THE
EMBO
JOURNAL

# Intragenic transcriptional interference regulates the human immune ligand MICA.

Da Lin, Thomas K. Hiron and Christopher A. O'Callaghan.

**Review timeline:**

| | |
|---|---|
| Submission date: | 19th April 2017 |
| Editorial Decision: | 20th July 2017 |
| Revision received: | 17th November 2017 |
| Editorial Decision: | 26th January 2018 |
| Revision received: | 8th Febuary 2018 |
| Accepted: | 9th Febuary 2018 |

Editor: Anne Nielsen

## Transaction Report:

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

1st Editorial Decision 20th July 2017

Thank you again for submitting your manuscript for consideration by the EMBO Journal and for sending me a preliminary point-by-point response outlining the data you could include to address the concerns raised by our three referees. I have now read it and discussed it with my colleagues in the editorial team.

The outcome is that I would like to invite you to submit a revised manuscript in which you address the comments made by all three reviewers along the lines discussed in your point-by-point response. I should add that it is EMBO Journal policy to allow only a single round of revision, and acceptance of your manuscript will therefore depend on the completeness of your responses in this revised version.

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

---
REFEREE REPORTS

Referee #1:

This manuscript describes a very clear case of promoter interference. Promoter interference is a mechanism of regulation of gene expression where a non-coding RNA initiated at an upstream promoter inhibits the activity of the proximal promoter simply by being made. This mechanism is well described in bacteria and in yeast, while relatively poorly documented in higher eukaryotes. Here, the authors show data suggesting that only the proximal "standard" promoter of the MICA gene (encoding a ligand for an immune receptor) is coding, while the transcript initiated at the distal promoter is unstable and never gets translated. While RNAi and overexpression experiments suggest that the upstream RNA does not have an effect in trans, CRISPR inactivation of the distal promoter elegantly demonstrated that this promoter has a negative effect on expression of the MICA gene.

 

Similarly, a series of transgenes carrying mutations either in the distal or in the proximal promoter confirmed that the distal promoter caused reduced transcription from the proximal promoter. These experiments also showed that insertion of a transcription terminator in between the two promoters rescued proximal promoter activity and reduced H3K36 methylation within this promoter. In parallel, the use of inducible constructs allowed the authors to demonstrate that distal promoter activity was inversely correlated with proximal promoter activity. Finally, the authors provide data suggesting that the distal promoter is positively regulated by IRF1, while repressed by E4BP4.
While many cases of promoter interference are to be found in the literature, the case of the MICA gene is particularly elegantly documented by the present study and it may be the first case of "true" promoter interference in human cells (although the study by Martianov et al, Nature, 2007 may claim seniority). Yet, I would recommend the following issues to be addressed:
Figure 1: the authors may conclude two quickly that the transcript initiated at the distal promoter is non-coding. Panels C and D only show correlations and panel J claims that the MICA-UT expression vector does not make protein in transfections, without demonstrating that the construct is functional. Testing the MICA-UT-Myc construct in reticulocyte lysates would be a minimum. A better experiment would be to examine whether the MICA-UT mRNA is to be found in RNA isolated from polysome fractions.
Figure 2: The depletion of MICA-UT with siRNAs (panel S2B) is of very poor efficiency and does not seem sufficient to claim that the MICA-UT does not have an effect in trans. May be the authors should draw this conclusion only after panel 2F. In addition, the experimental procedure in Panel S2A is unclear: are the peaks showing all the cells or just the transfected cells? Transfection efficiency needs to be estimated.
Figures 4 and 5: The approaches are somewhat outdated. Demonstrating that the distal promoter is activated by IRF1 and repressed by E4BP4 will require ChIP assays.

Minor points/suggestions:
- Figure 1A and B: Fantom5 CAGE data may allow to get a broader perspective on the respective promoter activity in a wide range of tissues.
- The CRISPR deletion in Figure 2 could have been nicely replaced by a SunTag-type approach.

Referee #2:

In this study, two closely located promoters of the human MICA gene are identified and thoroughly characterized in various human cell lines. The major finding is that the upstream promoter, encoding a non-coding RNA, regulates the transcriptional output of the second promoter through a novel mechanism of transcriptional interference in cis.
Several experiments dismiss putative regulation in trans from the upstream transcript. Promoter competition is also excluded via deletion of each promoter.
Using deletion approaches (both at the endogenous locus by crispr and in synthetic transgenes by BAC recombineering), the authors convincingly demonstrate that, the upstream promoter sequence is responsible for transcription inhibition of the standard promoter. To my mind the most convincing data are the doxycycline inducible reporters that confirm transcriptional interference from the upstream promoter. Finally the authors test the effect of two regulatory pathways of the upstream promoter (interferon-gamma pathway and interleukin-4 signaling) on levels of interference.
The manuscript is well written, the data clearly presented.
Overall, I believe that the majority of the experiments are well performed and the results are of sufficient novelty and broad interest for publication in EMBO Journal

However I do have 3 major comments and a few minor suggestions that hopefully would improve the clarity of the manuscript.

