

Clonal tracing with somatic epimutations reveals dynamics of blood aging

Corresponding Author: Dr Lars Velten

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Attachments originally included by the reviewers as part of their assessment can be found at the end of this file.

Version 0:

Reviewer comments:

Referee #1

(Remarks to the Author)

In this manuscript, Scherer et al. report a novel method, called EPI-clone to perform clonal tracking using CpG methylation patterns as "native marks" which allow to quantify clonal relationships between cells at single cell resolution. This method has the main advantage of being transgene-free and can therefore be, in principle, applied to any sample of interest including human samples, without a labelling step, and also provide cell type resolution.

The method is highly novel and places itself as a competitor to clonal tracking with mtDNA mutations (also a transgene-free method). One of the strengths of the manuscript is that it benchmarks EPI-clone with "ground-truth" measurements with the well established experimental barcoding system LARRY, of which one of the corresponding authors previously developed. The authors then go on to apply EPI-clone to steady-state haematopoiesis in aged mice as well as in elderly humans.

Overall, I find the method of great interest to the community. The paper gives good credit to previous work and the discussion comparing the method to previous ones is very good and very helpful. The list of limitations is also helpful and transparent. However the novel findings reported with the method are currently fairly descriptive and only incremental compared to those reported previously in the literature (see more below). In addition, the paper is very difficult to read (even for a reviewer with some level of bioinformatic analysis and clonal tracking experience) and the description of the method and of the experiments performed is not clear at all.

Major comments:

1. Lack of clarity on experiments performed:

This paper reports a new transgene-free clonal tracking method and therefore even within the brief format provided by the journal it is key that the rationale of the method and the experiments that benchmark it to transgene-based are explained clearly, also to an audience not expert in clonal tracking. But that is not the case.

a. First and foremost, I could not find a description of how LARRY barcodes are integrated with scTAMseq, in the main text, in the methods or even in the github pages. Previous work using LARRY (Weinreb et al, Science, 2020 and Rodriguez-Fraticelli, Nature, 2020) has read-out LARRY barcodes via their transcription and scRNAseq. Instead, scTAMseq reads out DNA exclusively (at least according to the references given). The authors write in the methods (line 664) that "using the stained cells that we used as input for scTAM-seq (Figure S1b), we additionally performed 10X Genomics Chromium Single Cell 3' for transcriptomic profiling of the cells". I understand from this that the cells were stained, sorted and then this pool was split into 2, a portion for scTAM and the other for 10X Genomics. However if that was how LARRY barcodes were read, the same cells did not have both mRNA and DNA profiled, how can they be linking LARRY barcode data to the methylation UMAP as in Figure 1c?

Given that LARRY integrates into the DNA, my guess is that LARRY barcodes are also read through scTAMseq by looking at the DNA insertion. This way each cell would have methylation data and a LARRY barcode. From inspecting the code, this seems the most likely scenario. However no methods as to how LARRY amplicons may be generated is given. This is a key experiment that validates the whole procedure, but not enough information is provided as to how it was performed.

b. Figure legends are really minimal, methods are choppy and overall schemes of how each step of the experiments and each step of the analysis are integrated is missing. The authors should also improve consistency in naming of experiments of features throughout the manuscript. For example, they often use barcodes in a general sense, but they should be more precise (cellular barcodes, LARRY barcodes...). Another example is: The authors list all the experiments performed in mice in line 108 but then they do not use consistent naming for each of the experiments later on, so it is confusing. At line 141 they talk about "native hematopoiesis", I guess this is experimental batch iii) "wildtype mouse without LARRY barcodes"?

2. The method of defining static CpGs as part of EPI-clone is elegant and the overlap with LARRY barcoding convincing. The fact that cellular state can be inferred from the methylation data alone is also convincing. However the biological conclusions from the experiments on ageing (Figure 4 and 5) are somewhat underwhelming in terms of novelty as well as in terms of how the data is presented.

a. The main message from the mouse aging study is that with ageing there are more expanded clones (from 40% to 78% of cells) and these tend to be larger. This message is novel to my knowledge. However the authors should improve the robustness of this finding.

- Figure 4a: quantification of the total number of clones would be useful. Are these conclusions based on a low number of clones or not? Is this total number also similar to that of other methods (Rodriguez-Fraticelli, 2018)?

