EMBO reports

The chromatin factor ROW cooperates with BEAF-32 in regulating long-range inducible genes

Neta Herman, Sebastian Kadener, and Sagiv Shifman
DOI: 10.15252/embr.202254720

Corresponding author(s): Sagiv Shifman (sagiv.shifman@mail.huji.ac.il) , Sebastian Kadener (skadener@brandeis.edu)

Review Timeline:

Submission Date:	24th Jan 22
Editorial Decision:	7th Mar 22
Revision Received:	7th Jul 22
Editorial Decision:	7th Sep 22
Revision Received:	19th Sep 22
Accepted:	29th Sep 22

Editor: Esther Schnapp

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. Depending on transfer agreements, referee reports obtained elsewhere may or may not be included in this compilation. Referee reports are anonymous unless the Referee chooses to sign their reports.)

Dear Prof. Shifman,

Thank you for the submission of your manuscript to EMBO reports. We have now received the full set of referee reports as well as referee cross-comments that are all pasted below.

As you will see, the referees acknowledge that the findings are potentially interesting. However, they have also several suggestions for how the study should be improved. I think all comments are valuable and should be addressed. Please let me know in case you disagree, and we can discuss the revisions further, also in a video chat, if you like.

I would thus like to invite you to revise your manuscript with the understanding that the referee concerns must be fully addressed and their suggestions taken on board. Please address all referee concerns in a complete point-by-point response. Acceptance of the manuscript will depend on a positive outcome of a second round of review. It is EMBO reports policy to allow a single round of major revision only and acceptance or rejection of the manuscript will therefore depend on the completeness of your responses included in the next, final version of the manuscript.

We realize that it is difficult to revise to a specific deadline. In the interest of protecting the conceptual advance provided by the work, we recommend a revision within 3 months (7th Jun 2022). Please discuss the revision progress ahead of this time with the editor if you require more time to complete the revisions.

IMPORTANT NOTE: we perform an initial quality control of all revised manuscripts before re-review. Your manuscript will FAIL this control and the handling will be DELAYED if the following APPLIES:
1) A data availability section providing access to data deposited in public databases is missing. If you have not deposited any data, please add a sentence to the data availability section that explains that.
2) Your manuscript contains statistics and error bars based on n=2. Please use scatter blots in these cases. No statistics should be calculated if n=2.

When submitting your revised manuscript, please carefully review the instructions that follow below. Failure to include requested items will delay the evaluation of your revision.

1) a .docx formatted version of the manuscript text (including legends for main figures, EV figures and tables). Please make sure that the changes are highlighted to be clearly visible.

2) individual production quality figure files as .eps, .tif, .jpg (one file per figure). See https://wol-prod-cdn.literatumonline.com/pb-assets/embo-site/EMBOPress_Figure_Guidelines_061115-1561436025777.pdf for more info on how to prepare your figures.

3) We replaced Supplementary Information with Expanded View (EV) Figures and Tables that are collapsible/expandable online. A maximum of 5 EV Figures can be typeset. EV Figures should be cited as 'Figure EV1, Figure EV2" etc... in the text and their respective legends should be included in the main text after the legends of regular figures.

- For the figures that you do NOT wish to display as Expanded View figures, they should be bundled together with their legends in a single PDF file called *Appendix*, which should start with a short Table of Content. Appendix figures should be referred to in the main text as: "Appendix Figure S1, Appendix Figure S2" etc. See detailed instructions regarding expanded view here: <https://www.embopress.org/page/journal/14693178/authorguide#expandedview>

- Additional Tables/Datasets should be labeled and referred to as Table EV1, Dataset EV1, etc. Legends have to be provided in a separate tab in case of .xls files. Alternatively, the legend can be supplied as a separate text file (README) and zipped together with the Table/Dataset file.

4) a .docx formatted letter INCLUDING the reviewers' reports and your detailed point-by-point responses to their comments. As part of the EMBO Press transparent editorial process, the point-by-point response is part of the Review Process File (RPF), which will be published alongside your paper.

5) a complete author checklist, which you can download from our author guidelines <https://www.embopress.org/page/journal/14693178/authorguide>. Please insert information in the checklist that is also reflected in the manuscript. The completed author checklist will also be part of the RPF.

6) Please note that all corresponding authors are required to supply an ORCID ID for their name upon submission of a revised manuscript (<https://orcid.org/>). Please find instructions on how to link your ORCID ID to your account in our manuscript tracking system in our Author guidelines <https://www.embopress.org/page/journal/14693178/authorguide#authorshipguidelines>

7) Before submitting your revision, primary datasets produced in this study need to be deposited in an appropriate public

database (see https://www.embopress.org/page/journal/14693178/authorguide#datadeposition). Please remember to provide a reviewer password if the datasets are not yet public. The accession numbers and database should be listed in a formal "Data Availability" section placed after Materials & Method (see also https://www.embopress.org/page/journal/14693178/authorguide#datadeposition). Please note that the Data Availability Section is restricted to new primary data that are part of this study. * Note - All links should resolve to a page where the data can be accessed. *
If your study has not produced novel datasets, please mention this fact in the Data Availability Section.

8) We would also encourage you to include the source data for figure panels that show essential data. Numerical data should be provided as individual .xls or .csv files (including a tab describing the data). For blots or microscopy, uncropped images should be submitted (using a zip archive if multiple images need to be supplied for one panel). Additional information on source data and instruction on how to label the files are available at <https://www.embopress.org/page/journal/14693178/authorguide#sourcedata>.

9) Our journal also encourages inclusion of *data citations in the reference list* to directly cite datasets that were re-used and obtained from public databases. Data citations in the article text are distinct from normal bibliographical citations and should directly link to the database records from which the data can be accessed. In the main text, data citations are formatted as follows: "Data ref: Smith et al, 2001" or "Data ref: NCBI Sequence Read Archive PRJNA342805, 2017". In the Reference list, data citations must be labeled with "[DATASET]". A data reference must provide the database name, accession number/identifiers and a resolvable link to the landing page from which the data can be accessed at the end of the reference. Further instructions are available at https://www.embopress.org/page/journal/14693178/authorguide#referencesformat

10) Regarding data quantification (see Figure Legends: https://www.embopress.org/page/journal/14693178/authorguide#figureformat)

The following points must be specified in each figure legend:

- the name of the statistical test used to generate error bars and P values,

- the number (n) of independent experiments (please specify technical or biological replicates) underlying each data point,

- the nature of the bars and error bars (s.d., s.e.m.),

- If the data are obtained from n {less than or equal to} 2, use scatter blots showing the individual data points.

Discussion of statistical methodology can be reported in the materials and methods section, but figure legends should contain a basic description of n, P and the test applied.

- Please also include scale bars in all microscopy images.

11) The journal requires a statement specifying whether or not authors have competing interests (defined as all potential or actual interests that could be perceived to influence the presentation or interpretation of an article). In case of competing interests, this must be specified in your disclosure statement. Further information: https://www.embopress.org/competing-interests

We would also welcome the submission of cover suggestions, or motifs to be used by our Graphics Illustrator in designing a cover.

As part of the EMBO publication's Transparent Editorial Process, EMBO reports publishes online a Review Process File (RPF) to accompany accepted manuscripts. This File will be published in conjunction with your paper and will include the referee reports, your point-by-point response and all pertinent correspondence relating to the manuscript.

You are able to opt out of this by letting the editorial office know (emboreports@embo.org). If you do opt out, the Review Process File link will point to the following statement: "No Review Process File is available with this article, as the authors have chosen not to make the review process public in this case."

I look forward to seeing a revised form of your manuscript when it is ready. Please use this link to submit your revision: https://embor.msubmit.net/cgi-bin/main.plex

Best regards,
Esther

Esther Schnapp, PhD
Senior Editor

Referee #1:

