THE EMBO JOURNAL

Manuscript EMBO-2012-83923

# Genome-wide Survey by ChIP-seq Reveals YY1 Regulation of LincRNAs in Skeletal Myogenesis

Leina Lu, Kun Sun, Xiaona Chen, Yu Zhao, Lijun Wang, Liang Zhou, Hao Sun and Huating Wang

*Corresponding author: Huating Wang, Chinese University of Hong Kong*

**Review timeline:**

| | |
|---|---|
| Submission date: | 15 November 2012 |
| Editorial Decision: | 04 January 2013 |
| Revision received: | 03 April 2013 |
| Editorial Decision: | 06 May 2013 |
| Revision received: | 04 July 2013 |
| Accepted: | 09 July 2013 |

## 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.)

*Editor: Anne Nielsen*

1st Editorial Decision 04 January 2013

Thank you for submitting your manuscript for consideration by The EMBO Journal and my apologies for the delay in communicating our decision to you. Your manuscript has now been seen by three referees whose comments are shown below.

As you will see, all referees express great interest in the findings reported in your manuscript, although they also raise a number of concerns that you will have to address in full before submitting a revised manuscript.
Referees 2 and 3 are positive about your findings and request reasonable experiments that should all be performed to strengthen the generality and impact of your findings. I would like to especially emphasize the regulation of all four Yam lncRNAs in muscle differentiation, the extent of cis-acting effects of Yam-1 on additional neighbouring genes and the functional relevance for Yam1 and miR715 in controlling myoblast differentiation in regenerating muscle.
Referee 1 raises the more critical issue of the statistical analysis and threshold values employed for the initial ChIP seq analysis. This is an important concern and you will therefore need to either significantly strengthen the analysis as requested by ref 1 or alternatively repeat the YY1-ChIP with a second antibody to YY1 to substantiate the relevance of the identified targets. This second option was suggested by referee 2 in the cross-referee commenting and I would strongly suggest you to perform this additional experiment.

 

Given the referees' positive recommendations, we offer you the opportunity to submit a revised version of the manuscript, addressing the comments of all three reviewers in full. I should add that it is EMBO Journal policy to allow only a single round of revision, and acceptance or rejection of your manuscript will therefore depend on the completeness of your responses to the full satisfaction of the referees in this revised version. Please do not hesitate to contact me if you have questions related to the review process and the requests made by the referees.

When preparing your letter of response to the referees' comments, please bear in mind that this will form part of the Peer-Review Process File, and will therefore be available online to the community. For more details on our Transparent Editorial Process, please visit our website: http://www.nature.com/emboj/about/process.html

We generally allow three months as standard revision time. As a matter of policy, competing manuscripts published during this period will not negatively impact on our assessment of the conceptual advance presented by your study. However, we request that you contact the editor as soon as possible upon publication of any related work, to discuss how to proceed.

Thank you for the opportunity to consider your work for publication. I look forward to your revision.

REFEREE REPORTS:

Referee #1:

The authors of this manuscript report ChIP-Seq datasets for the transcription factor YY1. In contrast to several previous reports,
YY1 was not found to repress transcription at H3K27me3 muscle loci.
Rather, YY1 activates transcription at many muscle loci. In addition, a muscle lincRNA (Yam-1) associates with YY1 adn acts as an inhibitor of myogenesis.

COMMENTS:

The major concern about this study is related to the. quality/statistical analysis of the ChIP-Seq datasets.

1. The authors indicate that the FDR employed for YY1 ChIP-seq analysis is 10(-5) using MACS. This in an unusually stringent parameter. While a 10(-5) value is commonly employed for p-value, more realistic FDR values are 1%-5%. Such stringent analysis may have lost thousands of YY1 peaks that would be considered "positive" using commonly accepted statistical parameters (as suggested in CHIP-seq guidelines and practices of the ENCODE and modENCODE consortia (Genome Re 2012, 22:1813-31).

