The Hda1 histone deacetylase limits divergent non-coding transcription and restricts transcription initiation frequency

Uthra Gowthaman, Maxim Ivanov, Isabel Schwarz, Heta Patel, Niels Müller, Desiré García-Pichardo, Tineke Lenstra, and Sebastian Marquardt

DOI: 10.15252/embj.2021108903

Corresponding author(s): Sebastian Marquardt (sebastian.marquardt@plen.ku.dk)

Review Timeline:

- Transfer from Review Commons: 7th Jun 21
- Editorial Decision: 7th Jul 21
- Revision Received: 31st Aug 21
- Accepted: 28th Sep 21

Editor: Stefanie Boehm

Transaction Report: This manuscript was transferred to The EMBO Journal following peer review at Review Commons.

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

Thank you again for transferring your revised Review Commons manuscript to The EMBO Journal, as well as sending the preliminary point-by-point response to the referee comments, which are again included below. In light of the referees’ comments and your response, we would like to invite you to prepare and submit a final revised manuscript.

As discussed, it will be important to thoroughly respond to each of the remaining concerns, in particular to incorporate the additional statistical analyses (ref#1- point 1, 2, 6) and to revise the manuscript to discuss referee #2’s concerns. We recognize that the analyses suggested by referee #3 and the potentially required follow-up validation would likely extend beyond the scope of this final round of revision. Please nonetheless also carefully respond to these points and include data where it is already available. In addition, please remember to explicitly state cell numbers and replicates in the respective figure legends.

Acceptance depends on the completeness of your responses included in the next, final version of the manuscript. We realize that lab work worldwide may currently still be affected by the COVID-19/SARS-CoV-2 pandemic and that an experimental revision can be delayed. We can extend the revision time when needed, and we have extended our ‘scooping protection policy’ to cover the period required for a full revision. However, it is important that you contact us as soon as possible in case there are any unexpected major delays.

Please do not hesitate to contact me if you have any questions regarding this final revision. Thank you again for the opportunity to consider your work for publication.

Kind regards,

Stefanie

Stefanie Boehm
Editor
The EMBO Journal
The study by Gowthaman et al. aims at better understanding the molecular mechanisms controlling divergent non-coding transcription (DNC). They describe a high-throughput yeast genetic screen using two strains in which two loci consisting of a coding and a divergent non-coding transcription unit were replaced by a bidirectional fluorescent reporter construct encoding mCherry in the coding and YFP in the non-coding direction. The two reporter strains were crossed with the yeast deletion library and mutants leading to increased or decreased YFP signal were selected as DNC repressors or activators. The two screens identified several common potential repressors and activators. Components of the Hda1C histone deacetylase complex were identified as DNC repressors in both screens. This phenomenon was confirmed genome-wide by performing NET-Seq in WT as well as hda1D and hda3D strains. This experiment allowed to identify 1517 DNC transcripts repressed by Hda1. Further analyses suggest that Hda1C represses DNC genome-wide independently of expression levels and that loss of Hda1 does not substantially affect coding transcription.

Live-cell imaging of transcription was then used to show that loss of Hda1 increases DNC transcription frequency rather than duration providing novel information on the link between DNC transcription initiation kinetics and chromatin regulation. Finally, using Chip-seq, the authors show that the level of acetylation over the divergent non-coding units is increased in the absence of Hda1 and some experiments suggest that H3K56 acetylation also contributes to DNC regulation, further strengthening the importance of elevated histone acetylation in efficient DNC.

Several components of the SWI/SNF chromatin remodeling complex were identified as activators confirming earlier observations (Marquardt et al., 2014). SAGA subunits were also among potential DNC activators, however these effects could not be confirmed through validation experiments. The authors conclude that DNC may be independent of specific activators and mainly due to transcriptional noise resulting from the adjacent NDR.

Overall this paper is well structured and clearly written. The genetic screen identifies novel factors involved in the regulation of DNC. The study demonstrates that the level of acetylation is a key regulator of divergent non-coding transcription and that histone deacetylation by Hda1 reduces the frequency of DNC initiation events. While this conclusion is strongly supported by NET-Seq and Chip-seq metagene analyses, the fluorescence mCherry and YFP values or qRP-PCR analyses of specific genes do not always behave as expected when looking at absolute values rather than mCherry/YFP or GCG1/SUT098 ratios, which is sometimes disturbing when reading the paper.

