EMBO reports

Global and precise identification of functional miRNA targets in mESCs by integrative analysis

Moritz Schaefer, Amena Nabih, Daniel Spies, Victoria Hermes, Maxime Bodak, Harry Wischnewski, Patrick Stalder, Richard Ngondo, Luz Liechti, Tatjana Sajic, Ruedi Aebersold, David Gatfield, and Constance Ciaudo
DOI: 10.15252/embr.202254762

Corresponding author(s): Constance Ciaudo (cciaudo@ethz.ch)

Review Timeline:

Submission Date:	1st Feb 22
Editorial Decision:	28th Feb 22
Revision Received:	1st Jun 22
Editorial Decision:	21st Jun 22
Revision Received:	27th Jun 22
Accepted:	30th Jun 22

Editor: Achim Breiling

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.)

1st Editorial Decision **28th Feb 2022**

Dear Prof. Ciaudo,

Thank you for the submission of your research manuscript to EMBO reports. We have now received the reports from the three referees that were asked to evaluate your study, which can be found at the end of this email.

As you will see, the referees think that these findings are of interest. Nevertheless, all referees have several comments and suggestions, indicating that a major revision of the manuscript is necessary to allow publication of the study in EMBO reports. As the reports are below, and all their points need to be addressed, I will not detail them here.

Given the constructive referee comments, we would like to invite you to revise your manuscript with the understanding that all referee concerns must be addressed in the revised manuscript or in the detailed point-by-point response. Acceptance of your manuscript will depend on a positive outcome of a second round of review. It is EMBO reports policy to allow a single round of revision only and acceptance of the manuscript will therefore depend on the completeness of your responses included in the next, final version of the manuscript.

Revised manuscripts should be submitted within three months of a request for revision. Please contact me to discuss the revision should you need additional time.

When submitting your revised manuscript, please also carefully review the instructions that follow below.

PLEASE NOTE THAT upon resubmission revised manuscripts are subjected to an initial quality control prior to exposition to re-review. Upon failure in the initial quality control, the manuscripts are sent back to the authors, which may lead to delays. Frequent reasons for such a failure are the lack of the data availability section (please see below) and the presence of statistics based on n=2 (the authors are then asked to present scatter plots or provide more data points).

When submitting your revised manuscript, we will require:

1) a .docx formatted version of the final manuscript text (including legends for main figures, EV figures and tables), but without the figures included. Please make sure that changes are highlighted to be clearly visible. Figure legends should be compiled at the end of the manuscript text.

2) individual production quality figure files as .eps, .tif, .jpg (one file per figure), of main figures and EV figures. Please upload these as separate, individual files upon re-submission.

The Expanded View format, which will be displayed in the main HTML of the paper in a collapsible format, has replaced the Supplementary information. You can submit up to 5 images as Expanded View. Please follow the nomenclature Figure EV1, Figure EV2 etc. The figure legend for these should be included in the main manuscript document file in a section called Expanded View Figure Legends after the main Figure Legends section. Additional Supplementary material should be supplied as a single pdf file labeled Appendix. The Appendix should have page numbers and needs to include a table of content on the first page (with page numbers) and legends for all content. Please follow the nomenclature Appendix Figure Sx, Appendix Table Sx etc. throughout the text, and also label the figures and tables according to this nomenclature.

For more details, please refer to our guide to authors:
http://www.embopress.org/page/journal/14693178/authorguide#manuscriptpreparation

Please consult our guide for figure preparation:
http://wol-prod-cdn.literatumonline.com/pb-assets/embo-site/EMBOPress_Figure_Guidelines_061115-1561436025777.pdf

See also the guidelines for figure legend preparation:
https://www.embopress.org/page/journal/14693178/authorguide#figureformat

3) 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.

4) a complete author checklist, which you can download from our author guidelines (https://www.embopress.org/page/journal/14693178/authorguide). Please insert page numbers in the checklist to indicate where the requested information can be found in the manuscript. The completed author checklist will also be part of the RPF.

Please also follow our guidelines for the use of living organisms, and the respective reporting guidelines:
http://www.embopress.org/page/journal/14693178/authorguide#livingorganisms

5) that primary datasets produced in this study (e.g. RNA-seq, ChIP-seq, structural and array data) are deposited in an appropriate public database. If no primary datasets have been deposited, please also state this in a dedicated section (e.g. 'No primary datasets have been generated and deposited'), see below.

See also: http://embor.embopress.org/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 & Methods) that follows the model below. This is now mandatory (like the COI statement). Please note that the Data Availability Section is restricted to new primary data that are part of this study.

Data availability

The datasets produced in this study are available in the following databases:

- RNA-Seq data: Gene Expression Omnibus GSE46843 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE46843)
- [data type]: [name of the resource] [accession number/identifier/doi] ([URL or identifiers.org/DATABASE:ACCESSION])

*** Note - All links should resolve to a page where the data can be accessed. ***

Moreover, I have these editorial requests:

6) We strongly encourage the publication of original source data with the aim of making primary data more accessible and transparent to the reader. The source data will be published in a separate source data file online along with the accepted manuscript and will be linked to the relevant figure. If you would like to use this opportunity, please submit the source data (for example scans of entire gels or blots, data points of graphs in an excel sheet, additional images, etc.) of your key experiments together with the revised manuscript. If you want to provide source data, please include size markers for scans of entire gels, label the scans with figure and panel number, and send one PDF file per figure.

7) Our journal 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: http://www.embopress.org/page/journal/14693178/authorguide#referencesformat

8) Regarding data quantification and statistics, can you please specify, where applicable, the number "n" for how many independent experiments were performed, if these were biological or technical replicates, the bars and error bars (e.g. SEM, SD) and the test used to calculate p-values in the respective figure legends. Please provide statistical testing where applicable, and also add a paragraph detailing this to the methods section. See: http://www.embopress.org/page/journal/14693178/authorguide#statisticalanalysis

9) Please also note our reference format:
http://www.embopress.org/page/journal/14693178/authorguide#referencesformat

10) We updated our journal's competing interests policy in January 2022 and request authors to consider both actual and perceived competing interests. Please review the policy https://www.embopress.org/competing-interests and add a statement declaring your competing interests. Please name that section 'Disclosure and Competing Interests Statement' and add it after the author contributions section.

I look forward to seeing a revised version of your manuscript when it is ready. Please let me know if you have questions or comments regarding the revision (which could also be discussed in person).

Please use this link to submit your revision: https://embor.msubmit.net/cgi-bin/main.plex

Yours sincerely,

Achim Breiling
Senior Editor
EMBO Reports

Referee #1:

In this manuscript, Schaefer and collaborators use an orthogonal approach that includes different complementary datasets (RNA-seq, AGO2 CLIP, and Ribo-seq) to determine the number of expressed genes regulated by miRNAs in mouse embryonic stem cells.

While other similar studies were published in the past, this one distinguishes itself by using different datasets obtained from mouse embryonic stem cells knockout for several key components of the miRNA pathway with the same genetic background. This approach is thus ideal for identifying "true" miRNA targets in mESCs. Though overall, I find this study quite compelling for better defining targets of miRNAs, some key issues must be addressed before publication:

1- As AGO1 and AGO2 are the most abundant Argonautes in mESCs, it is logical to use for this study Ago1/Ago2 KO mESCs to identify mRNAs targeted by miRNAs. On the other hand, integrating AGO2 CLIP data in their comparative analysis is somewhat complicated as AGO1 binding mRNA targets are missing. As this group already produced the Ago2KO mESC (Ngondo et al, Stem Cell Reports 2017), it will be more appropriate to use those cells to perform a comparative analysis with AGO2 CLIP data. In addition, this new analysis will also be extremely informative in providing an estimate on how many mRNAs are targeted by the AGO1/miRNA silencing complex, which will help to understand better the role of other Argonaute proteins in gene silencing in mammalian cells.

2- It is concerning to see from the PCA analysis that the RNA-seq data made with Dgcr8KO replicates are not clustering like other replicates' data. This can likely explain why the loss of Dgcr8 has no significant effect on ribosome occupancy of predicted miRNA targets when compared to data from Drosha, Ago1, and Ago2 KO mESCs. As Dgcr8 KO replicates are not comparable, the authors should avoid using data obtained with those cells in their integrative approach to avoid any misleading.

3- The functional analysis of Tfap4 as a new miRNA target should be improved by demonstrating that the miR-291a binding sites found in the 3'UTR are essential for Tfap4 regulation by altering the miRNA binding sites with CRISPR/Cas9. Overexpressing miR-290-295 and monitoring the effect on Tfap4 only provide indirect proof of miR-290/Tfap4 regulation.