2.The very low percentage of reads in peaks (Table1)for YY1, Ezh2, and H3K27me3 indicates a poor quality of the analyzed data.

3.The unusual statistical treatment of the data (point 1.), combined with the low read percentage (point 2), prompts these reviewers to ask how many reads were detected with IgG control ( i.e., background).

It is possible that the very low number of detected peaks for YY1 and EZH2 is due to very low quality of the data and/or data handling. It is therefore difficult to reach a firm conclusion in regards to the genome-wide distribution of YY1 and its overlap with EZh2.

Referee #2:

The manuscript by Lu et al. describes the genome-wide occupancy of YY1 using C2C12 myoblast

cell lines. They examined the occupancy of YY1 in myoblast and following their differentiation into myotubes. They found some interesting and unexpected findings.

Comments:

1-First, the authors found 1820 binding sites in MB but only 626 in MT; Surprisingly, The authors indicate that there is no overlap between the binding sites in the MB and MT. I cannot find any analysis of the binding sites in MB, there is no Suppl Fig.S1A. The authors need to provide the data and analysis for MT.

The authors then performed analysis on the binding sites and published transcripts levels following differentiation that leads them to conclude; that 1) there is large number of binding sites (475) in intergenic regions and 2) the number of targets are activated by YY1 in MB and silenced as MB become MT following loss of YY1. They also found very little overlap between YY1 and PRC2 chromatin residence, while a substantial overlap is found with MyoD occupancy.

2- I think ; the GO analysis is best represented in the Supplement perhaps with the exception of Figure 1F.

The authors then focus on the analysis of long noncoding RNAs by YY1 and report that YY1 occupies some 63 badly annotated lincRNAs by the Guttman paper. They validate YY1 occupancy of four of these lncRNAs, the are termed Yam-1 to Yam-4.

3- The authors need to present where are the binding sites of YY1 compared to Yam-1 to Yam-4. I cannot tell from Figure 2D, where Yam-1 is located and why is that the RNA-seq data (blue trace) and polymerase II and chromatin marks (green traces) are non-overlapping. If YY1 binding is to the distal elements as one imagines then there should be an overlap between YY1 sites and CBP/P300 sites in these regions. The authors need to extend their analysis to the four Yam, lncRNAs to substantiate YY1 regulation of lncRNAs as critical mode of action in muscle differentiation. I believe a more detailed analysis of the four YY1 regulated lncRNA genes in C2C12 cells will increase the generality and the impact of their observation.

It may also provide additional evidence for Yam-1 through Yam-4 regulation of critical differentiation genes in C2C12 cells.

4- The authors then assess the regulatory function of Yam-1 and find an in cis-regulation of miR-715/RN45s. While it is a bit puzzling as Yam-1 seems to be cytoplasmic; their finding of a lncRNA regulation of a microRNA expressing gene is very novel and the functional analysis seems to support their contention. It would be highly informative for the authors to assess the function of Yam-1 on a few other neighboring genes as a control to insure that miR-715 is indeed their main target.

5- It would also be highly informative to know whether Yam-1 to Yam-4 are evolutionarily conserved. Are there any human orthologs?

Referee #3:

In this manuscript Lu et al. integrated ChIP-seq analysis and biological assays in vitro and in vivo to uncover a novel regulatory network by which YY1 represses the myogenic program in undifferentiated myoblasts by activating lincRNA Yam-1-miR715, which in turn target Wnt7b.
The results are quite straightforward and the authors convincingly demonstrate the existence of a functional network linking YY1 and Wnt7b through linc Yam-1/miR715 expression.
The author analysis of YY1 ChIP-seq also shows an unexpected distribution pattern over the genome of muscle cells that is apparently in conflict with the common belief, with no association between YY1 and PRC2/H3K27me3 genome wide, and an association with MyoD. However, here the authors did not deepen their analysis by performing functional experiments in EzH2 and MyoD null cells. Therefore, their conclusion could be toned down or more critically discussed, at this stage.
Overall, I think this is a great paper that deserves to be published in high impact journal.