MAJOR POINTS

1-An important point that is never assessed by the authors is to check for levels of nascent transcripts, either by qPCR with specific primers or by Pol II ChIP-qPCR.
In none of the two elegant approaches employed (Crispr or Bac transgenes) do the author examine the phenotype in terms of nascent RNA, only protein levels are assessed.
Particularly for Figure 2, whereby crispr deletions are employed to delete the upstream promoter,

 

only effects on MICA protein levels are analyzed.
While, reduction in MICA protein is likely due to a reduction in transcription initiation from the standard promoter, this must be proven with an appropriate method.

2-The usage of multiple cell lines is somehow confusing and the choice of a particular cell line is not justified. Since the ratio of upstream transcript versus standard transcript is assessed in a myriad of human cell lines, it would have been logical to chose three extreme cases: - a cell line with high amounts of the standard transcript (for example Hela) - a cell line with the opposite ratio and -a cell line with similar levels of both transcripts. I would like to see either qPCR or western Blots like in Figure 1J in at least two other cell lines.

3- The ChIP experiments shown in the article should include additional controls (positive and negative). That H3K36me3 increase at the second promoter following doxycycline is interesting but it would be nice to see signals from primers pairs located upstream and downstream of the STD promoter. For the allele-specific restriction digestion combined with chromatin immunoprecipitation (ChIP), there is no control to show that the digestion is complete.

MINOR POINTS:
1- Conclusions on the existence of interference mechanism are mainly motivated by loss of function approaches whereby the upstream promoter is deleted.
Although not essential, it would be interesting to examine the effect of promoter swapping, and that of replacing the upstream promoter by the standard promoter. Would this still lead to transcriptional interference?

2- the nomenclature of genotypes is not easy to read in Fig. 2E, maybe keep only MICA004 and remove the proceeding numbers?

3- how long is the region deleted in Fig. 2A?

4- it would be helpful to add the size of the inserted BAC in Fig. 3A.

5- The promoter mapping figure is not very clear (Fig. 2S). The authors should include a little schematic.

6- the mathematical modeling presented in Figure 4D is obscure. It seems to me that this is a simple plotting of data and fitting of a curve. The authors should explain in the main text what do they mean by modeling?

7- In Fig. 2C-D, the authors should comment that the crispr deletion affects mostly the cell line in C and that the effect is less drastic in D.
The variability of transcriptional interference according to the cell line used (and this the ratio upstream transcript versus standard transcript) would be an interesting point to discuss.

8- The effect of termination site insertion (Fig. 3C) has somehow more effect on transcriptional interference than that of the upstream promoter deletion. Can the authors speculate why?

9- The authors should develop the argumentation that support the claim that interference depends on promoter strength rather than the nature of the promoter.

10- Abstract : What does 'nature of a promoter means', isn't the strength of a promoter highly dependent on promoter sequences and chromatin environment, and as such on its 'nature' ? The authors should change the phrasing both in the abstract and in the results Figure4 regarding the 'nature' of the promoter.

Referee #3:

The manuscript by Lin et al. describes studies on a single locus (MICA) and provide evidence for transcriptional interference as a gene regulatory mechanism. A limited mechanistic analysis

 

suggested that this mechanism mediates MICA transcription in response to certain cytokines. The authors propose that transcriptional interference between tandem intragenic promoters "may constitute a general mechanism with widespread importance in human transcriptional regulation". Overall, the results are clear and drilling into mechanisms underlying transcriptional interactions between neighboring genes and/or cis-elements at endogenous loci has potential to yield novel mechanistic and/or biological insights. Several major limitations of the current study exist, however, and the authors may want to develop one or more aspects of the work to extend the mechanistic findings and/or the relevance of the single-locus mechanism to the genome. I believe that such extensions would have potential to yield a study of higher importance/impact.

Major Comments:

1. While the authors suggest their findings with MICA represent a generally applicable mechanism, it is unclear whether this is indeed the case. How many IFNgamma or IL-4-regulated genes are controlled by this mechanism? How many dual promoter-linked genes in the human and/or mouse genome are regulated by this mechanism? It would be important to get a much better sense of what aspects of the study highlight general principles versus a locus-specific mechanism of which many locus-specific transcriptional mechanisms have been described.

2. Transcriptional interference has been studied for quite a few years in systems ranging from yeast to humans using diverse approaches including plasmids, transgenes and endogenous loci. While the current study attempts to derive mechanistic insights, the content described in the manuscript is not particularly surprising in this regard. The authors conclude that changes in a few histone marks are delayed relative to the interference, but this correlative analysis only measured a few marks at limited sites. The significance of the single-locus work would be increased considerably if a deeper mechanistic analysis was conducted, which would have potential to yield novel insights. This is critical, since MICA is a gene that has already been the subject of first-generation mechanistic analyses.

3. Fig. 1B - It would be informative to include a panel of primary human cells, rather than solely the cancer cell lines.

4. Fig. 5C - The gel shift analysis could be bolstered by endogenous locus ChIP studies. For example, it would be instructive to evaluate the assemble, disassembly and status of nucleoprotein complexes at key sites of the locus and neighboring loci at various stages of MICA gene regulation. For example, this could be conducted with wild type cells and cell lacking one or the other promoters.

5. If additional studies can be conducted to test the broader applicability of the current results and/or to incorporate new mechanistic insights, the results would best be summarized with a model to more effectively allow the reader to compare the results with the state of the art.