This manuscript presents some results suggesting that ROW protein is involved in the organization of chromosomal architecture in Drosophila. The authors aim to prove a hypothesis that ROW participates in the regulation of long-range transcription mediated by BEAF-32. To determine the functions of Row in vivo, the authors used two RNAi transgenic lines. They found that a partial inactivation of Row reduced viability and fertility. They also noticed that genes differentially expressed in the row knockdown flies, are associated with metabolism. They demonstrated that ROW physically interacts with BEAF-32. The authors have confirmed the previously published results that ROW binds predominantly to the promoters of housekeeping genes, and extensively co-localizes with its partner WOC. They also found a strong co-localization with the BEAF binding sites and TAD boundaries. They obtained some indirect evidence that ROW uses AT-hook domains to bind the AT-rich sequences in promoter regions. Based on this result, the authors speculated that Row can either bind to this DNA motif directly, or indirectly through the interaction with BEAF. Inactivation of row expression in the S2 cells showed a downregulation of genes that are the long-range targets of BEAF-32 and are bound indirectly by ROW. Based on these results, the authors conclude that BEAF-32 and ROW form chromatin loops that regulate gene transcription. The MS would present an important advancement in the field of chromatin architecture if the authors could prove that there are no other simple explanations for the obtained results. For example, ROW could bind to many promoters via zinc-fingers or by interaction with the DNA binding proteins. ROW is a component of the complex that is essential for the activation of transcription. Importantly, BEAF is not an essential protein for fly viability (Dev Biol. 2014 May 15;389(2):121-3). The model also suggests that interaction between BEAF and ROW is highly specific, but it is well known that BEAF can interact with many proteins. At the same time, the authors have only demonstrated a relatively weak interaction between BEAF and ROW in the S2 cells, in which the ROW-FLAG construct was over-expressed.
To obtain more convincing data confirming the proposed model, the authors should conduct additional experiments.
1. It would be important to identify those domains in BEAF and ROW that are involved in their interaction. Is the interaction between BEAF and ROW is exclusively strong, or ROW is just one of many partners of BEAF?
2. The available bioinformatic tools can't effectively predict which DNA motif can be specifically bound by the Cys2His2 zinc fingers. Thus, ROW could bind to at least some promoters by the Cys2His2 zinc fingers. To compare the role of AT-hook and zinc-fingers, the authors examined binding of the ROW-FLAG mutants (delta zinc-fingers and delta AT-hooks) in the S2 cells by ChIP-seq analysis. Alternatively, the authors could directly show that ROW specifically interacts with the AT-rich motifs using EMSA or a similar method.
3. The authors should examine whether the direct and indirect binding sites for BEAF and Row are indeed real. There is an ample evidence that active promoters are prone to the non-specific enrichment in the ChIP. The authors described successful inactivation of Row or BEAF by treatment of the S2 cells with dsRNA against row and beaf. Thus, the authors could do ChIP-seq using the Row or Beaf antibodies on the chromatin obtained from the RNAi treated S2 cells. These experiments would confirm the real BEAF and Row binding sites and demonstrate the interdependence of Row and BEAF binding to the chromatin, and the existence of the chromatin loops between the BEAF and Row sites.

Referee #2:

The chromatin factor ROW cooperates with BEAF-32 in regulating long-range inducible genes, Herman et al.

Herman et al. present an extensive investigation into the role of ROW, showing that it physically interacts with the insulator protein BEAF-32 as well as heterochromatin proteins HP1b/c, and that it binds to DNA via its AT-hook domains. Focusing on the relationship between BEAF-32 and ROW, they find that the two factors frequently co-bind at promoters with features of housekeeping gene promoters, and are enriched together at TAD boundaries. Knockdown of ROW and/or BEAF-32 in S2 cells affects few genes, and affected genes have fewer ROW and BEAF-32 motifs than expected, suggesting they are not bound directly by these factors even if they have ChIP-seq peaks. This leads the authors to propose a model of direct ROW and BEAF-32 binding to housekeeping gene promoters which form loops with the promoters of developmental genes, leading to indirect binding of BEAF-32 and ROW at these promoters. This is supported by enriched interactions between directly and indirectly bound promoters, based on aggregating Hi-C signal across promoter pairs.

I appreciated the thorough and well-written introduction. The experiments carried out are extensive and include multiple experimental systems (fly heads and cell culture). It is sometimes unclear in the text which experimental system was used. The manuscript might be simpler for the reader if some analyses/results were moved to supplementary figures.

Major comments

1. The claim that ROW binds to housekeeping gene promoters is supported by GO term enrichment, expression level, and expression variability across developmental stages. In my opinion only the last of these provides strong evidence for a gene to

be considered housekeeping. The authors could strengthen the claim that ROW binds specifically to housekeeping genes in a number of ways, for example looking at housekeeping enhancers identified by Zabidi et al. 2015 (PMID: 25517091) or the core promoter classes defined by Haberle et al. 2019 (PMID: 31092928) and enrichment of the associated motifs. It would also be helpful context to state the total number of genes classified as housekeeping in this study, and therefore what percentage of housekeeping genes are bound by ROW.

2. Since ROW binds to housekeeping genes which typically have high expression levels, when unbound genes are used as a control (e.g. in Figures 4H-I, 8D), a set of unbound genes with matched expression levels should be used to control for factors associated with higher expression.

3. It would be interesting, but not essential, to know whether the AT-mutant ROW protein is still associated indirectly with chromatin, for example by carrying out western blotting of soluble and chromatin-associated protein fractions. It would also be interesting but not essential to know whether the AT-hook is required for co-immuno-precipitation with BEAF-32 - this might provide insight into the nature of their interactions, e.g. whether this is chromatin-mediated or direct binding.

4. Is the enrichment of BEAF-32 and ROW at TAD boundaries (Fig 6) due to the fact that housekeeping gene promoters are enriched at TAD boundaries, or does the enrichment remain after controlling for this? This or other analyses would be required to support the statement in the discussion that the combination of ROW and BEAF-32 can better predict TAD boundaries, as this is not currently supported by the data shown.

5. Many more genes are affected by the in vivo KD of ROW than in S2 cells. Could the authors comment on this, e.g. could this be due to downstream effects of ROW targets on additional genes? It is also unclear in the main text which condition of the S2 cell knockdown experiments leads to the 51 affected genes - is this the double KD? Or the union of all affected genes across all KDs?

Minor comments

6. Fig 1F: The reduction in number of offspring in ROW KD is presented as novel here, but wouldn't this be expected based on the known phenotype of WOC mutants and the relationship between ROW and WOC?

7. Figure 2C: are the enriched GO terms presented here for all DE genes? Are there differences in the enriched terms for up and downregulated genes? This might be informative about the effects of ROW KD.

8. Fig 4: It would be helpful to comment on overlap of HP1b/c with ROW/WOC in the main text, as well as showing this in the figure.

9. The sentence at the top of page 13 describing expected MNase-seq patterns is unclear. It's also unclear from the main text which experimental system / condition the MNase-seq is performed in.

10. Page 16: a reference to support BEAF-32 binding to the identified TATCGA motif is missing.

11. Fig 6A: can the authors comment on why is there a periodicity in the BEAF-32 motif probability?

12. The analysis of BEAF-32/ROW peak overlap using existing data (pg 16-17) could be moved to supplementary information to simplify the manuscript. It would be less confusing for the reader to focus on the experiments and analysis in S2 cells in the main text.

13. There are very few punctate loops visible at the single-locus level in Hi-C data in Drosophila, compared to in mammals. While the analysis in Fig 8 clearly shows that there are enriched interactions in aggregate between direct and indirect ROW peaks at promoters compared to random unbound promoters, it is unclear whether these interactions are only apparent in aggregate or whether they overlap with clear punctate loops. An example of a locus showing such interactions or an analysis of overlaps with loops called by HICCUPS or a similar program would clarify the nature of these interactions.

14. The discussion makes some overly strong statements such as that ROW and BEAF-32 "bind to the same genomic positions, which are promoters of housekeeping genes located at the boundaries of TADs". While each of these promoter properties has been considered individually, without a comprehensive combined analysis it's not clear how many binding sites actually fulfil all these criteria and what fraction of TAD boundaries / housekeeping genes lack this binding.

Referee #3:

Main comments:

The manuscript presents the characterization of ROW complex highlighting novel co-factors that may contribute to its role in regulating gene expression. Of interest, such co-factors include HP1's and Beaf-32, a drosophila insulating protein. ROW co-localizes with such factors on promoters of genes and TAD borders. ROW influences the expression of multiple genes, as shown in flies and S2 cells. In the later, it can influence the expression of genes localizing on the other sides of loops, called long-range targets.

While multiple analyses allow a good characterization of ROW, it is not always clear to what extent its influence mechanistically depends on co-factors (notably HP1's or Baf-32). In particular, whether it requires such factors to bind is not clear, as to why it does not influence gene expression where it binds, and only through distances. This needs to be clarified.

Major points:
The FLAG mediated purification of ROW co-factors including HP1, Woc or additional co-factors - is of interest and deserves a better documentation. Figure 3B should include a western blot of these major co-factors to judge on the relative proportion of them in the -/+ Flag IPs.

2) It is not demonstrated that the lethality of flies repressing ROW is due to defaults in metabolism unless this can be restored by uncoupling such specific function of ROW on those genes. Only the most significant GOs are shown in Figure 2. The authors should also show additional gene ontologies beyond the top 5, e.g. with significant p-values. For example, it may be of interest to provide further evidence for such influence of metabolism by more specifically compensating metabolism through usage of alpha-amino acid intermediates, by challenging the oxidation-reduction stress response or the humoral immune system.

3) In the integration of differential gene expression with ChIP-seq data, the authors may consider a potential bias due to the detection of differentially expressed genes that are more often active, and the possibility that ROW/HP1-bound genes may be more globally inactive. If so, can they detect some specificity (e.g. by analyzing the intersection only among active genes also bound by ROW) ? The fact that 713 genes are repressed when targeted by ROW, instead of 959 by chance, does not automatically implies that the repression on 713 is indirect.