Besides those important issues, the authors can also push their analysis further to get more insights about miRNA/mRNA targets regulation:

-With all the data in hand, the authors can provide some interesting information about the miRNAs cooperation by determining whether the presence of multiple miRNA binding sites on mRNA correlates with an increase in AGO2 binding and a decrease in ribosome association. The authors mention the limitation of their approach in this manner (a section of the discussion that I really appreciated), but I think it is still feasible to get some helpful information on this.

-It will be informative to perform a GO term/pathways analysis on identified mRNA targets to determine whether specific pathway/cellular functions are targeted by miRNAs to cell behaviors such as maintaining pluripotency state of ESC.

Referee #2:

Rigorous linkage of a miRNA to a function in human cells is much more difficult are rare than is often assumed. This manuscript takes an integrative approach that includes data from multiple, strategically chosen, knockout cell lines to prioritize candidate genes for regulation by miRNAs. The authors present solid data that TFAP4 is regulated by miR-290-295 in mESC's. In itself, that is not a particularly major result. The value is in setting a standard for rigorous analysis of miRNAs.

Overall, this is a worthy paper that should eventually be published. I hope that it will be widely read and appreciated.

However, there are numerous issues that need to be addressed prior to publication.

Overall, it is essential that the authors restrict themselves to wording that is supported by their data. The entire paper sends a message that researchers need to work hard to link miRNAs to function, that message is undermined by poorly supported conclusions.

Page 2. This manuscript DOES NOT show that 9% of genes are functionally and directly regulated by miRNAs. This is the type of overstatement that has been made in papers describing miRNA action for the past twenty years. In reality, the authors have identified 9% of genes as CANDIDATES for regulation by miRNAs. There is a big difference between these two conclusions. The type of data presented in Figure 4 are necessary to even begin to believe that a proposed functional linkage is real.

Page 3. The introductory paragraph, in common with many other paragraphs, is too long. This is a difficult manuscript to understand because it is complex and nuanced. Long, overly-complicated paragraphs are an unnecessary hindrance for

readers.

Page 3. The introductory paragraph is also poorly chosen. It does nothing to focus the reader's attention on the gap in knowledge addressed by the manuscript. It merely summarizes information about miRNAs that could be found in any recent review. The paragraph should be retained (albeit broken up into small units). However a new paragraph(s) should be added to describe the real purpose of this manuscript - to provide a better strategy for solving the problem of linking miRNA expression to believable control of gene expression.

Page 5. Again, the notion that the data show that 9% of expressed genes are subject to direct and functional miRNA-regulation is not only incorrect, the notion works against the value of the rigorous analysis emphasized by the authors. One either approaches these numbers derived from RNAseq data with skepticism or one does not.

Page 6. It would be useful to have a scheme depicting the integrative experimental strategy as part of Figure 1. Figure 2A is too small. Indeed, it might be better to introduce the overall experimental scheme first as a better designed Figure 1A to orient the reader.

Page 6. The first paragraph on page 6 contains a lot of information and is an example of a paragraph that would be more understandable if it were sub-divided into more focused paragraphs.

Page 6. Figure 1C - This is not a very useful figure. A Venn diagram (or diagrams) showing the overlap would be better. It is dangerous to treat the RNAi factors as interchangeable.

Page 7. The published AGO2 binding data and the TargetScan predictions are introduced uncritically. How were AGO2 binds judged according to their significance? Why should the reader believe that the boundaries for judging significance were wisely chosen? Regarding TargetScan, predictive methods are notoriously inaccurate. I do not think that the authors conclusions will be much affected, but the way they introduce these methods strikes this reviewer as naïve. This would be a good place to reinforce the lesson that all of these methods have enormous limitations and are dangerous for the unwary.

Page 8. "High confidence" is not useful terminology here. These are the most highly ranked miRNA:mRNA pairs, and therefore the ones most likely to be worth investing time in experimental validation.

Page 8-9. The mass spec/ribosome profiling data adds little to the paper and is inadequately described. These are difficult experiments and merit much more attention to be believable. These weak data are not necessary and should be deleted.

Page 9. The paragraph describing the miR-290-295 cluster begins one of the most important sections in this paper. It is far too long and needs to be sub-divided.

Page 9. It is essential that the authors present compelling and transparent experimental data demonstrating the high expression of miR-290-295 in mECs. How many miRNAs are in this cluster? How many miRNAs are expressed per cell? They entire analysis pivots on believing that the concentration of miRNA is high enough to affect the biology of mECs. The authors need to make a stronger case for the physical basis of miR-290-295 action.

Page 10-11. What does the complementarity of miR-290-295 and Tfap4 look like? A physical interaction must underlie regulation. What is it? What does the AGO2 binding data look like? How many miR-290 binding sites are there? Are there binding sites for other well-expressed miRNAs. A much stronger case needs to be made and the authors need to think carefully about how to do this.

Note: It IS NOT necessary that this manuscript conclusively identify Tfap4 as a target. It is enough that the authors thoughtfully explore how difficult target identification might be even when they go to the great length of knocking out a miRNA cluster. Providing a false sense of certainty brings this work down to the level of the many poor papers that have been published on miRNAs.

Page 12-14. Without stronger reasons to believe that Tfap4 is a legitimate target, the analysis of downstream target genes carries little meaning.

In summary, this is a potentially important paper. It would be acceptable if it were revised to describe the exact conclusions the data support and no more. Unfortunately, as written, it falls into the common trap of feeling the need to provide some positive conclusion about the roles of miRNAs in cells. Perhaps the data support that conclusion, but I am unpersuaded. Either more transparent and thoughtfully described data needs to be added, or the text needs to be changed to reflect a more uncertain reality. The most important conclusion is that a layered approach is necessary to have any hope of assigning a function to a miRNA, and even then the connection is less then certain. While this is a lengthy review, I believe that these suggestions can be addressed through relatively straightforward changes to text and figures.

Minor comments:

1) "The extensive datasets developed in this study will support the development of improved predictive models for miRNA-mRNA functional interactions." I do not see the ways how it can be incorporated into predication models not generating the same datasets for different cell lines or cell context specific conditions. I think the authors mean to imply that the workflow they illustrate will be useful
2) This sentence requires following sentence with explanation what context specific factors authors talking about. "...relevant miRNA interactions could be partially attributed to the lack of incorporating important context-specific factors."
3) Interaction score calculation. Knowing that miRNAs manly function thru interaction with AGO proteins would last criterion calculating interaction score supposed to be not "higher than 0.5 in at least two miRNA_KO mutants", however specifically between Ago2&1_KO and another mutant. If gene does not show expression change in this KO cell it less likely to be regulated via RNAi. Also inclusion of targetscan prediction as a factor is questionable when it was discussed before as overestimating approach for miRNA-RNA binding identification.
4) "we observed seemingly functional interactions in regions outside of the 3'UTR (for example Figures S2E, F), which led us to exclude the 3'UTR-centric TargetScan context++ score model from the filtering, still leading to a surprisingly low number of 707 identified miRNA-targeted genes (Figure 2F)" What is distribution of binding sites? I am left guessing about the exact meaning of this sentence. Not transparent at all.
5) miRNA can act to repress translation without effecting stability therefore mass spectrometry and ribosomal profiling had focus not only on 707 genes which already known to be effected at mRNA level.
Work is poorly visualized with low transparancy:
• PCA figures have to be plotted with scale of X and Y axis, to have better insight about the distance among samples.
• Figure1 C would be better to plot as venn diagram. It would be more transparent and let to see exact overlap between different KO cell lines.
• Figure 2. The main panel F where 9% of direct regulated genes come from is covered by supplementary type panels B, C, D, E. To make Figure2 F more transparent would be great to have supportive panel showing where these 444 and 707 genes stand in context of other genes by expression and significant level. It could be visualized as volcano or plot or heat map etc.
• For this paper to have an impact, it needs to be understandable to a wide readership. The authors should take a step back and consider how to make their data both transparent and intuitive for a typical interested reader who does not consume large data sets regularly.

Referee #3:

In the manuscript titled "Integrative analysis allows a global and precise identification of functional miRNA target genes in mESCs " by Schaefer and colleagues, the authors describe an elegant multi-OMICs integrative approach to build a map of functional interaction of miRNAs-mRNAs in ESCs. Indeed, the simple miRNA expression analysis in specific cell types cannot fully reveal the functional miRNA-mRNA interactions leading to gene expression regulation. The work of Schaefer and colleagues elegantly addresses this point and aims to demonstrate that a multi-OMICs approach based on context-dependent datasets can allow to predict with a significant accuracy miRNA-mRNA functional interaction. Although the approach cannot be considered completely original compared to the papers cited by the authors and displays some limitations in the prediction of functional interaction, it definitely represents an important step forward compared to previous studies. The manuscript is well-written and methodological approaches and results are properly presented. One of the main limitations of this study is that the analysis is mainly based on RNA-seq and not on protein levels, but the authors extensively discuss this point and they also partially overcome this limitation using ribosome profiling. Also, the limitation on the sensitivity of the approach is well discussed and partially overcome by reducing thresholds.