---

 

**Summary of new experimental and genomic data**

- Polysome profiling of the *MICA* upstream and standard transcripts.
- New siRNA experiments with greater knockdown efficiency.
- New upstream transcript overexpression experiments.
- Expression studies with Suntag activation of upstream and downstream promoters.
- ChIP for Pol II Ser5 phosphorylation to study transcription initiation .
- ChIP for Spt16, H3, H4K20me3, H2A.Z with negative and positive controls
- ChIP for E4BP4 in primary B cells treated with IL-4.
- RT-qPCR of *MICA* upstream and standard transcripts in primary cells.
- Genome-wide analysis of CAGE-seq datasets from the FANTOM5 project.

**Referee 1**

Figure 1: the authors may conclude too quickly that the transcript initiated at the distal promoter is non-coding. Panels C and D only show correlations and panel J claims that the MICA-UT expression vector does not make protein in transfections, without demonstrating that the construct is functional. Testing the MICA-UT-Myc construct in reticulocyte lysates would be a minimum. A better experiment would be to examine whether the MICA-UT mRNA is to be found in RNA isolated from polysome fractions.

> We have now undertaken polysome fractionation and this is illustrated in Figure 1J and included in the first section of the Results. The result is clear and the standard downstream coding *MICA* transcript is enriched in the polysome-rich fraction similar to a *GAPDH* control and an *ITGB5* control (undertaken because ITGB5 encodes a membrane protein like MICA which is translated by ER-associated ribosomes). Conversely, the upstream transcript MICA-UT is located predominantly in the ribosome-free and monosome fractions. This confirms that it is poorly translated and is consistent with the prediction that it is non-coding and is subjected to nonsense-mediated decay.

Figure 2: The depletion of MICA-UT with siRNAs (panel S2B) is of very poor efficiency and does not seem sufficient to claim that the MICA-UT does not have an effect in trans. May be the authors should draw this conclusion only after panel 2F.

> We have changed the wording to reflect this.
>
> In addition to this we have further optimized the siRNA experiment by testing different siRNA transfection methods. We now achieve a ~70% knockdown of MICA-UT compared to the ~40% knockdown achieved previously using a different transfection reagent. We have replaced the original Figure S2B with the new result (Figure EV2B), which along with MICA-UT overexpression data, provides stronger evidence of the lack of *in trans* regulation.

In addition, the experimental procedure in Panel S2A is unclear: are the peaks showing all the cells or just the transfected cells? Transfection efficiency needs to be estimated.

> This figure shows all the cells. The transfection efficiency is high in these cells, but was not been quantified.

We have now repeated the experiment including co-transfection of a plasmid encoding a fluorescent marker, which allowed us to gate on transfected cells and so analyse MICA expression in this population of cells. These data are represented in Figure EV2A.

Figures 4 and 5: The approaches are somewhat outdated. Demonstrating that the distal promoter is activated by IRF1 and repressed by E4BP4 will require ChIP assays.

We already show ChIP for IRF1 (Figure 5B) in primary human arterial endothelial cells. We have now undertaken ChIP for E4BBP in primary human cells, which demonstrates IL-4-inducible binding of E4BP4 to the upstream promoter (Figure 6B).

Minor points/suggestions:
- Figure 1A and B: Fantom5 CAGE data may allow to get a broader perspective on the respective promoter activity in a wide range of tissues.

We have undertaken the analysis and discuss it below in our response to Referee 3 point 1.

- The CRISPR deletion in Figure 2 could have been nicely replaced by a SunTag-type approach.

The SunTag experiment is an excellent suggestion that complements the CRISPR deletion approach. We have now modified a recently reported all-in-one SunTag activation system (PMID: 28743878) and analysed MICA expression following transient transfection with a vector encoding GFP and SunTag-based activation. As predicted, a SunTag transcription activator targeting the upstream promoter downregulates cell surface expression of MICA from the downstream promoter. As a positive control, a SunTag transcription activator targeting the standard downstream promoter upregulated cell surface expression of MICA from the downstream promoter. This is now illustrated in Figure 2D and included in the second section of the Results.

**Referee 2**

1-An important point that is never assessed by the authors is to check for levels of nascent transcripts, either by qPCR with specific primers or by Pol II ChIP-qPCR.
In none of the two elegant approaches employed (Crispr or Bac transgenes) do the author examine the phenotype in terms of nascent RNA, only protein levels are assessed.
Particularly for Figure 2, whereby crispr deletions are employed to delete the upstream promoter, only effects on MICA protein levels are analyzed. While, reduction in MICA protein is likely due to a reduction in transcription initiation from the standard promoter, this must be proven with an appropriate method.

We agree with the reviewer that the likely explanation for the reduction in MICA protein is a reduction in transcription initiation and this would account for the changes in transcript levels we documented. These levels correlate well with protein expression.

We would like to have been able to assess nascent transcript levels directly, but in the current context, the assessment of nascent transcripts is confounded because the nascent upstream transcript and nascent downstream transcripts overlap. Any PCR or sequencing-based detection method for the unspliced downstream transcript will also detect the unspliced nascent upstream transcript. A subtraction approach would be unsatisfactory as the unspliced upstream transcript level will be high compared to the standard transcript level due to its very

short half-life. The same problem will arise with standard Pol II ChIP, which will detect polymerase activity relating to both transcripts.