4) It would be of interest to provide further expression analysis with respect to sub-categories of peaks define by co-occupancies of the various factors involved: not just ROW, but also Woc and HP1's, and see their influence on potential gene targets depending on co-occurence (e.g. with TAD boundaries). The authors may also consider re-analyzing some of their gene expression data from S2 cells according to each combination (as a function of Woc, HP1's).

5) The AT-hook mutant of ROW clearly impairs binding. Can the author exclude that this impairment due to DNA-binding as opposed to protein mis-folding and aggregation. This might be an alternative explanation for why binding is so severely compromised (Fig 5C and 7i). The unbound control is already lower indeed (Fig 5C).

6) In Fig 7F, can the authors provide motif search analysis among the 51 deregulated genes? Are these compatible with the proposed model?

7) DREF is the potential transcription factor accounting for activation. Can the authors exclude DREF from their motif search? Can their provide further support for a specific influence of gene expression through loops of genes with ROW/Beaf-32 from the TAD borders.

8) Is ROW depending on Beaf-32 binding at TSSs? or TAD borders? Or across loops to indirect targets?

9) How comes ROW cannot influence the expression of genes when it binds to their promoters, if it can influence those contacted through long-range interactions ? Is there a factor that prevent its influence in cis? This needs further clarification.
10) Further characterization of the influence of ROA and Beaf-32 seems needed. On direct or indirect sites, can Beaf-32 influences ROW binding, or vice-versa ? Are ROW proteins able to mediate loops without Beaf-32 (Fig 8D) ?

Minor points:

a) Some figure panels need further documentation. For example, the scales of Fig4E should be shown. The threshold used for GO p-values (Fig 2C, Fig 4L) must be indicated.
b) The authors should compare the influence of ROW on Beaf-32 binding (Fig 6F) by taking another factor / ChIP-seq as a negative control. Otherwise, it may just be an indirect influence of transcriptional activity of bias due to the more likely detection of peaks in open regions.
c) What was the rationale to state that 75% and 65% of genes are expected to be bound by ROW and Beaf-32 by chance, respectively (Figure 7D and E) ? Those numbers seem a little high.
d) The title needs clarification.

We would like to thank the reviewers for their constructive and valuable comments about this manuscript. We have given all the comments careful consideration and revised the manuscript accordingly. We feel that it is greatly improved and hope the reviewers find it suitable for publication. Please find below a point-by-point response to the comments.

Referee #1:

This manuscript presents some results suggesting that ROW protein is involved in the organization of chromosomal architecture in Drosophila. The authors aim to prove a hypothesis that ROW participates in the regulation of long-range transcription mediated by BEAF-32. To determine the functions of Row in vivo, the authors used two RNAi transgenic lines. They found that a partial inactivation of Row reduced viability and fertility. They also noticed that genes differentially expressed in the row knockdown flies, are associated with metabolism. They demonstrated that ROW physically interacts with BEAF-32. The authors have confirmed the previously published results that ROW binds predominantly to the promoters of housekeeping genes, and extensively co-localizes with its partner WOC. They also found a strong co-localization with the BEAF binding sites and TAD boundaries. They obtained some indirect evidence that ROW uses AT-hook domains to bind the AT-rich sequences in promoter regions. Based on this result, the authors speculated that Row can either bind to this DNA motif directly, or indirectly through the interaction with BEAF. Inactivation of row expression in the S2 cells showed a downregulation of genes that are the long-range targets of BEAF-32 and are bound indirectly by ROW. Based on these results, the authors conclude that BEAF-32 and ROW form chromatin loops that regulate gene transcription. The MS would present an important advancement in the field of chromatin architecture if the authors could prove that there are no other simple explanations for the obtained results. For example, ROW could bind to many promoters via zinc-fingers or by interaction with the DNA binding proteins. ROW is a component of the complex that is essential for the activation of transcription. Importantly, BEAF is not an essential protein for fly viability (Dev Biol. 2014 May 15;389(2):121-3). The model also suggests that interaction between BEAF and ROW is highly specific, but it is well known that BEAF can interact with many proteins. At the same time, the authors have only demonstrated a relatively weak interaction between BEAF and ROW in the S2 cells, in which the ROW-FLAG construct was over-expressed.
To obtain more convincing data confirming the proposed model, the authors should conduct additional experiments.

We thank the reviewerer for the positive and constructive feedback on our work, as well as for the thorough evaluation and the extremely valuable suggestions to improve our study.

1. It would be important to identify those domains in BEAF and ROW that are involved in their interaction. Is the interaction between BEAF and ROW is exclusively strong, or ROW is just one of many partners of BEAF?

Thank you for this comment. In the revised version of the manuscript we addressed this issue. Briefly, we performed a Co-IP using αFlag antibody in S2 cells

transfected with WT or AT-hook mutant ROW-FLAG tagged plasmids or with a plasmid containing only the ZNF domains of ROW (with FLAG-tag; see new Fig 5E). WOC, BEAF-32, and HP1c co-purified with WT and the AT-hook mutant of ROW, but not with the protein containing only the ZNF domains (Fig 5E). These results demonstrate that the AT-hook domains of ROW are not required for the binding to its main interactors and that the interactions are not dependent on the binding of ROW to DNA. In addition, the results show that the Zinc finger domains alone are not sufficient for the binding of ROW to its main partners. We show that ROW and BEAF-32 interact with each other, but we do not claim that the protein-protein interaction is exclusively strong compared to other interactors. It is clear from our findings that ROW interacts more strongly with WOC and HP1c. However, what we find important about the interaction between BEAF-32 and ROW is that they also co-localize at the same genomic sites. We found that 87.5% of gene promoters bound by BEAF-32 were also bound by ROW. In the revised manuscript, we also show that the binding of ROW to the chromatin is influenced by the presence of BEAF-32 (see updated Fig. 6E-F).

2. The available bioinformatic tools can't effectively predict which DNA motif can be specifically bound by the Cys2His2 zinc fingers. Thus, ROW could bind to at least some promoters by the Cys2His2 zinc fingers. To compare the role of AT-hook and zinc-fingers, the authors examined binding of the ROW-FLAG mutants (delta zinc-fingers and delta AT-hooks) in the S2 cells by ChIP-seq analysis. Alternatively, the authors could directly show that ROW specifically interacts with the AT-rich motifs using EMSA or a similar method.

Following this comment, we performed a DNA affinity pulldown assay. We incubated nuclear extract from S2 cells with a biotin-labeled AT-rich dsDNA probe or a control probe (49% AT bases). Then we performed a pulldown assay with streptavidin beads and western blot with ROW antibody. We found a strong enrichment for ROW in the pulldown with the AT-rich probes compared to no probe or with the control probe (Fig 5B). These results indicate that ROW bind directly to AT-rich sequences.

3. The authors should examine whether the direct and indirect binding sites for BEAF and Row are indeed real. There is an ample evidence that active promoters are prone to the non-specific enrichment in the ChIP. The authors described successful inactivation of Row or BEAF by treatment of the S2 cells with dsRNA against row and beaf. Thus, the authors could do ChIP-seq using the Row or Beaf antibodies on the chromatin obtained from the RNAi treated S2 cells. These experiments would confirm the real BEAF and Row binding sites and demonstrate the interdependence of Row and BEAF binding to the chromatin, and the existence of the chromatin loops between the BEAF and Row sites.

To determine the specificity of the ChIP-seq of ROW, we performed ChIP seq using fly heads of flies in which ROW was knockeddown. We found a substantial decrease in the ChIP-seq signal at the binding sites of ROW in these samples confirming the specificity of the ChIP (See Figure EV2A-B). Furthermore, we

performed ChIP-qPCR with ROW and BEAF-32 antibodies in S2 cells treated with dsRNA against row or BEAF-32. We examined by qPCR five promoter regions bound by both proteins. The knockdown of row and BEAF-32 resulted in a significant reduction in the ChIP enrichment at the examined promoters (Fig 6E-F). These results further indicate that the ChIP of ROW and BEAF-32 are specific. Interestingly, we found that the knockdown of BEAF-32 influenced the binding of ROW to the promoters, but not the opposite (Fig 6E-F). Three out of the promoter examined by ChIP qPCR are indirectly bound by BEAF-32 and ROW (without the sequence motifs). Those promoters showed similar results, indicating that indirect peaks of ROW and BEAF-32 are also specific. We thank the reviewer for raising this point which we believe have reveals further aspects of the regulation.

Referee #2:

The chromatin factor ROW cooperates with BEAF-32 in regulating long-range inducible genes, Herman et al.

Herman et al. present an extensive investigation into the role of ROW, showing that it physically interacts with the insulator protein BEAF-32 as well as heterochromatin proteins HP1b/c, and that it binds to DNA via its AT-hook domains. Focusing on the relationship between BEAF-32 and ROW, they find that the two factors frequently co-bind at promoters with features of housekeeping gene promoters, and are enriched together at TAD boundaries. Knockdown of ROW and/or BEAF-32 in S2 cells affects few genes, and affected genes have fewer ROW and BEAF-32 motifs than expected, suggesting they are not bound directly by these factors even if they have ChIP-seq peaks. This leads the authors to propose a model of direct ROW and BEAF-32 binding to housekeeping gene promoters which form loops with the promoters of developmental genes, leading to indirect binding of BEAF-32 and ROW at these promoters. This is supported by enriched interactions between directly and indirectly bound promoters, based on aggregating Hi-C signal across promoter pairs.