- Figure 4b, d, e and f: could the authors add statistics to each of these observations, given that the main novel conclusion of the paper is drawn from these?

- Line 237: "the large, low-output stem cells clones were mostly myeloid biased": is this qualitative statement only from 4i? Could the authors actually quantify this as again this is presented as a key finding?

- Importantly, while this reviewer appreciates the complexity of these experiments, the conclusions are drawn from one experiment with one young mouse and one old mouse. Is this really enough to make any conclusions? How variable may be the oligoclonality between old mice?

B. The authors also apply their method to investigate human aging. With this analysis they largely confirm a previous study reporting oligoclonality in individuals aged over 70 years old (Mitchell et al., Nature, 2022).

- Figure 5D: why were B and T cells removed from the analysis? Does EPI-clone work in these cells as well? Please comment in text or figure legend.

- Figure 5E: Is oligoclonality more marked at the CD34+ level than at the My level? What about lymphoid cells? This would be a novel finding not demonstrated in previous studies.

- The authors designed a panel allowing for the assessment of mutations in a quite large number of genomic regions (>140) usually associated with clonal hemopoiesis (CH). However no comment is made about these in the findings. Do expanded clones preferentially carry mutations in CH genes? Can the author address the relationship between individuals mutations and CH mutations? In other words, is EPI-clone more sensitive at finding clones than just using CH mutations?

3. EPI-clone seems to track clonal somatic epimutations that are stable over several months in mice, and likely many years in humans. The discussion has an interesting paragraph comparing EPI-clone with mtDNA tracking. Could the authors estimate the "timescale" of tracking of the two methods? Another unresolved question is how diverse is the functional behaviour of the clones expanded with age on shorter time scale (for example during differentiation). This is something that could be addressed by combining EPI-clone with shorter LARRY barcoding follow-ups.

Minor comments:

Line 141: the statement that clonal information was removed when three different experiments were integrated should be backed up by a figure.

Line 152: shouldn't this come earlier? This is how the clusters in Figure 2A have been annotated, correct? Very confusing

Line 217: it is not clear how was this done. Please also repeat which populations were sorted.

Figure 4H: it would be really useful to have the annotated map next to these clone examples for orientation

How are low-output and high-output defined? Is there a numeric criterium? Even if not, these should be explained in the main text, currently this is only relatively clear if one has read previous studies from the authors

Figure 4I: how do we distinguish low-output from high-output in these graphs? In the old mouse, is it the first split in the tree on the left that distinguishes them (low output on top, high output at the bottom).

(Remarks on code availability)

The authors have provided code in GitHub which is relatively clear and concise. While I did not try to run the code, it is useful that the authors provide the Seurat objects and overall it should be possible to reproduce the code. However there are no links provided to the raw data, which is not ideal. Most importantly the manuscript and/or the GitHub page both lack a schematic explanation of how the different pieces of data generated with which libraries are all integrated together. Consistency in terminology used in GitHub and in the methods of the manuscript could also be improved.

(Remarks to the Author)

The authors present EPI-Clone, which uses the scTAM-seq technology for lineage tracing purposes in addition to cell typing. The approach is based on the Mission Bio Tapestry system and provides amplicon readouts and methylation status of a pre-defined set of CpG sites which enable lineage assessment due to the stochastic changes at some of these sites, as well as general cell type due to the known status of the specific sites selected. While the technology itself is previously-described, the analysis and application is novel and serves as a powerful example of how DNA methylation can be leveraged for lineage tracing without the need for genetic manipulation. One notable weakness of the method is that a specific panel of CpG sites must be developed for any given target issue, limiting generalizability. The authors apply EPI-Clone to several systems, first validating alongside a LARRY genetic lineage tracing model, which is a key proof of principle followed by assessment in ageing mice and in human. The assessment of clonal lineages reveals the landscape within these systems which generally match with what is previously known, though approaching it from a new angle, and arguably providing improved power and resolution. That being said, there are not any big surprises or entirely novel findings in the analysis, which is largely descriptive. That being said, the method likely has the power to learn some really powerful biology, such as a large-scale analysis across many human samples as opposed to a vignette example.