I only suggest that the authors perform an experiment that functionally links more convincingly YY1-activated Yam1-miR715 with the inhibition of differentiation in myoblasts. They should show that overexpression of YY1 represses the expression of differentiation genes and that this repression is relieved by Yam1 siRNA and/or miR715 antagomyr. This can be done in C2C12 cells, although I would encourage the authors to try the same experiment in regenerating muscle in vivo, in a similar set as they have presented in Fig. 4.

Minor comments

1) The authors performed the ChIp-seq analysis in C2C12 in GM and DM 5 dyas. It would be informative to monitor the YY1 binding (persistence or disengagement?) at an intermediate stage (early differentiation) on some target gene by qPCR-ChIP.
2) In figure 2F and 4 the authors should show the morphology of muscles after injury in the different experimental conditions (by H&E staining) and provide a quantification of their ability to repair the injured muscle.

1st Revision - authors' response 03 April 2013

We would like to thank the reviewers for their time and efforts in reviewing our manuscript and providing us with insightful comments. We acknowledge that the changes made in response to your suggestions have solidified our conclusions and improved the overall quality of the study.

Response to Reviewer 1

*The major concern about this study is related to the quality/statistical analysis of the ChIP-Seq datasets.*

*1. The authors indicate that the FDR employed for YY1 ChIP-seq analysis is 10(-5) using MACS. This is an unusually stringent parameter. While a 10(-5) value is commonly employed for p-value, more realistic FDR values are 1%-5%. Such stringent analysis may have lost thousands of YY1 peaks that would be considered "positive" using commonly accepted statistical parameters (as suggested in CHIP-seq guidelines and practices of the ENCODE and modENCODE consortia (Genome Re 2012, 22:1813-31).*

We thank the reviewer for pointing out this obvious problem. We realized that our description of 10(-5) as FDR in the original Methods is not accurate. This value is actually q-value cutoff or adjusted p-value calculated by Benjamini-Hochberg procedure from MACS2. It corresponds to an empirical FDR value of 3.4% calculated by swapping balanced treatment and control dataset using a formula below:

FDR = (Number of control peaks <= q-value cutoff) / (Number of treatment peaks <= q-value cutoff).

As suggested by the reviewer, this number falls well into the 1-5% range from ENCODE guideline. By plotting empirical FDR vs q-value (revised Table S1), we believe q-value of 10(-5) is the optimal cutoff for analyzing YY1 ChIP-seq data.

*2. The very low percentage of reads in peaks (Table1)for YY1, Ezh2, and H3K27me3 indicates a poor quality of the analyzed data.*

We agree with the reviewer that very low percentage of reads in peaks or the Fraction of Reads in Peaks (FRiP) could indicate a poor quality of the data. However, we would like to point out that this may not be the case for factors that have very few true binding sites. As Landt et. al. pointed out (Genome Research, ChIP-seq guidelines and practices of the ENCODE and modENCODE consortia), "the 1% FRiP guidelines works well when there are thousands to tens of thousands of called occupancy sites in a large mammalian genome. However, passing this threshold does not automatically mean that an experiment is successful and a FRiP below the threshold does not automatically mean failure". In the case of our YY1 ChIP-seq, the total number of called peaks is comparatively small (~1800), which has led to a low FRiP. However, since FRiP values for our IgG controls are much lower (around 0.01%), the resultant enrichment turned out to be very high. This kind of examples can also be found in ENCODE, for example, a FRiP of <1.0% was obtained for ZNF274 due to its small number of true binding sites (Frietze et al., PONE 2010). In fact, based on the description of Landt et al., in ENCODE project, ~25% (265 out of 1052) of datasets have a FRiP enrichment less than 1% because a limited number of peaks are called using MACS with default parameters. This suggests that additional criteria other than FRiP may be used to assess the quality of the data.

