driving tTA expression, an ultrasensitive response is observed when the number of decoy sites becomes large enough. Moreover, when combined with positive feedback, the ultrasensitive response caused by decoy sites leads to bistability.

The question is interesting and the experiments are elegant. However, many details are unclear or missing, making the paper very difficult to understand and evaluate. The paper will be highly suitable for publication once the following concerns are addressed.

We thank the reviewer for these kind comments. We hope that our inclusion of the details of the model and fitting, in a new Model section as well as in the RESULTS section on p.7-10 will make the paper clearer and accessible.

Major comments:

(1) Three different methods are used to introduce tandem repeats and perhaps reporter genes into yeast cells: on high copy plasmids, on single copy centromeric plasmids, and by chromosomal integration. For each subsection and each figure panel, it should be meticulously and carefully described and specified whether: (1) the gene encoding tTA; (2) the tandem repeats; and (3) the promoter+reporter constructs are (1) on high copy plasmids; (2) on single copy centromeric plasmids; or (3) chromosomally integrated. If there are differences between the behaviors of these setups, they need to be carefully discussed and interpreted.

We have now explicitly stated in the figure legend and main text (p.6) that the gene encoding tTA is always integrated and driven by the MYO2 promoter. Furthermore, we have incorporated graphics in Figure 2 for each panel to make it clear whether the tandem repeats are on centromeric plasmids or chromosomally integrated. We have moved the high copy 2 micron plasmid data to Supplementary Figure 4, and explicitly describe how we determined copy number using a constitutively expressed RFP marker in that figure legend and in a section within the Supplementary Text titled “Estimating 2micron plasmid copy”. There are no other differences between these setups. The qualitative behaviors are very similar, but quantitative differences in fit parameters are provided in Table 1. The quantitative difference is summarized in a statement on p.10: “Higher copy plasmid-borne arrays tend to be more effective in decreasing expression”.

(2) In all the figures where the data and the mathematical model are co-plotted, the model is over-emphasized by the use of thick lines, and as a result the experimental data are obscured. The emphasis should be reversed on all plots, bringing the data to the forefront while plotting the modeling results with thin lines and dim colors.

We completely agree and have made the lines much thinner and in a dimmer but similar color.

(3) In all the figures where the active tTA concentration was titrated using doxycycline, the x-axis is showing tTA concentrations predicted by modeling. There should be corresponding figures, at least as Supplemental Data showing...
the reporter as a function of the inducer (actual experimental data on both axes). What would the model predict the reporter concentrations to be as a function of doxycycline?

We agree that showing the expression data versus doxycycline levels is important. We do this in main text in Figure 2B versus C and 2D versus E so the reader can directly compare data with active tTA levels predicted by modeling with the doxycycline data. For data in Figure 3, which is only reported versus active tTA levels, we provide the corresponding dose-response plots versus doxycycline in Supplementary Figure 9. We also provide the raw data with doxycycline concentrations as well as predicted tTA levels for all the data in the paper in electronic format.

(3) Cell population-level averaging combined with noise may make the dose-responses appear more graded than they actually are at the single cell level. Is this apparent in Figure 4, or can it be studied?

We have eliminated the data in Figure 4 because of the higher quality ChIP data that can more accurately measure tTA occupancy at low levels.

Although not in the paper, we can still address the reviewer’s question as it may be of interest. When tTA is fused to YFP, meanYFP fluorescence is elevated over background levels, but not by much. This makes it difficult to estimate low levels of tTA-YFP. We tried a different way to estimate tTA levels by introducing a second copy of the met-inducible promoter driving tTA-YFP, but now driving RFP. The thought was we could perhaps more sensitively measure RFP (which has a lower autofluorescent background) and assume RFP and tTA levels would be correlated in single cells. We found that at any given methionine level, the cells exhibited a whole range of RFP expression that did not correlate with tTA-YFP expression (likely because YFP expression was too low). It did correlate with the CFP expression driven by the tTA-promoter high levels of RFP expression. Unfortunately at lower levels, some cells were ON and some cells were OFF for CFP. This reflects the limited ability of using RFP to estimate tTA-YFP levels. It also shows how sensitive CFP expression is to tTA level. In summary, we can’t really detect what is going on at very low levels of tTA, so it remains possible that here the dose response is not linear. This does not change any conclusions of our study though, because the repeated decoy sites reduce target gene expression well past these low levels of tTA!

(4) On page 5, paragraph 2, a key statement is "This implies the degradation rate of tTA bound to the repeats is not significantly different than unbound tTA; otherwise, adding decoy sites would not decrease expression." While the authors refer to Burger, Walczak & Wolynes, it should be better explained in this manuscript and then computationally explored how the ultrasensitivity would be affected if the degradation of decoy-bound tTA is much larger/smaller than that of free tTA.

