THE EMBO JOURNAL

Genetic control of pluripotency epigenome determines differentiation bias in embryonic stem cells

Candice Byers, Catrina Spruce, Haley J. Fortin, Ellen I. Hartig, Anne Czechanski, Steven C. Munger, Laura G. Reinholdt, Daniel A. Skelly, Christopher L. Baker
DOI: 10.15252/embj.2021109445

Corresponding author: Christopher Baker (christopher.baker@jax.org)

Review Timeline:

Submission Date:	12th Aug 21
Editorial Decision:	8th Oct 21
Revision Received:	1st Nov 21
Accepted:	16th Nov 21

Editor: Daniel Klimmeck

Transaction Report:

(Note: Please note that the manuscript was previously reviewed at another journal and the reports were taken into account in the decision making process at The EMBO Journal. Since the original reviews are not subject to EMBO's transparent review process policy, the reports and author response cannot be published. With the exception of the correction of typographical or spelling errors that could be a source of ambiguity, letters and reports are not edited. Depending on transfer agreements, referee reports obtained elsewhere may or may not be included in this compilation. Referee reports are anonymous unless the Referee chooses to sign their reports.)

1st Editorial Decision **8th Oct 2021**

Dear Dr Baker,

Thank you again for the submission of your amended manuscript (EMBOJ-2021-109445) to The EMBO Journal. We have carefully assessed your manuscript and the point-by-point response provided to the referee concerns that were raised during review at a different journal. In addition, and as mentioned before, we decided to involve an arbitrating expert to evaluate the revised version of your work, with respect to technical robustness, conceptual advance and overall suitability of your work for publication in The EMBO Journal.

As you will see from the report provided below, the advisor is broadly in favour of the work stating the interest and value of your results and s/he is supportive of publication at The EMBO Journal. S/he also points to a number of minor amendments and experiments to complement the work and better distinguish it from the related, recently published studies.

We have discussed all those points carefully in the team and concluded that we are overall positive on the study, however, agree with the advisor that a more detailed presentation of the findings and revised discussion will be helpful to make this study amenable for The EMBO Journal at this stage. Also, the additional controls mentioned should be considered.

Based on the overall positive expert's view together with our own assessment, we decided to proceed with publication of your work at The EMBO Journal pending the above points related to the advisor's input could be conclusively addressed in a time frame of two weeks.

Once we have received the revised version, we should then be able to swiftly proceed with formal acceptance and expedited production of the manuscript.

Please submit a revised version of the manuscript using the link enclosed below, addressing the advisor's comments.

Further, I will share additional changes and comments from our production team during the next days to be considered.

As you might have seen on our web page, every paper at the EMBO Journal now includes a 'Synopsis', displayed on the html and freely accessible to all readers. The synopsis includes a 'model' figure as well as 2-5 one-short-sentence bullet points that summarize the article. I would appreciate if you could provide this figure and the bullet points.

Please contact me if you have any questions related, and note, that we would in principle like to proceed with this article as soon as possible.

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

Again, please contact me at any time if you need any help or have further questions.

Kind regards,

Daniel Klimmeck

Daniel Klimmeck PhD
Senior Editor
The EMBO Journal.

Arbitrating advisor's comments:

I read the manuscript. I will discuss mainly Figure5 and Figure 6 as these two figures contain new findings absent in previous publications (Skelly et al., 2020 and Ortmann et al., 2020). Overall, the approach is elegant, and the results are well-presented. I also appreciate that the author's computational analyses include both unbiased aspects (e.g.,

Figure 5A, 5C, Figure 6E) of the data and focused details on specific genomic loci.

In my view, the work can be positively considered at EMBO Journal. It is exciting to see the potential link between TRIM28 (at target gene regions) and KRAB-ZFP (as a trans-QTL expressing diffusible factor) that are involved in trans-GTL-gene regulation. Validation analysis is performed only with Chr4 KO and its targets, including genes in Chr18. However, based on the Figure 5D result, I expect that TRIM28 and KRAB-ZFP would be involved in another trans-QTL-gene regulation associated with chr5, 7, 12, and 13 as well. I understand that the authors saved a strong claim for future study, but they could write this link clearer in the manuscript.