*3.The unusual statistical treatment of the data (point 1.), combined with the low read percentage (point 2), prompts these reviewers to ask how many reads were detected with IgG control ( i.e., background). It is possible that the very low number of detected peaks for YY1 and EZH2 is due to very low quality of the data and/or data handling. It is therefore difficult to reach a firm conclusion in regards to the genome-wide distribution of YY1 and its overlap with EZh2.*

As stated above, we believe the quality of our data is reasonable. (1) Other than the arguments above, to convince the reviewer, we have performed additional ChIP-seq with the second antibody against YY1 (AB58066 from Abcam) as suggested by the editor and another reviewer. Peaks identified from this set of data largely overlapped with our original set obtained by using antibody SC-1703 from Santa Cruz. (Much fewer peaks were called from AB58066 (1061) than SC-1703 (1820), probably due to the fact that SC-1703 is a polyclonal antibody whereas AB58066 is monoclonal; Polyclonal antibodies have a higher rate of success in ChIP than monoclonals because polyclonals consist of a pool of antibodies against different epitopes.). (2) Furthermore, we added data from an independent biological replicate for SC-1703 and technical replicates for both antibodies. A very good overlapping was obtained for all these replicate pairs. (3) To make it more convincing, we performed Irreproducible Discovery Rate (IDR) analysis (Li Q., et al. in Ann. Appl. Stat. Volume 5, Number 3 (2011), 1752-1779.), which is considered as a sensitive quality measurement and one of the essential standards in ENCODE guidelines (Landt et. al. Genome Research). For SC-1703 biological replicates, at an IDR of 0.05 (suggested by Li Q., et al.), we obtained 537 bound regions identified in an IDR comparison between the two replicates (Nt) and 832 bound regions identified in an IDR comparison between two "pseudoreplicates" (Np) generated by pooling and randomly partitioning all available reads from replicates. The resultant ratio of Np/Nt is 1.5, which is less than 2, indicating a high reproducibility of these two biological replicates based on ENCODE guidelines. Using similar approach, Np/Nt ratios for technical replicates of SC-1703 and AB58066 are 0.99 and 0.96 (<2), respectively, indicating a high consistency between these replicates. Altogether, the above results convinced us that the quality of our data is reasonably good. (4) More confirming diagnostic comes from motif analysis as suggested by the ENCODE guidelines, the classical YY1 recognition motif can be easily derived from our data (Fig. 1F). (5) Lastly, we would like to point out a recent report on YY1 ChIP-seq performed in embryonic stem cells (Vella P. et. al. NAR, Dec 2011). A total number of 2459 binding peaks were obtained which is comparable to ours. Also, as we pointed out in the text, the non-overlapping distribution of YY1 and Ezh2 peaks were demonstrated ES cells (Mendenhall et. al. PLoS Genetics 2010 Dec e1001244 and Vella P. et. al. NAR, Dec 2011), which is in line with our main conclusion.

We thank again for the reviewer's critical comments and hope he/she finds our arguments acceptable. These results are now included in the revised Figure 1, Supplementary Figure S1 and Suppl. Table S1.

Response to Reviewer 2

*1. First, the authors found 1820 binding sites in MB but only 626 in MT; Surprisingly, The authors indicate that there is no overlap between the binding sites in the MB and MT. I cannot find any analysis of the binding sites in MB, there is no Suppl Fig.S1A. The authors need to provide the data and analysis for MT.*

We thank the reviewer for pointing out the missing figure. We now have added Suppl. Fig.S1G to show that no significant overlapping was found between the YY1 binding sites in the MB and MT. This may suggest distinct roles of YY1 in MBs and MTs.

*2. I think ; the GO analysis is best represented in the Supplement perhaps with the exception of Figure 1F.*

As suggested, we now removed the Ezh2 GO analysis to Supplemental Table S7.