I appreciated the thorough and well-written introduction. The experiments carried out are extensive and include multiple experimental systems (fly heads and cell culture). It is sometimes unclear in the text which experimental system was used. The manuscript might be simpler for the reader if some analyses/results were moved to supplementary figures.

We would like to thank the reviewer for the encouraging evaluation of our work and for the highly constructive and helpful comments. Beyond the changes listed below, we also rewrote parts of the manuscript to make it easier to read and follow.

Major comments

1. The claim that ROW binds to housekeeping gene promoters is supported by GO term enrichment, expression level, and expression variability across developmental

stages. In my opinion only the last of these provides strong evidence for a gene to be considered housekeeping. The authors could strengthen the claim that ROW binds specifically to housekeeping genes in a number of ways, for example looking at housekeeping enhancers identified by Zabidi et al. 2015 (PMID: 25517091) or the core promoter classes defined by Haberle et al. 2019 (PMID: 31092928) and enrichment of the associated motifs. It would also be helpful context to state the total number of genes classified as housekeeping in this study, and therefore what percentage of housekeeping genes are bound by ROW.

We thank the reviewer for raising this point. Following this suggestion, we identified a strong and significant overlap of the ROW binding sites with enhancers of housekeeping genes, and we also determined the percentage of genes bound by ROW that can be classified as housekeeping genes (64%) based on previously published definition (Corrales *et al*, 2017).

2. Since ROW binds to housekeeping genes which typically have high expression levels, when unbound genes are used as a control (e.g. in Figures 4H-I, 8D), a set of unbound genes with matched expression levels should be used to control for factors associated with higher expression.

Thank you for this important comment. In the revised manuscript, we used a set of unbound genes with matched expression levels as a control in Fig. 4H-I and in appendix Fig. S2.

3. It would be interesting, but not essential, to know whether the AT-mutant ROW protein is still associated indirectly with chromatin, for example by carrying out western blotting of soluble and chromatin-associated protein fractions. It would also be interesting but not essential to know whether the AT-hook is required for co-immuno-precipitation with BEAF-32 - this might provide insight into the nature of their interactions, e.g. whether this is chromatin-mediated or direct binding.

In the revised manuscript, we performed a Co-IP experiment using αFlag antibody in S2 cells transfected with WT or AT-hook mutant ROW-FLAG tagged plasmids or with a plasmid containing only the ZNF domains of ROW (with FLAG-tag; Fig. 5C and E). WOC, BEAF-32, and HP1c were co-purified with WT and AT-hook mutant of ROW, but not with the protein containing only the ZNF domains of ROW (Fig 5E). These results demonstrate that the AT-hook domains of ROW are not required for the binding to its main interactors and that the interactions are not dependent on the binding of ROW to DNA. In addition, the results show that the zinc finger domains alone are insufficient for binding ROW to its main partners.

4. Is the enrichment of BEAF-32 and ROW at TAD boundaries (Fig 6) due to the fact that housekeeping gene promoters are enriched at TAD boundaries, or does the enrichment remain after controlling for this? This or other analyses would be required to support the statement in the discussion that the combination of ROW and BEAF-

32 can better predict TAD boundaries, as this is not currently supported by the data shown.

The genomic distribution of BEAF-32, ROW, and housekeeping genes are correlated. We therefore performed multiple logistic regressions to examine the independent effect of BEAF-32, ROW, and housekeeping genes in explaining the location of TAD bounderies. In this model, the independent influnce of ROW and BEAF-32 on TAD bounderies remained, while the independent effect of housekeeping genes was abolished. This indicates that the enrichment of ROW and BEAF-32 at TAD bounderies is not because of their enrichment at promoters of housekeeping genes (Fig 6I). We also deleted the statement that combining ROW and BEAF-32 can better predict TAD boundaries.

5. Many more genes are affected by the in vivo KD of ROW than in S2 cells. Could the authors comment on this, e.g. could this be due to downstream effects of ROW targets on additional genes? It is also unclear in the main text which condition of the S2 cell knockdown experiments leads to the 51 affected genes - is this the double KD? Or the union of all affected genes across all KDs?

We agree with the interpretation of the reviewer and believe that this is likely the influence of downstream effects, but there could also be other explanations. For example, since the knockdown of row was performed from the beginning of development and we examined adult flies, the changes in gene expression could reflect the general unhealthy condition of the flies or changes in cell composition because of abnormal development. Furthermore, the differential expression in fly heads is across different types of cells, which could also add to the larger number of DE genes. The 51 differentially expressed genes in S2 cells are identified by comparing the untreated cells to all the different treatment groups (dsRNA against row, BEAF-32, or both genes). We have changed the text to make it clearer.

Minor comments

6. Fig 1F: The reduction in number of offspring in ROW KD is presented as novel here, but wouldn't this be expected based on the known phenotype of WOC mutants and the relationship between ROW and WOC?

We thank the reviewer for bringing this up. We added a sentence and references regarding those findings in WOC mutants.

7. Figure 2C: are the enriched GO terms presented here for all DE genes? Are there differences in the enriched terms for up and downregulated genes? This might be informative about the effects of ROW KD.

Following this comment, we now show the most significant GO term related to upregulated and downregulated genes separately (Fig. 2C-D) and the full lists (Table EV2-3).

8. Fig 4: It would be helpful to comment on overlap of HP1b/c with ROW/WOC in the main text, as well as showing this in the figure.

In the revised manuscript, we show the overlap between the peaks of all the proteins in the complex (see figure 4B) and describe the overlap of HP1b/c with ROW/WOC in the main text.

9. The sentence at the top of page 13 describing expected MNase-seq patterns is unclear. It's also unclear from the main text which experimental system / condition the MNase-seq is performed in.

We thank the reviewer for noticing this. Accordingly, we changed the text to clarify the expected MNase-seq patterns and the system used for performing the MNase-seq (Fly heads).

10. Page 16: a reference to support BEAF-32 binding to the identified TATCGA motif is missing.

Thanks, we added the relevant references.

11. Fig 6A: can the authors comment on why is there a periodicity in the BEAF-32 motif probability?

The periodicity in Fig.6A is caused by observing the probability of the BEAF-32 motif relative to the peak summit of ROW (and not relative to the peak summit of BEAF-32, Fig. EV4B). When the motif of ROW is examined relative to the peak summit of BEAF-32, the periodicity is also observed (Fig. EV4A-B). The motif of BEAF-32 is located to the left or right of the peak summit of ROW, and similarly, the motif of row is relative to the peak summit of BEAF-32. The fine periodicity indicates that the distance between ROW and BEAF-32 motifs varies.

12. The analysis of BEAF-32/ROW peak overlap using existing data (pg 16-17) could be moved to supplementary information to simplify the manuscript. It would be less confusing for the reader to focus on the experiments and analysis in S2 cells in the main text.

Thanks for this comment. The main texts and figures are now focused on the experiments and analysis performed in S2 cells.

13. There are very few punctate loops visible at the single-locus level in Hi-C data in Drosophila, compared to in mammals. While the analysis in Fig 8 clearly shows that there are enriched interactions in aggregate between direct and indirect ROW peaks at promoters compared to random unbound promoters, it is unclear whether these interactions are only apparent in aggregate or whether they overlap with clear punctate loops. An example of a locus showing such interactions or an analysis of overlaps with loops called by HICCUPS or a similar program would clarify the nature of these interactions.

As the reviewer indicated, there are only a few loops in *Drosophila* (KP *et al*, 2017). In addition, the majority of the loops are not at TAD borders (KP *et al*, 2017), while ROW and BEAF-32 peaks overlap TAD borders. Therefore, most of the long-range interactions between direct and indirect binding sites of ROW/BEAF-32 will not overlap identified loops. We observed overrepresented long-range interaction between ROW direct and indirect peaks that are shown by the aggregation plots (Fig. 8D and appendix Fig. S2) and by looking at overrepresented promoter interactions (Fig. 8E). Overrepresented long-range interactions between insulators at TAD borders were also shown by others (Liang *et al*, 2014; Heurteau *et al*, 2020). Following this comment, we use the term long-range interactions and not "loops" in the revised manuscript.

14. The discussion makes some overly strong statements such as that ROW and BEAF-32 "bind to the same genomic positions, which are promoters of housekeeping genes located at the boundaries of TADs". While each of these promoter properties has been considered individually, without a comprehensive combined analysis it's not clear how many binding sites actually fulfil all these criteria and what fraction of TAD boundaries / housekeeping genes lack this binding.