Regarding clearer writing, describing the results of Figure 6 could be better. For instance, "overlaps" (p16, line 357) can be better explained with scheme and numbers; "is bound by" (p16, line 363) is confusing; "cis-eQTL" (p17, line 365) seems trans-eQTL?

I also feel that the concept of trans-QTL could be better explained by adding a schematic model to the figure. For example, adding such a schematic for explaining Figure 5B as well as Figure 6E would be very helpful to conceptualize the author's idea/findings and deliver them to the readers.
Lastly, it will be nice to add more "control" analysis related to Figure 6H. For example, what happens to Chr12 QTL targets? Chr7 QTL targets? And other QTL targets? Are those gene expressions unchanged in Chr4 QTL KO?

****

Formatting changes required for the revised version of the manuscript:

>> Please add up to five keywords to your study.

>> Introduce ORCID IDs for the corresponding author (C.B.) via our online manuscript system. Please see below for additional information.

>> Please add a separate 'Statistical analysis' section to your manuscript, detailing the algorithms applied.

>> Introduce http links to the GEO database entries in the 'Data availability section'. Please make the data processing script code publicly available via github or similar.

>> Please specify individual author contributions for E.H. .

>> Provide all main and EV figures as individual high resolution .tiff files. When submitting the figures as individual files, the legends need to be removed from the figure files and added as a Figure Legends section to the manuscript file. Up to five supplemental figures can be made EV figures, with their legends in the manuscript after the main figure legends.

>> Appendix File with ToC: The two remaining supplementary figures should be added to appendix file, which is to be saved as a PDF with a ToC. Please change the nomenclature to 'Appendix Figure S1, S2...' and adjust references in the main text and legends.

>> There are 14 tables. Tables 1-11 should be renamed "Dataset EV1" etc. Tables 12-14 should be renamed "Table EV1" etc. Callouts need to be updated, all files need titles and legends added in a separate tab.

>> Rename the current 'Competing Interests' section to 'Conflicts of Interest'.

>> Please provide a filled author checklist for your study.

>> Recheck callouts and their correct order in the main text for Supplemental Figure 5.

>> The reference format needs to be corrected to EMBO Journal style and 10 author names before et al. .

****

Please note that as of January 2016, our new EMBO Press policy asks for corresponding authors to link to their ORCID iDs. You can read about the change under "Authorship Guidelines" in the Guide to Authors here: http://emboj.embopress.org/authorguide

In order to link your ORCID iD to your account in our manuscript tracking system, please do the following:

1. Click the 'Modify Profile' link at the bottom of your homepage in our system.
2. On the next page you will see a box half-way down the page titled ORCID*. Below this box is red text reading 'To Register/Link to ORCID, click here'. Please follow that link: you will be taken to ORCID where you can log in to your account (or create an account if you don't have one)
3. You will then be asked to authorise Wiley to access your ORCID information. Once you have approved the linking, you will be brought back to our manuscript system.

We regret that we cannot do this linking on your behalf for security reasons. We also cannot add your ORCID iD number manually to our system because there is no way for us to authenticate this iD number with ORCID.

Thank you very much in advance.

****

General instructions for preparing your revised manuscript:

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

2) individual production quality figure files as .eps, .tif, .jpg (one file per figure).

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

4) a complete author checklist, which you can download from our author guidelines (https://wol-prod-cdn.literatumonline.com/pb-assets/embo-site/Author Checklist%20-%20EMBO%20J-1561436015657.xlsx). Please insert information in the checklist that is also reflected in the manuscript. The completed author checklist will also be part of the RPF.

5) Please note that all corresponding authors are required to supply an ORCID ID for their name upon submission of a revised manuscript.

6) It is mandatory to include a 'Data Availability' section after the Materials and Methods. Before submitting your revision, primary datasets produced in this study need to be deposited in an appropriate public database, and the accession numbers and database listed under 'Data Availability'. Please remember to provide a reviewer password if the datasets are not yet public (see https://www.embopress.org/page/journal/14602075/authorguide#datadeposition).
In case you have no data that requires deposition in a public database, please state so in this section. Note that the Data Availability Section is restricted to new primary data that are part of this study.
*** Note - All links should resolve to a page where the data can be accessed. ***

7) Our journal encourages inclusion of *data citations in the reference list* to directly cite datasets that were re-used and obtained from public databases. Data citations in the article text are distinct from normal bibliographical citations and should directly link to the database records from which the data can be accessed. In the main text, data citations are formatted as follows: "Data ref: Smith et al, 2001" or "Data ref: NCBI Sequence Read Archive PRJNA342805, 2017". In the Reference list, data citations must be labeled with "[DATASET]". A data reference must provide the database name, accession

number/identifiers and a resolvable link to the landing page from which the data can be accessed at the end of the reference. Further instructions are available at .