To address this we undertook ChIP for Ser5-phosphorylated RNA Pol II which is associated with transcription initiation. Interposition of a transcription terminator between the two promoters caused a clear increase in the ChIP signal at the downstream promoter indicating that run-through transcription from the upstream promoter exerts an inhibitory effect on transcription initiation at the downstream promoter. This is now incorporated into the third section of the Results and Figure EV3D.

2-The usage of multiple cell lines is somehow confusing and the choice of a particular cell line is not justified. Since the ratio of upstream transcript versus standard transcript is assessed in a myriad of human cell lines, it would have been logical to chose three extreme cases: - a cell line with high amounts of the standard transcript (for example Hela) - a cell line with the opposite ratio and -a cell line with similar levels of both transcripts. I would like to see either qPCR or western Blots like in Figure 1J in at least two other cell lines.

We already show qPCR results for both the upstream transcript and the standard transcript in different cell types to demonstrate the wide range of expression levels of both transcripts and the clear correlation between the standard transcript and MICA surface expression. We have also now added data on a range of primary cells (Figure EV1B).

We tested five different MICA antibodies for western blotting (Santa cruz sc-23870, sc-5459; R&D Systems MAB1300, AF1300-SP; Abcam ab62540). Of these antibodies, only one produced a band consistent with the size and expression level of endogenous MICA (R&D Systems AF1300-SP). It also produced multiple background bands ranging from 20 to >100kDa, likely due to cross-reactivity because these bands are present at similar intensity in all the samples tested regardless of MICA standard or upstream transcript expression level. This precludes drawing any conclusion from western-blotting about the coding potential of MICA-UT directly.

However, we now have undertaken polysomal fractionation, which demonstrates that the upstream transcript is found in ribosome-free and monosomal fractions, consistent with it not being translated and undergoing likely nonsense-mediated decay (Figure 1J). In addition, experiments with tagged upstream transcript constructs further indicate that the upstream transcript is not translated (Figure EV1F).

3- The ChIP experiments shown in the article should include additional controls (positive and negative). That H3K36me3 increase at the second promoter following doxycycline is interesting but it would be nice to see signals from primers pairs located upstream and downstream of the STD promoter. For the allele-specific restriction digestion combined with chromatin immunoprecipitation (ChIP), there is no control to show that the digestion is complete.

We had done matched IgG negative controls for all the ChIP experiments previously presented and have now added these to the figures—the background levels are very low for ChIP done with Protein G dynabeads (Figures 4G-I, EV3E-K and EV4G-J).

For positive controls we have ChIP data for the endogenous MICA promoter for the doxycycline-inducible ChIP experiments. This was obtained using digestion with a restriction enzyme that specifically recognises a SNP in the transgenic allele in the PCR amplicon used for the ChIP. The ChIP signal from the endogenous allele does not change significantly with the experimental conditions unlike the signal changes seen with experimental manipulation of the

transgenic upstream promoter, demonstrating robustness of the ChIP experiments as well as the *in cis* nature of the regulation. These data have now been provided in Figures 4G-I, EV3E-K and EV4G-J.

As regards upstream and downstream sites, any ChIP site must contain a suitable SNP that allows digestion to distinguish the endogenous from the transgenic allele. It did not prove possible to identify suitable SNPs in appropriate locations. The closest upstream SNP is 2.7kb upstream of the standard promoter but only 150bp downstream from the upstream promoter and so cannot be used to provide reliable information about the downstream standard promoter. The only potentially usable downstream SNP is in the MICA exon 1, but the available enzyme was only able to digest the endogenous allele with relatively low efficiency and as our isogenic cell lines contain two copies of the endogenous allele and only one copy of the transgenic allele this was not a suitable basis for an informative ChIP study. Further, unlike the SNP which we did study, no suitable enzyme is available which can be used to digest the transgenic allele to examine endogenous allele status as a control at this SNP.

For the allele-specific ChIP that was used, the digestion efficiencies with BfaI (which digests the endogenous allele) and TspRI (which digests the transgenic allele) were 99.4% and 99.9% after careful optimisation of the reactions. Efficiencies were determined using dilutions of ChIP input control from cells homozygous for the endogenous or transgenic alleles. This information is now included in the Methods section.

MINOR POINTS:

1- Conclusions on the existence of interference mechanism are mainly motivated by loss of function approaches whereby the upstream promoter is deleted.

Although not essential, it would be interesting to examine the effect of promoter swapping, and that of replacing the upstream promoter by the standard promoter. Would this still lead to transcriptional interference?

Construction of the necessary multiple BAC clones with relevant controls would be a substantial undertaking and whilst this would be an interesting experiment we do not feel that it would add substantial value as we have already demonstrated that transcription interference occurs even with a different promoter (the doxycycline-inducible promoter) and so is not promoter-specific.

2- the nomenclature of genotypes is not easy to read in Fig. 2E, maybe keep only MICA004 and remove the proceeding numbers?

This has been done.

3- how long is the region deleted in Fig. 2A?

The sequences that was deleted was illustrated in Figure 2B. The sizes of deletions are 44bp for deletion #1 and 95bp for deletion #2.

4- it would be helpful to add the size of the inserted BAC in Fig. 3A.

This has been done.

5- The promoter mapping figure is not very clear (Fig. 2S). The authors should include a little schematic.

This has now been done as Figure EV2C and EV2D.