On the technical side, the experiments and analysis are sound. Additional detail on the specific readout and how those data are leveraged would be helpful in the main text. The github vignette documentation is transparent and clear, detailing exactly what analysis was performed which is important and should be commended. The repository was very helpful in deciphering exactly what analysis was performed and aided in review of the manuscript. An archived version should accompany any final release of the work.

The flow of the paper itself could use some work – there are multiple instances where later sections are referenced, forcing the reader to jump around, sometimes even having to jump to a later section / figure then jump to another later one, to then get back to the earliest point. Structuring the work in a way that does not require flipping back and forth would be a huge improvement.

Minor: The authors state that single-cell DNA methylation is limited to a few hundred cells; however, datasets from the BRAIN Initiative are in the 100's of thousands. There is also a commercially-available kit for single-cell methylation that has a throughput in the 10's of thousands per experiment. Though notably, coverage sparsity is an issue for all of those methods.

It is a bit unclear in the main text results what the specific readout is from the assay. 453 CpG sites are targeted and assessed using a methyl-sensitive restriction enzyme as opposed to bisulfite methods. Is the readout presence/absence? Or is there data on the actual methylation state? If it is presence/absence then dropout would contribute to the absence readout. The analysis suggests this is not the case, so clarification without having to read the scTAM paper would benefit the manuscript greatly.

Is "uMAP" different from "UMAP"? "UMAP" is generally the standard term.

Figure 1C – label it "LARRY Lineage Barcode" to be more descriptive

Figure 4 – the vast majority of cells profiled fall into clones with <75 cells. Is this due to an inability to properly assign? A limitation of the cell numbers profiled? Or expected and the focus is just on expanded clones? It is hard to know if one could even ascertain this.

(Remarks on code availability)

The code was provided and is detailed.

Referee #3

(Remarks to the Author)

The paper "Somatic Epimutations Enable Single-Cell Lineage Tracing in Native Haematopoiesis Across the Murine and Human Lifespan" by Michael Scherer et al. introduces a method called EPI-clone to identify both clonal identities and cell differentiation states in haematopoiesis in humans and mice. EPI-clone is based on the scTAM-seq method (Agostina Bianchi, et al., *Genome Biol* 23, 229, 2022), which allows for the simultaneous measurement of single-cell DNA methylation profiling and surface-protein expression. In this study, the authors use a specially designed set of primers to amplify selected CpG sites that are differentially methylated during different stages of haematopoiesis. They also use a set of oligo-tagged antibodies to identify cell surface protein markers associated with cell differentiation.

By studying murine haematopoiesis, the authors demonstrated that EPI-Clone can identify CpGs correlated with differentiation and clonal identity, referred to as Dynamic and Static CpGs, respectively. The accuracy of Static CpGs is compared to the ground truth orthogonal LARRY lentiviral barcodes, and the accuracy of Dynamic CpGs is assessed using cell surface protein expression. Static CpGs show high correlations with the ground truth LARRY lentiviral barcodes in blood stem and progenitor cell fate in transplantation experiments and slightly weaker correlations with mature immune cells. The applicability of this method is also demonstrated in aged vs young mice, and aged human male donors supporting EPI-clone's utility in important areas of biology (ageing and clonal expansion). The authors discuss this method as an alternative to synthetic barcoding systems (which are not always feasible, such as in human research) and argue its superiority compared to other approaches based on mitochondrial DNA mutations and WGS.

While this study presents fascinating findings with the potential to advance synthetic barcode-free lineage tracing in human

and murine models, there are some limitations that currently hinder its broader applicability. Were they to be addressed, one envisages this technique could be widely adopted by the scientific community, progressing the analysis of lineage relationships of tissue development and homeostasis, cancer, inflammation and numerous other fields.