*3. The authors need to present where are the binding sites of YY1 compared to Yam-1 to Yam-4. I cannot tell from Figure 2D, where Yam-1 is located and why is that the RNA-seq data (blue trace) and polymerase II and chromatin marks (green traces) are non-overlapping. If YY1 binding is to the distal elements as one imagines then there should be an overlap between YY1 sites and CBP/P300 sites in these regions.*

We apologize for unclear labeling. YY1 binding peaks were found at multiple sites both upstream of Yam-1 and in its gene body region. We focused on the upstream site as it is more likely to be regulating the promoter transcription. Pol II binding is usually associated with promoter region and H3K4me3 binding is found at TSS region, which is why they are found upstream of Yam-1 RNA-seq track but NOT overlapping. H3K36me3 is normally associated with active transcription region. This was not evident in the original submission due to the Y axis scaling. We now modified it to better illustrate the association of H3K36me3 with the gene body. We must point out that all these marks were also found downstream of Yam-1 gene, suggesting the possible presence of other transcripts within this region. The revised figure can be found in Fig. 1D. In addition, as the reviewer is curious about H3K18/27Ac marks, we overlapped YY1 ChIP-seq track with the publically available H3K18Ac mark which is normally associated with CBP/300; there seems to be some overlapping between H3K18Ac and YY1 binding. We did not include this track in the figure since this histone mark is not relevant in lincRNA identification. Similar analyses of the above bindings on Yam-2, Yam-3 and Yam-4 loci can be found on Suppl. Fig. S6. For Yam-2 and Yam-4, YY1 binding appears to in the promoter region and overlaps with PolII and H3K4me3 peaks; for Yam-3, it might be within the gene body. A complete picture of their genomic organization obtained by RACE cloning will allow us to define the exact positions of YY1 binding.

*4. The authors need to extend their analysis to the four Yam, lncRNAs to substantiate YY1 regulation of lncRNAs as critical mode of action in muscle differentiation. I believe a more detailed analysis of the four YY1 regulated lncRNA genes in C2C12 cells will increase the generality and the impact of their observation. It may also provide additional evidence for Yam-1 through Yam-4 regulation of critical differentiation genes in C2C12 cells.*

We thank the reviewer for this valuable suggestion. We now extended the analysis to Yam-2, 3 and 4 which are also associated by YY1. We performed a substantial number of experiments on Yam-2 to show that it is a pro-myogenic lincRNA during C2C12 differentiation. Loss-of-function assay was also performed on Yam-3 and Yam-4 to demonstrate that they are, like Yam-1 and Yam-2, also regulators of C2C12 differentiation. Due to the time limit, we were not able to do exhaustive investigations on them like what was done for Yam-1. We hope our data is sufficient to demonstrate

that all four Yams are novel functional lincRNAs during C2C12 differentiation. As the reviewer suggested, these extra data increased the generality and the impact of our observation. These results are now included in the revised Fig. 6 and Suppl. Fig. S6.

*5. The authors then assess the regulatory function of Yam-1 and find an in cis-regulation of miR-715/RN45s. While it is a bit puzzling as Yam-1 seems to be cytoplasmic;*

We thank the reviewer for pointing out the contradiction in terms of Yam-1 location and function. This has prompted us to critically evaluate our original FISH analysis and realized that the DNA probe used for FISH may not distinguish sense and antisense expressed genes, thus may have caused mislabeling of RN45S (anti-sense to Yam-1). We now repeated FISH with an RNA probe synthesized through in vitro transcription, which is specifically designed against Yam-1 but not RN45S. Indeed, with this probe, we observed both cytoplasmic and nuclear staining. To confirm the result, we also did cellular fractionation to isolate nuclear and cytoplasmic portion for qRT-PCR detection of Yam-1. Consistently, Yam-1 is found in both cytoplasmic and nuclear fractions. We now include the revised results in Fig. 3.