6- the mathematical modeling presented in Figure 4D is obscure. It seems to me that this is a simple plotting of data and fitting of a curve. The authors should explain in the main text what do they mean by modeling?

> This is explained in the Results section, which now reads “Mathematical analysis of the steady state response curve shows that the transcriptional activity arising from the standard promoter is in a simple reciprocal relationship with the transcriptional activity arising from the upstream promoter.”

7- In Fig. 2C-D, the authors should comment that the crispr deletion affects mostly the cell line in C and that the effect is less drastic in D.

The variability of transcriptional interference according to the cell line used (and this the ratio upstream transcript versus standard transcript) would be an interesting point to discuss.

> We have commented on this in the results section and added discussion of the variability in the first part of the discussion section.

8- The effect of termination site insertion (Fig. 3C) has somehow more effect on transcriptional interference than that of the upstream promoter deletion. Can the authors speculate why?

> This is now discussed in the main text. In brief, the likely explanation is that the transcription terminator prevents run through transcription with high efficiency (Figure 3C). In contrast, deletion of the core upstream promoter does not completely abolish transcription (Figure 3C), so some low level transcriptional activity persists and may contribute to transcriptional interference.

9- The authors should develop the argumentation that support the claim that interference depends on promoter strength rather than the nature of the promoter.

> The discussion now reads “Quantitative analysis using an experimental system in which the upstream promoter activity is tuneable under the control of doxycycline showed that the level of transcriptional interference observed was similar with either the native upstream promoter or with a heterologous promoter of equivalent strength. Therefore, transcription from an upstream promoter is sufficient for transcriptional interference; the degree of transcriptional interference depends on the strength of transcription from the upstream promoter, rather than the identity of the upstream promoter. ”

10- Abstract: What does 'nature of a promoter means', isn't the strength of a promoter highly dependent on promoter sequences and chromatin environment, and as such on its 'nature' ? The authors should change the phrasing both in the abstract and in the results Figure4 regarding the 'nature' of the promoter.

> This has been done. We agree that the word ‘nature’ is too vague in this context. The key point is that the strength of transcription from the upstream promoter determines the degree of transcriptional interference.

**Referee 3**

1. While the authors suggest their findings with MICA represent a generally applicable mechanism, it is unclear whether this is indeed the case. How many IFNgamma or IL-4-regulated genes are controlled by this mechanism? How many dual promoter-linked genes in the human and/or mouse genome are regulated by this mechanism? It would be important to get a much better sense of what aspects of the study highlight general principles versus a locus-specific mechanism of which many locus-specific transcriptional mechanisms have been described.

> Although we demonstrated that IFNgamma and IL-4 regulate MICA through transcription interference, transcription interference will be a gene-specific rather than stimulus-specific mechanism.
>
> We agree that it is valuable to survey the extent to which the findings in the MICA gene may be relevant to other genes. We have now undertaken a detailed analysis of whole genome data for this purpose. A particular problem with such datasets generated using RNA-seq is that it can be difficult to distinguish which promoter a transcript arises from and it is clear that dual promoter systems are common throughout the human genome. However, CAGE-seq data provides valuable clarity about which promoter a transcript originates from and provides an unbiased index of the level of transcription from each promoter. Time course data following a stimulus provides an opportunity to sample the trajectories of expression arising from each promoter in dual promoter sets and so identify patterns of transcription consistent with transcriptional interference.
>
> We analysed multiple human and mouse datasets with different stimuli for which reasonable time resolution was available. For each dataset we identified the subset of genes that contained a dual promoter system and for which there was expression from these promoters in that dataset. We then analysed the trajectories of each transcript over time to identify patterns consistent with transcriptional interference – that is, where there was a rise in the level of transcript arising from an upstream promoter associated with a fall in the level of transcript arising from the downstream transcript, or where there was a fall in the level of transcript arising from an upstream promoter associated with a rise in the level of transcript arising from the downstream transcript. Across multiple datasets we found a substantial number of cases where the pattern was similar to that seen with MICA and so was consistent with transcriptional interference. This was a common feature of both human and mouse datasets. These results are now included in the manuscript. The overall analysis is shown in Table EV1 and illustrative examples are provided in Figure 7 with coordinates of the illustrated promoter pairs in Table S5. Definitive proof that transcriptional interference operates on any individual gene will require substantial detailed experimental work as we present for *MICA*. However, we believe that this analysis demonstrates that transcriptional interference is likely to be involved in the regulation of multiple genes in higher eukaryotes.

2. Transcriptional interference has been studied for quite a few years in systems ranging from yeast to humans using diverse approaches including plasmids, transgenes and endogenous loci. While the current study attempts to derive mechanistic insights, the content described in the manuscript is not particularly surprising in this regard. The authors conclude that changes in a few histone marks are delayed relative to the interference, but this correlative analysis only measured a few marks at limited sites. The significance of the single-locus work would be increased considerably if a deeper mechanistic analysis was conducted, which would have potential to yield novel insights. This is

critical, since MICA is a gene that has already been the subject of first-generation mechanistic analyses.