1. Applicability Beyond HSPCs: The study is mostly restricted to HSPCs, and its applicability for analysis of downstream haematopoietic cell types and other tissues/ cell types is not clear. Below is a list of several suggestions but not all need to be addressed. For example, the authors could either focus on EPI-clone's utility beyond haematopoiesis (i.e. other tissue types), or deeper within haematopoiesis (including capturing clonality within more mature lineages and young human donors)

a) The authors could demonstrate the applicability of EPI-Clone for at least one non-hematopoietic tissue in humans and mice. This would demonstrate the versatility and robustness of the method across different tissue types. This might also help determine if there are overlapping (universal?) static CpGs.

b) The authors could also test EPI-Clone in a young human donor/s, as it is currently tested only in two aged donors where clonal haematopoiesis/oligoclonality is a factor (i.e., the number of clones is very limited while the size of clones is quite large). This would provide a more comprehensive understanding of the method's applicability across different age groups, not only aged humans.

c) As mentioned, the study does not cover many mature blood cell types. The authors could consider the performance of EPI-Clone for clonal identification (i.e., static CpGs) for lymphocytes, where T or B Cell receptor (TCR/BCR) repertoire could additionally provide a ground truth barcode after activation/expansion. Also, a broader range of myeloid subtypes in addition to monocytes (e.g. neutrophils, eosinophils, dendritic cells etc).

d) Although not essential, for human studies, an analysis of EPI-clone's utility using barcoded human hematopoietic stem and progenitor cells (HSPCs) could be conducted to compare its performance with the ground truth barcodes (e.g., LARRY). We are also concerned about the quality/resolution of data afforded by their antibody panel (see 3a)

2. Benchmarking EPI-Clone's Performance:

a) The paper argues its superiority over other approaches based on mitochondrial DNA mutations. A systematic comparison with mtDNA-based approaches such as Weng, C., et al. (Nature 627, 389–398, 2024) would provide a clearer picture of the relative strengths and weaknesses of EPI-Clone.

b) Although not essential, studying EPI-clone in a native setting in mice using one of the available in-vivo barcoding mouse models instead of transplanting barcoded cells into irradiated mice would provide a better benchmark in case transplantation is amplifying the reliability of the method and would help validate the method in a more physiologically relevant setting. To this point, when we visually assessed EPI-clone signatures in the authors' data set, we noted that the Transplantation experiments give visually 'stronger' clonal information compared to cells from a native setting (attached).

3. Technological Improvements for parallel cell type/state identification:

The current resolution of cell state, based on dynamic methylation sites and/or antibodies is quite restrictive for broad take up. a) For antibodies, is this due to sequencing depth and/or too small a panel? Could the authors generate uMAPs only using cell-surface protein expression to demonstrate the efficiency of the panel used for cell identification, or use a more comprehensive panel (e.g. TOTAL-seq)? This would ensure comprehensive haematopoietic cell type coverage and improve the accuracy of cell state identification (related to Point 1).

b) Regardless of 3a) capturing transcriptomes and static CpGs in the same single cells would massively increase EPI-clone's utility for the scientific community. Although the authors have performed scRNA-seq on the same samples, correlating these profiles demonstrates a similarity in the uMAP landscapes, but not for the same cells. Including scRNA-seq on the same cells would be highly advantageous, providing a more comprehensive view of cellular states and enhancing the interpretability of the results. The authors could consider practical ways to achieve this, perhaps the recent preprint from the Landau lab, or other approaches.

4. Algorithm Improvements:

a) Statistical Power: In the EPI-Clone algorithm for computing the minimum p-value for association with any protein, a minimum sample size of 3 cells per group can limit the statistical power of the Kolmogorov-Smirnov Test (ks.test). Given the noisy nature of protein expression data, the authors could consider increasing the sample size. This might enhance the reliability of the statistical associations identified.

b) Minimum p-values: When analysing the associations between methylation status and protein expression, the algorithm runs a Ks.test for each protein to obtain p-values for the significance of the association. For each methylation site, the minimum p-value across all proteins is selected to identify the most significant association. However, it is important to note that when p-values are extremely small, the exact differences between them may not be practically significant, as they all indicate a strong association. Therefore, while the minimum p-value is used to highlight the most significant association, and authors used this strategy to summarise the data, analysing top-n p-values can also be considered for a more comprehensive analysis.

c) Tied-values: In the context of the KS test, having many tied values can affect the test's performance. The frequency of tied values could change by applying different normalization methods to the raw data. The authors need to inspect this to ensure the test's performance is not affected. If necessary, they should consider alternative statistical tests. This would ensure the robustness of the statistical analyses.

d) Data Normalization: The normalization method provided in the code (ScaleData from Seurat) is robust and widely used for single-cell data. However, alternatives like SCTransform from Seurat might offer better performance for dealing with noisy protein data (Hafemeister, C., Satija, R. Genome Biol 20, 296, 2019). This might improve the reliability of the results.

e) Subclonal structure: Did the authors attempt to resolve clonal/subclonal relationships within a data set? This would be advantageous for the study of phylogeny and e.g. cancer evolution in the future (especially where CNVs are not prevalent).