Although the OMICS approach appears solid the experimental part of this work proves to be incomplete and must be improved to make stronger the results obtained by in silico analysis and to support some conclusions. After providing more functional data, the strength of this in silico approach can be truly appreciated. I believe that the work done by Schaefer and colleagues will be of broad interest for the scientific community once the authors have solved some concerns indicated below.

Major points:
- A such consistent and complex in silico analysis requires a more extensive validation of the predicted miRNA-mRNA functional interactions to convince about the strength and the efficacy of this approach. Beyond the data presented for TFAP4, that anyway (as discussed below) need to be increased, some more miRNA-mRNA functional interactions must be demonstrated on the basis of the in silico prediction.
- The integrative approach used, and the data shown in figure 4 indicate that TFAP4 can be a direct target of miR-290-295 cluster. Again, this functional observation is important to make more solid the conclusions based on the in silico approach. To demonstrate that TFAP4 is directly regulated by members of miR-290-295 cluster, luc assay using wt 3'UTR of TFAP4 and the relative miRNA binding site mutant coupled with overexpression of miR-290-295 cluster members must be reported.
-To understand whether the effect of miR-290-295 cluster members on TFAP4 expression is not context-dependent, i.e. not a mutant-specific effect, and that the in silico prediction worked well, the analysis of the level of TFAP4 in wt cells transfected with mimics for miR-290-295 cluster members must be shown.
- Since, as the authors mention, expression of members of miR-290-295 cluster decreases after the exit from pluripotency of mESCs, the analysis of the relative expression of these miRNAs and TFAP4 protein during the exit from the pluripotent state

can make stronger the direct correlation of these miRs and TFAP4.
- The phenotype of the clones newly generated (Ago2&1and miR-290-295) should be characterized and shown. Do Ago KO ESCs show the same phenotype described previously (Ngondo et al. 2018)? Are they able to maintain self-renewal upon several passages? This is important to understand whether the authors are actually comparing undifferentiated cell lines for all KO or any of them undergo aberrant differentiation in ESC culture conditions.
The same concerns the miR-290-295 KO line. The KO of miR-290-295 cluster was described only in mouse embryo (Medeiros et al., 2011). This is the first report of the miR-290-295 cluster KO ESC line. Thus, I believe that this line deserves a phenotype analysis. Self-renewal maintenance and ability to exit from the undifferentiated state of this KO cell line must be described since it can also reveal an interesting functional correlation of the mRNA targets predicted by the multi-OMICs approach.
- Figure 4A: Why the authors do not include the expression of TFAP4 in DGCR8 and Ago KO cells? Is TFAP4 comparable to wt in these cells? Is this a cell line dependent effect?
- To conclude that this work "identifies TFAP4 as a novel key transcription factor in stem cells" the phenotype of TFAP4 in ESCs should be characterized since its role is reported only in cancer cells. Thus at least the self-renewal maintenance upon TFAP4 suppression should be evaluated.

Responses to reviewers

We would like to thank the reviewers for their comments. We highlighted all the applied changes in red in the revised version of our manuscript.

Referee #1:

In this manuscript, Schaefer and collaborators use an orthogonal approach that includes different complementary datasets (RNA-seq, AGO2 CLIP, and Ribo-seq) to determine the number of expressed genes regulated by miRNAs in mouse embryonic stem cells.

While other similar studies were published in the past, this one distinguishes itself by using different datasets obtained from mouse embryonic stem cells knockout for several key components of the miRNA pathway with the same genetic background. This approach is thus ideal for identifying "true" miRNA targets in mESCs. Though overall, I find this study quite compelling for better defining targets of miRNAs, some key issues must be addressed before publication:

We thank the reviewer for noticing the novelty of our work.

1- As AGO1 and AGO2 are the most abundant Argonautes in mESCs, it is logical to use for this study Ago1/Ago2 KO mESCs to identify mRNAs targeted by miRNAs. On the other hand, integrating AGO2 CLIP data in their comparative analysis is somewhat complicated as AGO1 binding mRNA targets are missing. As this group already produced the Ago2KO mESC (Ngondo et al, Stem Cell Reports 2017), it will be more appropriate to use those cells to perform a comparative analysis with AGO2 CLIP data. In addition, this new analysis will also be extremely informative in providing an estimate on how many mRNAs are targeted by the AGO1/miRNA silencing complex, which will help to understand better the role of other Argonaute proteins in gene silencing in mammalian cells.

We thank the reviewer for the suggestion. Unfortunately we previously observed very few changes in gene expression in single *Ago2*_KO mESCs (Ngondo *et al*, 2018), which makes it complicated to integrate with the AGO2-HEAP data. Moreover, we have previously shown by AGO2- and AGO1-RIP small RNAseq that the majority of miRNAs loaded in AGO2 are also loaded in AGO1, suggesting that very few miRNAs-mRNA interactions are left unconsidered when integrating the AGO2-HEAP data with the *Ago2&1*_KO RNAseq (Ngondo et al., 2018). Due to this large overlap in loaded miRNAs, we now integrate AGO1 and AGO2-RIP small RNA-seq from (Ngondo et al., 2018) in our revised manuscript and show once again that they are very similar (Fig EV1E). Due to this large overlap, we have now also used AGO-loaded miRNAs rather than the expressed miRNAs in mESCs in our integrated analysis (see revised Fig 1A) to further enrich for potentially functional miRNA-mRNA interactions. We acknowledge that although the AGO/miRNA complexes may be the same, there are likely functional differences between AGO1- and AGO2/miRNA complexes, however, we chose to focus on the identification of generally potentially functional interactions for the purposes of this study.

Additionally, in order to reinforce the AGO2-HEAP data from (Li *et al*, 2020), we also used the AGO2-CLIP dataset from (Bosson et al.,2014) in our revised manuscript. This adds further validation of the miR-290 MREs found in the 3'UTR of *Tfap4* (Fig 3F).

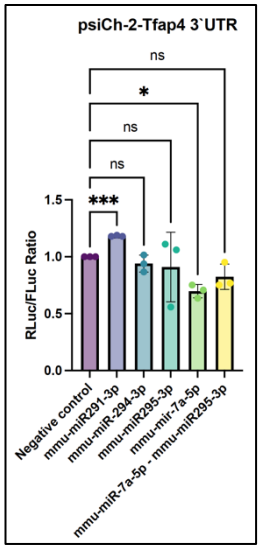
2- It is concerning to see from the PCA analysis that the RNA-seq data made with Dgcr8KO replicates are not clustering like other replicates' data. This can likely explain why the loss of Dgcr8 has no significant effect on ribosome occupancy of predicted miRNA targets when compared to data from Drosha, Ago1, and Ago2 KO mESCs. As Dgcr8 KO replicates are not

comparable, the authors should avoid using data obtained with those cells in their integrative approach to avoid any misleading.

We thank the reviewer for noticing this and have now removed the *Dgcr8*_KO samples from the manuscript. We are now using only WT, *Drosha*_KO, *Dicer*_KO and *Ago2&1*_KO mESC lines for our integrative analysis (Fig 1A).

3- The functional analysis of Tfap4 as a new miRNA target should be improved by demonstrating that the miR-291a binding sites found in the 3'UTR are essential for Tfap4 regulation by altering the miRNA binding sites with CRISPR/Cas9. Overexpressing miR-290-295 and monitoring the effect on Tfap4 only provide indirect proof of miR-290/Tfap4 regulation.

We agree with the reviewer that the suggested experiment would be a great addition to prove a direct link between the miR-290-295 cluster and *Tfap4* regulation. However, the generation and characterization of a new CRISPR mutant mESC line takes between 3 and 6 months. It was not possible in the time frame of this revision to perform such experiments for us. Additionally, we previously tried similar modifications of the 3'UTR of mRNAs and had difficulty to obtain stable engineered mRNAs.



In order to reinforce the link between the miR-290 family and *Tfap4* mRNA, we cloned the 3'UTR of *Tfap4* in a psiCHECK vector to perform luciferase assays. The results from the transfections of this plasmid with different miRNA mimics in HEK293T cells are presented in **Figure 1**. In HEK293T cells, we were not able to observe a significant regulation of *Tfap4* by miR-290 miRNAs. This could be due to the fact that these miRNAs are stem cells specific, and that other miRNAs may be acting on *Tfap4* in different cell types. Furthermore, we have previously faced many issues with performing luciferase assays in our system due to the low transfectability of mESCs.

Figure 1: Renilla to Firefly luciferase ratio in HEK293T cells upon transfection of the psiCHECK-3'UTR Tfap4 and different miRNA mimics.