In response to this comment, we now show that 59% of the genes bound by both ROW and BEAF-32 can be defined as housekeeping genes (2742 genes; Fig. EV4C). We found that 36% of the overlapped binding sites between ROW and BEAF-32 overlap TAD borders (Fig. EV4H). We also changed the statement to: "The two proteins interact and bind to overlapping genomic positions, many of which are promoters of housekeeping genes located at the boundaries of TADs."

Referee #3:

Main comments:
The manuscript presents the characterization of ROW complex highlighting novel co-factors that may contribute to its role in regulating gene expression. Of interest, such co-factors include HP1's and Beaf-32, a drosophila insulating protein. ROW co-localizes with such factors on promoters of genes and TAD borders. ROW influences the expression of multiple genes, as shown in flies and S2 cells. In the later, it can influence the expression of genes localizing on the other sides of loops, called long-range targets.

While multiple analyses allow a good characterization of ROW, it is not always clear to what extent its influence mechanistically depends on co-factors (notably HP1's or Baf-32). In particular, whether it requires such factors to bind is not clear, as to why it does not influence gene expression where it binds, and only through distances. This needs to be clarified.

Many thanks for these extremely valuable suggestions to improve our study. As detailed below, we addressed them in the revised manuscript and we believe we clarified all the pointed issues.

Major points:
1) The FLAG mediated purification of ROW co-factors including HP1, Woc or additional co-factors - is of interest and deserves a better documentation. Figure 3B should include a western blot of these major co-factors to judge on the relative proportion of them in the -/+ Flag IPs.

In the revised version, we present a new Co-IP experiment using αFlag antibody in S2 cells following a western blot (Fig. 5E).

2) It is not demonstrated that the lethality of flies repressing ROW is due to defaults in metabolism unless this can be restored by uncoupling such specific function of ROW on those genes. Only the most significant GOs are shown in Figure 2. The authors should also show additional gene ontologies beyond the top 5, e.g. with significant p-values. For example, it may be of interest to provide further evidence for such influence of metabolism by more specifically compensating metabolism through usage of alpha-amino acid intermediates, by challenging the oxidation-reduction stress response or the humoral immune system.

We understand the concern of the reviewer and agree that the statement was too strong. Accordingly, we deleted the sentence implying that the lethality of the flies with row knockdown is due to defaults in metabolism. In addition, we now show the most significant GO term related to upregulated and downregulated genes separately (Fig. 2C-D) and the full lists as well (Table EV2-3).

3) In the integration of differential gene expression with ChIP-seq data, the authors may consider a potential bias due to the detection of differentially expressed genes that are more often active, and the possibility that ROW/HP1-bound genes may be more globally inactive. If so, can they detect some specificity (e.g. by analyzing the intersection only among active genes also bound by ROW) ? The fact that 713 genes are repressed when targeted by ROW, instead of 959 by chance, does not automatically implies that the repression on 713 is indirect

We are sorry for this misunderstanding. The genes bound by ROW are globally active (Fig. 4). Therefore, both datasets (genes bound by ROW and differentially expressed genes) represent active genes. Additionally, the enrichment analysis was restricted only to genes that were included in the differential expression analysis (genes with a count per million (CPM) > 1 in at least three different samples), which would avoid such bias. Most of the differentially expressed genes in fly heads are not bound by ROW (n = 1322) (Fig. 4F); therefore, we suggest that most differentially expressed genes upon row knockdown may represent indirect effects of the knockdown of row. Furthermore, we show in the S2 cells that most differentially expressed genes bound by ROW lack the AT-rich sequence and the motif of BEAF-32 (Fig. 7I-J).

4) It would be of interest to provide further expression analysis with respect to sub-categories of peaks define by co-occupancies of the various factors involved: not just ROW, but also Woc and HP1's, and see their influence on potential gene targets depending on co-occurence (e.g. with TAD boundaries). The authors may also consider re-analyzing some of their gene expression data from S2 cells according to each combination (as a function of Woc, HP1's).

Following this suggestion, we analyzed the overlap of the differentially expressed genes in the fly heads with the peaks of ROW and the peaks of WOC and the HP1s. Additionally, we determined the number of genes upregulated and downregulated out of the genes differentially expressed and bound by all proteins. We summarize this in the main text of figure 4. We also determined the binding sites overlap between all the four proteins (Fig 4B) and the enrichment at promoter regions (EV2D) for better characterization of the complex. For the S2 cells, we performed ChIP-seq only for ROW and BEAF-32, which are analyzed together.

5) The AT-hook mutant of ROW clearly impairs binding. Can the author exclude that this impairment due to DNA-binding as opposed to protein mis-folding and aggregation. This might be an alternative explanation for why binding is so severely compromised (Fig 5C and 7i). The unbound control is already lower indeed (Fig 5C).

We thank the reviewer for pointing this out. Indeed, in the revise version of the manuscript, we performed a Co-IP experiment using αFlag antibody in S2 cells transfected with WT or AT-hook mutant ROW-FLAG tagged plasmids. WOC, BEAF-32, and HP1c were co-purified with WT and AT-hook mutant of ROW. The ability to interact with the other proteins indicates that the AT-hook mutant protein is not misfolded. In addition, we performed a DNA affinity pulldown assay. Briefly, we incubated nuclear extract from S2 cells with biotin-labeled AT-rich dsDNA probes or with control probes (49% AT bases), then performed a pulldown assay with streptavidin beads and western blot with ROW antibody. ROW was enriched in the pulldown with the AT-rich probes compared to the pulldown with no probe or with the control probe (Fig 5B). The results indicate that ROW bind specifically to AT-rich motifs.

7) DREF is the potential transcription factor accounting for activation. Can the authors exclude DREF from their motif search? Can their provide further support for a specific influence of gene expression through loops of genes with ROW/Beaf-32 from the TAD borders.

We thank the reviewer for bringing this up. Indeed, DREF binding motif is very similar to BEAF-32; however the overlap between DREF and ROW binding was not significant (Combined Score = 0.62, FDR = 0.068; Table EV7). We found an overlap between the differentially expressed genes in S2 cells with a previously published list of genes that were shown to be BEAF-32 long-range targets (depending on the

interaction of BEAF-32 with CP190) (Liang *et al*, 2014) (Fig. 7E). We also show that the majority of the differentially expressed genes in S2 cells do not contain the sequence motifs for BEAF-32 and ROW (Fig. 7I-J). These provide further support for a specific influence on gene expression through long-range contact rather than local regulation.

8) Is ROW depending on Beaf-32 binding at TSSs? or TAD borders? Or across loops to indirect targets?

In the revised manuscript, we indeed answered those question. More specifically, we added the results of ChIP-qPCR with ROW and BEAF-32 antibodies in S2 cells treated with dsRNA against row or BEAF-32. Cells not treated with dsRNA were used as a control. We then examined by qPCR promoter regions bound by both proteins. We found that in the knockdown of BEAF-32, the ChIP enrichment of ROW is significantly reduced compared to the control (Fig 6E; P = 4.5×10-5). Knockdown of ROW did not significantly change the ChIP enrichment of BEAF-32 (Fig 6F; P = 0.85). The results show that BEAF-32 can facilitate the binding of ROW to the specific promoters. Three out of the promoters we examined are indirectly bound by BEAF-32 and ROW (without the binding motif). Those promoters showed similar results. This indicates that BEAF-32 influence ROW binding also to the indirect targets.

9) How comes ROW cannot influence the expression of genes when it binds to their promoters, if it can influence those contacted through long-range interactions ? Is there a factor that prevent its influence in cis? This needs further clarification.

ROW binds directly to housekeeping genes and mostly does not influence their expression levels, and it binds indirectly to regulated (developmental) genes and affects their expression levels. Housekeeping genes and regulated genes have different transcription regulation mechanisms, including DNA elements, chromatin architecture, and cofactors (Jonge *et al*, 2017; Haberle *et al*, 2019; Zabidi *et al*, 2015). Regulated genes are very sensitive to changes in activator levels (Jonge *et al*, 2017), and they rely on long-distance enhancers, while housekeeping genes rely on spatial clustering but not on contacts with long-distance enhancers (Corrales *et al*, 2017). This was added to the discussion section.

10) Further characterization of the influence of ROA and Beaf-32 seems needed. On direct or indirect sites, can Beaf-32 influences ROW binding, or vice-versa ? Are ROW proteins able to mediate loops without Beaf-32 (Fig 8D) ?

As we mentioned above, in the revised manuscript, we performed ChIP-qPCR with ROW and BEAF-32 antibodies in S2 cells treated with dsRNA against row or BEAF-32, and we show that ROW binding to chromatin is facilitated by BEAF-32 but not

the other way around. It was previously shown that CP190 or Chromator mediates the long-range contacts through the interaction with BEAF-32 (Vogelmann *et al*, 2014) (Fig. 9). We show that the differential expressed genes in the S2 cells (with row and BEAF-32 knockdown), and especially the downregulated genes, are significantly enriched with genes that were differentially expressed in cells with a mutation that impairs the interaction between BEAF-32 and CP190 (and prevents the binding of BEAF-32 to the indirect peaks (Liang et al, 2014) (Fig. 7E). We also found that both direct and indirect peaks of ROW are associated with CP190.