5. Blind Design: Since LARRY barcodes are used as the ground truth, the authors could conduct the analysis in a manner where EPI-Clones are blind to the LARRY barcodes. If this approach was used in the current study, please discuss it in detail. Otherwise, for future experiments, we suggest implementing a blind design (train and test) to prevent potential biases

and circular logic in the analysis. This would enhance the robustness and credibility of the study's findings.

6. LARRY Barcoding: In our exploration of the LARRY barcodes across four mice, we found that mouse 1 and mouse 2 exhibit extremely similar LARRY barcodes, as do mouse 3 and mouse 4. This similarity is quite unexpected considering barcodes should be random. The authors should provide detailed information on the viral transduction process, transduction efficiency, in vitro expansion, transplantation procedures, and the preprocessing steps used for the LARRY barcodes to understand the observed similarities and ensuring the reliability of the barcoding results.

7. QC and methylation properties: A useful addition would be to include some QC metrics about the molecular and computational aspects of EPI-clone. In particular, what fraction of all static DNAm sites is detected per cell using scTAM-seq, and in terms of performance in the LARRY experiment, what fraction of the total must be captured to gain 'meaningful' clonal information using EPI-clone. And on that note, are some sites more informative than others? Is there a minimal set of sites that captures most information? What is the proportion of cells with Epiclone marks? Perhaps not all cells make random errors during methylation (a rough estimate of probability per cell division or something like that would be useful, if possible). And are some static CpG sites more likely to mutate?

Minor Comments:

1. Clarify Method Development: The abstract's description, "We develop EPI-clone, a droplet-based method for transgene-free lineage tracing, ..." is ambiguous and might lead readers to assume that scTAM-seq was developed in this study. Please clarify in the abstract and throughout the article where EPI-Clone is introduced, specifying that EPI-Clone builds upon the existing scTAM-seq technology.

2. Same Sample for scRNA-seq: In line 146, "Since we also performed scRNA-seq on the same samples ..."). This would benefit from further disambiguation throughout the article to avoid confusion whether it was the same single cells or 'different cells from the same sample'.

3. Literature Citations: Line 46: Perhaps refer to VanHorn S, Morris SA. Next-Generation Lineage Tracing and Fate Mapping to Interrogate Development. Dev Cell. 2021, or similar, and consider a broader range of relevant studies rather than a focus on those of the authors only.

4. Figure 1C Clarification: Clarify how the uMAP of DNA methylation is highlighted by LARRY barcodes, especially given that these analyses were not performed on the same single cells.

5. Figure 1H Clarification: Line 130: The description of Static CpGs stochastically gaining or losing methylation (preferentially losing, as shown in Figure 1h) requires more clarification. The current explanation seems counterintuitive and confusing. Please provide a detailed explanation of the methodology and findings to ensure clarity.

6. Quantitative Comparisons: The reliance on visual inspection of uMAPs and identifying a "similar landscape" is prevalent in this article to demonstrate the efficiency of EPI-Clone in identifying cell lineage and state. Whenever possible, use quantitative comparisons, such as a confusion matrix, to show the rates of true positive and true negative classifications. This would provide a more rigorous validation of the method.

7. Figure 3E Axis Labels: For Figure 3E, why the Y-axis is not labelled as TPR (True Positive Rate)? In a standard ROC plot (TPR vs. FPR), a higher AUC indicates better performance.

8. Figure 3F Overlaps: Figure 3F shows some CpG clusters overlapping with more than one LARRY barcode. Considering this analysis was run only on large clones ($n > 30$ cells), it is worth investigating why EPI-Clone could not classify correctly in these scenarios. Examine whether this issue arises from the panel design, the algorithm, or other factors.

