natureresearch

Peer Review Information

Journal: Nature Genetics
Manuscript Title: Prediction of histone post-translational modification patterns based on nascent transcription data
Corresponding author name(s): Dr Charles Danko

Reviewer Comments & Decisions:

Decision Letter, initial version:

5th Mar 2021

Dear Charles,

Your Article, "Interdependence between histone marks and steps in Pol II transcription" has now been seen by 3 referees. I apologize for the long review process. Despite our multiple chase emails, reviewer #4 has not submitted a timely report. We have now decided to proceed based on the other three reviews.

You will see from the reviewers' comments copied below that while they find your work of potential interest, they have raised quite substantial concerns that must be thoroughly addressed. In light of these comments, we cannot accept the manuscript for publication, but would be interested in considering a revised version that addresses these serious concerns.

Reviewer #1 seems positive about the work overall but notes several limitations, including data representation, performance metrics, and focus on heterochromatin/repressive marks.
Reviewer #2 says that the method seems reasonably accurate but that you would need to more carefully acknowledge and investigate the limitations of the imputation. Also, the software needs to be made fully accessible, including instructions on how to use it. Please see https://www.nature.com/nature-research/editorial-policies/reporting-standards#availability-of-computer-code
Reviewer #3 likes the work but feels that you should focus more on specific examples and perhaps remove others that are too preliminary; go deeper instead of broader. Ideally, the reviewer would like to see further validation (functional assays) to support your claims.
In sum, the reviewers feel that the approach seems promising but there are notable limitations in the analysis at this point.

We hope you will find the referees' comments useful as you decide how to proceed. If you wish to submit a substantially revised manuscript, please bear in mind that we will be reluctant to approach the referees again in the absence of major revisions.

If you choose to revise your manuscript taking into account all reviewer and editor comments, please highlight all changes in the manuscript text file. At this stage we will need you to upload a copy of the manuscript in MS Word .docx or similar editable format.

We are committed to providing a fair and constructive peer-review process. Do not hesitate to contact me if there are specific requests from the reviewers that you believe are technically impossible or unlikely to yield a meaningful outcome.

If revising your manuscript:

*1) Include a "Response to referees" document detailing, point-by-point, how you addressed each referee comment. If no action was taken to address a point, you must provide a compelling argument. This response will be sent back to the referees along with the revised manuscript.

*2) If you have not done so already please begin to revise your manuscript so that it conforms to our Article format instructions, available here.
Refer also to any guidelines provided in this letter.

*3) Include a revised version of any required Reporting Summary:
https://www.nature.com/documents/nr-reporting-summary.pdf
It will be available to referees (and, potentially, statisticians) to aid in their evaluation if the manuscript goes back for peer review.
A revised checklist is essential for re-review of the paper.

Please be aware of our guidelines on digital image standards.

You may use the link below to submit your revised manuscript and related files:

[REDACTED]

Note: This URL links to your confidential home page and associated information about manuscripts you may have submitted, or that you are reviewing for us. If you wish to forward this email to co-authors, please delete the link to your homepage.

If you wish to submit a suitably revised manuscript we would hope to receive it within 6 months. If you cannot send it within this time, please let us know. We will be happy to consider your revision so long as nothing similar has been accepted for publication at Nature Genetics or published elsewhere. Should your manuscript be substantially delayed without notifying us in advance and your article is eventually published, the received date would be that of the revised, not the original, version.

Please do not hesitate to contact me if you have any questions or would like to discuss the required revisions further.

Nature Genetics is committed to improving transparency in authorship. As part of our efforts in this direction, we are now requesting that all authors identified as 'corresponding author' on published papers create and link their Open Researcher and Contributor Identifier (ORCID) with their account on the Manuscript Tracking System (MTS), prior to acceptance. ORCID helps the scientific community achieve unambiguous attribution of all scholarly contributions. You can create and link your ORCID from the home page of the MTS by clicking on 'Modify my Springer Nature account'. For more information please visit please visit www.springernature.com/orcid.

Thank you for the opportunity to review your work.

Sincerely,

Tiago

Tiago Faial, PhD
Senior Editor
Nature Genetics
https://orcid.org/0000-0003-0864-1200

Reviewers' Comments:

Reviewer #1:
Remarks to the Author:
Wang, Chivu, and colleagues present an integrated analysis of the relationships between high-resolution transcription (as measured by PRO-seq) and various histone modifications. The main contribution is the demonstration that the distribution of many histone modifications can be imputed using machine learning models trained on PRO-seq data. The approach itself is quite straightforward; a support vector regression model is trained using vectors of PRO-seq read counts to predict histone modification read counts at the same location. But the work convincingly demonstrates that transcriptional data can be used to impute several active histone modifications, and can thereby be used to characterize chromatin state annotations. The approach is comprehensively evaluated and demonstrated to enable imputation across cell types and species.

Major comments:

1) While a large amount of evidence is presented to support the claim that dHIT can impute histone modification data, it is still difficult to get a sense of the approach's accuracy on various histone modifications and genomic locations. The favored assessments focus on correlations and visual representations (e.g., heatmaps). There are problems associated with both of these approaches.

Correlation is an imperfect metric of performance - it can be dominated by low-signal background regions and by outliers. And heatmaps can look convincing while obscuring incorrect predictions. Perhaps a more informative approach would be to calculate area under precision recall curves when predicting enriched domain/peak-level information in the various histone modification experiments. One could look at how imputation performance is measured in related work such as PREDICTD (Duham, et al. Nat Comms, 2018), ChromDragoNN (Nair, et al. Bioinformatics, 2019), or Avocado (Schreiber, et al. https://doi.org/10.1101/533273). I do understand that these other approaches have distinct motivations and training objectives, but their performance metrics are nonetheless relevant.

2) It would be very informative if performance evaluations can be presented separately for TSS-proximal regions and distal enhancers. Most presented heatmaps and metagenes focus on the TSS.

3) Further related to performance metrics, Figure 2E shows the results of Jaccard distances between chromHMM states found from real data and dHIT-imputed data. I was confused by these plots, as they seem to show the worst performance for the quiescent state and the strong transcriptions states. I would have naively thought that such states should be amongst the easiest to predict from PRO-seq (quiescent due to no transcriptional signal, and transcription states from the PRO-seq signal directly). Why is performance poor here?

4) It is clear from several figures that dHIT has poor performance in imputing histone modifications associated with repression or heterochromatin (i.e., H3K9me3, H3K27me3, H4K20me1, e.g., Fig 3C). This is not surprising, as such regions would contain little transcriptional activity. It is therefore perplexing that a substantial portion of the manuscript asks the reader to believe that H3K27me3 signals can be successfully imputed from PRO-seq. In several places, the manuscript states that success has been achieved in predicting polycomb enriched regions or bivalent gene, even when the presented imputed H3K27me3 signal is extremely different from the actual H3K27me3 ChIP data (e.g., Fig 2A, Supp Fig 11). We are shown dHIT-derived predictions of polycomb domains in several GMBs, but there are no validation experiments to assess the predictions.

5) Related to the previous point, the conclusion contains a section that relies on dHIT predictions to claim that H3K27me3 has a positive association with initiation. This is extremely speculative. Since dHIT has very poor performance in imputing H3K27me3, how can such sweeping biological conclusions be drawn from these analyses?

6) The second half of the manuscript, beginning with the Triptolide treatment experiments, seems very disconnected from the first half. The results are interesting, but mostly recapitulate the known relationships between transcription and the deposition of histone modifications. The results do not rely on or relate to the dHIT method, as they are mostly derived by examining the actual histone modification ChIP-seq data post-Trp. In fact, when dHIT is invoked, it probably should not be. For example, in the section where dHIT is used to predict ChIP-seq data on systematically varied initiation and pause-release rates; these results could have been more clearly interpreted if pausing indices were calculated directly from the PRO-seq and compared with the various histone modification ChIP-seq data directly. Similarly, in the section "chromatin accessibility is not sufficient for transcriptional initiation"; dHIT is not needed here - one could examine the PRO-seq data directly to see if PolII was at all DNaseI hypersensitive sites (and several others have already demonstrated that it is not).

7) It would be good to discuss or compare with other somewhat related work modeling PRO-seq data, such as the manuscripts from Azofeifa and others in Robin Dowell's lab, and NRSA from Yu Shyr's

group.

Minor comments:
8) Why are only two histone modifications assessed for mES cells (Fig. 1E)? There are many more available from mouse ENCODE and other labs.

9) Supp Fig 8d compares dHIT performance in a variety of cell types to a scheme that just directly transfers the training data from K562 cells. Have you tried comparing with the performance of a scheme that transfers the average histone modification signal from all other cell types (i.e., as described by Schreiber, et al. Genome Biology 2020)?

10) Please put x-axis and y-axis labels and scales on Supp Fig 1 plots. There is a strange vertical drop-off in midrange ChIP-seq signal in many of the plots (x-axis), suggesting that the plots do not show read counts.

Reviewer #2:
Remarks to the Author:
In the Wang et. al. manuscript, they train a support vector regression model that uses run-on sequencing as input and predicts the histone profile (marks, locations) genome wide. The predictions are reasonably accurate and dissection of how predictions change given distinct input transcription data (Figure 5) links particular marks to distinct stages of RNA polymerase activity. They further support the data by blocking transcription and assessing a number of marks acutely (<4 hours) afterwards. Overall this is a very thorough assessment of the relationship between transcription (using both simulated and experimental data) and histone marks. However, some concerns limited enthusiasm.

Major issues:
(*) The section on pause release, methylation and initiation (based on simulations) is really aimed at understanding what patterns the SVR is identifying. It should be couched as such and really it should precede the Trp experiment. It was disorienting for this section to follow the Trp experiment. Further, subsequently claiming that they have little idea what patterns are driving the model is an over-statement, as this section begins to dissect it. Though, I do agree that this section alone falls short of fully understanding their black box model.

(*) The experiment to block transcription is a nice, strong validation of directionality — i.e. that goes to causality rather than just correlation. That said, it is somewhat complicated by the half-life of these marks — which is only commented on for the one that looks unchanged. Additionally, what is the evidence that small molecule Trp has absolutely no impact on histone marks or accessibility. Plus, Trp is quite toxic — so are the 4 hour cells just quite sick/dying and that is responsible for the change in histone patterns?

(*) The paper needs to more carefully look at the mistakes made by their imputation strategy. They clearly are aware some regions may be more difficult than others as evidenced by them removing regions of poor mappability, black list regions and the 2% of identified "spikes" as these are likely poorly predictable for technical regions. This makes sense, but have they examined the error characteristics of their imputed patterns for the remainder of the data? The focus on large windows for their correlation coefficients (SFig4, SFig 8 use 10kb, Fig 1E 1kb) provide nice summaries but say little about whether the errors are randomly distributed, punctate, prone to particular regions, or what. For example — SFig 3 — some of the patterns (H3K27ac, H3K4me2 and H3K4me3) it looks like the imputed data tends to be wider and darker — are these the general trends of the errors made?