Minor points:

a) Some figure panels need further documentation. For example, the scales of Fig4E should be shown. The threshold used for GO p-values (Fig 2C, Fig 4L) must be indicated.

As suggested by the reviewer, we provide further documentation of the figures.

b) The authors should compare the influence of ROW on Beaf-32 binding (Fig 6F) by taking another factor / ChIP-seq as a negative control. Otherwise, it may just be an indirect influence of transcriptional activity of bias due to the more likely detection of peaks in open regions.

In the revised manuscript, we provide an alternative experiment. We performed ChIP-qPCR with ROW and BEAF-32 antibodies in S2 cells treated with dsRNA against row or BEAF-32 (Fig. 6E-F). This experiment better predicts the influence of ROW and BEAF-32 on each other. Cells not treated with dsRNA were used as a control. In addition, enrichment was calculated relative to IgG control. We also show in those experiments that our ChIP is specific.

c) What was the rationale to state that 75% and 65% of genes are expected to be bound by ROW and Beaf-32 by chance, respectively (Figure 7D and E) ? Those numbers seem a little high.

The percentages are high because ROW and BEAF-32 bind many expressed genes in S2 cells (ROW: 6299 genes; BEAF-32: 5318 genes).

d) The title needs clarification.

We apologize, but we were not sure what needed clarification in the title. We will be happy to change it if the reviewer provides further feedback.

References:

Corrales M, Rosado A, Cortini R, Van Arensbergen J, Van Steensel B & Filion GJ

(2017) Clustering of Drosophila housekeeping promoters facilitates their expression. Genome Res 27: 1153–1161

Haberle V, Arnold CD, Pagani M, Rath M, Schernhuber K & Stark A (2019) Transcriptional cofactors display specificity for distinct types of core promoters. Nature 570: 122–126

Heurteau A, Perrois C, Depierre D, Fosseprez O, Humbert J, Schaak S & Cuvier O (2020) Insulator-based loops mediate the spreading of H3K27me3 over distant micro-domains repressing euchromatin genes. Genome Biol

Jonge WJ, O'Duibhir E, Lijnzaad P, Leenen D, Groot Koerkamp MJ, Kemmeren P & Holstege FC (2017) Molecular mechanisms that distinguish TFIID housekeeping from regulatable SAGA promoters. EMBO J 36: 274–290

KP E, EL A & RD K (2017) Polycomb-mediated Chromatin Loops Revealed by a Subkilobase-Resolution Chromatin Interaction Map. Proc Natl Acad Sci U S A 114

Liang J, Lacroix L, Gamot A, Cuddapah S, Queille S, Lhoumaud P, Lepetit P, Martin PGP, Vogelmann J, Court F, et al (2014) Chromatin Immunoprecipitation Indirect Peaks Highlight Long-Range Interactions of Insulator Proteins and Pol II Pausing. Mol Cell 53: 672–681

Vogelmann J, Le Gall A, Dejardin S, Allemand F, Gamot A, Labesse G, Cuvier O, Nègre N, Cohen-Gonsaud M, Margeat E, et al (2014) Chromatin Insulator Factors Involved in Long-Range DNA Interactions and Their Role in the Folding of the Drosophila Genome. PLoS Genet

Zabidi MA, Arnold CD, Schernhuber K, Pagani M, Rath M, Frank O & Stark A (2015) Enhancer-core-promoter specificity separates developmental and housekeeping gene regulation. Nature 518: 556–559

Dear Sagiv,

Thank you for your patience while your revised manuscript was peer-reviewed. Given that referee 3 still raises major concerns, I asked for cross-comments from the other referees and we have received the latest cross-comments only today, hence the delay. All comments are pasted below.

Given referees 1 and 2's cross-comments, I am happy to invite you to revise your manuscript once more, and to address referee 3's concerns along the lines suggested by referee 2. Please co-submit a point-by-point response with your final manuscript.

When all comments will have been successfully addressed, we can proceed with the official acceptance of your manuscript.

A few editorial requests will also need to be addressed:

- Not more than 5 keywords may be added to the ms file. Please correct.

- Please complete the author contributions/credits in our online ms handling system.

- Please add the subheading 'Expanded View Figure Legends'

- I attach to this email a related ms file with comments by our data editors. Please address all comments in the final ms.

I would like to suggest a few minor changes to the abstract, that needs to be written in present tense:

Insulator proteins located at the boundaries of topological associated domains (TAD) are involved in higher-order chromatin organization and transcription regulation. However, it is still not clear how long-range contacts contribute to transcriptional regulation. Here we show that Relative-of-WOC (ROW) is essential for the long-range transcription regulation mediated by the Boundary Element-Associated Factor of 32kD (BEAF-32). We find that ROW physically interacts with heterochromatin proteins (HP1b and HP1c) and the insulator protein BEAF-32. These proteins interact at TAD boundaries where ROW, through its AT-hook motifs, binds AT-rich sequences flanked by BEAF-32 binding sites and motifs. Knockdown of row downregulates genes that are long-range targets of BEAF-32 and bound indirectly by ROW (without binding motif). Analyses of high-throughput chromosome conformation capture (Hi-C) data reveal long-range interactions between promoters of housekeeping genes bound directly by ROW and promoters of developmental genes bound indirectly by ROW. Thus, our results show cooperation between BEAF-32 and the ROW complex, including HP1 proteins, to regulate the transcription of developmental and inducible genes through long-range interactions.

Please let me know whether you agree with these changes.

EMBO press papers are accompanied online by A) a short (1-2 sentences) summary of the findings and their significance, B) 2-3 bullet points highlighting key results and C) a synopsis image that is exactly 550 pixels wide and 200-600 pixels high (the height is variable). You can either show a model or key data in the synopsis image. Please note that text needs to be readable at the final size. Please send us this information along with the revised manuscript.

I look forward to seeing a final version of your manuscript as soon as possible. Please use this link to submit your revision: https://embor.msubmit.net/cgi-bin/main.plex

Best regards,
Esther

Esther Schnapp, PhD
Senior Editor
EMBO reports

Referee #1:

In general, the authors have satisfactorily answered the raised questions.
I have no further criticisms of this manuscript and believe it is suitable for publication.

Referee #2:

The authors have addressed all our questions satisfactorily.

A couple of minor points remain to be addressed in the revised version of the manuscript.

1. Could the authors please explain/comment re the BEAF-32 band that appears in the ZNFs lane (last lane) in Figure 5E?

2. The multiple logistic regression analysis in Fig 6 G/I is not well explained.

Referee #3:

Major points:

1) Whether the influence of Row occurs through the proposed model is still not demonstrated. Though it is of interest to show a dependence of Row binding on Baf. This does not mean that Row regulates gene expression through looping. The authors use a simple enrichment test with a list provided in another article, as a list of genes being deregulated through looping. But this does not seem to be sufficient to demonstrate anything.

2) The fact that 713 genes are repressed when targeted by ROW, instead of 959 by chance, does not automatically implies that the repression on 713 is indirect, through a loop. So, while Row is associated with active genes, this is still not shown.

3) The previous point was not addressed: « How comes ROW cannot influence the expression of genes when it binds to their promoters, if it can influence those contacted through long-range interactions ?».
The authors suggest some interesting literature. They also report some interaction of Row with CP190, a factor involved in looping. Yet this is their central point in the paper, and the mechanism should be demonstrated.

4) In the Figure 3 (co-IP), a loading control is missing. The list of peptides obtained from mass-spec should be better quantified and compared to the one obtained with IgG control.

5) The previous message was that the lethality of flies repressing ROW is due to defaults in metabolism. Now the authors remove such statement but then what was the interest of Gene ontology search?

Cross-comments referee 1:

I agree with the main points raised by Referee #3 that the authors provide only indirect evidence that the ROW protein is essential for the organization of long-distance interactions with BEAF/CP190. It also remains unclear how strong and specific the detected interactions are between proteins that should form a chromatin loop according to the model. At the same time, many manuscripts published in leading journals have similar problems with "mechanistic insights" and direct experimental demonstration of the role of a particular protein in the organization of the chromatin loop. I believe that current state of the art makes it technically exceedingly unlikely that the authors could satisfy these valid points. It seems likely that the main problem is that many tested and as yet unexplored proteins can contribute additively/cooperatively to the formation of chromatin loops, making it difficult to directly demonstrate the role of each individual protein in this process. In addition, each transcription factor usually has other functions, such as the formation of open chromatin in regulatory elements, participation in the recruitment of various transcriptional activation and repression complexes, chromatin remodeling complexes, etc.