The paper was submitted to Review Commons and a revised version has been submitted to EMBO J. The revised version of the paper by Gowthaman now clarifies a certain number of questions by additional analyses and figures as well as text modifications. Overall the data and conclusions are more convincing. However there are still few points that should be addressed.

Major comments:

Figure S3A-C: the authors conclude that low, medium and highly expressed DNCs are similarly affected by loss of Hda1 or Hda3. However, they just compared NETseq signal with no statistics. It would be worth defining the fold change between WT and hda1D or hda3D including some statistical analyses; this fold change could be substantially different considering low or highly expressed DNCs, or the other non-coding transcripts examined in new Fig. S3D.

More generally this reviewer is wondering whether the sensitivity to Hda1 is really specific to DNCs, or whether Hda1 would more generally contribute to the repression of lowly expressed transcripts, whether coding or non-coding. For example, GCG1 and ORC2 are significantly increased in hda3D and hda1D respectively (Fig. S3F-I).

It would therefore be interesting to perform these fold-change analyses not only for non-coding transcripts, but also for coding RNAs. The authors could define whether lowly expressed coding genes are more prone to repression by Hda1C by stratifying coding genes according to NET-Seq expression levels in the presence or absence of Hda1 or Hda3. These additional analyses would avoid establishing an inaccurate bias towards DNCs.

Figure S3D: it is not clear why the NET-Seq signal is rather high at the TSS of these non-coding transcripts, while it is really low at the TSS of coding RNAs and DNCs in Figure S3ABC. Some explanation should be provided.

Figure S5: why is mcherry and not YFP mostly increasing in the H3WTDhda1, H3K56Q or the double mutants? This again indicates a stronger repressive effect of Hda1 on coding rather than non-coding transcription.
Minor comments:

Abstract: "mutations in the Hda3 subunit increased (not reduced) the frequency of DNC transcription
In their response to the reviewers, the authors say that the Table S3 provides the fold-change values for all DNC transcripts in
mutant compared to wild type, but the fold-change is not indicated.

Referee #2:

Response to the revision of Gowthama et al.,

Promotors are frequently transcribed in both directions. The divergent, 'upstream' transcript is frequently unstable. Transcription
initiation is regulated through the acetylation of promoter-proximal nucleosomes, where HDAC-dependent deacetylation of
histones typically represses transcription initiation.
The current manuscript addresses the question whether initiation of coding and divergent, non-coding (DNC) transcription is
regulated by the same factors. Previously Marquardt and others showed that H3K56ac-mediated histone exchange has a
differential effect on coding and DNC transcription.
Using a clever reporter system, the authors screened for positive and negative regulators that preferentially affect DNC
transcription. They discover the Hda1 deacetylase complex as a DNC-biased repressor and diverse HATs as DNC-biased
activators. The role of activators could not be validated, presumably due to high variability of the system.
Focusing on Hda1c the authors present data suggesting a larger effect of Hda1c on 'upstream' nucleosomes associated with
DNC transcription than in coding transcription. Genome-wide NET-seq mapping was consistent with this differential regulation.
Life cell imaging of one specific case argues that Hda1-mediated repression reduced the time between initiation events.
The authors employ state of the art methods and in general the data are of very good quality. The effect size is small, which
raises the broader question whether the results, while statistically significant is biological relevant.

The authors addressed some of the points I raised in my previous evaluation to my satisfaction, but I want to maintain that some
of the issues are not due to misunderstandings, and thus are maintained.

- If an experiment involves analysing 50,000 cells and a median signal is derived, then N=1. This is basis statistics. It becomes a
good experiment when N=3 and you can compare three medians and see whether the data are reproducible. I can see the point
that for an exploration one conducts one large experiment, to derive candidates, which are then validated independently. The
fact that for the SAGA subunits the validation was apparently not possible, illustrates this point. For any experiment, where
quantitative data are produced as a result, the number of biological experiments (different days) must be stated in the legends.
- The point about live cell images is whether they were all done on one special day, or collected on multiple independent
occasions.
- I also maintain that the single molecule analysis has limited impact. Unfortunately, the coding GCG1 transcription was low, so
could not be well compared to DNC SUT089. Of course SUT089 can be measured as such (I do understand that), but it was not
designed as a stand-alone experiment in the planning, you declared it to be one, when half of it did not work. The weakness is,
as I stated, that you have only one experiment, for which half did not work.
In summary, this is a high-quality study that presents the results of a genome-wide screen that will be of interest to colleagues in
the narrower field. Due to the small effects the results may appeal less to a general readership.