9. Minimal Panel Size: The authors could conduct additional analysis using the current data to determine the minimal size of the primer panel (for CpG sites) required to achieve similar resolution in cell lineage and cell state identification with EPI-Clone. This analysis will help identify the smallest set of CpG sites necessary for effective lineage tracing and state identification, potentially making the method easier to apply in various research settings. Demonstrating that EPI-Clone can maintain high accuracy with a reduced panel would significantly enhance its utility and accessibility.

10. uMAP Annotations: Wherever uMAP plots are used, please specify the type of data used to generate the plot (e.g., DNA methylation, cell surface protein expression, etc.). This will provide clarity and help readers understand the context and significance of each plot.

Version 1:

Reviewer comments:

Referee #1

(Remarks to the Author)

The manuscript by Scherer et al. has been revised with the addition of a large amount of data, which, in my opinion, cover all areas of improvement raised by the reviewers. The authors have made the manuscript substantially stronger and their new experiments now contribute to provide interesting novel insights into haematopoiesis, not reported before with other techniques.

Specifically, the authors:

- Expanded the study on primary human hematopoietic samples from 2 to 13 samples. This part brings very interesting new data that 1) further characterises clonal expansions without driver mutations from what previous reported; 2) links clones traced by EPI-clones to clonal hematopoiesis (CH) mutations.
- Included a replicate of their studies of mouse aging, performed transplantations and improved the quantification and statistical analysis of the mouse data.
- Provided a direct experimental comparison of EPI-Clones to mtDNA tracing, which is most useful to the field given recent

controversies on the mtDNA method.

- Combined EPI-Clones with single cell targeted RNAseq in the same cell, being able to provide transcriptional information on specific clones.
- Demonstrated the broader utility of their method by performing lineage tracing of endothelial cells.
- Substantially improved the clarity in which the method is reported in the main text and with much helpful graphical representations of their experiments (Fig.S1A).

Overall, I believe this method will be of great use to the lineage tracing community. In addition, they report novel findings of high relevance to the hematology field, namely: 1) in mice aged HSC-expanded clones contribute little to transplantation, with the bulk of hematopoietic reconstitution provided by non expanded clones; 2) in humans expanded clones can be observed from age 50 onwards, do correspond to clones carrying CH mutations but also include driver-less clones. CH clone tend to be larger, more enriched in HSC and more lineage biased than driver-less clones.

Minor:

Figures 4I-J: please add the number of observations per tertile and clarify in the figure legend that T1 is the tertile with lowest mature output (I) or with smaller immature clone size (J).

(Remarks on code availability)

I have reviewed the GitHub page, which seems complete and well laid out. In addition, I am pleased that the authors have deposited all datasets.

Referee #2

(Remarks to the Author)

The authors have addressed all of the minor comments, primarily pertaining to clarifications and additional detail in the manuscript. The authors also apply the method to an additional set of samples - 13 bone marrow samples.

The authors make the case for novel biological discovery and there is some there, but still completely descriptive. The characterization of these populations is a bit of stretch when assessing <500 CpG sites.

Overall, the manuscript is very well put together, has appropriate statistics and detail, and the method appears to be of high quality.

(Remarks on code availability)

Code is provided and well annotated.

Referee #3

(Remarks to the Author)

In this revision the authors have sufficiently addressed our experimental, computational and biological queries. In the spectrum of methods available that balance cost/throughput vs information content, epi-clone is demonstrably powerful in both uncoupling and separately reporting on clonal origin (using static CpGs) and lineage state (using dynamic CpGs, CITE-seq, and targeted mRNA expression). They have added both depth (more patients, CH mutation analysis, more mice, lymphoid cells) to the study of haematopoiesis and breadth (assessment of endothelial lineage clonal origins). The comparison with mtDNA variants shows the superiority of EPI-Clones in tracking human progenitors for in vivo lineage tracing. Furthermore, the presence of EPI-Clones in multiple tissues highlight the biological relevance and wide applicability of this methods. This reviewer thinks that EPI-Clone, and more in general the use of DNA methylation in human lineage tracing, represent a crucial advancement in the field and will be very useful to answer many research questions in the future. Below are minor suggestions/questions that, if addressed, we believe would enhance remaining queries about the study.