Also I'm a little concerned about the fact that they claim that poorly performing windows are enriched for CTCF when they use CTCF sites in their MNase normalization strategy. Doesn't the use of CTCF peaks to identify untranscribed regions make it impossible for them to then claim that that CTCF windows are the most poorly predicted?

(*) Abstract argues that their accuracy is on par with replicates but they never explicitly look at replicate variability.

(*) Their approach for training used a lot of heuristics (how they defined informative positions, how they pseudo-randomly select examples). Is this done merely to balance positive/negative regions, to reduce the overall burden of training data, because it worked, or what? Minimal justification is provided for what is actually a ton of heuristics. Leaves the reader wondering why a lot and whether these heuristics drastically influence the results. I'm fine (in general) with tuning a model as long as the methodology behind such tuning is conveyed with reasoning to the reader, especially in the scenario where the model isn't being built to generalize but rather to prove a point.

(*) Availability on the software isn't specified.

Minor issues:
(*) The H3K27me3 discrepancy is quite interesting. They clearly identify distinct patterns, but do not clearly show that their approach could predict the more punctate pattern. So the "both of which appear linked to features of active transcription" is an over statement. Would be interesting to follow this mark through differentiation. May speak to roles of H3K27me3 in differentiation?

(*) By reference they refer to a collection of K562 PRO and assorted ChIP marks from ENCODE, but this paper would be more cohesive if they included some assurance on the quality of these datasets. Alternatively it might be relevant to know whether quality of the data influences the ability of the model to identify these patterns?

(*) Does the registry of the non-overlapping windows matter in their ChIP predictions? Windowing methods always have edge effects — which are admittedly minimized at smaller window sizes. But at the larger sizes (500, 5kb) does an offset in the window start size relabel any regions? Likewise I don't recall them mentioning how they combine windows yet the evaluation windows are quite large (10kb and 1kb) compared to the 10 bp windows on which they impute/predict.

(*) I have a general issue with saying PRO-seq throughout the paper when reality is this is applicable to several run-on /nascent assays. Perhaps a more general term is preferred? In methods section they say "PRO-seq, GRO-seq, or ChROseq; henceforth referred to simply as PRO-seq" — and this would be fine if this statement had been in the main text.

(*) It is unclear why they are using hg19 when hg38 has been available since December of 2013. Justification of hg19 is warranted. However, it is unlikely that a shift to hg38 will alter any results, so this is perceived as a minor concern.

(*) The sentence, “Signal on the lower end was better spread out using data imputed from PRO-seq, possibly making use of the greater dynamic range of PRO-seq over ChIP-seq.” is completely opaque. What does “lower end” here refer to explicitly — ChIP or PRO? Are they trying to say that PRO predicted a broader dynamic range then ChIP? If so, it’s possibly true but somewhat dismissive of one of the error types seen by the model.

(*) Text refers to SFig 1a-b but there are no a/b labels on SFig 1.

(*) “As TRP does not affect engaged RNA polymerase, we observe a clearing wave of Pol II ~100kb from the TSSs on long genes at 1 h (Fig 4B)…” But this is not shown in Fig 4B.

(*) Does “local environment” actually just mean TFs? “Our analysis supports a model in which both chromatin accessibility and the local environment are important factors to facilitate transcription initiation by Pol II”

(*) On SFig 11 some of the labels were overlapping and hard to read.

(*) Figure SFig 12 is completely uninterpretable other than to look impressive. Can these be systematically classified as wide and punctate? Can the two types be shown as distinct figures to make each tract more readable? Fewer tracts? Meta-genes across multiple cell lines?

(*) Color schemes on many of the figures is pseudo-random. For example, H3K36me3 is shown in Figure 1 and SFig17 as a mid-tone green. But in Figure 3 and SFig 15 it’s a light green and H3K4me1 is the mid-tone green. A similar green is used in the correlation grids to mean Pearson’s but then Spearman’s and the jsd values are both purple. So every figure panel required the reader to figure out what the color scheme was now. Admittedly, they are showing a tremendous amount of data across the figures and the bouncing around of colors is likely, to some extent, unavoidable — but every effort to make colors and symbols standardized throughout the paper would help the reader with the disorienting nature of having to figure out how to interpret every figure with its own unique color scheme.

(*) Also — on these measures of correlation/similarity there is an alpha value being used to scale the “heat maps” but no key is given.

Reviewer #3:
Remarks to the Author:
This article from Wang et al. describes a computational framework they call dHIT that uses machine learning to impute histone modification landscapes across the genome using nascent RNA data such as PRO-seq. The model was extensively trained and tested in a number of cell types and across species and appears to do a reasonable job of predicting the locations and levels of a number of histone marks that are often used to define chromatin state. In the cell type used for training, Pearson’s correlation between prediction and actual data range from mediocre (0.37 for H3K27me3) to quite good (~0.7 for H3K27ac, H3K36me3, H3K9ac, and H3K4me2/ me3). The imputed locations are considerably more ‘smeary’ than real data and lack the positional information on nucleosomes gleaned from experimental data, but define general regions where histones are likely to bear a certain modification. Weaknesses

of the model are in predicting heterochromatin and repressed Polycomb regions, ZNF genes and repeat elements.

Together, these findings indicate that nascent transcription data can be used to predict areas of activity in the genome, such as active promoters and enhancers, and the histone marks associated with activity (acetylation, H3 K4methylation, H3K36 methylation). Whereas this is probably not surprising, I appreciate the point being made here, which is that PRO-seq or a related nascent RNA assay is a much more efficient way to characterize a cell type than is 20 ChIP-seq assays.

The ability of nascent transcription levels to predict active histone marks also supports the growing body of data showing that these histone marks reflect transcription, rather than dictating or regulating transcription. Previous work from the Spicuglia lab (numerous papers, should be cited) Lis lab (Core et al), Adelman lab (Henriques et al, 2018, should be cited) have shown correlation between levels of active histone marks like H3K4me3 and transcription activity at both promoters and enhancers. The current work extends these studies, going beyond correlation to determine causality, using transcription inhibition with Triptolide. As predicted based on prior work in yeast (Howe lab; Martin et al., cited), loss of transcription causes loss of active histone marks H3K27ac and H3K4me3. Interestingly, the H3K36me3 mark and H3K4me1 turnover more slowly and are not as temporally dependent on transcription. This is a nice set of experiments that will hopefully help drive home the point that histone modifications aren't directive for activity, nor do they bookmark regions for future activity.

Oddly, after providing some of the cleanest evidence yet that histone modifications, in particular H3K27ac and H3K4 methylation, reflect transcription rather than controlling it, the authors then delve into a section wherein they investigate "whether each histone modification facilitates either initiation or pause release". This section of the manuscript is very weak, and I remain unconvinced by these simulations that histone modifications 'facilitate' either initiation or pause release. I strongly suggest that this section of the manuscript be dramatically strengthened or (preferably) removed.

Finally, the authors work to demonstrate that not all accessible chromatin regions are sites of transcription initiation. This too has already been described in the literature, and it is known that DNase or ATAC-seq accessible sites include CTCF-bound loop anchors that are not transcriptionally active (Higgs, Buenrostro). However, this is probably the clearest description of this finding that I know of, and I appreciate the bigger commentary the authors are trying to make. However, Figure 6 is currently such a jumble of small panels that the main point does not come across clearly. I recommend that Figure 5 be removed and the authors use this space to expand Figure 6. This would allow them to better document the absence of transcription initiation at accessible regions, showing heatmaps and larger figures that bring this point home more clearly.

Overall, dHIT seems like a powerful tool and the take home message that histone modifications are not directive for transcription, but instead reflect transcription activity is important. However, the model does have weaknesses that should be acknowledged, and I have several specific concerns, as outlined below:

Major concerns:
1) dHIT doesn't perform nearly as well at predicting regions of gene inactivity or repressed chromatin domains. This is perhaps not surprising, since it is based on nascent RNA sequencing. It would be helpful if the authors could comment on which histone ChIP-seq assays might complement PRO-seq

and dHIT to give this fuller picture of chromatin? Could one do PRO-seq/dHIT and H3K9me3 plus H3K27me3 ChIP-seq to achieve this? This manuscript would be stronger if the authors could provide some insights into which repressive marks one should investigate by ChIP-seq to get a comprehensive picture of the chromatin landscape.

2) The authors observe a rapid loss of H3K27ac and H3K4me3 upon inhibition of transcription. Is this due to rapid deacetylation/ demethylation or histone turnover? They appear to argue for deacetylation rather than turnover, but this is not demonstrated. I suggest that the authors perform a simple assay to test this, using deacetylase inhibitors in Triptolide treated cells to confirm that acetylation is retained under these conditions. Whereas ChIP-seq would be optimal here, even western blots would help make this argument more compelling. This small experiment could go a long way to develop the model for how transcription stimulates deposition or retention of active chromatin marks.

3) Figure 5 shows a number of correlations between histone acetylation or methylation and simulations of initiation and pause release. These are nice correlations but don't speak in a clear way to function, and thus the conclusions such as 'methylation works at the stage of transcription initiation' appear unfounded. To support these comments, the authors could treat cells with inhibitors of acetylation/ deacetylation or methylation/ demethylation, or work in cells with catalytically inactive methyltransferases. In these conditions, one could test the authors conjecture that methylation or acetylation directly 'work' at a specific step in the transcription cycle. Such concrete experiments testing the simulation would be required to support the authors conclusions about function.

Minor comments:
1) The jumbled and small nature of many figure panels makes this manuscript more difficult to read than optimal.
2) Figure 4B. The butterfly D. iulia doesn't look like the picture shown. That appears to be a monarch?

Reviewer #4:
None

Author Rebuttal to Initial comments

We have made substantial changes to our manuscript, incorporating comments and suggestions from our three reviewers. We were happy to see that reviewers were unanimously excited about the work presented in our original manuscript, saying, for example, that our “work convincingly demonstrates that transcriptional data can be used to impute several active histone modifications, and can thereby be used to characterize chromatin state annotations"; and that our work “is a very thorough assessment of the relationship between transcription (using both simulated and experimental data) and histone marks". We were also pleased to have a large number of highly constructive comments that have contributed significantly to improving our revised manuscript. We believe that we have fully addressed the reviewers’ thoughtful comments in the accompanying revision.

Changes of particular note include:

1. We removed Figure 5 and used the remaining space to expand Figure 6, making the panels easier for readers to digest.
2. We included alternative performance metrics for dHIT, including MSE quantification at different subsets of genomic sites, as well as ROC and PRC curves for the recovery of peak calls.
3. We performed experiments to demonstrate that the rapid removal of H3K27ac is explained, at least in part, by removing the histone mark, rather than evicting histones carrying the mark. This result implies a tight balance between the deposition and removal of H3K27ac that follows active transcription.

Please find below a point-by-point breakdown addressing the reviewers’ comments.