We previously showed that the reintroduction of miR-291a-3p and miR-291a-5p in *miR-290-290*_KO mESCs can rescue the protein level of TFAP4 (revised Fig 3I). In order to reinforce this observation, we have performed a similar experiment in *Drosha*_KO mESCs and again observed a rescue of TFAP4 levels upon miR-291a-3p and miR-291a-5p mimic transfection (revised Fig EV3G).

Finally, previous work by (Ma et al., 2018) showed that *TFAP4* is repressed by the human miRNA hsa-miR-302c-3p in colorectal cancer cells. We show in our revised manuscript that the seeds of mmu-miR-291a-3p and hsa-miR-302c-3p are identical pointing to a strong conservation across species and that the MRE for miR-291a-3p is also conserved in the human 3'UTR of *TFAP4*. In addition, we reanalyzed AGO2 CLIP-seq data from hESCs (Lipchina *et al*, 2011) and observed several reads mapping to the hsa-miR-302c-3p binding site in the human 3'UTR of *TFAP4*, further suggesting a direct interaction between the two. We have now added these data in the revised Fig 3J, as we believe they further support a direct interaction between the miR-290-295 family and *Tfap4* in mESCs.

We hope that these novel experiments will convince the reviewer of the regulation of *Tfap4* by the miR-290 cluster in mESCs.

Besides those important issues, the authors can also push their analysis further to get more insights about miRNA/mRNA targets regulation:

We thank the reviewer for this suggestion and have now added several novel analyses to better characterize miRNA-mediated regulation in mESCs and the nature of the potential miRNA-mRNA interactions that we identify with our integrative analysis. We present these data in revised Fig 2A-C and Fig EV2F-J.

-With all the data in hand, the authors can provide some interesting information about the **miRNAs cooperation** by determining whether the presence of multiple miRNA binding sites on mRNA correlates with an increase in AGO2 binding and a decrease in ribosome association. The authors mention the limitation of their approach in this manner (a section of the discussion that I really appreciated), but I think it is still feasible to get some helpful information on this.

In order to address this comment, we examined the correlation between number of MREs per gene and indicators of functionality of these MREs such as AGO2-HEAP peaks and ribosome occupancy, which we present in revised Fig 2A-C. Although many of our analyses support a cooperative binding effect of miRNAs, some analyses do not. In particular, there was virtually no correlation between the number of predicted MREs (from TargetScan) and the mean intensity of the HEAP peaks (i.e. peak size) (**Figure 2**).

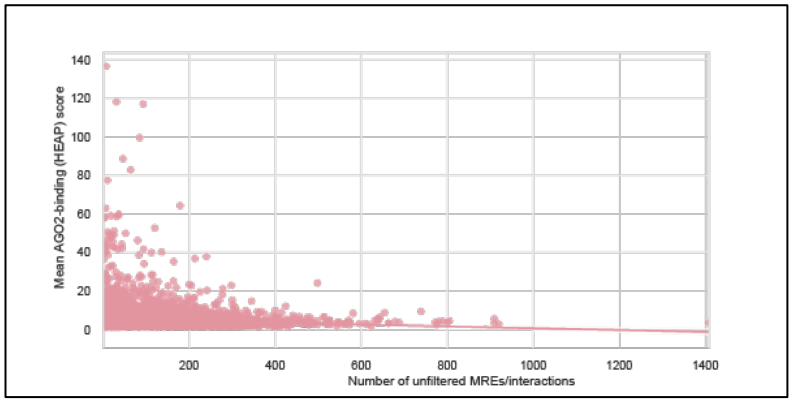


Figure 2: (No) correlation between the mean AGO2-binding peak intensities (HEAP) and the number of unfiltered MREs/interactions per gene. Pearson correlation coefficient: -0.08315614767291646

Furthermore, when comparing mean HEAP peak sizes for interactions from our integrative approach on a per gene basis, we saw virtually no correlation between HEAP peak size and number of interactions. (pearson corr: 0.09; again, each dot is one of our 759 genes) (**Figure 3**).

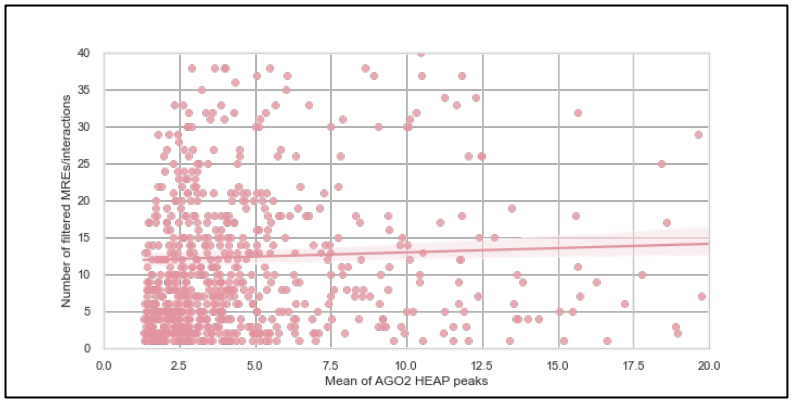
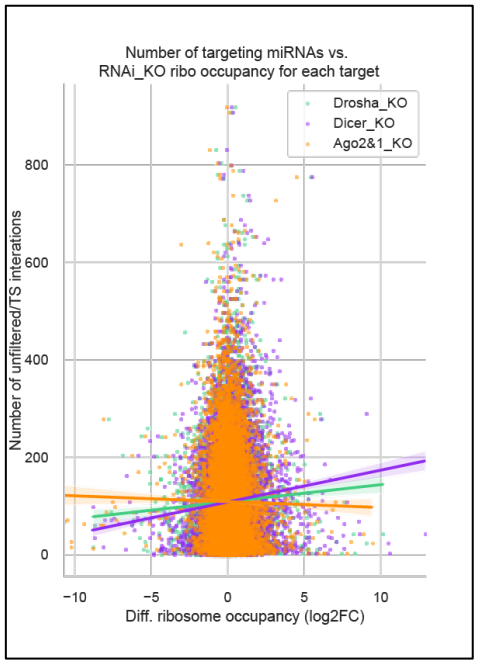


Figure 3: (No) correlation between mean AGO2-binding signal (HEAP) and the number of filtered (by integrative approach of the paper) MREs/interactions



Finally, ribosome occupancy did not increase with increasing number of interactions in *miRNA*_KO mESC lines when using TargetScan and miRNA-expression for filtering interactions.

Figure 4: Correlation plot for differential ribosome occupancy vs. number of unfiltered miRNA interactions (from TargetScan (TS)). *Drosha* pearsonr: 0.036 *Dicer* pearsonr: 0.079 *Ago2&1* pearsonr: -0.012

However, we show in Fig 2C, that there is a relatively pronounced correlation, when we filter for interactions using our integrative approach.

-It will be informative to perform a GO term/pathways analysis on identified mRNA targets to determine whether specific pathway/cellular functions are targeted by miRNAs to cell behaviors such as maintaining pluripotency state of ESC.

We thank the reviewer for this suggestion and have now performed such analysis that we present in our revised Fig 1C.

Referee #2:

Rigorous linkage of a miRNA to a function in human cells is much more difficult are rare than is often assumed. This manuscript takes an integrative approach that includes data from multiple, strategically chosen, knockout cell lines to prioritize candidate genes for regulation by miRNAs. The authors present solid data that TFAP4 is regulated by miR-290-295 in mESC's. In itself, that is not a particularly major result. The value is in setting a standard for rigorous analysis of miRNAs.

We thank the reviewer for noticing the quality of our work.

Overall, this is a worthy paper that should eventually be published. I hope that it will be widely read and appreciated.

However, there are numerous issues that need to be addressed prior to publication.

Overall, it is essential that the authors restrict themselves to wording that is supported by their data. The entire paper sends a message that researchers need to work hard to link miRNAs to function, that message is undermined by poorly supported conclusions.

Page 2. This manuscript DOES NOT show that 9% of genes are functionally and directly regulated by miRNAs. This is the type of overstatement that has been made in papers describing miRNA action for the past twenty years. In reality, the authors have identified 9%

of genes as **CANDIDATES** for regulation by miRNAs. There is a big difference between these two conclusions. The type of data presented in Figure 4 are necessary to even begin to believe that a proposed functional linkage is real.

We agree with the reviewer about the use of "CANDIDATES" and now rigorously update the wording in our revised manuscript. Additionally, in order to reinforce our previous data, we now present more validation of candidate genes (Fig 2E-F and Fig EV3F).

Page 3. The introductory paragraph, in common with many other paragraphs, is too long. This is a difficult manuscript to understand because it is complex and nuanced. Long, overly-complicated paragraphs are an unnecessary hindrance for readers.