8) We would also encourage you to include the source data for figure panels that show essential data. Numerical data can be provided as individual .xls or .csv files (including a tab describing the data). For 'blots' or microscopy, uncropped images should be submitted (using a zip archive or a single pdf per main figure if multiple images need to be supplied for one panel). Additional information on source data and instruction on how to label the files are available at .

9) We replaced Supplementary Information with Expanded View (EV) Figures and Tables that are collapsible/expandable online (see examples in https://www.embopress.org/doi/10.15252/embj.201695874). A maximum of 5 EV Figures can be typeset. EV Figures should be cited as 'Figure EV1, Figure EV2" etc. in the text and their respective legends should be included in the main text after the legends of regular figures.

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

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

10) When assembling figures, please refer to our figure preparation guideline in order to ensure proper formatting and readability in print as well as on screen:
http://bit.ly/EMBOPressFigurePreparationGuideline

Please remember: Digital image enhancement is acceptable practice, as long as it accurately represents the original data and conforms to community standards. If a figure has been subjected to significant electronic manipulation, this must be noted in the figure legend or in the 'Materials and Methods' section. The editors reserve the right to request original versions of figures and the original images that were used to assemble the figure.

11) For data quantification: please specify the name of the statistical test used to generate error bars and P values, the number (n) of independent experiments (specify technical or biological replicates) underlying each data point and the test used to calculate p-values in each figure legend. The figure legends should contain a basic description of n, P and the test applied. Graphs must include a description of the bars and the error bars (s.d., s.e.m.).

Further information is available in our Guide to Authors: https://www.embopress.org/page/journal/14602075/authorguide

The revision must be submitted online within 90 days; please click on the link below to submit the revision online before 6th Jan 2022.

Link Not Available

I read the manuscript. I will discuss mainly Figure5 and Figure 6 as these two figures contain new findings absent in previous publications (Skelly et al., 2020 and Ortmann et al., 2020). Overall, the approach is elegant, and the results are well-presented. I also appreciate that the author's computational analyses include both unbiased aspects (e.g., Figure 5A, 5C, Figure 6E) of the data and focused details on specific genomic loci.

In my view, the work can be positively considered at EMBO Journal. It is exciting to see the potential link between TRIM28 (at target gene regions) and KRAB-ZFP (as a trans-QTL expressing diffusible factor) that are involved in trans-GTL-gene regulation. Validation analysis is performed only with Chr4 KO and its targets, including genes in Chr18. However, based on the Figure 5D result, I expect that TRIM28 and KRAB-ZFP would be involved in another trans-QTL-gene regulation associated with chr5, 7, 12, and 13 as well. I understand that the authors saved a strong claim for future study, but they could write this link clearer in the manuscript.

Indeed, our current hypothesis is that these KZFP clusters are driving the differences in chromatin and gene expression we see at all or most of QTL hotspots identified here, and potentially developmental phenotypes broadly between mouse strains. As the manuscript is written, there is a full paragraph in the discussion dedicated to highlighting the connection between developmental phenotypes, location of our molecular and other physiological QTL, and location and implication of KZFPs. This paragraph explicitly states in part, “Further, putative regulatory elements targeted by all six QTL hotspots were enriched for binding by TRIM28 (Fig. 5D), which is recruited to chromatin through interaction with KZFPs (Friedman et al., 1996). Additionally, the effect of the QTL was found to be dominantly repressive in F1 hybrids, consistent with the function of KZFP/TRIM28 complexes formation of heterochromatin in trans. Notably, a single KZFP contained within the Chr 13 QTL hotspot interval was shown to be causal in the progression of a lupus phenotype (Treger et al., 2019). And while the other studies outlined above largely have not pinpointed causal factors, the overlapping molecular and physiological QTL harbor clusters of newly emergent murine KZFPs (Bruno et al., 2019; Kauzlaric et al., 2017). This provides exciting future work into assigning causality to a rapidly evolving gene family whose divergence in different strain backgrounds may account for evolution of regulatory function that shapes development and disease (Elmer & Ferguson-Smith, 2020).” We have updated the topical sentence of this paragraph to say, “Several lines of evidence support that the QTL discovered in this study are of significant developmental importance and are driven by variable KZFPs.” Plus the subheading for the final figure (Figure 6) is titled, “KRAB zinc-finger proteins are implicated as trans acting factors underlying QTL”. Therefore, we feel we have fairly strongly stated the link in the manuscript, without overstating what is currently largely correlative evidence (with possible exception of Chr 4).