Referee #3:

In this manuscript, Gowthaman et al describe the results and follow up of their screen aimed at identifying regulators of divergent
noncoding (DNC) transcription in S. cerevisiae. From this screen, they identify Hda1C as a repressor of DNC transcription, and
perform follow experiments to support and detail this finding. In addition to RTqPCR to confirm the reporter and endogenous
changes, the authors perform NET-seq to look at global DNC alteration upon Hda1C subunit deletion and identify a number of
non-coding transcripts with altered expression levels. In addition, the authors perform live cell imaging to demonstrate that there
is a modest restriction of initiation frequency when one of the subunits of Hda1C is deleted. Finally, the authors explore changes
to pan-H3 acetylation and the genetic overlap between Hda1C and H3K56ac demonstrating independent genetic pathways, but
overall increases in H3 acetylation over DNCs when Hda1C is deleted.

The authors addressed all my major concerns in the revision of this manuscript.
I am sorry for bringing up something new on the revision, but re-reading the manuscript made me think about whether there is
anything in common with the DNC transcripts identified in the NETseq data as being regulated by hda1del. Examining the FPKM
(Fig S3) and correlating with the H3ac (Fig SSD) was great, but I am curious about more information. For instance, is there any
GO term enrichment in the corresponding mRNA, any characteristics of the -1 nucleosome (H2AZ occupancy?), any sequence
characteristics/motifs of the promoter for the DNCs, any characteristic NDR size? These are just a few thoughts, but perhaps
there are other defining characteristics that could be quickly examined using public datasets. I realize some of these analyses
could take a few hours to complete, but I do think that if something were identified (or also if things were found to not be connected), it would be of great interest.

***
Rev_Com_number: RC-2021-00750
New_manu_number: EMBOJ-2021-108903
Corr_author: Marquardt
Title: Hda1C limits divergent non-coding transcription and restricts transcription initiation frequency
The study by Gowthaman et al. aims at better understanding the molecular mechanisms controlling divergent non-coding transcription (DNC). They describe a high-throughput yeast genetic screen using two strains in which two loci consisting of a coding and a divergent non-coding transcription unit were replaced by a bidirectional fluorescent reporter construct encoding mCherry in the coding and YFP in the non-coding direction. The two reporter strains were crossed with the yeast deletion library and mutants leading to increased or decreased YFP signal were selected as DNC repressors or activators. The two screens identified several common potential repressors and activators. Components of the Hda1C histone deacetylase complex were identified as DNC repressors in both screens. This phenomenon was confirmed genome-wide by performing NET-Seq in WT as well as hda1D and hda3D strains. This experiment allowed to identify 1517 DNC transcripts repressed by Hda1. Further analyses suggest that Hda1C represses DNC genome-wide independently of expression levels and that loss of Hda1 does not substantially affect coding transcription. Live-cell imaging of transcription was then used to show that loss of Hda1 increases DNC transcription frequency rather than duration providing novel information on the link between DNC transcription initiation kinetics and chromatin regulation. Finally, using Chip-seq, the authors show that the level of acetylation over the divergent non-coding units is increased in the absence of Hda1 and some experiments suggest that H3K56 acetylation also contributes to DNC regulation, further strengthening the importance of elevated histone acetylation in efficient DNC. Several components of the SWI/SNF chromatin remodeling complex were identified as activators confirming earlier observations (Marquardt et al., 2014). SAGA subunits were also among potential DNC activators, however these effects could not be confirmed through validation experiments. The authors conclude that DNC may be independent of specific activators and mainly due to transcriptional noise resulting from the adjacent NDR. Overall this paper is well structured and clearly written. The genetic screen identifies novel factors involved in the regulation of DNC. The study demonstrates that the level of acetylation is a key regulator of divergent non-coding transcription and that histone deacetylation by Hda1 reduces the frequency of DNC initiation events. While this conclusion is strongly supported by NET-Seq and Chip-seq metagene analyses, the fluorescence mCherry and YFP values or qRP-PCR analyses of specific genes do not always behave as expected when looking at absolute values rather than mCherry/YFP or GCG1/SUT098 ratios, which is sometimes disturbing when reading the paper.