*6. Their finding of a lncRNA regulation of a microRNA expressing gene is very novel and the functional analysis seems to support their contention. It would be highly informative for the authors to assess the function of Yam-1 on a few other neighboring genes as a control to insure that miR-715 is indeed their main target.*

As suggested, we now assessed the function of Yam-1 on 12 neighboring genes and found in addition to miR-715, five other genes nearest to Yam-1 were also affected by Yam-1 knockdown. However, we noticed that Yam-1 knockdown showed the strongest effect on miR-715 expression (>60% decrease) compared to other neighbor genes, suggesting miR-715 is the main target. These results are now included in the revised Suppl. Fig. S5A.

7. *It would also be highly informative to know whether Yam-1 to Yam-4 are evolutionarily conserved. Are there any human orthologs?*

As suggested by the reviewer, we have searched for homologous sequences in human genome and found Yam-1 and Yam-3 are highly conserved on human chromosome 16 and 22, respectively. Moreover, RNA-seq reads have been found on these two regions from ENCODE data in some human cell lines, suggesting that these two Yams are not only conserved between mouse and human but also expressed in many human cells. These results are now included in the revised Suppl. Fig.S7.

Response to Reviewer 3

*1. The author analysis of YY1 ChIP-seq also shows an unexpected distribution pattern over the genome of muscle cells that is apparently in conflict with the common belief, with no association between YY1 and PRC2/H3K27me3 genome wide, and an association with MyoD. However, here the authors did not deepen their analysis by performing functional experiments in EzH2 and MyoD null cells. Therefore, their conclusion could be toned down or more critically discussed, at this stage.*

We agree with the reviewer on this point. So we now revised the discussion (page 20) to tone it down. And the original Table S9 on YY1 and MyoD overlapping analysis was removed.

*2. Overall, I think this is a great paper that deserves to be published in high impact journal.*

*I only suggest that the authors perform an experiment that functionally links more convincingly YY1-activated Yam1-miR715 with the inhibition of differentiation in myoblasts. They should show that overexpression of YY1 represses the expression of differentiation genes and that this repression is relieved by Yam1 siRNA and/or miR715 antagomyr. This can be done in C2C12 cells, although I would encourage the authors to try the same experiment in regenerating muscle in vivo, in a similar set as they have presented in Fig. 4.*

We thank the reviewer for his/her positive comments. We totally agree that the suggested rescue experiments will convincingly show the functional link. As suggested, we have now performed transfection with YY1+Yam1 siRNA and YY1+anti-miR-715 in C2C12 cells. The results showed that siYam-1 or anti-miR-715 could functionally rescue the inhibitory effect of YY1 overexpression on C2C12 myogenic differentiation. In addition, we also performed the above experiment in vivo in CTX induced regenerating muscles by electroporating YY1 + siYam1 or YY1+anti-miR715 oligos. A combination of siYY oligos+ Yam-1 expression plasmid was also conducted along. In keeping with the C2C12 result, we found that siYam-1 or anti-miR-715 could rescue the inhibitory effect of YY1 on muscle regeneration as measured by a relevant RNA marker; also, Yam-1 expression restored the promoting effect of siYY1. These results are now included in the revised Fig. 3M, 5I as well as Suppl. Fig. S4C-D and Suppl. Fig. S5B-C.

*3. The authors performed the ChIP-seq analysis in C2C12 in GM and DM 5 days. It would be informative to monitor the YY1 binding (persistence or disengagement?) at an intermediate stage (early differentiation) on some target gene by qPCR-ChIP.*

We thank the reviewer for the suggestion. We now added ChIP-PCR for an intermediate stage (DM 1day). A variety of YY1 binding patterns is observed on 10 target genes. It is persistently high on some genes while continuously decreased on others. Interestingly, on a few (4) genes the binding dropped to the lowest on Day1 (disengagement) but increased again on Day 5. This suggested a dynamic regulatory role of YY1 on different target genes. These results are now included in the revised Suppl. Fig. S1H.