> We have now undertaken additional experimental work to derive further mechanistic insight. We assessed additional histone marks including pan-H3, H4K20me3 and histone H2A.Z. We also studied the FACT histone chaperone subunit Spt16. Various other analyses were attempted, but proved uninformative for technical reasons.
>
> Of the markers tested, Spt16, H3K36me3 and H4K20me3 occupancy at the downstream standard MICA promoter correlates with upstream promoter activity. However, unlike H3K36me3 and H4K20me3 which increase gradually in the doxycycline-inducible model, the increase of Spt16 level occurs rapidly and follows the same time-course as the increase in upstream promoter activity and the effect of transcriptional interference. This is consistent with FACT-facilitated nucleosomal remodelling playing a central role in transcriptional interference. These results are now included in the manuscript and illustrated in Figures EV3K and 4I, and graphically in the model (Figure 8).
>
> With reference to first-generation work, there has been only limited work done on transcriptional regulation of MICA previously. Work on *MICA* by other groups has generally used RT-PCR primers that would have measured both the upstream transcript and the standard downstream transcript and in the light of our current findings this complicates the interpretation of some previous studies.

3. Fig. 1B - It would be informative to include a panel of primary human cells, rather than solely the cancer cell lines.

> We have now included analysis of the two transcripts in a panel of primary cells (Figure EV1B).

4. Fig. 5C - The gel shift analysis could be bolstered by endogenous locus ChIP studies. For example, it would be instructive to evaluate the assembly, disassembly and status of nucleoprotein complexes at key sites of the locus and neighboring loci at various stages of MICA gene regulation. For example, this could be conducted with wild type cells and cell lacking one or the other promoters.

> We have added ChIP analysis for primary B cells treated with IL-4 (Figure 6B). ChIP for IRF1 was already shown in Figure 5B.
>
> In addition, as outlined above (point 2), we have now undertaken studies of additional nucleosomal marks and factors involved in nucleosomal remodelling during transcriptional interference-mediated *MICA* gene regulation.

5. If additional studies can be conducted to test the broader applicability of the current results and/or to incorporate new mechanistic insights, the results would best be summarized with a model to more effectively allow the reader to compare the results with the state of the art.

> We have now generated a model to summarise the mechanism of transcriptional interference in MICA (Figure 8).

2nd Editorial Decision 26th January 2018

Thank you for submitting a revised version of your manuscript and my apologies for the extended duration of the re-review period. Your study has now been seen by all three original referees and their comments are shown below.

As you will see, refs #1 and #2 both find that all criticisms have been sufficiently addressed and recommend the manuscript for publication, while ref #3 remains critical about the overall advance provided. However, given the strong support from referees #1 and #2 - and the additional analysis that has been included in the revised version of the manuscript - we have decided to overrule the concerns from ref #3. I would therefore invite you to submit a final version of the study in which you elaborate on the figure legends and data descriptions as suggested by ref #2. In addition, please address the following editorial points:

-> include the running title in the manuscript doc file
-> move the EV table legends from the manuscript file to a separate tab in the corresponding excel files
-> Appendix tables should be included in the Appendix PDF. Alternatively, you can turn them into EV tables and update the callouts in the manuscript
-> The number of replicas used for calculating statistics and the nature of the error bars need to be indicated in the figure legend wherever applicable. We noticed that this is currently missing for Fig 1B, G, H and for Fig 4B-I

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

---

REFEREE REPORTS

Referee #1:

The authors have thoroughly addressed the points that I had raised. The new Figure 7 clearly increases the outreach of the manuscript.

Referee #2:

The revised manuscript has been significantly improved. I particularly liked the new mechanistic data concerning FACT. Although raised by another referee, I personally don't feel that the analysis of the CAGE data was essential, nor that it should be part of the main figures. Indeed this analysis is descriptive and presenting after a detailed mechanistic study on MICA is somehow strange.
However, I do feel that most of the comments raised by the referees have been addressed. I am thus in favor for its publication in EMBO journal.
I have a few minor suggestions that could potentially improve the clarity of the manuscript.

1-I think that the new data presented in FigureEV2 A-B should be included in the main Figure2 since it shows an important result, namely that the upstream promoter does not regulate MICA expression via a trans mechanism.

2-several figures should be clarified by adding extra captions/information in the legend.
Examples :
a- Figure EV3-D, we don't know which ChIP results are depicted, a simple Pol II Ser-5P title (as for panels E to K) would suffice

b-Figure 1 G and H, what is the difference? a short name of the cell line would indicate the difference between the panels

c-Figure2 D : since an elegant Suntag system is used, it would be useful to describe it in the figure, rather than vague 'upstream promoter activation'. Also in the main text, the activators tethered with

 

the suntag system are not described.

d- Figure 2E: I still believe that genotypes like MICA004 or MICA 010 are not self-explanatory.

Referee #3:

The authors have revised the manuscript based on prior recommendations. Some of these revisions have appropriately addressed prior issues. Others remain to be addressed.

Prior query 1: One of the prior important points involved extending the single-locus mechanistic analysis to multiple genes - and ideally genome-wide. The authors conducted an analysis of existing datasets and include a tabulation of potentially relevant expression data. However, the question of whether mechanistic principles developed from the single-locus study inform mechanisms controlling other genes in the genomes remains uncertain - and it does not appear that this has been investigated.

Prior query 2: The authors conducted a limited number of additional correlative studies, but did not address the query with mechanistic analyses.