We agree with the reviewer and have reformatted our introduction paragraph.

Page 3. The introductory paragraph is also poorly chosen. It does nothing to focus the reader's attention on the gap in knowledge addressed by the manuscript. It merely summarizes information about miRNAs that could be found in any recent review. The paragraph should be retained (albeit broken up into small units). However a new paragraph(s) should be added to describe the real purpose of this manuscript - to provide a better strategy for solving the problem of linking miRNA expression to believable control of gene expression.

We agree with the reviewer and have reformatted our introduction paragraph.

Page 5. Again, the notion that the data show that 9% of expressed genes are subject to direct and functional miRNA-regulation is not only incorrect, the notion works against the value of the rigorous analysis emphasized by the authors. One either approaches these numbers derived from RNAseq data with skepticism or one does not.

We thank the reviewer and have now rewritten this part about candidate genes.

Page 6. It would be useful to have a scheme depicting the integrative experimental strategy as part of Figure 1. Figure 2A is too small. Indeed, it might be better to introduce the overall experimental scheme first as a better designed Figure 1A to orient the reader.

We agree with the reviewer and now present our integrative analysis in Fig 1A, with a few changes requested by other reviewers.

Page 6. The first paragraph on page 6 contains a lot of information and is an example of a paragraph that would be more understandable if it were sub-divided into more focused paragraphs.

We thank the reviewer and have now reformatted this part of the manuscript.

Page 6. Figure 1C - This is not a very useful figure. A Venn diagram (or diagrams) showing the overlap would be better. It is dangerous to treat the RNAi factors as interchangeable.

We thank the reviewer for this suggestion and now present Venn diagrams for up and downregulated genes in *miRNA*_KO mESC lines in revised Fig EV1D.

Page 7. The published AGO2 binding data and the TargetScan predictions are introduced uncritically. How were AGO2 binds judged according to their significance? Why should the reader believe that the boundaries for judging significance were wisely chosen? Regarding TargetScan, predictive methods are notoriously inaccurate. I do not think that the authors conclusions will be much affected, but the way they introduce these methods strikes this

reviewer as naïve. **This would be a good place to reinforce the lesson that all of these methods have enormous limitations and are dangerous for the unwary.**

We agree with the reviewer and have now reinforced the lesson about limitations of all approaches in our revised manuscript.

Page 8. “High confidence” is not useful terminology here. These are the most highly ranked miRNA:mRNA pairs, and therefore the ones most likely to be worth investing time in experimental validation.

We have changed the text according to the reviewer’s suggestion.

Page 8-9. The mass spec/ribosome profiling data adds little to the paper and is inadequately described. These are difficult experiments and merit much more attention to be believable. These weak data are not necessary and should be deleted.

Following requests from other reviewers, we have kept these data in our manuscript and described them better. We have also used them to assess combinatory effects of miRNAs as suggested by reviewer 1 (Revised Fig 2).

Page 9. The paragraph describing the miR-290-295 cluster begins one of the most important sections in this paper. It is far too long and needs to be sub-divided.

We have changed the text according to the reviewer’s suggestion.

Page 9. It is essential that the authors present compelling and transparent experimental data demonstrating the high expression of miR-290-295 in mECs. How many miRNAs are in this cluster? How many miRNAs are expressed per cell? They entire analysis pivots on believing that the concentration of miRNA is high enough to affect the biology of mECs. The authors need to make a stronger case for the physical basis of miR-290-295 action.

We thank the reviewer for this suggestion and have now added a better description of the cluster (Revised Fig 3J) and its expression in mESCs (Revised Fig EV3B). In addition, we have added several validations of mRNA expression by qRT-PCR in *miR-290*_KO mESCs in Fig EV3F for known targets and non-targets of the miR-290 miRNA cluster.

Page 10-11. What does the complementarity of miR-290-295 and Tfap4 look like? A physical interaction must underlie regulation. What is it? What does the AGO2 binding data look like? How many miR-290 binding sites are there? Are there binding sites for other well-expressed miRNAs. A much stronger case needs to be made and the authors need to think carefully about how to do this.

We thank the reviewer for highlighting this point and now present a novel panel (Fig 3F) illustrating the complementarity of miR-290 cluster with the 3’UTR of *Tfap4* and the nature of the AGO2 HEAP peaks observed in mESCs. We also show similar patterns of AGO2 binding in human ESCs in previously published CLIP-seq data, as well as the conservation of the miR-291-3p MRE in the human *TFAP4* 3’UTR (Fig 3J).

Note: It IS NOT necessary that this manuscript conclusively identify Tfap4 as a target. It is enough that the authors thoughtfully explore how difficult target identification might be even when they go to the great length of knocking out a miRNA cluster. Providing a false sense of certainty brings this work down to the level of the many poor papers that have been published on miRNAs.

We thank the reviewer for this remark. Nevertheless, we have added another set of experiments to reinforce our data on the regulation of *Tfap4* by the miR-290 family in Figures 3 and 4 as per the requests of (see other reviewer comments) and we clearly state the limitation of our study in the revised version of the manuscript.

Page 12-14. Without stronger reasons to believe that Tfap4 is a legitimate target, the analysis of downstream target genes carries little meaning.

Please see our novel experiments presented in Figures 3 and 4 reinforcing our point that *Tfap4* is a novel important target of miRNAs in mESCs.

In summary, this is a potentially important paper. It would be acceptable if it were revised to describe the exact conclusions the data support and no more. Unfortunately, as written, it falls into the common trap of feeling the need to provide some positive conclusion about the roles of miRNAs in cells. Perhaps the data support that conclusion, but I am unpersuaded. Either more transparent and thoughtfully described data needs to be added, or the text needs to be changed to reflect a more uncertain reality. The most important conclusion is that **a layered approach is necessary to have any hope of assigning a function to a miRNA**, and even then the connection is less then certain. While this is a lengthy review, I believe that these suggestions can be addressed through relatively straightforward changes to text and figures.

Minor comments:
1) "The extensive datasets developed in this study will support the development of improved predictive models for miRNA-mRNA functional interactions." I do not see the ways how it can be incorporated into predication models not generating the same datasets for different cell lines or cell context specific conditions. I think the authors mean to imply that the workflow they illustrate will be useful
We remove this sentence in the revised version of our manuscript.

2) This sentence requires following sentence with explanation what context specific factors authors talking about. "...relevant miRNA interactions could be partially attributed to the lack of incorporating important context-specific factors."
We thank the reviewer and rewrite this sentence in our revised manuscript.

3) Interaction score calculation. Knowing that miRNAs mainly function through interaction with AGO proteins would last criterion calculating interaction score supposed to be not "higher than 0.5 in at least two miRNA_KO mutants", however specifically between Ago2&1_KO and another mutant. If gene does not show expression change in this KO cell it less likely to be regulated via RNAi. Also inclusion of targetscan prediction as a factor is questionable when it was discussed before as overestimating approach for miRNA-RNA binding identification.
During revisions, we have carefully reassessed the rule set to define our miRNA target predictions. Indeed, TargetScan appears to be quite permissive, which is why we implemented a TargetScan score threshold to be necessary (not sufficient) for considering a gene as predicted target. Further necessities for genes to be considered as positive predictions are a minimal (>10CPM) miRNA loading in AGOs, a called AGO2-binding peak in the (HEAP data from (Li *et al*, 2020)) and a “common” upregulation in at least two of the three mutants. Here, we applied an adjusted p-value threshold of 0.2, accounting for the often-subtle regulation effects of miRNAs, and believe that false positives introduced by this loose threshold are largely filtered by the other three integrated data sets.
While our approach tries to maximize the certainty about whether each given gene is directly targeted by miRNAs, we do recognize that more experimental validations may be required to reach absolute certainty.

All modifications to the integrative analysis have been annotated properly and transparently in the Material and Method section.

4) "we observed seemingly functional interactions in regions outside of the 3'UTR (for example Figures S2E, F), which led us to exclude the 3'UTR-centric TargetScan context++ score model from the filtering, still leading to a surprisingly low number of 707 identified miRNA-targeted genes (Figure 2F)" What is distribution of binding sites? I am left guessing about the exact meaning of this sentence. Not transparent at all.
We thank the reviewer for noticing the unclarity of this sentence and now present data demonstrating the localization of the miRNA/mRNA interaction from our integrative approach in Fig EV2H.
In accordance with reviewer 1, we have added a thorough characterization of the identified miRNA targets, including the distribution of binding sites (Fig 2 and EV2). The paragraph describing the 759 targets has been modified to make more precise and clear statements.