Cross-comments referee 2:

Comments #1 and #3 ask for mechanistic insight and demonstration of the proposed hypotheses that ROW mediates long-range interactions. While this a reasonable endeavour, in my opinion this is probably out of the scope of this manuscript, since this would require the generation of significant datasets (eg, Hi-C data on ROW depleted samples) which might not be straightforward to produce and might not fully address the reviewer's concerns. It might be enough if the authors would make text edits to soften statements around the causal implication in looping pointing that future work will be needed to demonstrate the hypothesis.

I think comment #2 is based on a misunderstanding by the reviewer. The authors do not claim that 713 genes are indirectly regulated (or repressed) by ROW (this was a misunderstanding already from the original review). The authors claim that since the majority (65%) of differentially expressed genes are not bound by ROW, that these unbound genes might be indirectly regulated. Unfortunately, the authors did not clarify the reviewer's misunderstanding enough in the rebuttal, but in my opinion, this should be straightforward for the authors to address.

Comment #4 should also be straightforward to address by the authors since I am not sure what loading control the reviewer

would expect for an anti-FLAG other than the input sample (which is already included in the figure). The additional analysis asked in this comment should also be easy to address.

Comment #5 is not particularly helpful. The reviewer had originally asked for a softened interpretation of the GO analysis results, which the authors tried to implement. Of course, GO analysis has its limitations, doesn't imply causality, etc, but in my opinion this would be straightforward to address by text edits in the discussion to include a link pointing towards an effect in the metabolic state of these mutants.

We thank the reviewers for the very helpful and constructive review of our work.

Referee #1:

In general, the authors have satisfactorily answered the raised questions.
I have no further criticisms of this manuscript and believe it is suitable for publication.

Referee #2:

The authors have addressed all our questions satisfactorily.

A couple of minor points remain to be addressed in the revised version of the manuscript.

1. Could the authors please explain/comment re the BEAF-32 band that appears in the ZNFs lane (last lane) in Figure 5E?

We are not sure about it. Since this band is not observed in the input and is not in the right size of BEAF-32, we assume this is an unspecific binding.

2. The multiple logistic regression analysis in Fig 6 G/I is not well explained.

We added a sentence to the results to better explain the logistic regression analysis: "The logistic regression model can distinguish between proteins that are likely to influence the TAD boundaries (influence coefficient $\hat{\beta} > 0$) and proteins that don't influence the boundaries but are enriched because they colocalize with the influential proteins (influence coefficient $\hat{\beta} = 0$)". A more detailed description of the method and model can be found in the original paper (Mourad & Cuvier, 2016).

Answers to referee 3's concerns along the lines suggested by referee 2

Referee #3:

Major points:

1) Whether the influence of Row occurs through the proposed model is still not demonstrated. Though it is of interest to show a dependence of Row binding on Baf. This does not mean that Row regulates gene expression through looping. The authors use a simple enrichment test with a list provided in another article, as a list of genes being deregulated through looping. But this does not seem to be sufficient to demonstrate anything.

A set of observations supports our model that ROW influences developmental/inducible genes through long-range interactions that are mediated by BEAF-32 and its interactors (CP190 or Chromator). First, we show that ROW and

BEAF-32 physically interact and bind a similar set of promoters. Second, we found that, similar to BEAF-32, ROW has direct and indirect targets. The direct targets tend to be housekeeping genes with binding sequences of ROW (AT-rich) and BEAF-32 (CGATA motifs). The indirect targets are developmental/inducible genes without the sequence motifs of ROW and BEAF-32. Third, we show that indirect peaks result from ROW direct binding to AT-rich sequences. Fourth, we show using Hi-C data that there are interactions between the direct and indirect targets of ROW. Fifth, previous studies demonstrated that BEAF-32 could participate in DNA looping via the interaction with CP190 (Vogelmann et al., 2014). We show that the genes downregulated in S2 cells with row knockdown are also differentially expressed when BEAF-32 was mutated to prevent the interaction with CP190 and the binding of BEAF-32 to the indirect peaks via loops (Liang et al., 2014). We also found that both direct and indirect peaks of ROW are associated with CP190. Our results suggest that ROW is involved in the long-range regulation previously demonstrated for BEAF-32.

2) The fact that 713 genes are repressed when targeted by ROW, instead of 959 by chance, does not automatically implies that the repression on 713 is indirect, through a loop. So, while Row is associated with active genes, this is still not shown.

We found in both the in vivo and in vitro studies that the majority of differentially expressed genes are not bound directly by ROW. We are sorry for the continuous misunderstanding, but we don't claim that the 713 genes are indirectly regulated. 2035 genes are differentially expressed in the heads of flies with row knockdown (not repressed; ~50% are upregulated and ~50% downregulated). Only 35% of them are bound by ROW, which is even less than what is expected by chance. This is clearly stated in the following sentence: "Surprisingly, out of 2035 differentially expressed genes, only 713 genes (35%) were bound by ROW, which is significantly lower than what is expected by chance". In the in vitro experiments with S2 cells, we show that the majority of differentially expressed genes have a ChiP-seq peak of ROW, but the signal is relatively low, and most genes are without ROW or BEAF-32 motifs (much less than expected by chance). This indicates that the binding of ROW to the promoters of most differentially expressed genes is the outcome of indirect binding.

3) The previous point was not addressed: « How comes ROW cannot influence the expression of genes when it binds to their promoters, if it can influence those contacted through long-range interactions ?».
The authors suggest some interesting literature. They also report some interaction of Row with CP190, a factor involved in looping. Yet this is their central point in the paper, and the mechanism should be demonstrated.

This is a very interesting question, which we addressed in the discussion by pointing out the major differences between promoters of housekeeping genes and inducible genes and the known reliance of inducible genes on long-range interactions. There is also the possibility of redundancy in the system restricted to housekeeping genes. Experimentally testing it is out of the scope of this manuscript. We added to the discussion that future work would be needed to demonstrate the mechanism for this hypothesis.

4) In the Figure 3 (co-IP), a loading control is missing. The list of peptides obtained from mass-spec should be better quantified and compared to the one obtained with IgG control.

The loading control for the anti-FLAG is the input sample, which is already included in the figure. The peptide number in the IgG control is zero; therefore, it is not possible to quantify the peptide number relative to the control.

5) The previous message was that the lethality of flies repressing ROW is due to defaults in metabolism. Now the authors remove such statement but then what was the interest of Gene ontology search?

Since gene ontology enrichment does not imply causality, we changed the statement in the results based on the previous comment by this reviewer. The gene ontology analysis shows enrichment of metabolic processes and is reported in the results section. This may be of interest to other researchers, especially since other proteins in the complex were also linked to metabolism regulation (e.g., Mills BB, Thomas AD, Riddle NC. HP1B is a euchromatic Drosophila HP1 homolog with links to metabolism. PloS one. 2018 Oct 22;13(10):e0205867.).

Prof. Sagiv Shifman
Department of Genetics, The Institute of Life Sciences, The Hebrew University of Jerusalem
The Alexander Silberman Inst. of Life Sciences
Dept. of Genetics, Room 2-317
Jerusalem, Givat Ram 91904
Israel

Dear Prof. Shifman,

I am very pleased to accept your manuscript for publication in the next available issue of EMBO reports. Thank you for your contribution to our journal.

Given that you as corresponding author are in Israel, you are eligible for publication of your article in the open access format in a way that is free of charge for the authors. Please contact either the administration at your institution or our publishers at Wiley (emboreports@wiley.com) for further questions. The general information of Wiley's agreements with Israel can be found here: https://authorservices.wiley.com/author-resources/Journal-Authors/open-access/affiliation-policies-payments/malmad-agreement.html

At the end of this email I include important information about how to proceed. Please ensure that you take the time to read the information and complete and return the necessary forms to allow us to publish your manuscript as quickly as possible.

As part of the EMBO publication's Transparent Editorial Process, EMBO reports publishes online a Review Process File to accompany accepted manuscripts. As you are aware, this File will be published in conjunction with your paper and will include the referee reports, your point-by-point response and all pertinent correspondence relating to the manuscript.

If you do NOT want this File to be published, please inform the editorial office within 2 days, if you have not done so already, otherwise the File will be published by default [contact: emboreports@embo.org]. If you do opt out, the Review Process File link will point to the following statement: "No Review Process File is available with this article, as the authors have chosen not to make the review process public in this case." Please note that the author checklist will still be published even if you opt out of the transparent process.

Thank you again for your contribution to EMBO reports and congratulations on a successful publication. Please consider us again in the future for your most exciting work.

Best regards,
Esther

Esther Schnapp, PhD
Senior Editor
EMBO reports

****

THINGS TO DO NOW:

Once your article has been received by Wiley for production, the corresponding author will receive an email from Wiley's Author Services system which will ask them to log in and will present them with the appropriate license for completion.

You will receive proofs by e-mail approximately 2-3 weeks after all relevant files have been sent to our Production Office; you should return your corrections within 2 days of receiving the proofs.

Please inform us if there is likely to be any difficulty in reaching you at the above address at that time. Failure to meet our deadlines may result in a delay of publication, or publication without your corrections.