The authors apparently added a model at Fig. 8, but their figures are not labeled. What I believe to be Fig. 8 could be viewed as an incremental advance in existing knowledge. Thus, if this model indeed summarizes the results from the single-locus studies, this reaffirms the vital need to extend the analysis to address whether the findings unveil a: (i) broadly applicable mechanism; or (ii) a single locus mechanism that informs important biology or pathology. The current manuscript does not appear to represent either of these important outcomes.

 

**Editorial comments**

-> include the running title in the manuscript doc file

This has been done.

-> move the EV table legends from the manuscript file to a separate tab in the corresponding excel files

The legend to EV Table 1 has been moved out of the manuscript file and is now in the excel file.

-> Appendix tables should be included in the Appendix PDF. Alternatively, you can turn them into EV tables and update the callouts in the manuscript

The tables have been included in an Appendix PDF.

-> The number of replicas used for calculating statistics and the nature of the error bars need to be indicated in the figure legend wherever applicable. We noticed that this is currently missing for Fig 1B, G, H and for Fig 4B-I

This has been done for all figures.

**Referee 2**

1-I think that the new data presented in FigureEV2 A-B should be included in the main Figure2 since it shows an important result, namely that the upstream promoter does not regulate MICA expression via a trans mechanism.

This has been done. In order to make space in this figure we have moved the panel with sequence information about the CRISPR deletion to Figure EV2C.

2-several figures should be clarified by adding extra captions/information in the legend.

We have reviewed the figures and made some minor alterations as suggested by the referee.

Examples :

a- Figure EV3-D, we don't know which ChIP results are depicted, a simple Pol II Ser-5P title (as for panels E to K) would suffice

This has been added.

b-Figure 1 G and H, what is the difference? a short name of the cell line would indicate the difference between the panels

This has been added.

c-Figure2 D : since an elegant Suntag system is used, it would be useful to describe it in the figure, rather than vague 'upstream promoter activation'. Also in the main text, the activators tethered with the suntag system are not described.

This has been done in the figure, figure legend and results section. We have also made a new diagram (Figure 2E) to describe the method. In the previous version we used the term Suntag activator to describe a dCas9 activator in general, but the precise system we used is a recently developed variant of this approach which is more accurately known as a SAM activator. The experimental details were already correctly described and referenced in the Methods section, but we have now edited the relevant sentences in the results section to clarify this and provide more details of the tethered activators as suggested by the author.

d- Figure 2E: I still believe that genotypes like MICA004 or MICA 010 are not self-explanatory.

We agree and have edited this figure accordingly. To improve clarity we have provided an explanatory key to Figure 2G and have removed details of the SNP from this Figure. The use of the SNP is now explained in the legend to figure 2I.

Accepted 9th Febuary 2018

Thank you for submitting the final revision of your manuscript, I am pleased to inform you that your study has now been officially accepted for publication in The EMBO Journal.

 

EMBO PRESS

YOU MUST COMPLETE ALL CELLS WITH A PINK BACKGROUND ↓

PLEASE NOTE THAT THIS CHECKLIST WILL BE PUBLISHED ALONGSIDE YOUR PAPER

| | |
|---|---|
| Corresponding Author Name: | Chris O'Callaghan |
| Journal Submitted to: | EMBO Journal |
| Manuscript Number: | EMBOJ-2017-97138 |

**Reporting Checklist For Life Sciences Articles (Rev. June 2017)**

This checklist is used to ensure good reporting standards and to improve the reproducibility of published results. These guidelines are consistent with the Principles and Guidelines for Reporting Preclinical Research issued by the NIH in 2014. Please follow the journal's authorship guidelines in preparing your manuscript.

## A- 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.
- figure panels include only data points, measurements or observations that can be compared to each other in a scientifically meaningful way.
- graphs 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 and any statistical test employed should be justified
- Source Data should be included to report the data underlying graphs. Please follow the guidelines set out in the author ship 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.

Any descriptions too long for the figure legend should be included in the methods section and/or with the source data.

**In the pink boxes below, please ensure that the answers to the following questions are reported in the manuscript itself. Every question should be answered. If the question is not relevant to your research, please write NA (non applicable). We encourage you to include a specific subsection in the methods section for statistics, reagents, animal models and human subjects.**

## B- Statistics and general methods

Please fill out these boxes ↓ (Do not worry if you cannot see all your text once you press return)

| | |
|---|---|
| 1.a. How was the sample size chosen to ensure adequate power to detect a pre-specified effect size? | NA |
| 1.b. For animal studies, include a statement about sample size estimate even if no statistical methods were used. | NA |
| 2. Describe inclusion/exclusion criteria if samples or animals were excluded from the analysis. Were the criteria pre-established? | NA |
| 3. Were any steps taken to minimize the effects of subjective bias when allocating animals/samples to treatment (e.g. randomization procedure)? If yes, please describe. | NA |
| For animal studies, include a statement about randomization even if no randomization was used. | NA |
| 4.a. Were any steps taken to minimize the effects of subjective bias during group allocation or/and when assessing results (e.g. blinding of the investigator)? If yes please describe. | NA |
| 4.b. For animal studies, include a statement about blinding even if no blinding was done | NA |
| 5. For every figure, are statistical tests justified as appropriate? | Yes |
| Do the data meet the assumptions of the tests (e.g., normal distribution)? Describe any methods used to assess it. | Yes |
| Is there an estimate of variation within each group of data? | Yes |
| Is the variance similar between the groups that are being statistically compared? | Yes |