- Since submission of the revision, the MethylTree paper in Nature Methods proposes a new method to provide clonal and even subclonal resolution of cell relationships, including unbiased transcriptional state. Please briefly compare the method, for throughput, cost, pros vs cons in discussion.
- Please provide a version of heatmap (Reviewer Fig C: CpG sites vs single-cells matrix) sorted but not averaged per clone, in addition. This can be in SI. It would allow a visualisation of the raw data underlying single cell epi-clone features, which is missing to date – even if it doesn't look pretty.
- Fig 5C: Dnmt3a, the most frequently mutated gene in CH is also involved in CpG methylation. Is it conceivable that mutations in Dnmt3a lead therefore to particularly distinguishable Epiclone signatures (as suggested by UMAP clustering). Conversely could Epiclone signatures allow for biological insights into the epigenomic effects of Dnmt3a mutations at the clonal level? As a potential analysis to explore this, using the existing data, could the authors pool all cells from all Dnmt3a clones (in Fig 5), and create a clustered heatmap of Epiclone signatures (rows: cells, columns: CpG sites), where the rows are annotated/coloured by Dnmt3a mutation, and the columns are annotated/coloured whether the CpG site is a known Dnmt3a target site. ?
- Fig 6G: proportions of cells per clone positive for 7076A>G appear tri-modal, i.e., close to 0%, 50% and 100%. Especially 50% is intriguing. Do the authors have an hypothesis to explain this observation?
- Regarding Figure S12C-E:
The authors stated, "The DNA methylation amplicons achieved a somewhat weaker cell state resolution, but displayed less technical variation." Instead of this qualitative description, it would be more useful to include a quantitative metric, such as the Adjusted Rand Index (ARI), to compare resolution more clearly.

Shalin Naik, Tom Weber, Esmaeel Azadian, Sara Tomei

(Remarks on code availability)

We were able to reproduce all the figures, the code is well-written, and the instructions are clear enough for users to follow without issues.

Open Access This Peer Review File is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

In cases where reviewers are anonymous, credit should be given to 'Anonymous Referee' and the source.

The images or other third party material in this Peer Review File are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>

Referee #1 (Remarks to the Author):

In this manuscript, Scherer et al. report a novel method, called EPI-clone to perform clonal tracking using CpG methylation patterns as "native marks" which allow to quantify clonal relationships between cells at single cell resolution. This method has the main advantage of being transgene-free and can therefore be, in principle, applied to any sample of interest including human samples, without a labelling step, and also provide cell type resolution.

The method is highly novel and places itself as a competitor to clonal tracking with mtDNA mutations (also a transgene-free method). One of the strengths of the manuscript is that it benchmarks EPI-clone with "ground-truth" measurements with the well established experimental barcoding system LARRY, of which one of the corresponding authors previously developed. The authors then go on to apply EPI-clone to steady-state haematopoiesis in aged mice as well as in elderly humans.

Overall, I find the method of great interest to the community. The paper gives good credit to previous work and the discussion comparing the method to previous ones is very good and very helpful. The list of limitations is also helpful and transparent. However the novel findings reported with the method are currently fairly descriptive and only incremental compared to those reported previously in the literature (see more below). In addition, the paper is very difficult to read (even for a reviewer with some level of bioinformatic analysis and clonal tracking experience) and the description of the method and of the experiments performed is not clear at all.

We appreciate the reviewer's positive assessment of our work and the constructive criticism. We now worked on improving the overall flow of the manuscript. To that end, we have re-structured figure 1&2, we have included a new section, "methods summary", with the goal of making our methods description easier to follow, and we have also improved the description of the methods and experiments in the main results section.