Reviewer #1:

Remarks to the Author:
Wang, Chivu, and colleagues present an integrated analysis of the relationships between high-resolution transcription (as measured by PRO-seq) and various histone modifications. The main contribution is the demonstration that the distribution of many histone modifications can be imputed using machine learning models trained on PRO-seq data. The approach itself is quite straightforward; a support vector regression model is trained using vectors of PRO-seq read counts to predict histone modification read counts at the same location. But the work convincingly demonstrates that transcriptional data can be used to impute several active histone modifications, and can thereby be used to characterize chromatin state annotations. The approach is comprehensively evaluated and demonstrated to enable imputation across cell types and species.

Response: We thank the reviewer for their constructive and thoughtful comments which have contributed substantially to improving our revised manuscript.

Major comments:

1) While a large amount of evidence is presented to support the claim that dHIT can impute histone modification data, it is still difficult to get a sense of the approach's accuracy on various histone modifications and genomic locations. The favored assessments focus on correlations and visual representations (e.g., heatmaps). There are problems associated with both of these approaches. Correlation is an imperfect metric of performance - it can be dominated by low-signal background regions and by outliers. And heatmaps can look convincing while obscuring incorrect predictions. Perhaps a more informative approach would be to calculate the area under precision-recall curves when predicting enriched domain/peak-level information in the various histone modification experiments. One could look at how imputation performance is measured in related work such as PREDICTD (Duham, et al. Nat Comms, 2018), ChromDragoNN (Nair, et al. Bioinformatics, 2019), or Avocado (Schreiber, et al.
https://doi.org/10.1101/533273). I do understand that these other approaches have distinct motivations and training objectives, but their performance metrics are nonetheless relevant.

Response: We added precision recall curves (PRC) for active (H3K27ac, H3K4me1, H3K4me1, H3K36me3) and repressive histone marks (H3K9me3, H3K27me3) to the revised manuscript.

We used a similar setup as Nair et. al. (Bioinformatics, 2019), in which we divided the holdout chromosome into 500 bp non-overlapping windows from which we exacted (presumptive) ground truth labels using cell type specific peak calls generated by ENCODE. We generated PRCs by thresholding the imputed histone modification signal intensity to divide the same windows into those predicted to be enriched/ not enriched for each histone mark. Although we favor PRCs because of substantial class imbalance between true positive and true negative windows, we also show ROC curves generated using the same strategy. Finally, to provide additional context for the PRC (or ROC curves) that we expect to achieve when applying this performance evaluation to experimental data, we have also included PRC/ ROC curves for the same histone modifications using an experimental dataset. All analyses focus on the holdout chromosome (chr21) in the holdout cell type (GM12878).

This new performance metric shows that we are able to predict histone modifications, with the notable exception of H3K9me3, with nearly the same fidelity as experimental measurements, consistent with our prior analysis. All of these ROC/ PRC curves, as well as the area under the curve, are shown in the revised **Sup Fig 2**.

2) It would be very informative if performance evaluations can be presented separately for TSS-proximal regions and distal enhancers. Most presented heatmaps and metagenes focus on the TSS.

Response: In response to this comment, and a related comment by reviewer 2, we now present performance metrics separately for specific genomic regions. We adopted performance metrics similar to those presented by Durham et. al. (Nat. Com., 2018) and Schreiber et. al. (Nat. Com., 2020), in which mean squared error is computed in different genomic regions, including the top

1% of imputed windows (MSEimp); and the top 1% of experimental windows (MSEobs). Additionally, we added two independent definitions of promoter and enhancer, using either proximity to gene annotations (GENCODE) or the stability of the transcription unit produced by each annotation (following the nomenclature detailed in Core and Martins *et al*., 2014). None of these performance metrics appear to identify major discrepancies between different genomic regions that are not present in previous work by Durham or Schreiber.

We note that Durham and Schreiber observed a substantially lower error for MSE global, rather than the peak regions, reflecting the ease of imputing background signals using other ChIP-seq datasets. For dHIT, we found that MSE global is within the same ballpark as the various peak regions. Our intuition is that this reflects variation in the experimental ChIP-seq background signal that is difficult to accurately measure using PRO-seq data. For a more thorough discussion of the types of errors made by dHIT in background regions, please see our comment to Reviewer #2 (comment #3) and the new **Supplementary Note 1**.

These new performance metrics are depicted in **Sup Fig 5**, which complement the global correlations, Jaccard similarity, and mean absolute deviation values that we previously included for each histone mark.

3) Further related to performance metrics, Figure 2E shows the results of Jaccard distances between chromHMM states found from real data and dHIT-imputed data. I was confused by these plots, as they seem to show the worst performance for the quiescent state and the strong transcriptions states. I would have naively thought that such states should be amongst the easiest to predict from PRO-seq (quiescent due to no transcriptional signal, and transcription states from the PRO-seq signal directly). Why is performance poor here?

Response: Fig. 2e uses Jaccard as a similarity index, using the definition implemented by BedTools (see: https://bedtools.readthedocs.io/en/latest/content/tools/jaccard.html). Based on this definition, Jaccard will range between 0 and 1, with higher values indicating higher similarity between two BED files. Therefore, as predicted by the reviewer, quiescent and strong transcription states (along with active TSS and weak polycomb) are some of the states which can be predicted most accurately using PRO-seq. More difficult states to predict include active or weak enhancers - which often swap labels among other enhancer states in both experimental and imputed data (as shown in **Supplementary Fig 16**) - and heterochromatin/ posed enhancer, which are not predicted well using PRO-seq. Notably, these states also have the weakest correspondence when ChromHMM is run between biological replicates.

We have clarified the writing and figure caption to more accurately explain how we use the Jaccard metric. In particular, we believe the revier's confusion was caused by the word "difference" in the Results section. That sentence now reads as follows:

> "The Jaccard similarity index between imputed and experimental data were highly correlated with those observed between other ChIP-seq datasets (Pearson's R = 0.92; Fig. 2E, Supplementary Fig. 14)."

4) It is clear from several figures that dHIT has poor performance in imputing histone modifications associated with repression or heterochromatin (i.e., H3K9me3, H3K27me3, H4K20me1, e.g., Fig 3C). This is not surprising, as such regions would contain little transcriptional activity. It is therefore perplexing that a substantial portion of the manuscript asks the reader to believe that H3K27me3 signals can be successfully imputed from PRO-seq. In several places, the manuscript states that success has been achieved in predicting polycomb enriched regions or bivalent gene, even when the presented imputed H3K27me3 signal is extremely different from the actual H3K27me3 ChIP data (e.g., Fig 2A, Supp Fig 11). We are shown dHIT-derived predictions of polycomb domains in several GMBs, but there are no validation experiments to assess the predictions.

Response: Our expectation when starting this project was that we would not be able to predict repressive marks at all. As expected, dHIT has no predictive power at all for H3K9me3. However, results for H3K27me3 are a bit more nuanced, and we have made several changes to make sure this nuance comes across to the reader in the revised manuscript.

We do think our H3K27me3 model learns signals that are pretty close to the experimental data in K562, GM12878, CD4+ T-cells and other somatic cell types. There are several existing (and newly added!) pieces of evidence for this claim: First, correlations between experimental and imputed H3K27me3 are largely within the range observed between experimental datasets in K562 and GM12878 (see Pearson's and Spearman's correlations in **Supplementary Fig. 11**). Second, peaks of H3K27me3 are predicted with an accuracy that is only slightly lower than experimental data and much better than random guessing (see the revised **Supplementary Fig. 2**). Third, examination of H3K27me3 on the genome browser, in K562 cells (see **Fig. 1B** and **Fig. 2A**), shows that experimental data is broadly distributed across large genomic regions that have low transcription levels. Boundaries of the broad H3K27me3 domains are predicted relatively well by the imputation. We do believe this represents a partial success and we have clarified where our models work fairly well in the revised manuscript by emphasizing the concordance in cell types where H3K27me3 imputation works well.

However, our capacity to predict K27me3 breaks down in certain cell types, particularly stem cells (both IPS and ESCs). The reason for this breakdown in accuracy is that H3K27me3 is distributed in a very different way in (for example) mESCs than it is in K562: experimental data shows a punctate pattern where peaks frequently occur near the promoter of genes with low transcription levels. We believe this difference in distribution likely reflects a fundamental difference in the biology of H3K27me3 between different cell types.

To explore this biological difference in more detail, we have added new analyses to the revised manuscript and made several changes to the text in order to clarify when H3K27me3 imputation works well and where it does not. First, we have included an additional main figure panel to show that imputation does, to a large degree, predict H3K27me3 domains in K562 and GM12878. Second, we have added a new analysis that systematically characterizes the degree to which 86 ENCODE or Roadmap datasets in different cell lines reproduce the "dispursed" or "focal" pattern. Our analysis (see the revised **Supplementary Fig. 14**) shows that ESCs and IPSs tend to have a focal pattern, whereas somatic cell types are dispersed.

5) Related to the previous point, the conclusion contains a section that relies on dHIT predictions to claim that H3K27me3 has a positive association with initiation. This is extremely speculative. Since dHIT has very poor performance in imputing H3K27me3, how can such sweeping biological conclusions be drawn from these analyses?

Response: We agree with the reviewer's comment that our data suggesting that H3K27me3 has a positive association with Pol II initiation rates is highly speculative at this point. Based on this comment, and following useful suggestions from Reviewer 3, we have removed this section of the manuscript and the associated figure from the revised paper.

6) The second half of the manuscript, beginning with the Triptolide treatment experiments, seems very disconnected from the first half. The results are interesting, but mostly recapitulate the known relationships between transcription and the deposition of histone modifications. The results do not rely on or relate to the dHIT method, as they are mostly derived by examining the actual histone modification ChIP-seq data post-Trp.

Response: We have refined the connection between dHIT and the triptolide experiments in the revised manuscript by rearranging the order in which sections are presented, removing the simulation studies (as noted above and below), and refining the transition text that connects them.

Briefly, we think dHIT and the triptolide experiments are conceptually related enough that they are stronger when presented together in the same paper. Our analysis of dHIT shows clearly just how strong the correspondence between histone modifications and transcription actually is. The strength of this correlation serves as a useful backdrop to motivate exploring what the nature of any causal link might be. This motivates the experiments in which we ask whether transcription is necessary for histone modifications.

Specific changes to the manuscript include changes to the transition between dHIT and the tryptolide experiments (see the section "Transcription is required for promoter-associated histone modification") as well as the discussion (see section "Active histone modifications as essential cogs, rather than causes, of transcription").

In fact, when dHIT is invoked, it probably should not be. For example, in the section where dHIT is used to predict ChIP-seq data on systematically varied initiation and pause-release rates; these results could have been more clearly interpreted if pausing indices were calculated directly from the PRO-seq and compared with the various histone modification ChIP-seq data directly.

Response: In response to this comment (and a related comment by Reviewer #3, comment #3), we removed the section in which we simulate transcription initiation and pause release rates from the revised manuscript. We used the additional space that this change freed up to expand our analysis of primary PRO-seq and ChIP-seq data after blocking transcription using triptolide and improve the transition between different sections of the manuscript. We believe that these changes make the revised manuscript feel like a more cohesive, self-contained story.