5) miRNA can act to repress translation without effecting stability therefore mass spectrometry and ribosomal profiling had focus not only on 707 genes which already known to be effected at mRNA level.
Work is poorly visualized with low transparancy:
• PCA figures have to be plotted with scale of X and Y axis, to have better insight about the distance among samples.
We have changed the plot according to the reviewer's suggestion.

• Figure1 C would be better to plot as venn diagram. It would be more transparent and let to see exact overlap between different KO cell lines.
We have changed the plot according to the reviewer's suggestion.

• Figure 2. The main panel F where 9% of direct regulated genes come from is covered by supplementary type panels B, C, D, E. To make Figure2 F more transparent would be great to have supportive panel showing where these 444 and 707 genes stand in context of other genes by expression and significant level. It could be visualized as volcano or plot or heat map etc.
As requested by reviewer 1 as well, we have added an extensive characterization of the 759 identified miRNA targets in Fig 2 and EV2, including an expression distribution (Fig EV2F). We did abstain from adding a significance and/or upregulation distribution, as this is very close to our selection criteria and therefore does not add much value to the manuscript. For transparency, we have added the significance and upregulation-distribution here (**Figures 5 and 6**).

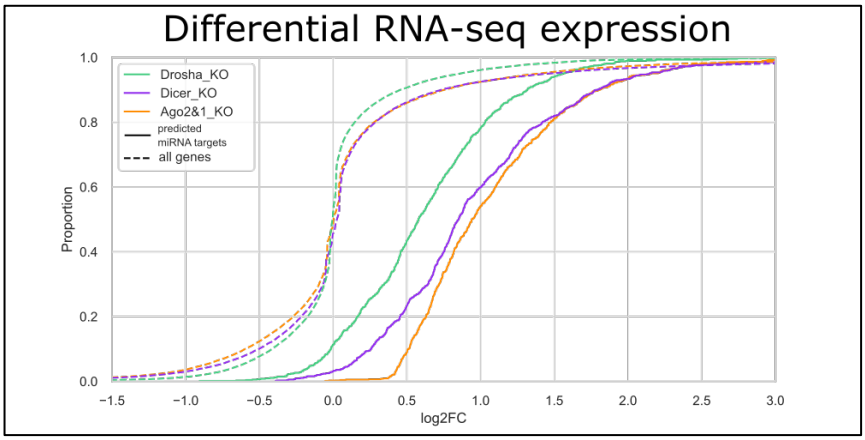


Figure 5: Differential expression distribution (log2FC) of 759 predicted miRNA targets in *miRNA*_KO mutants, versus all genes.

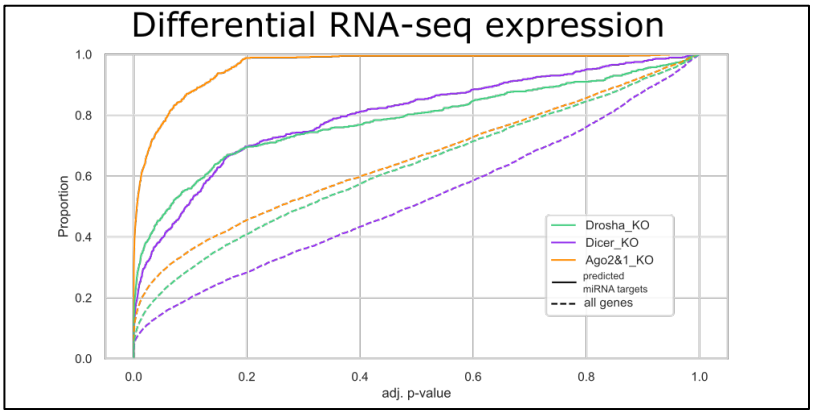


Figure 6: Differential expression distribution (adjusted p-value) of 759 predicted miRNA targets in *miRNA*_KO mutants, versus all genes.

• For this paper to have an impact, it needs to be understandable to a wide readership. The authors should take a step back and consider how to make their data both transparent and intuitive for a typical interested reader who does not consume large data sets regularly.

We thank the reviewer for all their comments and hope that our revised manuscript properly addresses all the points raised.

Referee #3:

In the manuscript titled "Integrative analysis allows a global and precise identification of functional miRNA target genes in mESCs " by Schaefer and colleagues, the authors describe an elegant multi-OMICs integrative approach to build a map of functional interaction of miRNAs-mRNAs in ESCs. Indeed, the simple miRNA expression analysis in specific cell types cannot fully reveal the functional miRNA-mRNA interactions leading to gene expression regulation. The work of Schaefer and colleagues elegantly addresses this point and aims to demonstrate that a multi-OMICs approach based on context-dependent datasets can allow to predict with a significant accuracy miRNA-mRNA functional interaction. Although the approach cannot be considered completely original compared to the papers cited by the authors and displays some limitations in the prediction of functional interaction, it definitely represents an important step forward compared to previous studies. The manuscript is well-written and methodological approaches and results are properly presented. One of the main limitations of this study is that the analysis is mainly based on RNA-seq and not on protein levels, but the authors extensively discuss this point and they also partially overcome this limitation using ribosome profiling. Also, the limitation on the sensitivity of the approach is well discussed and partially overcome by reducing thresholds.

We thank the reviewer for noticing the quality and importance of our work.

Although the OMICS approach appears solid the experimental part of this work proves to be incomplete and must be improved to make stronger the results obtained by in silico analysis and to support some conclusions. After providing more functional data, the strength of this in silico approach can be truly appreciated. I believe that the work done by Schaefer and colleagues will be of broad interest for the scientific community once the authors have solved some concerns indicated below.

Major points:

- A such consistent and complex in silico analysis requires a more extensive validation of the predicted miRNA-mRNA functional interactions to convince about the strength and the efficacy of this approach. Beyond the data presented for TFAP4, that anyway (as discussed below) need to be increased, some more miRNA-mRNA functional interactions must be demonstrated on the basis of the in silico prediction.

We thank the reviewer for this suggestion and as replied previously for reviewer 1, we have added several new data to reinforce the validation of our predicted targets.

In addition to the full proteome and the ribosome profiling experiments, we have now used the updated 2020 miRTarbase database (Huang *et al*, 2020) to assess how many of our candidates were previously validated in other cellular contexts. As shown in Fig 2E, almost half of our candidates are also present in the miRtarbase database.

In addition, we also now add experimental validations at protein (Fig 2F) and RNA levels (Fig EV3F).

We hope that these additional validations will convince the reviewer of the validity of our integrated approach.

- The integrative approach used, and the data shown in figure 4 indicate that TFAP4 can be a direct target of miR-290-295 cluster. Again, this functional observation is important to make more solid the conclusions based on the in silico approach. To demonstrate that TFAP4 is directly regulated by members of miR-290-295 cluster, luc assay using wt 3'UTR of TFAP4 and the relative miRNA binding site mutant coupled with overexpression of miR-290-295 cluster members must be reported.

We agree with the reviewer and as explained before we have performed such an experiment in the past without great success (please see explanations to reviewer1).

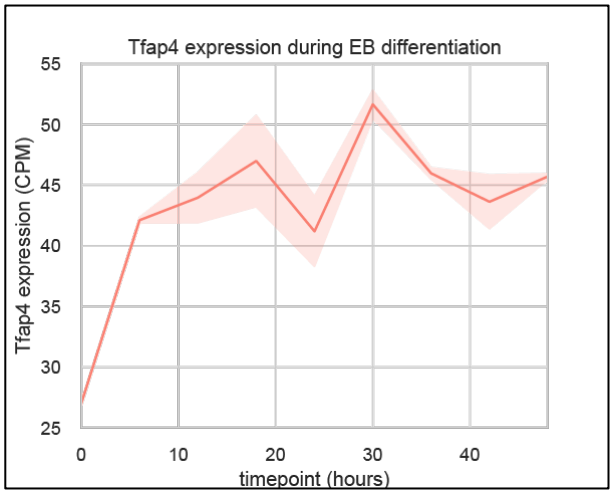
We hope that the additional experiments performed in *Drosha*_KO mESCs (Fig EV3G) and the conservation of the interaction in hESCs (Fig 3J) will suffice to convince the reviewer of the interaction between the miR-290 family and *Tfap4*.

-To understand whether the effect of miR-290-295 cluster members on TFAP4 expression is not context-dependent, i.e. not a mutant-specific effect, and that the in silico prediction worked well, the analysis of the level of TFAP4 in wt cells transfected with mimics for miR-290-295 cluster members must be shown.

We thank the reviewer for this suggestion. However, as shown now in Fig 3H the level of TFAP4 in WT cells is very low and sometimes even undetectable by Western blot. Therefore, instead of performing this experiment in WT cells, as we have done it in *miR290*_KO mESCs, (Fig 3I) and we have repeated it in *Drosha*_KO mESCs (Fig EV3G). In both cell lines, we were able to rescue TFAP4 expression upon mimics transfection.