The paper was submitted to Review Commons and a revised version has been submitted to EMBO J. The revised version of the paper by Gowthaman now clarifies a certain number of questions by additional analyses and figures as well as text modifications. Overall the data and conclusions are more convincing. However there are still few points that should be addressed.

Thank you for helping us to present the conclusions more convincingly.

Major comments:
#1.1 Figure S3A-C: the authors conclude that low, medium and highly expressed DNCs are similarly affected by loss of Hda1 or Hda3. However, they just compared NETseq signal with no statistics. It would be worth defining the fold change between WT and hda1D or hda3D including some statistical analyses; this fold change could be substantially different considering low or highly expressed DNCs, or the other non-coding transcripts examined in new Fig. S3D.

Thank you for this suggestion. We present a revised table (Dataset EV3) that includes the fold change values and annotation of expression strength group. We performed the additional statistical analyses for both the coding and DNC transcripts (see also response for #1.2).

#1.2 More generally this reviewer is wondering whether the sensitivity to Hda1 is really specific to DNCs, or whether Hda1 would more generally contribute to the repression of lowly expressed transcripts, whether coding or non-coding. For example, GCG1 and ORC2 are significantly increased in hda3D and hda1D respectively (Fig. S3F-I). It would therefore be interesting to perform these fold-change analyses not only for non-coding transcripts, but also for coding RNAs. The authors could define whether lowly expressed coding genes are more prone to repression by Hda1C by stratifying coding genes according to NET-Seq expression levels in the presence or absence of Hda1 or Hda3. These additional analyses would avoid establishing an inaccurate bias towards DNCs.

Thank you for the suggestions, we added the missing statistics to the revised manuscript Fig EV3D. The p-value (Mann-Whitney test) for differences of fold change values between the stratified groups for the mutants compared to wild type in Fig EV3A-C. It appears that the fold change of coding transcription in hda mutants vs WT is similar irrespective of the transcription levels. In the hda1Δ, the moderately expressed DNC transcripts are affected more than the highly expressed DNC regions. Whereas hda3Δ affects the low and moderately expressed transcripts. We believe the effect could be explained by the differential action of the Hda1C subunits. We have clarified this in the revised manuscript (line 215).

#1.3 Figure S3D: it is not clear why the NET-Seq signal is rather high at the TSS of these non-coding transcripts, while it is really low at the TSS of coding RNAs and DNCs in Figure S3ABC. Some explanation should be provided.

Thank you for spotting this. There is a simple explanation for this signal. The genomic windows shown on the metagene plot were defined as [TSS-100 bp, TSS+500 bp] relative to the TSS of novel non-coding transcripts outside of DNC loci. The relatively high signal in this area comes from the upstream genes. The yeast genome is dense, and many genes start immediately downstream of another gene with little or no intergenic spacer (and sometimes even with an overlap). We have repeated the metagene plot (RFigure 1), now trimming the windows by upstream features (with 10 bp offset). With these new genomic windows, the upstream peak of NET-seq signal has disappeared. We believe this representation is an improvement compared to previous version and have replaced the figure in the revised manuscript (Fig EV3E).
Figure 1. Metagene plot of NET-seq data for non-coding RNA that are not classified as DNC. The plot shows genomic windows [TSS -100 bp, TSS + 500 bp] relative to the annotated transcript start site (TSS) of ncRNA transcripts.

#1.4 Figure S5: why is mcherry and not YFP mostly increasing in the H3WTdha1, H3K56Q or the double mutants? This again indicates a stronger repressive effect of Hda1 on coding rather than non-coding transcription.