> Similarly, in the section "chromatin accessibility is not sufficient for transcriptional initiation"; dHIT is not needed here - one could examine the PRO-seq data directly to see if PolII was at all DNaseI hypersensitive sites (and several others have already demonstrated that it is not).

Response: In response to this comment we have tried to better articulate the advantages of using the imputation in this task.

We absolutely agree with the reviewer that simple approaches examining whether Pol II was found at all DNase-I hypersensitive sites (HS) have (mostly) supported the conclusion that not all DNase-I HS are transcribed. We have added citations to additional prior studies in our revision. However, we know of at least one recent paper that has arrived at the opposite conclusion, namely that transcription is found at *all* DNase-I HS (see Young et. al. (Genome Biology, 2017). While we do believe this paper was flawed, it nevertheless received quite a bit of attention when it came out. Therefore, we believe that revisiting this question with an additional lens (i.e., imputation) provides some benefit in convincing the community, especially those who do not follow this discussion all that closely, that *not* all DNase-I HS are transcribed.

We also believe that there are some unique advantages to using dHIT to address this question. Previous work has largely relied on either comparing peak calls between DNase-I HS and transcribed regions or taking subsets of DNase-I HS and examining the amount of transcription. Both of these strategies are potentially problematic in several ways: either false negatives or false positives in peak calling could result in an incorrect answer without careful filtering; likewise, aligning PRO-seq signal on a subset of DNase-I HS that are candidates for being *not* transcribed is potentially fraught with circular logic, or could artificially reveal transcription at a small subset of the DNase-I HS included in the analysis. Our view is that inappropriate use of heuristics in defining regions to analyze is what has led to the observations by other groups that *all* DNase-I HS are transcribed and the confusion that has followed.

We believe that the primary benefit of using dHIT is that we can identify regions where there is a *huge* mismatch between transcription and DNase-I hypersensitivity in a relatively unbiased way. By selecting windows in which we observe that DNase-I HS experimental signal is strong, but for which we find no evidence of imputed accessibility, provides an alternative (and arguably more principled) way to identify DNase-I HS that are not transcribed. This strategy allows us to work around some of the limitations of the alternative approaches noted above.

In the revised manuscript, we have clarified the rationale for revisiting this problem by more clearly articulating the discrepancy between previous papers. We have also provided a more articulate motivation for the use of imputation in this task. See, for instance, the discussion (section titled “*Chromatin accessibility: Necessary, but not sufficient, for transcription*”):

> Unlike previous work, which relied on arbitrary heuristics to select sites with or without evidence for transcription, dHIT allowed us to directly identify candidate DNase-I accessible regions with a typically large imbalance between experimental and predicted transcription.

7) It would be good to discuss or compare with other somewhat related work modeling PRO-seq data, such as the manuscripts from Azofeifa and others in Robin Dowell's lab, and NRSA from Yu Shyr's group.

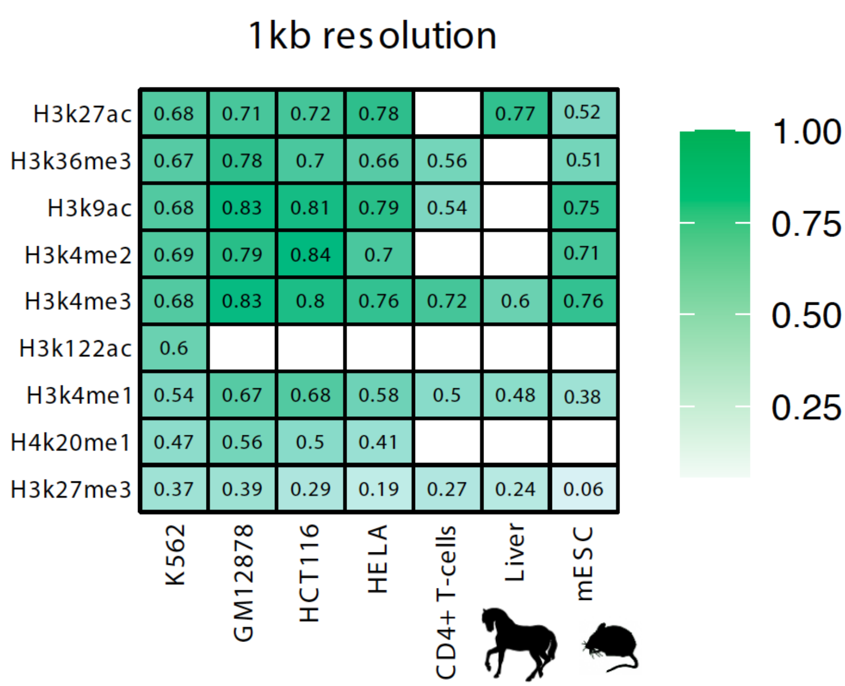
Response: We fully agree with this point. The excellent work from the Dowell and Shyr labs are now cited and discussed in the discussion section. See in the discussion (section titled "*dHIT: A powerful tool for genome annotation*"):

> Tools such as NRSA and Tfit leverage similar information, such as the shape or density distribution of nascent transcription, to annotate functional elements in eukaryotic genomes.

Minor comments:

8) Why are only two histone modifications assessed for mES cells (Fig. 1E)? There are many more available from mouse ENCODE and other labs.

Response: We have imputed five additional histone modifications in mESCs (H3K27ac, H3K36me3, H3K9ac, H3K4me2, and H3K4me1). Imputation was compared to data obtained by the mouse ENCODE project. Correlations and other QC for these additional histone modifications are now shown in the revised **Fig. 1e** and **Supplementary Fig. 10A,C,D**, copied for the convenience of the reviewer below.



9) Supp Fig 8d compares dHIT performance in a variety of cell types to a scheme that just directly transfers the training data from K562 cells. Have you tried comparing with the performance of a scheme that transfers the average histone modification signal from all other cell types (i.e., as described by Schreiber, et al. Genome Biology 2020)?

Response: We have added an additional 'straw-man' benchmark that uses the average histone modification signal from all other cell types, as suggested by the reviewer. We believe our changes in the revision emphasize that dHIT excels at picking out cell-type specific differences in histone modification signals that are important for characterizing the biology of new cell types.

(Schreiber et al. 2020) argues, very convincingly in our opinion, that machine learning models trained on genomic data can effectively learn the average signal intensity across cell types in the training data for a specific locus. While this is, to some extent, what machine learning models are designed to do, models that learn the average signal intensity are arguably not particularly useful in a biological setting for a number of reasons described in detail by the original authors.

We don't think this particular pitfall affects dHIT. Our rationale for this includes several observations we made in the original manuscript: (1) We only used one cell type for training, which prevents the model from memorizing the average signal across multiple cell types in a particular locus, (2) Comparing to training data copied from K562 (as described in the original manuscript) will let us determine whether we are simply copying the training data, and (3) We evaluated models using data held out from both a different holdout chromosome and a holdout cell type. (Full credit to Schreiber et. al. (2020): many of our decisions in designing the benchmarks in our manuscript were made based in part on hearing about their work before it was posted to bioRxiv).

In the revised manuscript, we have also added an additional benchmark designed to further make a case that dHIT is learning features that allow it to generalize across cell types. Schreiber et. al. (2020) suggests a performance metric that examines genomic loci that have a high degree of variability across cell types. In the revised manuscript, we compared the average signal to the imputation in regions on the holdout chromosome that have GM12878-specific signal for each histone modification. In general, dHIT performs well in this task, achieving correlations that are not significantly different from its global performance. By contrast, the average signal across cell types has a correlation of ~0 for most histone marks at these loci, as expected. We think this benchmark is a useful addition to the revised manuscript because it demonstrates that dHIT is useful for characterizing the regulatory features that separate different cell types.

These points have been added to the revised manuscript. See especially the revised section on the generalization across cell-types (results, section titled: "*Active histone modifications have a similar relationship to transcription across mammalian cells* "):

> Finally, cell-type specific signal differences were predicted with reasonably high accuracy (Pearson's R = 0.44-0.70 for active marks; Supplementary Fig. 10f), providing additional confidence that dHIT was not simply learning the average signal intensity of histone modification.

10) Please put x-axis and y-axis labels and scales on Supp Fig 1 plots. There is a strange vertical drop-off in midrange ChIP-seq signal in many of the plots (x-axis), suggesting that the plots do not show read counts.

Response: We have added axes labels to the scatterplots in the revised **Supplementary Fig. 1**. All values represent normalized read counts.

Our intuition is that the vertical drop-off reflects differences in the way that background signal is distributed between PRO-seq and ChIP-seq assays. In ChIP-seq, there is substantial background pulldown of DNA due to non-specific binding of DNA to beads, tubes, tips or other sources of contamination. This background signal can vary due to a variety of technical and biological factors (e.g., mappability; copy number alterations in some cell lines; etc.).

In PRO-seq, the background is generally much lower and distributed in a very different way than ChIP-seq. This is both because the signal is derived from RNA, rather than DNA, and also because background tends to be much lower thanks in large part to more affinity purification steps. We think this difference in the background distribution between assays makes it more difficult to predict the precise number of reads using PRO-seq in regions that do not have signal. We have commented on this point in the revised **Supplementary Note 1**.

Reviewer #2:

Remarks to the Author:
In the Wang et. al. manuscript, they train a support vector regression model that uses run-on sequencing as input and predicts the histone profile (marks, locations) genome wide. The predictions are reasonably accurate and dissection of how predictions change given distinct input transcription data (Figure 5) links particular marks to distinct stages of RNA polymerase activity. They further support the data by blocking transcription and assessing a number of marks acutely (<4 hours) afterwards. Overall this is a very thorough assessment of the relationship between transcription (using both simulated and experimental data) and histone marks. However, some concerns limited enthusiasm.

Response: We thank this reviewer for their constructive comments. This reviewer's comments motivated us to reorganize the manuscript in a more logical fashion, add additional control experiments to rule out cell viability as an explanation for histone modification loss following triptolide treatment, and add additional performance metrics to our evaluation of dHIT. We believe these changes have significantly improved the revised manuscript.

Major issues:
(*) The section on pause release, methylation and initiation (based on simulations) is really aimed at understanding what patterns the SVR is identifying. It should be couched as such and really it should precede the Trp experiment. It was disorienting for this section to follow the Trp experiment. Further, subsequently claiming that they have little idea what patterns are driving the model is an over-statement, as this section begins to dissect it. Though, I do agree that this section alone falls short of fully understanding their black box model.

Response: We agree with this comment. Additionally, we also agree with a related comment made by Reviewer #3, who suggested that we remove this section from the manuscript entirely. In response to these constructive suggestions, we have removed the section which describes the Pol II simulation studies as well as the accompanying Figure (previously Fig. 5) from the revised manuscript.

(*) The experiment to block transcription is a nice, strong validation of directionality — i.e. that goes to causality rather than just correlation. That said, it is somewhat complicated by the half-life of these marks — which is only commented on for the one that looks unchanged.