## C- Reagents

| | |
|---|---|
| 6. To show that antibodies were profiled for use in the system under study (assay and species), provide a citation, catalog number and/or clone number, supplementary information or reference to an antibody validation profile. e.g., Antibodypedia (see link list at top right), 1DegreeBio (see link list at top right). | See methods section. |

## USEFUL LINKS FOR COMPLETING THIS FORM

http://www.antibodypedia.com
http://1degreebio.org
http://www.equator-network.org/reporting-guidelines/improving-bioscience-research-repo

http://grants.nih.gov/grants/olaw/olaw.htm
http://www.mrc.ac.uk/Ourresearch/Ethicsresearchguidance/Useofanimals/index.htm
http://ClinicalTrials.gov
http://www.consort-statement.org
http://www.consort-statement.org/checklists/view/32-consort/66-title

http://www.equator-network.org/reporting-guidelines/reporting-recommendations-for-tun

http://datadryad.org

http://figshare.com

http://www.ncbi.nlm.nih.gov/gap

http://www.ebi.ac.uk/ega

http://biomodels.net/

http://biomodels.net/miriam/
http://jjj.biochem.sun.ac.za
http://oba.od.nih.gov/biosecurity/biosecurity_documents.html
http://www.selectagents.gov/

| | |
|---|---|
| 7. Identify the source of cell lines and report if they were recently authenticated (e.g., by STR profiling) and tested for mycoplasma contamination. | ATCC with regular mycoplasma testing. |

* for all hyperlinks, please see the table at the top right of the document

## D- Animal Models

| | |
|---|---|
| 8. Report species, strain, gender, age of animals and genetic modification status where applicable. Please detail housing and husbandry conditions and the source of animals. | NA |
| 9. For experiments involving live vertebrates, include a statement of compliance with ethical regulations and identify the committee(s) approving the experiments. | NA |
| 10. We recommend consulting the ARRIVE guidelines (see link list at top right) (PLoS Biol. 8(6), e1000412, 2010) to ensure that other relevant aspects of animal studies are adequately reported. See author guidelines, under 'Reporting Guidelines'. See also: NIH (see link list at top right) and MRC (see link list at top right) recommendations. Please confirm compliance. | NA |

## E- Human Subjects

| | |
|---|---|
| 11. Identify the committee(s) approving the study protocol. | South Central - Hampshire B Research Ethics Committee |
| 12. 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. | Informed consent was obtained for all subjects who donated blood and experiments conformed to the principles set out in the WMA Declaration of Helsinki and the Department of Health and Human Services Belmont Report. |
| 13. For publication of patient photos, include a statement confirming that consent to publish was obtained. | NA |
| 14. Report any restrictions on the availability (and/or on the use) of human data or samples. | NA |
| 15. Report the clinical trial registration number (at ClinicalTrials.gov or equivalent), where applicable. | NA |
| 16. 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. | NA |
| 17. 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. | NA |

## F- Data Accessibility

| | |
|---|---|
| 18: Provide a "Data Availability" section at the end of the Materials & Methods, listing the accession codes for data generated in this study and deposited in a public database (e.g. RNA-Seq data: Gene Expression Omnibus GSE39462, Proteomics data: PRIDE PXD000208 etc.) Please refer to our author guidelines for 'Data Deposition'.<br><br>Data deposition in a public repository is mandatory for:<br>a. Protein, DNA and RNA sequences<br>b. Macromolecular structures<br>c. Crystallographic data for small molecules<br>d. Functional genomics data<br>e. Proteomics and molecular interactions | This has been done. Sequences for MICA promoter haplotypes have been submitted to Genbank with accession numbers KF724603, KF724624 and KF724587 as discussed in the Materials and Methods section. |
| 19. Deposition is strongly recommended for any datasets that are central and integral to the study; please consider the journal's data policy. If no structured public repository exists for a given data type, we encourage the provision of datasets in the manuscript as a Supplementary Document (see author guidelines under 'Expanded View' or in unstructured repositories such as Dryad (see link list at top right) or Figshare (see link list at top right). | NA |
| 20. Access to human clinical and genomic datasets should be provided with as few restrictions as possible while respecting ethical obligations to the patients and relevant medical and legal issues. If practically possible and compatible with the individual consent agreement used in the study, such data should be deposited in one of the major public access-controlled repositories such as dbGAP (see link list at top right) or EGA (see link list at top right). | NA |
| 21. Computational models that are central and integral to a study should be shared without restrictions and provided in a machine-readable form. The relevant accession numbers or links should be provided. When possible, standardized format (SBML, CellML) should be used instead of scripts (e.g. MATLAB). Authors are strongly encouraged to follow the MIRIAM guidelines (see link list at top right) and deposit their model in a public database such as Biomodels (see link list at top right) or JWS Online (see link list at top right). If computer source code is provided with the paper, it should be deposited in a public repository or included in supplementary information. | NA |

## G- Dual use research of concern

| | |
|---|---|
| 22. Could your study fall under dual use research restrictions? Please check biosecurity documents (see link list at top right) and list of select agents and toxins (APHIS/CDC) (see link list at top right). According to our biosecurity guidelines, provide a statement only if it could. | Not to the best of our knowledge. |