- Since, as the authors mention, expression of members of miR-290-295 cluster decreases after the exit from pluripotency of mESCs, the analysis of the relative expression of these miRNAs and TFAP4 protein during the exit from the pluripotent state can make stronger the direct correlation of these miRs and TFAP4.

We thank the reviewer for this suggestion. It is well known that miR-290-295 cluster expression decreases upon mESC differentiation (Ciaudo *et al*, 2009). In order to assess the expression of *Tfap4* upon exit from pluripotency, we extracted data from a previously published paper (Gloss *et al*, 2017) and were able to show an increase in *Tfap4* expression



upon exit from pluripotency (**Figure 7**). Although this observation is only correlative, it is indicative of the repressive effect of the miR-290 family on *Tfap4*. Since we have now added more evidence for a regulation of miR-290 on *Tfap4*, we did not include this correlative data to our revised manuscript. However, we are happy to do so if the reviewer thinks it is important.

Figure 7: Expression of *Tfap4* as measured by RNA-seq during an embryoid body differentiation protocol. First 48 hours, indicative of the pluripotency exit, are shown.

- The phenotype of the clones newly generated (Ago2&1and miR-290-295) should be characterized and shown. Do Ago KO ESCs show the same phenotype described previously (Ngondo et al. 2018)? Are they able to maintain self-renewal upon several passages? This is important to understand whether the authors are actually comparing undifferentiated cell lines for all KO or any of them undergo aberrant differentiation in ESC culture conditions.

We thank the reviewer for this suggestion. In fact, both cell lines characterization have been already published (*Ago2&1*_KO recently in (Müller *et al*, 2022) and *miR-290*_KO in (Wang *et al*, 2017). We have previously shown that the *Ago2&1*_KO mESC line is locked in pluripotency like the other *miRNA*_KO mESC lines (Müller *et al*, 2022; Bodak *et al*, 2017; Cirera-Salinas *et al*, 2017). Wang et al. (2017) have characterized the *miR-290*_KO mESCs and shown that they are pluripotent and do not undergo spontaneous differentiation. We have also clarified these points in the text of our revised manuscript.

The same concerns the miR-290-295 KO line. The KO of miR-290-295 cluster was described only in mouse embryo (Medeiros et al., 2011). This is the first report of the miR-290-295 cluster KO ESC line.
To generate the *miR-290*_KO mESCs in the same genetic background than the other *miRNA*_KO mESCs, we used previously published sgRNAs from (Wang *et al*, 2017). Our characterization of this cell line is similar to the one published so we did not add more information and now refer to the paper more clearly in our revised manuscript.

Thus, I believe that this line deserves a phenotype analysis. Self-renewal maintenance and ability to exit from the undifferentiated state of this KO cell line must be described since it can also reveal an interesting functional correlation of the mRNA targets predicted by the multi-OMICs approach.
Please see (Wang *et al*, 2017).

- Figure 4A: Why the authors do not include the expression of TFAP4 in DGCR8 and Ago KO cells? Is TFAP4 comparable to wt in these cells? Is this a cell line dependent effect?
We thank the reviewer for this comment and have added the expression of TFAP4 in all *miRNA*_KO and *miR290*_KO mESC lines in revised Fig 3H.

- To conclude that this work "identifies TFAP4 as a novel key transcription factor in stem cells" the phenotype of TFAP4 in ESCs should be characterized since its role is reported only in cancer cells. Thus at least the self-renewal maintenance upon TFAP4 suppression should be evaluated.
We thank the reviewer for this comment and have added more experiments on TFAP4 in Fig 4 and EV4.

We wanted to assess whether the restoration of the proper expression of TFAP4 in *Drosha*_KO mESCs could partially rescue some common stem cell phenotypes that are perturbed in the *miRNA*_KO lines such as proliferation or heterogeneity of the NANOG transcription factor. To do so, we have generated a double mutant mESC line, where we knock out *Tfap4* in the *Drosha*_KO background and characterized it (Fig 4D-E and EV4B-E). Unfortunately, the depletion of *Tfap4* in a *Drosha*_KO background is not sufficient to rescue the proliferation defect of *Drosha*_KO cells, or the homogenous expression of the NANOG TF.

References:

Bodak M, Cirera-Salinas D, Yu J, Ngondo RP & Ciaudo C (2017) Dicer , a new regulator of pluripotency exit and LINE-1 elements in mouse embryonic stem cells. *FEBS Open Bio* 7: 204–220

Bosson ADD, Zamudio JRR & Sharp PAA (2014) Endogenous miRNA and Target Concentrations Determine Susceptibility to Potential ceRNA Competition. *Mol Cell* 56: 1–13

Ciaudo C, Servant N, Cognat V, Sarazin A, Kieffer E, Viville S, Colot V, Barillot E, Heard E & Voinnet O (2009) Highly dynamic and sex-specific expression of microRNAs during early ES cell differentiation. *PLoS Genet* 5: e1000620

Cirera-Salinas D, Yu J, Bodak M, Ngondo RP, Herbert KM & Ciaudo C (2017) Noncanonical function of DGCR8 controls mESC exit from pluripotency. *J Cell Biol* 216: 355–366

Gloss BS, Signal B, Cheetham SW, Gruhl F, Kaczorowski DC, Perkins AC & Dinger ME (2017) High resolution temporal transcriptomics of mouse embryoid body development reveals complex expression dynamics of coding and noncoding loci. *Sci Rep* 7: 1–11

Huang HY, Lin YCD, Li J, Huang KY, Shrestha S, Hong HC, Tang Y, Chen YG, Jin CN, Yu Y, *et al* (2020) MiRTarBase 2020: Updates to the experimentally validated microRNA-target interaction database. *Nucleic Acids Res* 48: D148–D154

Li X, Pritykin Y, Concepcion CP, Lu Y, La Rocca G, Zhang M, King B, Cook PJ, Au YW, Popow O, *et al* (2020) High-Resolution In Vivo Identification of miRNA Targets by Halo-Enhanced Ago2 Pull-Down. *Mol Cell* 79: 167–179

Lipchina I, Elkabetz Y, Hafner M, Sheridan R, Mihailovic A, Tuschl T, Sander C, Studer L & Betel D (2011) Genome-wide identification of microRNA targets in human ES cells reveals a role for miR-302 in modulating BMP response. *Genes Dev* 25: 2173–2186

Müller M, Schaefer M, Fäh T, Spies D, Hermes V, Ngondo RP, Peña-Hernández R, Santoro R & Ciaudo C (2022) Argonaute proteins regulate a specific network of genes through KLF4 in mouse embryonic stem cells. *Stem Cell Reports* 17: 1–11

Ngondo RP, Cirera-Salinas D, Yu J, Wischnewski H, Bodak M, Vandormael-Pournin S, Geiselmann A, Wettstein R, Luitz J, Cohen-Tannoudji M, *et al* (2018) Argonaute 2 Is Required for Extra-embryonic Endoderm Differentiation of Mouse Embryonic Stem Cells. *Stem Cell Reports* 10: 1–16

Wang X-WW, Hao J, Guo W-TT, Liao L-QQ, Huang S-YY, Guo X, Bao X, Esteban MA & Wang Y (2017) A DGCR8-Independent Stable MicroRNA Expression Strategy Reveals Important Functions of miR-290 and miR-183–182 Families in Mouse Embryonic Stem Cells. *Stem Cell Reports* 9: 1618–1629

Dear Prof. Ciaudo,

Thank you for the submission of your revised manuscript to our editorial offices. I have now received the reports from the three referees that were asked to re-evaluate your study, you will find below. As you will see, the referees now support publication of the study in EMBO reports. However, referees #1 and #2 have remaining points and suggestions to improve the manuscript I ask you to address in a final revised manuscript. I would support the suggestion of referee #2 to shorten and simplify the manuscript and to break some paragraphs apart to make them more accessible. Please also provide a final p-b-p-response addressing the remaining points of the referees.

Moreover, I have these editorial requests:

- Please provide a more active title (with not more than 100 characters). How about:
Global and precise identification of functional miRNA targets in mESCs by integrative analysis

- Please reduce the number of keywords to five.

- Please provide the abstract written in present tense.

- We updated our journal's competing interests policy in January 2022 and request authors to consider both actual and perceived competing interests. Please review the policy https://www.embopress.org/competing-interests and update your competing interests if necessary. Please name this section 'Disclosure and Competing Interests Statement' and put it after the Acknowledgements section.

- Please order the manuscript sections like this, using these names:
Title page - Abstract - Key Words - Introduction - Results - Discussion - Materials and Methods - Data availability section - Acknowledgements - Author contributions - Disclosure and Competing Interests Statement - References - Figure legends - Expanded View Figure legends

- We do not accept Adobe Illustrator figures. Please provide individual production quality figure files as .eps, .tif, or .jpg (one file per figure), of main figures and EV figures.