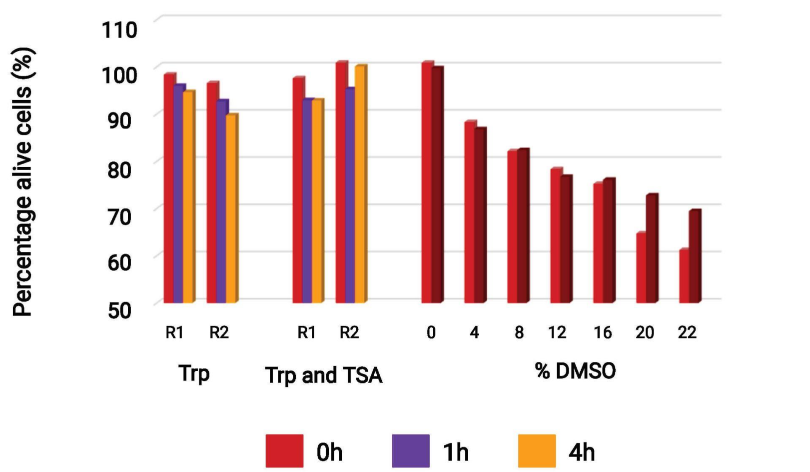
Response: In the revised manuscript, we now directly test the hypothesis that changes in H3K27ac reflect the rapid turnover of histone marks by eraser enzymes, rather than histone depletion near the transcription start site (see also our response to Reviewer #3, point 2). These new findings implicate the rapid half-life of H3K27ac in the depletion of histone modifications following Trp. We comment on this in the revised manuscript (see especially the Results section "*Transcription is required for promoter-associated histone modification*").

Additionally, what is the evidence that small molecule Trp has absolutely no impact on histone marks or accessibility. Plus, Trp is quite toxic — so are the 4 hour cells just quite sick/dying and that is responsible for the change in histone patterns?

Response: We have conducted several experiments designed to evaluate the effects of triptolide on cell viability or aspects of promoter/ enhancer molecular biology:

- **Analysis of chromatin accessibility following triptolide treatment.** We performed ATAC-seq following a time course of triptolide treatment in K562 cells. We observed only modest changes in chromatin accessibility genome-wide (a modest increase). This data is presented in the revised **Figures 5 and 6**.

- **Experiments to evaluate the effects of triptolide on cell viability.** We also examined whether cytotoxic effects of triptolide impact cell viability in the 1-4 hour time window examined in our experiments. We found that triptolide, as well as a dual treatment of Triptolide and Trichostatin A (included in response to comments by Reviewer #3, point #2), did not affect cell viability at 4 hour time points. In contrast, high concentrations of DMSO, a positive control, had a large impact on cell viability. This data can be found in **Supplementary Figure 24**, and is depicted below for the convenience of the reviewer. We conclude that the effect of triptolide on histone modifications can not be explained simply by cell death.



Supplementary Figure 24. Cytotoxicity measurements for Triptolide, and Triptolide - Trichostatin A dual treatment in K562 cells
Bar plots display absorbance quantified at 590nm for AlmarBlue dye incubated with K562 cells during Triptolide, or Triptolide and Trichostatin A treatments. Two technical replicates were averaged for each time point. R1 and R2 define separate biological replication of the experiment.

Finally, we also wish to add that triptolide is widely used in the literature to block transcription. Our application, including the concentrations and time points that we have used, are consistent with previous papers from multiple groups, including work by Jonkers et al., 2011 (Lis), Tettey et al, 2019 (Conaway), and others.

(*) The paper needs to more carefully look at the mistakes made by their imputation strategy. They clearly are aware some regions may be more difficult than others as evidenced by them removing regions of poor mappability, black list regions and the 2% of identified "spikes" as these are likely poorly predictable for technical regions. This makes sense, but have they examined the error characteristics of their imputed patterns for the remainder of the data? The focus on large windows for their correlation coefficients (SFig4, SFig 8 use 10kb, Fig 1E 1kb) provide nice summaries but say little about whether the errors are randomly distributed, punctate, prone to particular regions, or what. For example — SFig 3 — some of the patterns (H3K27ac, H3K4me2 and H3K4me3) it looks like the imputed data tends to be wider and darker — are these the general trends of the errors made?

Response: We have added **Supplementary Note 1** to the revised manuscript which lists and provides discussion on systematic differences that we have noticed between experimental and dHIT imputed ChIP-seq data. Several of the more important points are listed here for the benefit of the reviewer:

- **Poor prediction in background regions:** Several quality control metrics show evidence of a weak correlation between dHIT and ChIP-seq signal in background windows. Errors in the background signal reflect small differences in the predicted and experimental read counts at individual genomic positions in background regions, which add up when summed over large window sizes. We think the most likely explanation is that these errors reflect differences in the way ChIP-seq and PRO-seq assays capture background. In ChIP-seq, there is substantial background pulldown of DNA due to non-specific binding of DNA to beads, tubes, tips or other sources of contamination. This background signal varies across the genome due to a variety of technical and biological factors (e.g., mappability; copy number alterations in some cell lines; etc.). In PRO-seq, the background is generally much lower and distributed in a very different way than ChIP-seq. Differences in the background distribution reflect fundamental differences in the assays: PRO-seq signal is derived from RNA (ChIP-seq is DNA), and PRO-seq has less background signal because the assay has three affinity purification steps (ChIP-seq has one). Differences in the background distribution between assays make it more difficult for dHIT to predict the number of ChIP-seq reads in regions that do not have much signal, and this is especially notable in larger window sizes where small differences add up to a larger aggregate signal.

- **Regions of focal amplification:** We have noted systematic error arising from regions with very high copy number in cells with abnormal karyotypes (e.g., K562). We attribute this difference to increased DNA content causing a large increase in ChIP-seq background and signal, but is more difficult to detect using PRO-seq. A great example of this error type is the first part of chromosome 21 in K562 cells: this region is amplified in K562 (based on the amount of background signal in ChIP-seq data), but the increased signal is not identified in the imputation. This error type has a fairly large effect on many of the quality control metrics in cell types with abnormal karyotypes.

- **Differences in the distribution of ChIP-seq signal within peaks:** Although dHIT captured most of the variation within experimental ChIP-seq peaks, the imputed signal was often spread over larger regions than experimental signal, as noted by the reviewer (see especially **Fig. 1C, Supplementary Fig. 4, Supplementary Fig. 7**). This may indicate either systematic biological variation in the distance between Pol II and marked nucleosomes or it may reflect uncertainty in the model due to noise in either PRO-seq or ChIP-seq assays.

- **Clear disagreement between imputed and experimental data:** We identified a handful of cases where there are clearly defined peaks in the imputed histone modification data, but not the experimental data (or vice versa). Many of these examples are located in intergenic regions, and cannot be explained by signal in gene bodies or other adjacent regulatory elements. An outstanding example of this type of error at the *CERK* promoter is shown in **Supplementary Fig. 12**.

  We were initially very excited about manipulations we could perform on these candidate differences. However, when we used ChIP-seq to examine H3K27ac in our own cell stocks (which were more closely matched to those used for PRO-seq), we found that nearly all of the aberrant/ missing peaks in the imputation were found in our own ChIP-seq data. This finding indicates that systematic biological differences in expression between different stocks of K562 cells, environmental differences, or other biological factors explain most cases where there is a clear disagreement between the imputation and experimental signal.

Also I'm a little concerned about the fact that they claim that poorly performing windows are enriched for CTCF when they use CTCF sites in their MNase normalization strategy. Doesn't the use of CTCF peaks to identify untranscribed regions make it impossible for them to then claim that that CTCF windows are the most poorly predicted?

Response: We have clarified in the revised manuscript that we only use CTCF sites to normalize new experimental MNase ChIP-seq data in the triptolide time course.

We do not use CTCF to normalize ChIP-seq signal prior to training or benchmarking dHIT for either chromatin accessibility or histone modifications. In all cases, imputation models were trained using library depth normalized read counts provided by ENCODE. To clarify this point in the revised manuscript, we have changed the section titles throughout the Methods section. For example, see Methods, section titled "*Data processing for newly collected MNase ChIP-seq, CUT&RUN, ATAC-seq, and PRO-seq*", which emphasizes that the section applies to newly collected data only.

Additionally, we have added new language to the Methods section which describes training dHIT (see the subsection titled *Training dataset*) to clarify that we used counts normalized only for sequencing depth for all dHIT training and evaluation tasks:

> "All training and validation analyses used sequencing depth normalized read counts, where possible using bigWig or bedGraph files provided by the original authors as input."

(*) Abstract argues that their accuracy is on par with replicates but they never explicitly look at replicate variability.

Response: We show a comparison of the accuracy of biological replicates and dHIT imputation in the revised **Supplementary Fig. 2** and **Supplementary Fig. 11**. These plots show that dHIT imputation is frequently within the range of values observed between pairs of experimental datasets for multiple performance metrics (Pearson and Spearman's correlation, mean absolute deviation, and peak identification). For some marks that we can impute with particularly high accuracy, dHIT estimates are on the high end of the range of correlations produced from experimental datasets, especially for small window sizes (10-100 bp). We have emphasized this analysis in the revised methods section of the manuscript.

(*) Their approach for training used a lot of heuristics (how they defined informative positions, how they pseudo-randomly select examples). Is this done merely to balance positive/negative regions, to reduce the overall burden of training data, because it worked, or what? Minimal justification is provided for what is actually a ton of heuristics. Leaves the reader wondering why a lot and whether these heuristics drastically influence the results. I'm fine (in general) with tuning a model as long as the methodology behind such tuning is conveyed with reasoning to the reader, especially in the scenario where the model isn't being built to generalize but rather to prove a point.

Response: In response to this comment, we have expanded our description of the selection of heuristics in the revised methods section. Briefly, most of the heuristics used in our present manuscript were systematically optimized for the classification of transcription initiation regions (TIRs) with dREG (Danko et. al. (2015), Nature Methods; Wang et. al. (2019) Genome Research). Since the imputation of histone modifications relies on signals in the PRO-seq data that are similar to those used by dREG, we used the values that were optimal for dREG without modification. Below we include a brief description for the benefit of the reviewer:

- The number of windows, window sizes, and the data transformation strategy were optimized for the classification of TIRs using dREG over a grid of reasonable values (See especially: **Supplementary Table 2** of Danko et. al. (2015) Nature Methods). These values were not changed in the present manuscript.

- Heuristics to identify "informative positions", which can intuitively be thought of as genomic positions with evidence of transcription nearby, were chosen by Danko et. al. (2015). In this analysis, Danko et. al. (2015) selected informative positions as a way to optimize the tradeoff between the number of positions analyzed and the fraction of real TIRs that were scored. We (previously) reasoned that the optimal heuristics would minimize the number of sites that we would have to score (to improve run time), but would score at least one site near every TIRs. We defined the sensitivity for TIRs as the fraction of all GRO-cap peaks (extended by 500 bp and merged) that we recovered. We optimized these values using a K562 dataset with ~40M mapped reads (a subsampled version of the deeply sequenced K562 dataset, G1, see **Supplemental Table 1**), because this is a reasonably standard sequencing depth for a PRO-seq library for which we would like to achieve a high sensitivity for TIR discovery. As above, we used the heuristic values selected for dREG without modification in dHIT.