All further communications concerning your paper should quote reference number EMBOR-2022-54720V3 and be addressed to emboreports@wiley.com.

Should you be planning a Press Release on your article, please get in contact with emboreports@wiley.com as early as

possible, in order to coordinate publication and release dates.

EMBO Press Author Checklist

Corresponding Author Name: Sagiv Shifman
Journal Submitted to: EMBO Reports
Manuscript Number: EMBOR-2022-54720V2

USEFUL LINKS FOR COMPLETING THIS FORM
The EMBO Journal - Author Guidelines
EMBO Reports - Author Guidelines
Molecular Systems Biology - Author Guidelines
EMBO Molecular Medicine - Author Guidelines

Reporting Checklist for Life Science Articles (updated January 2022)
This checklist is adapted from Materials Design Analysis Reporting (MDAR) Checklist for Authors. MDAR establishes a minimum set of requirements in transparent reporting in the life sciences (see Statement of Task: 10.31222/osf.io/9sm4x). Please follow the journal's guidelines in preparing your manuscript.
Please note that a copy of this checklist will be published alongside your article.

Abridged guidelines for figures

1. Data
The data shown in figures should satisfy the following conditions:
- the data were obtained and processed according to the field's best practice and are presented to reflect the results of the experiments in an accurate and unbiased manner.
- ideally, figure panels should include only measurements that are directly comparable to each other and obtained with the same assay.
- plots include clearly labeled error bars for independent experiments and sample sizes. Unless justified, error bars should not be shown for technical replicates.
- if n<5, the individual data points from each experiment should be plotted. Any statistical test employed should be justified.
- Source Data should be included to report the data underlying figures according to the guidelines set out in the authorship guidelines on Data Presentation.

2. Captions
Each figure caption should contain the following information, for each panel where they are relevant:
- a specification of the experimental system investigated (eg cell line, species name).
- the assay(s) and method(s) used to carry out the reported observations and measurements.
- an explicit mention of the biological and chemical entity(ies) that are being measured.
- an explicit mention of the biological and chemical entity(ies) that are altered/varied/perturbed in a controlled manner.
- the exact sample size (n) for each experimental group/condition, given as a number, not a range;
- a description of the sample collection allowing the reader to understand whether the samples represent technical or biological replicates (including how many animals, litters, cultures, etc.).
- a statement of how many times the experiment shown was independently replicated in the laboratory.
- definitions of statistical methods and measures:
  - common tests, such as t-test (please specify whether paired vs. unpaired), simple χ2 tests, Wilcoxon and Mann-Whitney tests, can be unambiguously identified by name only, but more complex techniques should be described in the methods section;
  - are tests one-sided or two-sided?
  - are there adjustments for multiple comparisons?
  - exact statistical test results, e.g., P values = x but not P values < x;
  - definition of 'center values' as median or average;
  - definition of error bars as s.d. or s.e.m.

Please complete ALL of the questions below.
Select "Not Applicable" only when the requested information is not relevant for your study.

Materials

Newly Created Materials	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
New materials and reagents need to be available; do any restrictions apply?	Yes	Materials and Methods

Antibodies	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
For **antibodies** provide the following information: - Commercial antibodies: RRID (if possible) or supplier name, catalogue number and or/clone number - Non-commercial: RRID or citation	Yes	Materials and Methods

DNA and RNA sequences	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
Short novel DNA or RNA including primers, probes: provide the sequences.	Yes	Materials and Methods

Cell materials	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
Cell lines: Provide species information, strain. Provide accession number in repository **OR** supplier name, catalog number, clone number, and/**OR** RRID.	Yes	Materials and Methods
Primary cultures: Provide species, strain, sex of origin, genetic modification status.	Not Applicable	
Report if the cell lines were recently **authenticated** (e.g., by STR profiling) and tested for mycoplasma contamination.	Not Applicable	

Experimental animals	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
Laboratory animals or Model organisms: Provide species, strain, sex, age, genetic modification status. Provide accession number in repository **OR** supplier name, catalog number, clone number, **OR** RRID.	Yes	Materials and Methods
Animal observed in or captured from the field: Provide species, sex, and age where possible.	Not Applicable	
Please detail **housing and husbandry conditions**.	Yes	Materials and Methods

Plants and microbes	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
Plants: provide species and strain, ecotype and cultivar where relevant, unique accession number if available, and source (including location for collected wild specimens).	Not Applicable	
Microbes: provide species and strain, unique accession number if available, and source.	Not Applicable	

Human research participants	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
If collected and within the bounds of privacy constraints report on age, sex and gender or ethnicity for all study participants.	Not Applicable	

Core facilities	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
If your work benefited from core facilities, was their service mentioned in the acknowledgments section?	Not Applicable	

Design

Study protocol	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
If study protocol has been **pre-registered, provide DOI in the manuscript**. For clinical trials, provide the trial registration number **OR** cite DOI.	Not Applicable	
Report the **clinical trial registration number** (at ClinicalTrials.gov or equivalent), where applicable.	Not Applicable	

Laboratory protocol	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
Provide DOI OR other citation details if **external detailed step-by-step protocols** are available.	Not Applicable	

Experimental study design and statistics	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
Include a statement about **sample size** estimate even if no statistical methods were used.	Yes	Materials and Methods and Figure legends
Were any steps taken to minimize the effects of subjective bias when allocating animals/samples to treatment (e.g. **randomization procedure**)? If yes, have they been described?	Not Applicable	
Include a statement about **blinding** even if no blinding was done.	Not Applicable	
Describe **inclusion/exclusion criteria** if samples or animals were excluded from the analysis. Were the criteria pre-established? If sample or data points were omitted from analysis, report if this was due to attrition or intentional exclusion and provide justification.	Not Applicable	
For every figure, are **statistical tests** justified as appropriate? Do the data meet the assumptions of the tests (e.g., normal distribution)? Describe any methods used to assess it. Is there an estimate of variation within each group of data? Is the variance similar between the groups that are being statistically compared?	Yes	Materials and Methods

Sample definition and in-laboratory replication	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
In the figure legends: state number of times the experiment was **replicated** in laboratory.	Yes	Figure legends
In the figure legends: define whether data describe **technical or biological replicates**.	Yes	Figure legends

Ethics

Ethics	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
Studies involving **human participants**: State details of **authority granting ethics approval** (IRB or equivalent committee(s), provide reference number for approval.	Not Applicable	
Studies involving **human participants**: Include a statement confirming that **informed consent** was obtained from all subjects and that the experiments conformed to the principles set out in the WMA Declaration of Helsinki and the Department of Health and Human Services Belmont Report.	Not Applicable	
Studies involving **human participants:** For publication of **patient photos**, include a statement confirming that consent to publish was obtained.	Not Applicable	
Studies involving experimental **animals**: State details of **authority granting ethics approval** (IRB or equivalent committee(s), provide reference number for approval. Include a statement of compliance with ethical regulations.	Not Applicable	
Studies involving **specimen and field samples:** State if relevant **permits** obtained, provide details of authority approving study; if none were required, explain why.	Not Applicable	

Dual Use Research of Concern (DURC)	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
Could your study fall under dual use research restrictions? Please check biosecurity documents and list of **select agents and toxins** (CDC): https://www.selectagents.gov/sat/list.htm	Not Applicable	
If you used a select agent, is the security level of the lab appropriate and reported in the manuscript?	Not Applicable	
If a study is subject to dual use research of concern regulations, is the name of the **authority granting approval and reference number** for the regulatory approval provided in the manuscript?	Not Applicable	

Reporting

The MDAR framework recommends adoption of discipline-specific guidelines, established and endorsed through community initiatives. Journals have their own policy about requiring specific guidelines and recommendations to complement MDAR.

Adherence to community standards	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
State if relevant guidelines or checklists (e.g., **ICMJE, MIBBI, ARRIVE, PRISMA**) have been followed or provided.	Not Applicable	
For **tumor marker prognostic studies,** we recommend that you follow the **REMARK** reporting guidelines (see link list at top right). See author guidelines, under 'Reporting Guidelines'. Please confirm you have followed these guidelines.	Not Applicable	
For **phase II and III randomized controlled trials**, please refer to the **CONSORT** flow diagram (see link list at top right) and submit the CONSORT checklist (see link list at top right) with your submission. See author guidelines, under 'Reporting Guidelines'. Please confirm you have submitted this list.	Not Applicable	

Data Availability

Data availability	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
Have **primary datasets** been deposited according to the journal's guidelines (see 'Data Deposition' section) and the respective accession numbers provided in the Data Availability Section?	Yes	Data Availlability section
Were **human clinical and genomic datasets** deposited in a public access-controlled repository in accordance to ethical obligations to the patients and to the applicable consent agreement?	Not Applicable	
Are **computational models** that are central and integral to a study available without restrictions in a machine-readable form? Were the relevant accession numbers or links provided?	Not Applicable	
If publicly available data were reused, provide the respective **data citations in the reference list**.	Yes	Materials and Methods and main text