*4. In figure 2F and 4 the authors should show the morphology of muscles after injury in the different experimental conditions (by H&E staining) and provide a quantification of their ability to repair the injured muscle.*

As suggested, we now added morphological evidence by counting the fibers with centrally localized nuclei (CLN) in H&E stained muscles to provide a quantitative measurement on their ability to repair the injured muscle. These results are now included in the revised Fig. 4D.

2nd Editorial Decision 06 May 2013

Thank you for submitting a revised version of your manuscript. It has now been seen by all three original referees whose comments are shown below.

As you will see, while referees #2 and #3 find that the main criticisms have been addressed in the revised manuscript, referee #1 remains more critical in light of the apparent contradiction between

previous work describing YY1 as a repressor and your data arguing for transcriptional activation as the main role for YY1. This concern is partially reflected in the comments from ref #3 and you will therefore need to extensively discuss these apparent discrepancies as well as present possible explanations for the largely non-overlapping pattern for YY1 and PRC2 before submitting a revised version of the manuscript. Furthermore, ref#3 asks you to provide a more thorough assessment for muscle regeneration by staining for MyHC, a measure that we find will strengthen the physiological aspects of your manuscript further.

Thank you again for giving us the chance to consider your manuscript for The EMBO Journal, I look forward to your revision.

REFEREE REPORTS:

Referee #1:

The lack of YY1 transcriptional repression remains puzzling.
It has been reported that, in the same experimental setting- i.e, C2C12 cells, YY1 represses transcription of several genes ( Blatter et al. Cell Metabolism 2012, 4:505-517). Consistent with a repressive function of YY1 in C2C12 cells, YY1knock-down leads to upregulation of genes of the insulin/IGF pathway. Modulation of the genes controlled by YY1 in C2C12 cells was confirmed in muscle-specific YY1 knock-out animals.Importantly, YY1-mediated transcriptional repression was significantly blunted in C2C12 cells
depleted of Pc2 (PRC1) or Ezh2 (PRC2). The transcriptional repression ( in addition to its transcriptional activation properties) exerted by YY1 justifies its acronym (Ying-Yang).

Referee #2:

The authors have responded to my concerns and I deemed the manuscript fit for publication.

Referee #3:

The revised manuscript is significantly improved, as the authors have satisfactorily responded to most of the reviewer concerns. It remains some minor concerns that I highly suggest to address in order to further improve the manuscript.
1) One issue that still remains unsolved relates to the detail of YY1-bound genes and the intersection with PRC2-repressed genes. The authors should try to intersect the YY1-bound genes also with publicly available PRC2-target genes in C2C12 cells (e.g. EzH2-bound and H3K27me3 enriched genes from ChIP-seq performed in other labs) and possibly show the list the YY1-PRC2 co-bound genes in a table.
2) In my first revision it escaped one concern that I apologize for raising only now. However, for the sake of the technical and conceptual consistency of this manuscript I suggest that the author addressed it. It regards the effect of YY1/Yam1 axis during muscle regeneration. The count of centrally nucleated fibers cannot be a readout of regeneration efficiency, in the absence of a clear evidence of regenerating myofibers. Therefore, I invite the authors to perform the staining for perinatal and embryonic MyHC that are more reliable markers of regeneration outcome.

2nd Revision - authors' response 04 July 2013

We would like to thank the reviewers for their time and efforts in reviewing our manuscript and providing us with insightful comments. We acknowledge that the changes made in response to your suggestions have solidified our conclusions and improved the overall quality of the study.