- It seems separate callouts for panels 1E and 4E are missing. Please check and make sure that all figure panels are called out separately and sequentially.

- Please make sure that the number "n" for how many independent experiments were performed, their nature (biological versus technical replicates), the bars and error bars (e.g. SEM, SD) and the test used to calculate p-values is indicated in the respective figure legends (main, EV and Appendix figures), and that statistical testing has been done where applicable. Please avoid phrases like 'independent experiment', but clearly state if these were biological or technical replicates. Please add complete statistical testing to all diagrams (for main, EV and Appendix figures). Please also indicate (e.g. with n.s.) if testing was performed, but the differences are not significant.

- Please add scale bars of similar style and thickness to all the microscopic images (main and EV figures), using clearly visible black or white bars (depending on the background). Please place these in the lower right corner of the images. Please do not write on or near the bars in the image but define the size in the respective figure legend. Presently, for most images the scale bars that are too small/thin and have text nearby. Please check.

- EV Tables 1-7 are too large to be displayed as tables. These are datasets. Please name this Dataset EVx and upload them as dataset file (excel file) with a title and a legend on the first TAB. Finally, please change the callouts for these items to Dataset EVx throughout the manuscript text.

- Table EV8 is then Table EV1. Please name this Table EV1, update the callouts and upload it with a title and a legend in the first TAB.

- Please remove all the table legends from the manuscript text file.

- Please ZIP together all source data for one figure (main and EV figures) and upload one folder per figure.

- Please remove the sentence 'Source data are available online for this figure' from the figure legends. The SD will be linked to each figure directly and this can be clearly seen by the reader.

- Finally, please find attached a word file of the manuscript text (provided by our publisher) with changes we ask you to include in your final manuscript text, and some queries, we ask you to address. Please provide your final manuscript file with track

changes, in order that we can see any modifications done.

In addition, I would need from you:
- a short, two-sentence summary of the manuscript (around 35 words).
- three to four short bullet points highlighting the key findings of your study
- a schematic summary figure (in jpeg or tiff format with the exact width of 550 pixels and a height of not more than 400 pixels) that can be used as a visual synopsis on our website.

Please provide these in two separate files (text and figure).

I look forward to seeing the final revised version of your manuscript when it is ready. Please let me know if you have questions regarding the revision.

Please use this link to submit your revision: https://embor.msubmit.net/cgi-bin/main.plex

Best,

Achim Breiling
Senior Editor
EMBO Reports

Referee #1:

I appreciate the effort made by the authors to answer reviewers' comments. This revised version of the manuscript is substantially improved, but it still needs to include a discussion about the following two concerns I raised before its publication.

-The authors should mention in the text the limitation of their system to directly test the contribution of the miR-290-295 cluster on Tfap4 regulation. That will clarify why the authors cannot use a state-of-the-art approach to test the regulation of Tfap4 by those microRNAs directly, and alternative methods are needed. I also think that adding the luciferase assay included in the rebuttal will be informative (this will illustrate well why the biological context does matter when we assess microRNA/target regulation).

-It will also be essential to include in the manuscript their findings about the absence of correlation between AGO2-binding signal and the number of MRE/interactions per gene. That will be highly informative for the community (and especially for non-specialists) as RBP CLIP-data and putative binding sites are often seen as similar readouts.

Referee #2:

I appreciate that the manuscripts received feedback from three reviewers and that my suggestion to simplify the manuscript by removing weaker data that was not central to the main point was not taken. The manuscript remains too complex, but I accept that is the authors and editors decision. I have two requests.

Please quantitate, on a per cell basis, the number of miR-290 family members per cell. Citing an old paper is not sufficient. The value in this paper is that it offers lessons to researches, miRNA per cell is an important lesson.

Also, some of the paragraphs have become longer and even more complex. The authors should consider breaking these paragraphs apart to make them more accessible.

Referee #3:

Although the authors have answered to most of my concerns, some requests remained not fully address. However, I can understand the difficulties that the authors may have encountered in performing some experiments (i.e. ESC transfection). Thus, I believe that the quality of the manuscript is significantly improved, enough to be acceptable for publication.

Responses to reviewers

We would like to thank the reviewers for their comments. We highlighted all the applied changes in red in the revised version of our manuscript.

Referee #1:

I appreciate the effort made by the authors to answer reviewers' comments. This revised version of the manuscript is substantially improved, but it still needs to include a discussion about the following two concerns I raised before its publication.

We thank the reviewer for noticing the improvements made.

-The authors should mention in the text the limitation of their system to directly test the contribution of the miR-290-295 cluster on Tfap4 regulation. That will clarify why the authors cannot use a state-of-the-art approach to test the regulation of Tfap4 by those microRNAs directly, and alternative methods are needed. I also think that adding the luciferase assay included in the rebuttal will be informative (this will illustrate well why the biological context does matter when we assess microRNA/target regulation).

We thank the reviewer for these suggestions and have added a novel paragraph in our discussion about the limitations of our system. Additionally, we have added the luciferase assays performed in HEK293T cells as Appendix Fig S3.

-It will also be essential to include in the manuscript their findings about the absence of correlation between AGO2-binding signal and the number of MRE/interactions per gene. That will be highly informative for the community (and especially for non-specialists) as RBP CLIP-data and putative binding sites are often seen as similar readouts.

We have also applied this recommendation and have added an additional Appendix Fig S1 with a figure illustrating the absence of a clear correlation between AGO2-binding signal and the number of MREs per gene.

Referee #2:

I appreciate that the manuscripts received feedback from three reviewers and that my suggestion to simplify the manuscript by removing weaker data that was not central to the main point was not taken. The manuscript remains too complex, but I accept that is the authors and editors decision. I have two requests.

Please quantitate, on a per cell basis, the number of miR-290 family members per cell. Citing an old paper is not sufficient. The value in this paper is that it offers lessons to researches, miRNA per cell is an important lesson.

We have now performed a quantification of the number of miR-290 family members per cell according to the method used in (Gu *et al*, 2016) and added a novel figure in Appendix Fig S2.

Also, some of the paragraphs have become longer and even more complex. The authors should consider breaking these paragraphs apart to make them more accessible.

We thank the reviewer for this suggestion and have now introduced more breaks to many paragraphs.

Referee #3:

Although the authors have answered to most of my concerns, some requests remained not fully address. However, I can understand the difficulties that the authors may have encountered in performing some experiments (i.e. ESC transfection). Thus, I believe that the quality of the manuscript is significantly improved, enough to be acceptable for publication.

We thank the reviewer for noticing the quality of our manuscript and as mentioned above, we have added an additional paragraph in our discussion about the limitations of our system.

References:

Gu KL, Zhang Q, Yan Y, Li TT, Duan FF, Hao J, Wang XW, Shi M, Wu DR, Guo WT, *et al* (2016) Pluripotency-associated miR-290/302 family of microRNAs promote the dismantling of naive pluripotency. *Cell Res* 26: 350–366

Prof. Constance Ciaudo
ETHZ
D-Biol
IMHS HPL G32.1
Otto-Stern-Weg 7
Zurich 8093
Switzerland

Dear Prof. Ciaudo,

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.

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."

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.

Yours sincerely,

Achim Breiling
Editor
EMBO Reports

****

THINGS TO DO NOW:

Please note that you will be contacted by Wiley Author Services to complete licensing and payment information. The required 'Page Charges Authorization Form' is available here: https://www.embopress.org/pb-assets/embo-site/er_apc.pdf - please download and complete the form and return to embopressproduction@wiley.com

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-2022-54762V3 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: Constance Ciaudo	**USEFUL LINKS FOR COMPLETING THIS FORM**
Journal Submitted to: Embo Reports	The EMBO Journal - Author Guidelines
Manuscript Number: EMBOR-2022-54762V2	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	Material and Methods, Addgene

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	Material 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	Material 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	Material 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.	Yes	Material and Methods, cells were tested for mycoplasma

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.	Not Applicable	
Animal observed in or captured from the field: Provide species, sex, and age where possible.	Not Applicable	
Please detail **housing and husbandry conditions**.	Not Applicable	

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?	Yes	Acknowledgments

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.	Yes	Wettstein et al., 2016 DOI: 10.1007/7651_2015_213; Bodak et al., 2016 DOI: 10.1007/978-1-4939-3372-3_16

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	Material and Methods, Figures and Source data
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	No blinding was done
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	Material and Methods, Figures

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	Figures, Source data
In the figure legends: define whether data describe **technical or biological replicates**.	Yes	Figures

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 availability section and Reagent_Tools Table
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	Reference list and Reagents_Tools Table