- The main modifications we made to heuristics in the new manuscript were related to the composition of the training dataset. Unlike dREG, we trained dHIT to impute every 10 bp regardless of whether there was evidence of PRO-seq data in the surrounding region or not (i.e., whether that 10bp window is an 'informative position'). This change required dHIT to estimate the signal density in background regions, and therefore we included non-informative positions into our training dataset. Since we assume that the background does not have much signal (and therefore, not much variation in the PRO-seq data), these can be a small portion of our training sample. We selected 2% of training data to be in this group. We selected 5% to be near GRO-cap transcription start sites and left the remaining 93% of training examples to be informative positions that have PRO-seq signal nearby, but are not found in peaks. Notably, these values were selected in large part because we have previously noticed that using an unbalanced set of positive and negative sites performed best for dREG (see: Wang et. al. (2019) Genome Research). We did analyze different compositions of the training dataset in preliminary testing, but did not notice much difference in performance unless the training dataset wildly underrepresented informative positions.

Finally, we wish to emphasize that all of the heuristics noted above were optimized and fixed using an analysis of K562 cells before examining the holdout test datasets in GM12878 and other cell types. Thus, we believe the performance in GM12878 and other cell types represents a bona-fide out-of-sample test.

We have clarified all of this information in the Methods section, especially Methods, under the section *SVR feature vector*:

> We have previously optimized the number of window sizes and the window sizes for optimal classification of TIRs using dREG [53,87]. Since the imputation of histone modifications uses signals in the PRO-seq data that are similar to dREG, we used the values that were optimal for dREG without modification. Like for dREG, we passed data from windows at multiple size scales, including 10, 25, 50, 500, and 5,000 bp windows (n = 10, 10, 30, 20, and 20 windows, respectively), representing read data as far as 100 KB from the genomic region in question. PRO-seq data was standardized across each length scale in a similar fashion as we use for dREG [87], using a logistic function, $F(t)$, to transform raw read counts using two free parameters, α and β.

And Methods, under the section *Selecting training positions:*

> We defined regions of potential PRO-seq signal, which we call "informative positions" using the same heuristics we described previously for dREG[87]. Each window was defined as an "informative position" when the window had more than 3 reads within 100 bp on the single strand or at least one read within 1000 bp on both the positive and negative strands. These heuristics were selected as a way to optimize the tradeoff between the number of positions analyzed and the fraction of real TIRs that were scored based on the overlap with GRO-cap peaks.

(*) Availability on the software isn't specified.

Response: We have specified where to obtain dHIT software and the pre-trained models in the revised manuscript (see Methods, Training dHIT SVRs to predict histone marks using PRO-seq, GRO-seq or ChRO-seq data, Overview).

The software and analysis scripts are all publicly available on GitHub, under this repository: https://github.com/Danko-Lab/histone-mark-imputation. We have implemented relatively simple R commands to either use the models we trained, or to use the basic framework that we developed to train a new model. Basic documentation for using the dHIT R package is provided on that GitHub page as well.

Minor issues:

(*) The H3K27me3 discrepancy is quite interesting. They clearly identify distinct patterns, but do not clearly show that their approach could predict the more punctate pattern. So the “both of which appear linked to features of active transcription” is an over statement. Would be interesting to follow this mark through differentiation. May speak to roles of H3K27me3 in differentiation?

Response: We agree that we do not show the punctate pattern can be predicted using transcription. In the revised manuscript we have removed the sentence indicated by the reviewer.

We also agree that it would be interesting to follow the mark through differentiation. We have added new analyses to confirm that the punctate pattern is found primarily in ESCs and IPS cells, while the dispersed pattern is found in somatic, fully differentiated cell types and adult tissues. Partially differentiated cell types appear to have an intermediate pattern that lies somewhere between the punctate and broad patterns. The new analysis, shown in the revised **Supplementary Figure 14**, further supports the reviewer’s proposal that H3K27me3 distribution may have a role during differentiation.

(*) By reference they refer to a collection of K562 PRO and assorted ChIP marks from ENCODE, but this paper would be more cohesive if they included some assurance on the quality of these datasets. Alternatively it might be relevant to know whether quality of the data influences the ability of the model to identify these patterns?

Response: We have examined the quality metrics of both PRO-seq and ChIP-seq data in additional detail in the revised manuscript. For PRO-seq data, we used PEPPRO, a QC pipeline by the Gurtin and Sheffield labs (Smith et. al. (2021) Genome Biology) to obtain several quality control parameters for each dataset. These are presented in the revised **Supplementary Table 3**. For ChIP-seq data, all of our main analyses in ENCODE cell lines use datasets that passed ENCODE 2 data quality standards (Landt et. al. (2012) Genome Research). We have updated the Methods section to state how our quality control measurements were completed. Finally, we comment in the results section that we do indeed find, as expected, that poor quality data tends to have lower correspondence between imputed and experimental data (see the Results section, titled "*Active histone modifications have a similar relationship to transcription across mammalian cells*"):

> Lower correlations were generally observed when the experimental ChIP-seq data (certain CD4+ T-cell datasets) or the PRO-seq data (e.g., HeLa) had fewer sequenced reads or lower values in other data quality metrics (Supplementary Table 3).

(*) Does the registry of the non-overlapping windows matter in their ChIP predictions? Windowing methods always have edge effects — which are admittedly minimized at smaller window sizes. But at the larger sizes (500, 5kb) does an offset in the window start size relabel any regions? Likewise I don't recall them mentioning how they combine windows yet the evaluation windows are quite large (10kb and 1kb) compared to the 10 bp windows on which they impute/predict.

Response: In most cases, our metrics comparing experimental and imputed data use continuous error values, rather than picking thresholds to make peak calls. We think this strategy focusing on continuous error values is likely to be slightly more robust to edge effects than alternatives that use peaks. Additionally, to provide context on the best accuracy that we could theoretically achieve, we compare dHIT accuracy metrics to those observed between experimental datasets. In these analyses we always use the same windows, in the same register, and therefore any edge effects will affect both the imputation and experimental comparisons.

Finally, during the revision we added new analyses designed to estimate the accuracy of peak calling at the suggestion of Reviewer #1 (see comment #1). We do think it is likely that edge effects decrease the perceived performance in this task. In an effort to avoid the influence of edge effects in this analysis, we have excluded windows which lie on the edge of a peak from analysis. Likewise, we also compare the performance of dHIT with experimental data using the same windows.

Thus, although we do agree with the reviewer that edge effects could affect our performance metrics, especially when we are using larger windows, we think the revised manuscript provides enough information for readers to draw conclusions about accuracy.

(*) I have a general issue with saying PRO-seq throughout the paper when reality is this is applicable to several run-on /nascent assays. Perhaps a more general term is preferred? In methods section they say "PRO-seq, GRO-seq, or ChROseq; henceforth referred to simply as PRO-seq" — and this would be fine if this statement had been in the main text.

Response: To avoid confusion, we included a statement in the main text suggested by the reviewer. See especially the results section:

“dHIT uses the distribution of RNA polymerase, measured using either of the related methodologies PRO-seq, GRO-seq, or ChRO-seq data (henceforth referred to simply as PRO-seq), to impute the level of histone modifications genome-wide. “

(*) It is unclear why they are using hg19 when hg38 has been available since December of 2013. Justification of hg19 is warranted. However, it is unlikely that a shift to hg38 will alter any results, so this is perceived as a minor concern.

Response: We have added a justification for using hg19 in the Methods section.

Like so many others, our choice to work in hg19 coordinates was one of convenience. When we began this project ~4 years ago, all of the ENCODE data used in training and evaluating models was available in hg19, but not hg38, coordinates. Switching to hg38 at that time would have required that we re-analyze the ENCODE data ourselves (which would have been a huge effort), or use liftOver to convert across assemblies (which we do not believe to be an ideal solution). Likewise, we also had all of the PRO-seq data mapped in hg19 coordinates already. Finally, as the reviewer notes, we did not believe that updating to hg38 would affect the results in any meaningful way. Therefore, we made the decision to work in hg19 because we did not believe the amount of time it would take to update the coordinate system would result in any real benefit.

(*) The sentence, “Signal on the lower end was better spread out using data imputed from PRO-seq, possibly making use of the greater dynamic range of PRO-seq over ChIP-seq.” is completely opaque. What does “lower end” here refer to explicitly — ChIP or PRO? Are they trying to say that PRO predicted a broader dynamic range then ChIP? If so, it’s possibly true but somewhat dismissive of one of the error types seen by the model.

Response: We agree with the reviewer and have revised the indicated sentence.

(*) Text refers to SFig 1a-b but there are no a/b labels on SFig 1.

Response: We have fixed the caption to **Supplementary Fig. 1**.

(*) “As TRP does not affect engaged RNA polymerase, we observe a clearing wave of Pol II ~100kb from the TSSs on long genes at 1 h (Fig 4B)…” But this is not shown in Fig 4B.

Response: We thank the reviewer for catching this! We fixed the figure labels in the main text.

(*) Does “local environment” actually just mean TFs? “Our analysis supports a model in which both chromatin accessibility and the local environment are important factors to facilitate transcription initiation by Pol II”

Response: We have changed the term “local environment”, which we agree was unnecessarily vague, to explicitly name transcription factors, PIC machinery, and chromatin remodelers as important determinants of transcription initiation by Pol II. The revised sentence reads:

> “Our analysis supports a model in which both chromatin accessibility and other aspects of the local chromatin environment, including transcription factors, pre-initiation complex machinery, chromatin remodelers, and other transcription related proteins, are all necessary to facilitate transcription initiation by Pol II. “

(*) On SFig 11 some of the labels were overlapping and hard to read.

Response: We enlarged and reorganized the labels on this (and other) figures to make them easier to read.

(*) Figure SFig 12 is completely uninterpretable other than to look impressive. Can these be systematically classified as wide and punctate? Can the two types be shown as distinct figures to make each tract more readable? Fewer tracts? Meta-genes across multiple cell lines?

Response: We agree with the reviewer that a systematic characterization of punctate and broad H3K27me3 patterns would improve the paper. In response, we have replaced the old **Supplementary Figure 12** with a new figure depicting the systematic classification and analysis of ENCODE and Roadmap datasets.

We obtained data from 86 H3K27me3 datasets from the Roadmap Epigenome Project (Data sources listed in **Supplementary Table 4**). We classified these datasets as focused or dispersed using a principal component analysis, designed to separate samples based on their focal enrichment on principle component 1 (PC1) (**Supplementary Figure 14A**). To examine our hypothesis that undifferentiated cell types tend to have the focal pattern of H3K27me3, we classified the 86 cell types into five classes based on the cell or tissue of origin used in the H3K27me3 ChIP-seq experiment: Primary/Adult, Fetal tissue, Multipotent, Pluripotent, and Other cell types. We asked whether there were significant differences in PC1 score between cell types using a two-sided Wilcoxon rank sum test: Pluripotent (punctate H3K27me3) *versus* Primary/Adult (dispersed H3K27me3) (**Supplementary Figure 14C**). The revised manuscript describes these analyses in the results and methods sections.