The isogenic control “H3 wild type” carries only one copy of the two genes coding for H3, which has a general effect on transcription. Essentially, the genetic background of the yeast synthetic histone mutant collection sensitizes for a decreased ratio of mCherry/YFP. This result is also included in table S2, where deletions of the histone genes HHT2 (H3) and HHF2 (H4) are listed as shared repressors in both screens. Hda1C mutations show the increased ratio in the sensitized “H3 wild type” background, but not in backgrounds we tested that contain a wild-type dosage of histone genes. The loss of Hda1, as a deacetylase complex indeed might influence the coding transcription. If the effect was only on the coding transcription, the NET-seq would clarify the notion. Since by NET-seq we observe an increased signal in hda1Δ vs WT on DNC, it appears that the Hda1 effects on DNC at endogenous loci is more clear than potential effects on mRNA expression. These experiments only give us the information whether Hda1C acts alongside or independent of the H3K56ac pathway and not the specifications of the directionality. In general, the fluorescence ratio offers the most insightful information on the effects of mutations compared to wild type on the effects of expression from the promoters in this assay.
Referee #2 (Report for Author)

Response to the revision of Gowthama et al.,

Promoters are frequently transcribed in both directions. The divergent, 'upstream' transcript is frequently unstable. Transcription initiation is regulated through the acetylation of promoter-proximal nucleosomes, where HDAC-dependent deacetylation of histones typically represses transcription initiation. The current manuscript addresses the question whether initiation of coding and divergent, non-coding (DNC) transcription is regulated by the same factors. Previously Marquardt and others showed that H3K56ac-mediated histone exchange has a differential effect on coding and DNC transcription.

Using a clever reporter system, the authors screened for positive and negative regulators that preferentially affect DNC transcription. They discover the Hda1 deacetylase complex as a DNC-biased repressor and diverse HATs as DNC-biased activators. The role of activators could not be validated, presumably due to high variability of the system.

Focusing on Hda1c the authors present data suggesting a larger effect of Hda1c on 'upstream' nucleosomes associated with DNC transcription than in coding transcription. Genome-wide NET-seq mapping was consistent with this differential regulation. Life cell imaging of one specific case argues that Hda1-mediated repression reduced the time between initiation events. The authors employ state of the art methods and in general the data are of very good quality. The effect size is small, which raises the broader question whether the results, while statistically significant is biological relevant.

Thank you for appreciating our choice of state-of-the-art methods, and the very good quality of the analyses. Higher-order mutants in pathways repressing DNC are often not viable, or show growth defects. It remains an open question if this reflects the outcome of genome-wide DNC regulation, or if some specific DNCs may be responsible. We agree that clarifying the purpose of DNC in cells is a fascinating question, but this is beyond the scope of this manuscript.
The authors addressed some of the points I raised in my previous evaluation to my satisfaction, but I want to maintain that some of the issues are not due to misunderstandings, and thus are maintained.

We are pleased that we could address some points to the satisfaction of this reviewer.

#2.1- If an experiment involves analysing 50,000 cells and a median signal is derived, then N=1. This is basis statistics. It becomes a good experiment when N=3 and you can compare three medians and see whether the data are reproducible. I can see the point that for an exploration one conducts one large experiment, to derive candidates, which are then validated independently. The fact that for the SAGA subunits the validation was apparently not possible, illustrates this point. For any experiment, where quantitative data are produced as a result, the number of biological experiments (different days) must be stated in the legends.

We thank the reviewer #2 for bringing up this point. We agree that N describes the number of observations, and we can see two possible interpretations for the screen data in our manuscript. Since we have not performed technical replicates of the same screen, we can agree that agree N=1 is appropriate to describe the number of technical repeat measurements of the entire mutant collection. However, we measure fluorescence in up to 50,000 individual cells, which can be appropriately regarded as N=50,000 (individual observations of biological events). In modern high-throughput applications we often have one biological sample with thousands or even million observations (cells, patient sample, DNA fragments, time series measurements etc). The more observations, the more confident we are that the measured parameters of the data distribution (e.g. median) are close to the true parameters of the general population. A median derived from 50,000 individual measurements that were measured once has a high quality. We also present two separate screens to support the conclusions. It is certainly interesting to debate if one interpretation of N is strictly correct here, but we feel that is most adequately discussed outside the peer review setting. To avoid confusion, we added this missing information on technical and biological repeats more accessibly in the revised manuscript (line 140) and the figure legends.

For the validation, we introduced the mutations into FPR strain independently of the screen. For these analyses we generated medians from three flow cytometry runs of independent yeast transformation events. The flow cytometry analyses of SAGA mutants actually validated the effect we expected based on the screens. We did not pursue these mutants further based on results of an orthogonal method measuring RNA expression.

#2.2- The point about live cell images is whether they were all done on one special day, or collected on multiple independent occasions.