We now show in Supplementary Figure 1 the impact that different ratios of degradation rates for decoy bound and free tTA would have on dose-response curves, using the model in Figure 1. We also discuss this in the main text in the latter part of the MODEL section (p.5)

(6) A similar statement is reiterated in the Discussion: 1) decoy-bound tTA is degraded at the same rate as unbound tTA, 2) the effective binding avidity of tTA to the decoys is several orders of magnitude higher than to the promoter, and 3) the sequestration effect is limited by the number of available decoy sites. What happens to the ultrasensitivity and the overall dose-response if these conditions are not satisfied? This should be explored computationally.

As per the answer to the previous question, we have addressed (1) in the latter part of the MODEL section and in Supplementary Figure 1. We address (2) in Supplementary Figure 1 as well as Figure 1B and C, where we compare the differences when the binding affinity of tTA to the decoys is and is not several orders of magnitude higher than the promoter. One could certainly make a number of plots for different ratios, but we do not think this will add much to the paper. We have addressed (3) in Figure 1B&C as we show what occurs to the dose response if N/T changes (the number of available decoy sites). We also provide our model in SBML format for anyone to explore what happens when these parameters are changed.
Minor comments:

(1) It should be stated whether the individual TetO site sequences in the promoters were completely identical with the sequences in the repeats. They are identical in sequence and we reinforce this by explicitly stating in the first paragraph of the discussion: “These results are surprising given the tetO sites in both regions are identical in sequence.”

(2) Page 3 and onwards: "TR's" is a possessive form, which is grammatically incorrect here. We have corrected this grammatical error.

(3) Page 7: "1xtetO promoter response in the absence of feedback is linear": "linear" is probably an overstatement. We have changed “linear” to “gradual and nearly linear” in this statement. We have measured a linear response from a 1xtetO promoter when it is integrated at HIS3 in To et al 2011, which we cited, but given the uncertainty at the highest tTA levels, (zero dox), we agree with the reviewer that saying linear is an overstatement.

(4) Figure 6: Four illustrations are needed in panel A to correspond to the four lines in panel B. We have eliminated Figure 6.

(5) Could it be discussed why the model breaks down at high tTA? We now explicitly discuss how our model can never capture the decreased gap in reduction of target gene expression seen at zero doxycycline (p.8 first paragraph). We further discuss this in the Supplementary Text in the section labeled “Model fitting and behavior at high tTA levels”. We argue that tTA levels may be different in strains with and without decoy sites, but only at high tTA levels, based on ChIP data. Inasmuch as promoter occupancy reflects active tTA levels, it increases linearly with our estimate active tTA level except at the zero dox level. However, this is still speculative – we do not know why.

(6) There is something wrong with the caption of Figure 2E. We have moved the old Figure 2E panel into Supplementary Figure 4 and changed the caption.

(7) Page 11, last paragraph: the following are further interesting examples of molecular titration resulting in ultrasensitivity that the Authors may consider citing: ATc titration by cytoplasmic TetR: PNAS 106(13), 5123-8 (2009) and Nucl. Acids Res. 38(8), 2712-26 (2010); sigma factor titration by anti-sigma factor: Phys Biol. 7(3):036005 (2010).

We have added the Physical Biology reference in the last paragraph of the Discussion section within the main text.
Thank you again for submitting your work to Molecular Systems Biology. We have now heard back from two of the three referees who agreed to evaluate this study, and have decided to render a decision now to avoid further delay. As you will see, the referees feel that the new evidence presented in this revised work have satisfied their concerns, and they are now generally supportive. They have, however, a series of suggestions for modifications, which we would ask you to carefully address in a final revision of the present work.

Most importantly, Reviewer #2 indicates that care should be taken not to imply that bimodality always leads to bistability, and notes that the data only currently supports bimodality. Claims should be correspondingly modified in the abstract and the rest of the manuscript as necessary. Purely as an aside, the editor notes recent results in our journal demonstrating global bimodality in gene expression (Hebenstreit et al, 2011).

The reviewers also felt that this work would benefit from a thorough check for proper grammar and clear sentence construction.

Molecular Systems Biology general requires that underlying numeric data are provided for key experiments, especially for experimental data that are used to fit or support mathematical models. In addition to our capacity to host datasets in our supplementary information section, we allow authors provide 'source data' files for specific figure panels (e.g. <http://tinyurl.com/365zpej>). This sort of figure-associated data may be particularly appropriate for this Figure 2-4. Please see our Instructions of Authors for more details on preparation and formatting of figure source data (<http://www.nature.com/msb/authors/index.html#a3.4.3>).

For formatting reasons, we also that you make the "Model" section a subsection of the Results.

Thank you for submitting this paper to Molecular Systems Biology.

Yours sincerely,

Editor - Molecular Systems Biology
msb@embo.org

Referee comments:

Reviewer #2 (Remarks to the Author):

In my judgment, the authors have done a thorough job of addressing the three reviewers' comments. In particular, I salute them for adding the chromatin IP data. In addition, the relation between theory and experimental data is now more clearly explained.

Unfortunately (and inexplicably), the text is still extremely sloppy, to a degree that some sentences are very hard to understand. Here are a number of examples:

1) "Based on both modeling and chromatin IP measurements, the altered response is attributed to ... Chromatin IP measurements of TF occupancy confirm this interpretation" (abstract, page 2).