(*) Color schemes on many of the figures is pseudo-random. For example, H3K36me3 is shown in Figure 1 and SFig17 as a mid-tone green. But in Figure 3 and SFig 15 it’s a light green and H3K4me1 is the mid-tone green. A similar green is used in the correlation grids to mean Pearson's but then Spearman's and the jsd values are both purple. So every figure panel required the reader to figure out what the color scheme was now. Admittedly, they are showing a tremendous amount of data across the figures and the bouncing around of colors is likely, to some extent, unavoidable — but every effort to make colors and symbols standardized throughout the paper would help the reader with the disorienting nature of having to figure out how to interpret every figure with its own unique color scheme.

Response: We have changed colors in the revised figures throughout the manuscript. We think our revisions do a better job of matching colors across the paper.

(*) Also — on these measures of correlation/similarity there is an alpha value being used to scale the "heat maps" but no key is given.

Response: We have added a color key to the figures that include heat maps.

Reviewer #3:

Remarks to the Author:

This article from Wang et al. describes a computational framework they call dHIT that uses machine learning to impute histone modification landscapes across the genome using nascent RNA data such as PRO-seq. The model was extensively trained and tested in a number of cell types and across species and appears to do a reasonable job of predicting the locations and levels of a number of histone marks that are often used to define chromatin state. In the cell type used for training, Pearson's correlation between prediction and actual data range from mediocre (0.37 for H3K27me3) to quite good (~0.7 for H3K27ac, H3K36me3, H3K9ac, and H3K4me2/ me3). The imputed locations are considerably more 'smeary' than real data and lack the positional information on nucleosomes gleaned from experimental data, but define general regions where histones are likely to bear a certain modification. Weaknesses of the model are in predicting heterochromatin and repressed Polycomb regions, ZNF genes and repeat elements.

Together, these findings indicate that nascent transcription data can be used to predict areas of activity in the genome, such as active promoters and enhancers, and the histone marks associated with activity (acetylation, H3 K4methylation, H3K36 methylation). Whereas this is probably not surprising, I appreciate the point being made here, which is that PRO-seq or a related nascent RNA assay is a much more efficient way to characterize a cell type than is 20 ChIP-seq assays.

The ability of nascent transcription levels to predict active histone marks also supports the growing body of data showing that these histone marks reflect transcription, rather than dictating or regulating transcription. Previous work from the Spicuglia lab (numerous papers, should be cited) Lis lab (Core et al), Adelman lab (Henriques et al, 2018, should be cited) have shown correlation between levels of active histone marks like H3K4me3 and transcription activity at both promoters and enhancers. The current work extends these studies, going beyond correlation to determine causality, using transcription inhibition with Triptolide. As predicted based on prior work in yeast (Howe lab; Martin et al., cited), loss of transcription causes loss of active histone marks H3K27ac and H3K4me3. Interestingly, the H3K36me3 mark and H3K4me1 turnover more slowly and are not as temporally dependent on transcription. This is a nice set of experiments that will hopefully help drive home the point that histone modifications aren't directive for activity, nor do they bookmark regions for future activity.

Response: We appreciate the reviewer's supportive and highly constructive summary of our manuscript. We have added additional citations to the work described by the reviewer, especially work from Spicuglia and Adleman labs (as well as additional references by Dowell).

Oddly, after providing some of the cleanest evidence yet that histone modifications, in particular H3K27ac and H3K4 methylation, reflect transcription rather than controlling it, the authors then delve into a section wherein they investigate "whether each histone modification facilitates either initiation or pause release". This section of the manuscript is very weak, and I remain unconvinced by these simulations that histone modifications 'facilitate' either initiation or pause release. I strongly suggest that this section of the manuscript be dramatically strengthened or (preferably) removed.

Response: We agree with the reviewer's points - we have removed the aforementioned section, and the accompanying figure, from the revised manuscript.

Finally, the authors work to demonstrate that not all accessible chromatin regions are sites of transcription initiation. This too has already been described in the literature, and it is known that DNase or ATAC-seq accessible sites include CTCF-bound loop anchors that are not transcriptionally active (Higgs, Buenrostro). However, this is probably the clearest description of this finding that I know of, and I appreciate the bigger commentary the authors are trying to make. However, Figure 6 is currently such a jumble of small panels that the main point does not come across clearly. I recommend that Figure 5 be removed and the authors use this space to expand Figure 6. This would allow them to better document the absence of transcription initiation at accessible regions, showing heatmaps and larger figures that bring this point home more clearly.

Overall, dHIT seems like a powerful tool and the take home message that histone modifications are not directive for transcription, but instead reflect transcription activity is important. However, the model does have weaknesses that should be acknowledged, and I have several specific concerns, as outlined below:

Response: We want to thank the reviewer for their thoughtful comments. Comments left by this reviewer were essential in the reorganization of our paper and our figures. Moreover, they helped us design experiments to understand the relationship between remodelers and the removal of active histone modifications from chromatin after Triptolide treatments.

Major concerns:

1) dHIT doesn't perform nearly as well at predicting regions of gene inactivity or repressed chromatin domains. This is perhaps not surprising, since it is based on nascent RNA sequencing. It would be helpful if the authors could comment on which histone ChIP-seq assays might complement PRO-seq and dHIT to give this fuller picture of chromatin? Could one do PRO-seq/dHIT and H3K9me3 plus H3K27me3 ChIP-seq to achieve this? This manuscript would be stronger if the authors could provide some insights into which repressive marks one should investigate by ChIP-seq to get a comprehensive picture of the chromatin landscape.

Response: We have added text to the discussion section about which additional molecular assays would provide a complementary view on genome function.

Briefly, we think that the most important mark to add is probably H3K9me3, because there are no circumstances in which that mark is accurately predicted using transcription. For H3K27me3, the importance of measuring this mark depends on the quality of information about cell state: Fully differentiated, somatic cell types can be predicted with surprisingly high accuracy using dHIT (see our response to reviewer #1 comment #4 for a summary of the evidence on this point). For instance, we achieve reasonably good predictions for horse liver, K562, GM12878, CD4+ T-cells, and others. However, because H3K27me3 is distributed in two very distinct patterns in different cell types, and additionally because we cannot currently predict the patterns of H3K27me3 in embryonic or partially differentiated cell types, experimental measurements of K27me3 provide new information in many projects. Finally, in our view, it is also useful to measure chromatin accessibility, because not all open chromatin regions are actively transcribed.

This information was added to the discussion (see section titled *dHIT: A powerful tool for genome annotation*), which reads as follows:

> Chromatin state annotations made using PRO-seq data could provide an efficient path to genome annotation, especially when complemented by experimental data for which dHIT provides incomplete or inaccurate information (e.g., H3K9me3, H3K27me3, and open chromatin). Our view is that, depending on the problem at hand, dHIT/ PRO-seq would be complemented best by the addition of experimental H3K9me3 (which we are not able to predict at all), followed by ATAC-seq (which adds the position of candidate insulators) and H3K27me3.

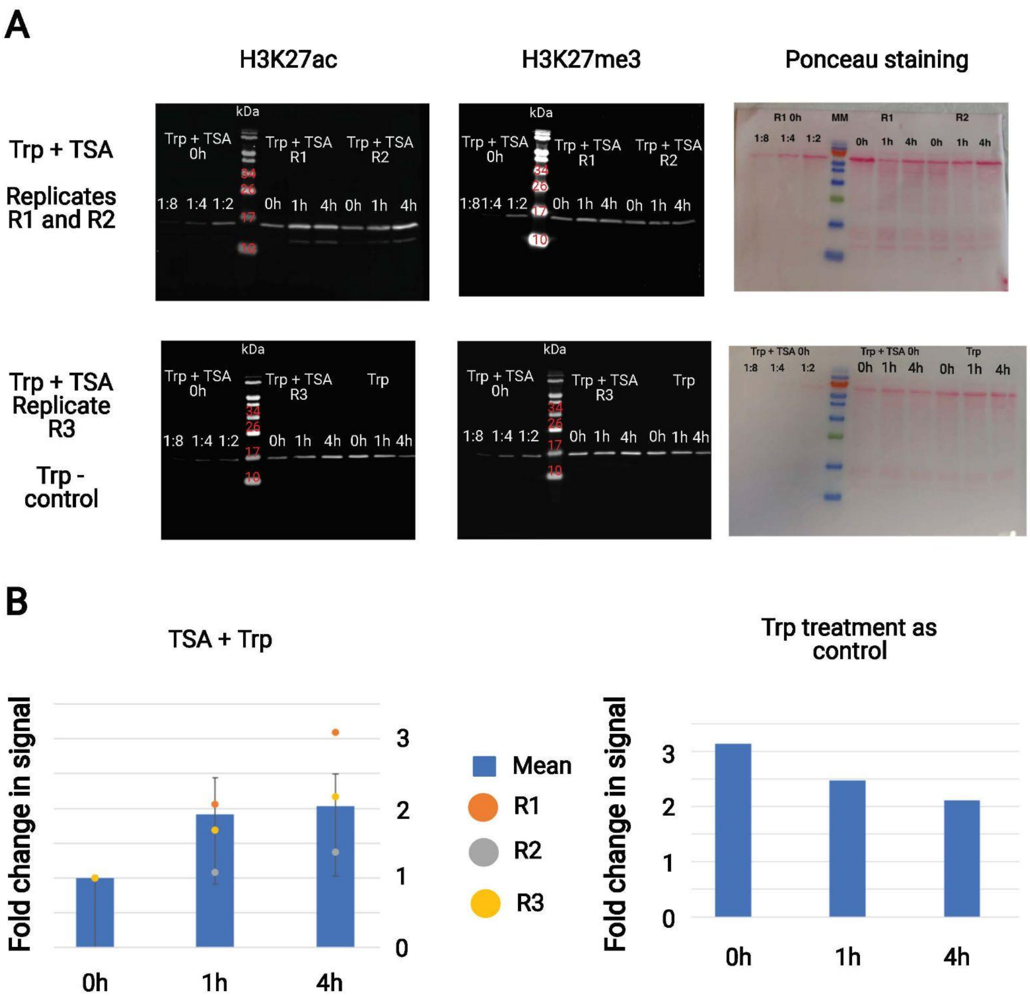
2) The authors observe a rapid loss of H3K27ac and H3K4me3 upon inhibition of transcription. Is this due to rapid deacetylation/ demethylation or histone turnover? They appear to argue for deacetylation rather than turnover, but this is not demonstrated. I suggest that the authors perform a simple assay to test this, using deacetylase inhibitors in Triptolide treated cells to confirm that acetylation is retained under these conditions. Whereas ChIP-seq would be optimal here, even western blots would help make this argument more compelling. This small experiment could go a long way to develop the model for how transcription stimulates deposition or retention of active chromatin marks.

Response: We performed the experiment suggested by the reviewer to test whether rapid deacetylation or histone turnover best explains the rapid loss of H3K27ac.