Response to Reviewer 1

*The lack of YY1 transcriptional repression remains puzzling.*
*It has been reported that, in the same experimental setting- i.e, C2C12 cells, YY1 represses transcription of several genes ( Blatter et al. Cell Metabolism 2012, 4:505-517). Consistent with a repressive function of YY1 in C2C12 cells, YY1knock-down leads to upregulation of genes of the insulin/IGF pathway. Modulation of the genes controlled by YY1 in C2C12 cells was confirmed in muscle-specific YY1 knock-out animals. Importantly, YY1-mediated transcriptional repression was significantly blunted in C2C12 cells depleted of Pc2 (PRC1) or Ezh2 (PRC2). The transcriptional repression ( in addition to its transcriptional activation properties) exerted by YY1 justifies its acronym (Ying-Yang).*

We thank the reviewer for pointing out this puzzling issue. We totally agree with the reviewer that YY1 represses many genes including several muscle loci and muscle relevant miRNA loci as well as genes of the insulin/IGF pathway. In fact, all the previous studies on YY1 function in muscle cells identified it mainly as a repressive factor. However, we must correct by saying that "the lack of YY1 transcriptional repression" was not our conclusion. In fact, our results showed that a total of 230 YY1 bound genes were found to be up-regulated in MTs versus MBs (Suppl. Table S4), suggesting YY1 may act to repress these genes in the MBs. However, a larger number (472) were found to be highly expressed in MBs but down-regulated in MTs, suggesting YY1 may be necessary for activation of these genes in the MBs. We would like to stress that YY1 plays dual roles but not just repressive roles as previously known. We realized that our original writing may be slightly misleading. So following the suggestions by the reviewer and Editor, we now expanded the discussion on page 19 to explain the apparent contradiction between previous work describing YY1 as a repressor and our data arguing for transcriptional activation as the main role for YY1. Also, the aforementioned paper (Blatter eta l. Cell Metabolism 2012, 4: 505-517) was now cited on page 4.

Response to Reviewer 3

*1. One issue that still remains unsolved relates to the detail of YY1-bound genes and the intersection with PRC2-repressed genes. The authors should try to intersect the YY1-bound genes also with publicly available PRC2-target genes in C2C12 cells (e.g. EzH2-bound and H3K27me3 enriched genes from ChIP-seq performed in other labs) and possibly show the list the YY1-PRC2 co-bound genes in a table.*

We agree with the reviewer on this point. Originally, in Fig. 1J we have intersected YY1 and Ezh2/H3K27me3 binding peaks to show there is little co-occupancy. Now to gain more insights to the detail of the genes, we analyzed the co-occupied genes to discover that YY1 and Ezh2 do not co-bind many genes. To strengthen the results, according to the reviewer's suggestion, we also intersected our YY1 dataset with the published Ezh2 and H3K27me3 ChIP-seq datasets (Ezh2 dataset from Mousavi K. et al. Molecular Cell 2012; two H3K27me3 datasets from the above paper and Asp P. et al. PNAS 2011), which led to the same findings. These results are now included in the revised manuscript and Suppl. Table S6. The largely non-overlapping pattern for YY1 and PRC2 is surprising compared to what we knew of YY1 and Ezh2 function on a few muscle loci, but not unexpected according to recent genome-wide studies on Ezh2 recruitment. In ES cells, researchers discovered that Ezh2 occupancy is largely independent with YY1. This is also because YY1 plays dual roles both and is not just repressing transcription. As suggested by the Editor, we now expanded the point in discussion part on page 19 and 21.

2. *In my first revision it escaped one concern that I apologize for raising only now. However, for the sake of the technical and conceptual consistency of this manuscript I suggest that the author addressed it. It regards the effect of YY1/Yam1 axis during muscle regeneration. The count of centrally nucleated fibers cannot be a readout of regeneration efficiency, in the absence of a clear evidence of regenerating myofibers. Therefore, I invite the authors to perform the staining for perinatal and embryonic MyHC that are more reliable markers of regeneration outcome.*

We agree with the reviewer that staining for perinatal or embryonic MyHC is a more reliable marker. We have performed the experiments and showed consistent result with CNL counting. The results are now included in the revised Fig. 4E.