If, to begin with, the observation is "based on... chromatin IP measurements", how do chromatin IP
measurements then "confirm this interpretation"?

2) "...TF binding to repeated decoy binding sites is much stronger versus the promoter" (page 4).

3) "We ... introduced arrays containing various numbers of contiguous tetO sites either single copy centromeric plasmids, high copy 2 plasmids, or by genomic integration" (page 7).

4) "Larger contiguous tetO arrays are less effective but multiple non-contiguous tetO arrays" (section title, page 10).

5) "This mechanism requires the number of sequestered tTA bound to decoys should inversely correlate with..." (page 11).

6) Beyond the textual sloppiness, some of the figures are also very hard to interpret. For example, many panels contain an inset with numbers and colors (1B, 1C, 2B-E, ...) but no clear indication is given as to the meaning of these numbers.

7) Finally, on a more substantive issue, the authors repeatedly refer to the emergence of "bistability" (abstract and elsewhere). However, their data (page 13 and Figure 5) only demonstrates bimodality, not bistability. As the authors probably know, demonstrating bistability requires showing that each of the alternative states is locally stable, e.g. by showing hysteresis. Bimodality, on the other hand, can easily arise even in the absence of bistability, via a range of mechanisms.

Reviewer #3 (Remarks to the Author):

I would like to thank the Authors for the revisions made to this manuscript, which has improved considerably. It should be an interesting contribution to Molecular Systems Biology.

Some parts may still need some clarification, such as:

What is the intuitive explanation for the observation that the addition of decoy sites has no effect on gene expression noise (Supp. Figure S8)? It would be interesting to see these noise plots as function of dox rather than YFP reporter.

"repeats versus the promoter (Figure 4B)" - I think this figure shows promoter versus repeats.

Some issues with the grammar require attention, such as: "promoter-bound TF's", "dose-response of curve".

2nd Revision - authors' response 22 February 2012
I am pleased to submit a final revision of our manuscript entitled, “A regulatory role for repeated decoy transcription factor binding sites in target gene expression”. Contents include the main text (39,805 characters including spaces, 5 figures and 1 table), a supplementary section (including text, 9 supplemental figures and 2 supplemental tables), two XML model files in SBML format, and an excel file containing figure source data to support the conclusions of the main text. As was pointed out by both reviewers, the text contained many grammatical errors and typos. We have edited both the manuscript and supplemental text thoroughly correcting both grammar and making stylistic changes to enhance readability. These include not just the specific concerns listed by Reviewers 2 and 3 but many other small changes.

We have also addressed a substantive comment by Reviewer 2, which is that the claim of bistability was overreaching:

7) Finally, on a more substantive issue, the authors repeatedly refer to the emergence of "bistability" (abstract and elsewhere). However, their data (page 13 and Figure 5) only demonstrates bimodality, not bistability. As the authors probably know, demonstrating bistability requires showing that each of the alternative states is locally stable, e.g. by showing hysteresis. Bimodality, on the other hand, can easily arise even in the absence of bistability, via a range of mechanisms.

We agree with the comment and no longer make the definitive claim of the decoy sites introducing bistability in the network. In the text, we suggest an alternate mechanism – a slowly fluctuating upstream factor may be “read out” by a more switch-like downstream promoter response caused by the combination of the positive feedback loop and the decoy sites (that is still deterministically monostable). After some brainstorming, we could not come up with additional mechanisms that we could not rule out immediately.

Reviewer 3 also had a comment related to our measurements of noise in gene expression:

What is the intuitive explanation for the observation that the addition of decoy sites has no effect on gene expression noise (Supp. Figure S8)? It would be interesting to see these noise plots as function of dox rather than YFP reporter.

We have now added a paragraph in the supplemental text to provide an intuitive explanation that is by no means definitive. Because the burst frequency of the 7xtetO promoter is low (0.1 – 1 burst per 15 minutes over the range of expression we analyzed), it is possible that there are several tTA / decoy transactions (and several nuclear tTA import and degradation events) between bursts. Then gene expression time averages specific tTA fluctuations brought about by the decoy sites. We have plotted noise versus expression in Supp. Figure 8 to control for the fact that noise increases with decreased expression (free tTA levels). We were interested in whether decoy sites increase noise further than what is expected by the decreased average free tTA levels. We have provided the raw data in a figure source data file, for anyone to replot versus dox.

We hope these changes have adequately addressed your and the reviewers’ concerns. Please do let me know if you require any additional information.
Acceptance letter

27 February 2012

Thank you again for sending us your revised manuscript. We are now satisfied with the modifications made and I am pleased to inform you that your paper has been accepted for publication.

Before we can send this work to production, we have some remaining content issue that we would ask you address:

Molecular Systems Biology generally requires that biological models are deposited in a public repository, such as BioModels. Please submit your SBML files to BioModels, and then incorporate the resulting accession numbers into the Methods section of this work.

Proofs will be forwarded to you 2-3 weeks after we are able to send this work to production.

Thank you very much for submitting your work to Molecular Systems Biology.

Sincerely,

Editor - Molecular Systems Biology
msb@embo.org