We treated cells with a mixture of 500nM triptolide and 250nM trichostatin A (TSA), a deacetylase inhibitor, or (to reproduce our earlier findings) 500nM triptolide alone. First, to alleviate concerns raised by reviewer #2, we confirmed that neither drug (or their combination) was cytotoxic at the concentration and time points used in these experiments (**Supplementary Figure 24**).

We then measured the loss in H3K27ac after 1h and 4h of treatment by Western blotting. In the presence of triptolide alone, we found that H3K27ac was depleted from chromatin after both 1 and 4 hours of treatment, providing additional replication for the observations made in our original manuscript. In cells that were treated with both triptolide and TSA, H3K27ac was not removed - if anything the amount of H3K27ac on chromatin increased following 1-4 hours of triptolide + TSA (**Supplementary Figure 23**). The increase in H3K27ac likely reflects histone acetyltransferases continuing to modify histones in a manner that is independent of Pol II (perhaps directed by transcription factors, as observed for HSF1 by (Vihervaara et al. 2017) and other authors). In contrast, H3K27me3, a negative control, showed no evidence of changes across the time course (**Supplementary Figure 23**).

These results support a model in which histone deacetylation is responsible for the rapid loss in H3K27ac observed after the addition of triptolide. We have added this information into the main text and supplement. The main Supplementary Figure supporting these findings is reproduced below:



Supplementary Figure 23. Western blots of H3K27ac after Triptolide (Trp) and Trichostatin A (TSA) treatment.
(A) Each western blot depicts the abundance of chromatin bound H3K27ac or H3K27me3 during the indicated incubation time point of Triptolide, or Triptolide and Trichostatin dual treatment. Each blot represents a different experiment. A dilution series of the untreated samples was used as standard curve to quantify changes in signal. Ponceau staining of membranes imaged are also depicted as total protein loading control.
(B) Quantification of H3K27ac/H3K27me3 signals of the western blot in (A). H3K27me3 was used as loading control.

3) Figure 5 shows a number of correlations between histone acetylation or methylation and simulations of initiation and pause release. These are nice correlations but don't speak in a clear way to function, and thus the conclusions such as 'methylation works at the stage of transcription initiation' appear unfounded. To support these comments, the authors could treat cells with inhibitors of acetylation/ deacetylation or methylation/ demethylation, or work in cells with catalytically inactive methyltransferases. In these conditions, one could test the authors conjecture that methylation or acetylation directly 'work' at a specific step in the transcription cycle. Such concrete experiments testing the simulation would be required to support the authors conclusions about function.

Response: We agree with the reviewer that the evidence that methylation/ acetylation contributes to either initiation or elongation is not very strong at this point. Following this reviewer's earlier suggestion, we removed the figure which shows the simulation studies (originally Figure 5) from the revised manuscript. Instead, we used the additional space to break Fig. 6 into separate figures which more clearly depict the relationship between transcription, chromatin accessibility, and histone H3 deposition.

Minor comments:

1) The jumbled and small nature of many figure panels makes this manuscript more difficult to read than optimal.

Response: To address the reviewer's comment, we expanded some of the main figures and moved multiple panels to the supplement. We have also increased the font size on many of the main and supplementary panels.

2) Figure 4B. The butterfly D. iulia doesn't look like the picture shown. That appears to be a monarch?

Response: We updated the butterfly picture in Fig. 4 to *D. iulia*.

natureresearch

Decision Letter, first revision:

Our ref: NG-A56764R

27th Oct 2021

Dear Charles,

Thank you for submitting your revised manuscript "Interdependence between histone marks and steps in Pol II transcription" (NG-A56764R). It has now been seen by the original referees and their comments are below. The reviewers find that the paper has improved in revision, and therefore we'll be happy in principle to publish it in Nature Genetics, pending minor revisions to comply with our editorial and formatting guidelines.

Since the current version of your manuscript is in a PDF format, please email us (genetics@us.nature.com) a copy of the file in an editable format (Microsoft Word) - we cannot proceed with PDFs at this stage.

We will then be performing detailed checks on your paper and will send you a checklist detailing our editorial and formatting requirements soon afterwards. Please do not upload the final materials and make any revisions until you receive this additional information from us.

Thank you again for your interest in Nature Genetics. Please do not hesitate to contact me if you have any questions.

Congratulations!

Sincerely,

Tiago

Tiago Faial, PhD
Senior Editor
Nature Genetics
https://orcid.org/0000-0003-0864-1200

Reviewer #1 (Remarks to the Author):

The authors have addressed all of my previous comments, and the manuscript is substantially improved. I apologize for having incorrectly assumed that Jaccard was being used as a distance metric (rather than similarity) in my previous review.

Reviewer #2 (Remarks to the Author):

The authors have addressed my concerns. The revised manuscript is clear, concise, and tells a cohesive story around the relationship between RNA pol II and histone marks. The inclusion of Supplementary Note on the typical errors of dHIT is particularly appreciated.

Minor issues:
(*) The section "Chromatin accessibility at transcription start sites does not depend on transcription" refers to Figure 6K-L whereas Figure 6 only has panels A-E.
(*) Figure 5 H&I are somewhat redundant with Figure 6 A&B. Perhaps the Figure 5 panels could be moved to the supplement in order to enlarge the remainder of Figure 5.
(*) I would hope that Tables 1-3 would be provided in a computer-readable (i.e. parse-able) format such as a tab-delimited file or similar. Printed tables (as presented in the supplement) are far less useful.

Reviewer #3 (Remarks to the Author):

The authors have done a great job addressing my concerns. I fully support publication.
This manuscript has a take-home message that is timely and important to the field. I will start teaching this paper in my graduate course as soon as it is published!

Author Rebuttal, first revision:

Responses to reviewers:

Minor issues:
(*) The section "Chromatin accessibility at transcription start sites does not depend on transcription" refers to Figure 6K-L whereas Figure 6 only has panels A-E.

Response: Fixed!

(*) Figure 5 H&I are somewhat redundant with Figure 6 A&B. Perhaps the Figure 5 panels could be moved to the supplement in order to enlarge the remainder of Figure 5.

Response: To address this comment, we have enlarged the font in Fig. 5 so that it can be read and interpreted more easily.

We agree with the reviewer that it would be very nice to enlarge the remainder of Figure 5. However, the revised supplement is packed with large numbers of figure panels, and we are not sure that it is the best solution to move panels there.

We also believe that the content of Fig. 5 H&I is important to show readers. The primary goal of Fig. 5 H&I is to show that changes in chromatin accessibility and H3 composition following Trp (which are also shown in Fig. 6) occur specifically at promoter and enhancer regions, but not at CTCF-bound DNase-I accessible sites. We feel that there is value in showing that changes are specific to transcribed regions.

With all of this said, if the editorial staff have suggestions on where to show these panels, or other suggestions on figure panel arrangement, we are very happy to make additional changes!

(*) I would hope that Tables 1-3 would be provided in a computer-readable (i.e. parse-able) format such as a tab-delimited file or similar. Printed tables (as presented in the supplement) are far less useful.

Response: We agree. Supplementary Tables 1-3 are provided in Excel format in our final submission. Both Excel and CSV files can also be found in our GitHub repo, here: https://github.com/alexachivu/dHITpaper_2021

Final Decision Letter:

In reply please quote: NG-A56764R1 Danko

24th Jan 2022

Dear Charles,

I am delighted to say that your manuscript, entitled "Prediction of histone post-translational modification patterns based on nascent transcription data", has been accepted for publication in an upcoming issue of Nature Genetics.

Over the next few weeks, your paper will be copy-edited to ensure that it conforms to Nature Genetics style. Once your paper is typeset, you will receive an email with a link to choose the appropriate publishing options for your paper and our Author Services team will be in touch regarding any additional information that may be required.

After the grant of rights is completed, you will receive a link to your electronic proof via email with a request to make any corrections within 48 hours. If, when you receive your proof, you cannot meet this deadline, please inform us at rjsproduction@springernature.com immediately.

You will not receive your proofs until the publishing agreement has been received through our system.

Due to the importance of these deadlines, we ask that you please let us know now whether you will be difficult to contact over the next month. If this is the case, we ask you provide us with the contact information (email, phone and fax) of someone who will be able to check the proofs on your behalf, and who will be available to address any last-minute problems.

Your paper will be published online after we receive your corrections and will appear in print in the next available issue. You can find out your date of online publication by contacting the Nature Press Office (press@nature.com) after sending your e-proof corrections. Now is the time to inform your Public Relations or Press Office about your paper, as they might be interested in promoting its publication. This will allow them time to prepare an accurate and satisfactory press release. Include your manuscript tracking number (NG-A56764R1) and the name of the journal, which they will need when they contact our Press Office.

Before your paper is published online, we shall be distributing a press release to news organizations worldwide, which may very well include details of your work. We are happy for your institution or funding agency to prepare its own press release, but it must mention the embargo date and Nature Genetics. Our Press Office may contact you closer to the time of publication, but if you or your Press Office have any inquiries in the meantime, please contact press@nature.com.

Acceptance is conditional on the data in the manuscript not being published elsewhere, or announced in the print or electronic media, until the embargo/publication date. These restrictions are not intended to deter you from presenting your data at academic meetings and conferences, but any enquiries from the media about papers not yet scheduled for publication should be referred to us.

Please note that <i>Nature Genetics</i> is a Transformative Journal (TJ). Authors may publish their research with us through the traditional subscription access route or make their paper immediately open access through payment of an article-processing charge (APC). Authors will not be required to make a final decision about access to their article until it has been accepted. Find out more about Transformative Journals

<B>Authors may need to take specific actions to achieve compliance with funder and institutional open access mandates. For submissions from January 2021, if your research is supported by a funder that requires immediate open access (e.g. according to Plan S principles) then you should select the gold OA route, and we will direct you to the compliant route where possible. For authors selecting the subscription publication route our standard licensing terms will need to be accepted, including our self-archiving policies. Those standard licensing terms will supersede any other terms that the author or any third party may assert apply to any version of the manuscript.

Please note that Nature Research offers an immediate open access option only for papers that were first submitted after 1 January, 2021.

If you have any questions about our publishing options, costs, Open Access requirements, or our legal forms, please contact ASJournals@springernature.com

If you have posted a preprint on any preprint server, please ensure that the preprint details are updated with a publication reference, including the DOI and a URL to the published version of the article on the journal website.

To assist our authors in disseminating their research to the broader community, our SharedIt initiative provides you with a unique shareable link that will allow anyone (with or without a subscription) to read the published article. Recipients of the link with a subscription will also be able to download and print the PDF.

As soon as your article is published, you will receive an automated email with your shareable link.

You can now use a single sign-on for all your accounts, view the status of all your manuscript submissions and reviews, access usage statistics for your published articles and download a record of your refereeing activity for the Nature journals.

An online order form for reprints of your paper is available at https://www.nature.com/reprints/author-reprints.html. Please let your coauthors and your institutions' public affairs office know that they are also welcome to order reprints by this method.

Sincerely,

naturerearch

Tiago

Tiago Faial, PhD
Senior Editor
Nature Genetics
https://orcid.org/0000-0003-0864-1200