Regarding clearer writing, describing the results of Figure 6 could be better. For instance, "overlaps" (p16, line 357) can be better explained with scheme and numbers; "is bound by" (p16, line 363) is confusing;

We have attempted to clarify our language through these sections, specifically addressing the two points above. We are not entirely sure what additional numbers to provide that were not already in the manuscript. We stated the coordinates and size of the QTL interval (Chr 4: 143,302,047-148,864,661, 5.6 Mb), the coordinates and size of the targeted genomic deletion (Chr 4: 145,383,917-147,853,435, 2.47 Mb), and the number of KZFP encoding genes that were deleted (21).

"cis-eQTL" (p17, line 365) seems trans-eQTL?

Here we did indeed mean cis-QTL; however, based on the reviewers question we have now tried to clarify the importance and meaning of this distinction in the manuscript and moved this observation

towards the end of the paragraph to try to capture how these genes could be causal mediators of the QTL. The expression of the genes encoding KZFPs at the location where the QTL maps to (ie. Chr4) are regulated through local variation that impacts their expression in a manner suggestive that they may be the mediators of the distal changes in chromatin accessibility (i.e. trans-QTL). This is because they are higher expressed when the QTL haplotype on Chr 4 is B6, which coincides with higher H3K9me3 and lower target gene expression. As suggested by the reviewer, a schematic model may help visualize this molecular chain of causation and is now included as Figure 6I.

I also feel that the concept of trans-QTL could be better explained by adding a schematic model to the figure. For example, adding such a schematic for explaining Figure 5B as well as Figure 6E would be very helpful to conceptualize the author's idea/findings and deliver them to the readers.

Thank you for this suggestion, we have now added several models to better conceptualize the genetic regulation computationally summarized in Figures 5C and 6I.

Lastly, it will be nice to add more "control" analysis related to Figure 6H. For example, what happens to Chr12 QTL targets? Chr7 QTL targets? And other QTL targets? Are those gene expressions unchanged in Chr4 QTL KO?

To test specificity of Chr 4 gene regulation we have extended our enrichment analysis of gene set overlap to all QTL as controls. This found that the Chr 4 QTL targets were enriched in the genes that are differentially expressed in between the WT and Chr 4 KO mESCs. This has been added to the Results on page 18.

Dear Dr Baker,

Thank you for submitting the revised version of your manuscript. I have now evaluated your amended manuscript and concluded that the remaining minor concerns have been sufficiently addressed.

Thus, I am pleased to inform you that your manuscript has been accepted for publication in the EMBO Journal.

Please note that it is EMBO Journal policy for the transcript of the editorial process (containing referee reports and your response letter) to be published as an online supplement to each paper.

Also, in case you might NOT want the transparent process file published at all, you will also need to inform us via email immediately. More information is available here: http://emboj.embopress.org/about#Transparent_Process

---

Please note that in order to be able to start the production process, our publisher will need and contact you regarding the following forms:

- PAGE CHARGE AUTHORISATION (For Articles and Resources)
http://onlinelibrary.wiley.com/journal/10.1002/(ISSN)1460-2075/homepage/tej_apc.pdf

- LICENCE TO PUBLISH (for non-Open Access)

Your article cannot be published until the publisher has received the appropriate signed license agreement. Once your article has been received by Wiley for production you will receive an email from Wiley's Author Services system, which will ask you to log in and will present them with the appropriate license for completion.

- LICENCE TO PUBLISH for OPEN ACCESS papers

Authors of accepted peer-reviewed original research articles may choose to pay a fee in order for their published article to be made freely accessible to all online immediately upon publication. The EMBO Open fee is fixed at $5,200 (+ VAT where applicable).