The live-cell imaging of the strains were carried out in multiple occasions and resulted in multiple videos. Of the movies recorded in different days, individual cells were analyzed, and the data was pooled for better statistics. We have clarified this more in the Material & Methods section of the revised manuscript (line 811).

#2.3- I also maintain that the single molecule analysis has limited impact. Unfortunately, the coding GCG1 transcription was low, so could not be well compared to DNC SUT089. Of course
SUT089 can be measured as such (I do understand that), but it was not designed as a stand-alone experiment in the planning, you declared it to be one, when half of it did not work. The weakness is, as I stated, that you have only one experiment, for which half did not work.

I hope our response to #2.2 has clarified that we performed more than one experiment. We agree that it would have been fantastic to obtain additional insight into the regulation of shared promoters and effect of hda1Δ the acquisition of GCG1 live-cell imaging data. However, this manuscript focuses on DNC, so we would like to point out that this approach actually worked well to fully address the objective of capturing DNC. It also helped to verify the effect of hda1Δ on DNC region by an additional approach. The live-cell imaging experiment was also the key that connects the effect of hda1Δ on DNC regulation and initiation parameters. We believe the presentation of our data is more complete if we maintain the GCG1 analyses in the manuscript.

In summary, this is a high-quality study that presents the results of a genome-wide screen that will be of interest to colleagues in the narrower field. Due to the small effects the results may appeal less to a general readership.

Thank you for characterizing this manuscript as a high-quality study.

Referee #3 (Report for Author)

In this manuscript, Gowthaman et al describe the results and follow up of their screen aimed at identifying regulators of divergent noncoding (DNC) transcription in S. cerevisiae. From this screen, they identify Hda1C as a repressor of DNC transcription, and perform follow experiments to support and detail this finding. In addition to RTqPCR to confirm the reporter and endogenous changes, the authors perform NET-seq to look at global DNC alteration upon Hda1C subunit deletion and identify a number of non-coding transcripts with altered expression levels. In addition, the authors perform live cell imaging to demonstrate that there is a modest restriction of initiation frequency when one of the subunits of Hda1C is deleted. Finally, the authors explore changes to pan-H3 acetylation and the genetic overlap between Hda1C and H3K56ac demonstrating independent genetic pathways, but overall increases in H3 acetylation over DNCs when Hda1C is deleted.

The authors addressed all my major concerns in the revision of this manuscript.

Thank you for helping us to improve our manuscript. We are very pleased that we could address all major concerns.

#3.1 I am sorry for bringing up something new on the revision, but re-reading the manuscript made me think about whether there is anything in common with the DNC transcripts identified in the NETseq data as being regulated by hda1Δ. Examining the FPKM (Fig S3) and correlating with the H3ac (Fig S5D) was great, but I am curious about more information. For instance, is there any GO term enrichment in the corresponding mRNA, any characteristics of the -1 nucleosome (H2AZ occupancy?), any sequence characteristics/motifs of the promoter for the DNCs, any characteristic NDR size? These are just a few thoughts, but perhaps there are other
defining characteristics that could be quickly examined using public datasets. I realize some of these analyses could take a few hours to complete, but I do think that if something were identified (or also if things were found to not be connected), it would be of great interest.

We are happy to see that the revisions stimulated the curiosity of this reviewer. Thank you for these helpful suggestions. We are actively trying to identify patterns that could explain Hda1C targeting or effect. However, fully resolving this question will require major analyses with a timeline of months or even years rather than a few hours as the reviewer rather optimistically suggests, at least for us. We present the H2A.Z signal at the DNC loci (n=1517) using ChIP-seq tracks of WT samples from Gu et al., 2015 (PMID 25765960). The DNC loci were classified into “upregulated” and “downregulated” by the hda1Δ based on the FPKM ratios of hda1Δ vs WT samples. There were 364 down- and 1153 upregulated DNC transcripts in the hda1Δ. We observe high occupancy of H2A.Z at DNC loci that are upregulated in hda1Δ. We believe that this knowledge is interesting and hope reviewer #3 can see that it will be more helpful to present these new analyses as part of separate study.