We offer two licenses for Open Access papers, CC-BY and CC-BY-NC-ND.
For more information on these licenses, please visit: http://creativecommons.org/licenses/by/3.0/ and
http://creativecommons.org/licenses/by-nc-nd/3.0/deed.en_US

- PAYMENT FOR OPEN ACCESS papers

You also need to complete our payment system for Open Access articles. Please follow this link and select EMBO Journal from the drop-down list and then complete the payment process: https://authorservices.wiley.com/bauthor/onlineopen_order.asp

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

On a different note, I would like to alert you that EMBO Press is currently developing a new format for a video-synopsis of work published with us, which essentially is a short, author-generated film explaining the core findings in hand drawings, and, as we believe, can be very useful to increase visibility of the work. This has proven to offer a nice opportunity for exposure i.p. for the first author(s) of the study. Please see the following link for representative examples and their integration into the article web page:
https://www.embopress.org/video_synopses
https://www.embopress.org/doi/full/10.15252/embj.2019103932

Please let me know, should you be interested to engage in commissioning a similar video synopsis for your work. According operation instructions are available and intuitive.

Finally, we have noted that the submitted version of your article is also posted on the preprint platform bioRxiv. We would appreciate if you could alert bioRxiv on the acceptance of this manuscript at The EMBO Journal in order to allow for an update

of the entry status. Thank you in advance!

If you have any questions, please do not hesitate to call or email the Editorial Office.

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

Kind regards,

Daniel Klimmeck

Daniel Klimmeck, PhD
Senior Editor
The EMBO Journal
EMBO
Postfach 1022-40
Meyerhofstrasse 1
D-69117 Heidelberg
contact@embojournal.org
Submit at: http://emboj.msubmit.net

EMBO PRESS

YOU MUST COMPLETE ALL CELLS WITH A PINK BACKGROUND ↓

PLEASE NOTE THAT THIS CHECKLIST WILL BE PUBLISHED ALONGSIDE YOUR PAPER

Corresponding Author Name: Christopher L Baker
Journal Submitted to: EMBO Journal
Manuscript Number: EMBOJ-2021-109445R

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

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

A- Figures

1. Data

The data shown in figures should satisfy the following conditions:

- the data were obtained and processed according to the field's best practice and are presented to reflect the results of the experiments in an accurate and unbiased manner.
- figure panels include only data points, measurements or observations that can be compared to each other in a scientifically meaningful way.
- graphs include clearly labeled error bars for independent experiments and sample sizes. Unless justified, error bars should not be shown for technical replicates.
- if n< 5, the individual data points from each experiment should be plotted and any statistical test employed should be justified
- Source Data should be included to report the data underlying graphs. Please follow the guidelines set out in the author ship guidelines on Data Presentation.

2. Captions

Each figure caption should contain the following information, for each panel where they are relevant:

- a specification of the experimental system investigated (eg cell line, species name).
- the assay(s) and method(s) used to carry out the reported observations and measurements
- an explicit mention of the biological and chemical entity(ies) that are being measured.
- an explicit mention of the biological and chemical entity(ies) that are altered/varied/perturbed in a controlled manner.

- the exact sample size (n) for each experimental group/condition, given as a number, not a range;
- a description of the sample collection allowing the reader to understand whether the samples represent technical or biological replicates (including how many animals, litters, cultures, etc.).
- a statement of how many times the experiment shown was independently replicated in the laboratory.
- definitions of statistical methods and measures:
  - common tests, such as t-test (please specify whether paired vs. unpaired), simple $\chi 2$ tests, Wilcoxon and Mann-Whitney tests, can be unambiguously identified by name only, but more complex techniques should be described in the methods section;
  - are tests one-sided or two-sided?
  - are there adjustments for multiple comparisons?
  - exact statistical test results, e.g., P values = x but not P values < x;
  - definition of 'center values' as median or average;
  - definition of error bars as s.d. or s.e.m.