**Figure 2.** Metagene plot representing the wild type H2A.Z signal at DNC loci (ChIP-seq data from Gu et al., 2015. PMID 25765960). The genomic intervals cover the first 500 bp of the coding gene (scaled to 100 bins), the first 1 kb of the DNC transcript (scaled to 200 bins), and the variable length gap between the coding TSS and the DNC TSS (scaled to 50 bins). The solid lines represent the transcription start site (TSS) of DNC transcript and host genes, respectively. The orange and blue signal peaks represent the DNC loci downregulated (n=364) and upregulated (n=1153) in hda1Δ.
Dear Sebastian,

Thank you again for submitting the final revised version of your manuscript. I am pleased to inform you that we have now accepted it for publication in The EMBO Journal.

Your article will be processed for publication in The EMBO Journal by EMBO Press and Wiley, who will contact you with further information regarding production/publication procedures and license requirements.

Should you be planning a Press Release on your article, please get in contact with embojournal@wiley.com as early as possible, in order to coordinate publication and release dates.

Congratulations on your successful publication, and thank you again for this contribution to The EMBO Journal! Please continue to consider EMBO Journal for your work in the future.

Kind regards,

Stefanie

Stefanie Boehm
Editor
The EMBO Journal

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

Referee #1:

The authors have addressed all my concerns in the revised version of their manuscript, which should now be acceptable for publication.

***
Rev_Com_number: RC-2021-00750
New_manu_number: EMBOJ-2021-108903R
Corr_author: Marquardt
Title: The Hda1 histone deacetylase limits divergent non-coding transcription and restricts transcription initiation frequency
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 authority guidelines in preparing your manuscript.

### B- Statistics and general methods

The experiments include adequate technical and biological repeats to perform the necessary statistical analysis and tests.

The vast majority of p-values obtained in this study were calculated by non-parametric tests (Fisher’s exact test, Spearman correlation test, bootstrap). The only parametric test was Student’s t-test. It was used for analysis of validation of flow cytometry in mutants, qPCR and smFISH data. For validation of mutants and qPCR the number of replicates (n=3) was not enough to reject the assumption on Normal distribution of the measurement error.

#### 2a. How was the sample size chosen to ensure adequate power to detect a pre-specified effect size?

The experiments include adequate technical and biological repeats to perform the necessary statistical analysis and tests.

#### 2b. For animal studies, include a statement about sample size estimate even if no statistical methods were used.

NA

#### 2c. Describe inclusion/exclusion criteria if samples or animals were excluded from the analysis. Were the criteria pre-established?

NA

#### 2d. 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

#### 2e. For animal studies, include a statement about randomization even if no randomization was used.

NA

#### 2f. 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

#### 2g. For animal studies, include a statement about blinding even if no blinding was done.

NA

#### 2h. For every figure, are statistical tests justified as appropriate?

No

#### 2i. Do the data meet the assumptions of the tests (e.g., normal distribution)? Describe any methods used to assess if.

Yes

#### 2j. Is there an estimate of variation within each group of data?

The barplots show standard error of the mean (SEM). Variance or standard deviation were not explicitly calculated for the barplots.
### 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), IDeepevalve (see link list at top right).

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

| Dual use research of concern | Provided in the reagents_tools excel file. |
|-----------------------------|-----------------------------------------|
| Data Accessibility          | Provided.                                |
| Human Subjects              | Provided.                                |
| Animal Models               | Provided.                                |
| Select agents and toxins    | Provided.                                |

1. Is the variance similar between the groups that are being statistically compared? Yes. Since the groups on barplots have the same number of measurements, the values of SEM are proportional to the values of SD within each barplot. The analysis of SEM shows that there were no outlier groups with unexpectedly high variance.

### D- Animal Models

8. Identify the committee(s) approving the study protocol.

9. For experiments involving live vertebrates, include a statement of compliance with ethical regulations and identify the committee(s) approving the experiments.

| Animal Models | NA |
|---------------|----|

### E- Human Subjects

11. Identify the committee(s) approving the study protocol.

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.

### F- Data Accessibility

16. 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 GEO/3942, Antibody data: PRIDE P0000036 etc.) Please refer to our author guidelines for 'Data Deposition'.

18. Deposition is strongly recommended for any datasets that are central and integral to the study.

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).

### G- Dual use research of concern

21. Could your study fall under dual use research restrictions? Please check biosafety documents (see link list at top right) and list of select agents and toxins (APHS/CDC) (see link list at top right). According to our biosafety guidelines, provide a statement only if it could.

| Dual use research of concern | NA |
|-----------------------------|----|

For all hyperlinks, please see the table at the top right of the document.