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

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

B- Statistics and general methods

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

1.a. How was the sample size chosen to ensure adequate power to detect a pre-specified effect size?	No sample size calculation was performed; biological replicates and triplicates were shown sufficient to capture variability by PCA. For genetic mapping, 33 BXD lines were deemed sufficienct to map molecular traits (i.e. gene expression by RNA-seq or chromatin accessibility by ATAC-seq) based on prior experience (Baker et al. Genetics. 2019).
1.b. For animal studies, include a statement about sample size estimate even if no statistical methods were used.	N/A
2. Describe inclusion/exclusion criteria if samples or animals were excluded from the analysis. Were the criteria pre-established?	No data were excluded
3. Were any steps taken to minimize the effects of subjective bias when allocating animals/samples to treatment (e.g. randomization procedure)? If yes, please describe.	No
For animal studies, include a statement about randomization even if no randomization was used.	N/A
4.a. Were any steps taken to minimize the effects of subjective bias during group allocation or/and when assessing results (e.g. blinding of the investigator)? If yes please describe.	No
4.b. For animal studies, include a statement about blinding even if no blinding was done	N/A
5. For every figure, are statistical tests justified as appropriate?	Yes
Do the data meet the assumptions of the tests (e.g., normal distribution)? Describe any methods used to assess it.	For data presented in Figure panels 3F, 5F, and 5I t-tests normality was expected based on type of experimental measurement. For data presented in Figure panel 6E normality was visuallized using QQ plots and non-parametric test was applied.
Is there an estimate of variation within each group of data?	Individual replicates are indicated in all relevant figure panels along with mean +/- standard error of the mean.
Is the variance similar between the groups that are being statistically compared?	For data presented in Figure panels 3F, 5F, and 5I t-tests were performed with unequal variance.

C- Reagents

USEFUL LINKS FOR COMPLETING THIS FORM

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

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

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

http://datadryad.org

http://figshare.com

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

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

http://biomodels.net/

http://biomodels.net/miriam/
http://jij.biochem.sun.ac.za
https://osp.od.nih.gov/biosafety-biosecurity-and-emerging-biotechnology/
http://www.selectagents.gov/

6. To show that antibodies were profiled for use in the system under study (assay and species), provide a citation, catalog number and/or clone number, supplementary information or reference to an antibody validation profile. e.g., Antibodypedia (see link list at top right), 1DegreeBio (see link list at top right).	Chromatin immunoprecipitation was performed using antibodies against P300 (12 µl, Bethyl A300-358A), POU5F1 (20 µl, Cell Signaling Tech Oct-4A rabbit mAb, 5677S), and TRIM28 (10 µl, Abcam 201C, ab22553). Primary antibodies used for FACS analysis were α-EpCAM (dilution: 1/10,000, abcam ab71916) and α-SOX1 (dilution: 1/300, R&D Systems AF3369).
7. Identify the source of cell lines and report if they were recently authenticated (e.g., by STR profiling) and tested for mycoplasma contamination.	All embryonic stem cells were derived for use in this study and tested negative for mycoplasma.

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

D- Animal Models

8. Report species, strain, gender, age of animals and genetic modification status where applicable. Please detail housing and husbandry conditions and the source of animals.	All mice used to derive embryonic stem cells were obtained from The Jackson Laboratory (Bar Harbor, ME) including C57BL/6J (stock number 000664), DBA/2J (stock number 100006), and BXD recombinant inbred lines (see Extended Data Table 12).
9. For experiments involving live vertebrates, include a statement of compliance with ethical regulations and identify the committee(s) approving the experiments.	All animal experiments were approved by the Animal Care and Use Committee of The Jackson Laboratory (summary #04008; PI: Christopher L Baker).
10. We recommend consulting the ARRIVE guidelines (see link list at top right) (PLoS Biol. 8(6), e1000412, 2010) to ensure that other relevant aspects of animal studies are adequately reported. See author guidelines, under 'Reporting Guidelines'. See also: NIH (see link list at top right) and MRC (see link list at top right) recommendations. Please confirm compliance.	I confirm compliance with the published guidelines for all mouse experiments.

E- Human Subjects

11. Identify the committee(s) approving the study protocol.	N/A
12. Include a statement confirming that informed consent was obtained from all subjects and that the experiments conformed to the principles set out in the WMA Declaration of Helsinki and the Department of Health and Human Services Belmont Report.	N/A
13. For publication of patient photos, include a statement confirming that consent to publish was obtained.	N/A
14. Report any restrictions on the availability (and/or on the use) of human data or samples.	N/A
15. Report the clinical trial registration number (at ClinicalTrials.gov or equivalent), where applicable.	N/A
16. For phase II and III randomized controlled trials, please refer to the CONSORT flow diagram (see link list at top right) and submit the CONSORT checklist (see link list at top right) with your submission. See author guidelines, under 'Reporting Guidelines'. Please confirm you have submitted this list.	N/A
17. For tumor marker prognostic studies, we recommend that you follow the REMARK reporting guidelines (see link list at top right). See author guidelines, under 'Reporting Guidelines'. Please confirm you have followed these guidelines.	N/A

F- Data Accessibility

18: Provide a "Data Availability" section at the end of the Materials & Methods, listing the accession codes for data generated in this study and deposited in a public database (e.g. RNA-Seq data: Gene Expression Omnibus GSE39462, Proteomics data: PRIDE PXD000208 etc.) Please refer to our author guidelines for 'Data Deposition'. Data deposition in a public repository is mandatory for: a. Protein, DNA and RNA sequences b. Macromolecular structures c. Crystallographic data for small molecules d. Functional genomics data e. Proteomics and molecular interactions	Original data discussed in this study have been deposited in NCBI's Gene Expression Omnibus and are accessible through GEO Series accession number GSE164935 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE164935). Along with raw sequencing data, processed data tables in the accession include normalized read counts for gene expression for ESC and EpiLC samples; normalized read counts along with peak intervals for chromatin accessibility and ChIP factor occupancy for ESC and EpiLC samples; matrices produced by CellRanger for expression libraries for single cell RNA-seq from EBs as well as single cell barcode table displaying EB samples associated with unique lipid-modified oligos after demultiplexing following MULTIseq pipeline. H3K4me3 ChIP-seq data for B6, D2, BXD75, and BXD87 were collected previously (Baker et al, 2019a) and are available through GEO accession GSE113192 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE113192). RNA-seq and histone modification ChIP-seq data for mESCs from Chr 4 knock-out and wild-type cells lines were published previously (Wolf et al, 2020) and are available through GEO accession GSE115291 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE115291). RNA-seq data for mESCs grown in different media formulations to access pluripotency spectrum were published previously (Hackett et al, 2017) and are available through GEO accession GSE98517 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE98517). Results in Figure 1 and EV1 for differential gene expression analysis is available as Dataset EV1. Results in Figure 2 and EV2 and Appendix Fig. S1 for defining gene modules is available as Datasets EV2 and EV3. Results in Figure 3 and EV3 for single cell RNA-seq and FACS analysis is available as Dataset EV4 and Table EV3. Results in Figure 4 and EV4 for QTL analysis is available as Datasets EV5-8. Results in Figure 5E for locus overlap enrichment analysis is available as Dataset EV10. Results in Figure 5F for ChIP-seq occupancy is available as Dataset EV11.
19. Deposition is strongly recommended for any datasets that are central and integral to the study; please consider the journal's data policy. If no structured public repository exists for a given data type, we encourage the provision of datasets in the manuscript as a Supplementary Document (see author guidelines under 'Expanded View' or in unstructured repositories such as Dryad (see link list at top right) or Figshare (see link list at top right).	Genomic data (RNA-, ATAC-, ChIP-seq) and single-cell RNA-seq have all been deposited in GEO as indicated above.
20. Access to human clinical and genomic datasets should be provided with as few restrictions as possible while respecting ethical obligations to the patients and relevant medical and legal issues. If practically possible and compatible with the individual consent agreement used in the study, such data should be deposited in one of the major public access-controlled repositories such as dbGAP (see link list at top right) or EGA (see link list at top right).	N/A
21. Computational models that are central and integral to a study should be shared without restrictions and provided in a machine-readable form. The relevant accession numbers or links should be provided. When possible, standardized format (SBML, CellML) should be used instead of scripts (e.g. MATLAB). Authors are strongly encouraged to follow the MIRIAM guidelines (see link list at top right) and deposit their model in a public database such as Biomodels (see link list at top right) or JWS Online (see link list at top right). If computer source code is provided with the paper, it should be deposited in a public repository or included in supplementary information.	N/A

G- Dual use research of concern

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