Regarding the novelty of our biological findings, we have **substantially expanded** the human study: We now characterize bone marrow from 13 donors, including seven carriers of clonal hematopoiesis (CH) mutations (revised figure 5 and new figures 6, S12, S13). While the existence of clonal expansions without known driver mutation in bone marrow from elderly has been described previously, these clones have never been functionally characterized and compared to CH clones, due to a lack of suitable methods. EPI-clone enables this comparison and thereby allows us to put CH in perspective to clones without a known driver mutation ("non-CH"). We summarize the biological novelty of these findings as follows:

(Discussion, lines 459-468)

Our data further put CH mutations into a perspective with clonal expansions without known driver: CH clones are more strongly biased towards the myeloid lineage and towards an expansion of stem cells, but together with "non-CH" clones form part of a spectrum of age-related clonal expansions that display similar functional properties. In aged mice, we similarly detected large HSC-expanded clones that had reduced regenerative capacity. Together with recent transplantation studies of human HSCs⁶⁸ this suggests conservation of the processes that drive hematopoietic ageing and decline in clonal complexity, and highlights that CH mutations might not be the main driver of this process. Epidemiological studies found an increased mortality risk in carriers of driver-free expanded clones⁴⁵. These results call for a broader investigation of age-related decline in clonality instead of a strict focus on CH.

This finding is further in line with our main insight from the mouse study, that is, an age-related accumulation of clones with an expansion at the level of stem cells. For the latter observation, we now include biological replicates, and we provide, for the first time, a comparison of pre- and post-transplant hematopoiesis by tracking clones from the native setting through transplantation (revised Figure 4, new figure S11). We summarize the biological novelty of this finding as follows:

(Results, lines 303-308)

In summary our data demonstrate age-related loss of clonal complexity in murine aging that is accompanied by an emergence of HSC-expanded clones with low engraftment ability. We propose that these rare but expanded clones drive the increase in stem cell number and decrease in output that had typically been associated with aged hematopoiesis in transplantation^{15,59,60} and Cre-lox native lineage tracing studies⁶¹. Our transplant data support that HSCs that do not expand with age persist and drive regeneration.

We finally highlight that:

(Discussion, 448-457)

Our study is the first to demonstrate that both native human and murine hematopoiesis shifts from highly polyclonal to oligoclonal blood production, and it is the first study to investigate clone function in these two species using a coherent, unified method. Expanded clones in mouse tended to be more numerous, but individually smaller, and poorly contribute to hematopoiesis in transplants. This observation appears in line with a larger and more polyclonal stem cell compartment in human, but a much longer period of clonal selection and drift. In our human data, oligoclonal blood production becomes detectable at an age of ~50 years, and it manifests itself as an inevitable and potentially clock-like process after the age of 60.

Furthermore, we now include data on EPI-Clone applications to endothelial cells and mature immune cells (new figure 3g-j, S5-7). We demonstrate, from a technical perspective, that EPI-clone can be combined with targeted RNA-seq from the same single cell (Figure 6c,d, S14), and we provide a direct experimental comparison to mitochondrial lineage tracing (Figure 6e-g, S15). We believe that these additions further strengthen the importance of EPI-clone as a tool for understanding clonal dynamics in native hematopoiesis.

Major comments:

1. Lack of clarity on experiments performed:

This paper reports a new transgene-free clonal tracking method and therefore even within the brief format provided by the journal it is key that the rationale of the method and the experiments that benchmark it to transgene-based are explained clearly, also to an audience not expert in clonal tracking. But that is not the case.

We would like to thank the reviewer for pointing out the missing clarity in descriptions of the experimental procedures and appreciate the constructive criticism provided below. We have restructured figure 1&2, added a methods summary, and organized the overview of the different experiments better. Specific points are addressed in a point-by-point answer below.

a. First and foremost, I could not find a description of how LARRY barcodes are integrated with scTAMseq, in the main text, in the methods or even in the github pages. Previous work using LARRY (Weinreb et al, Science, 2020 and Rodriguez-Fraticelli, Nature, 2020) has read-out LARRY

barcodes via their transcription and scRNAseq. Instead, scTAMseq reads out DNA exclusively (at least according to the references given). The authors write in the methods (line 664) that "using the stained cells that we used as input for scTAM-seq (Figure S1b), we additionally performed 10X Genomics Chromium Single Cell 3' for transcriptomic profiling of the cells". I understand from this that the cells were stained, sorted and then this pool was split into 2, a portion for scTAM and the other for 10X Genomics. However if that was how LARRY barcodes were read, the same cells did not have both mRNA and DNA profiled, how can they be linking LARRY barcode data to the methylation UMAP as in Figure 1c?

Given that LARRY integrates into the DNA, my guess is that LARRY barcodes are also read through scTAMseq by looking at the DNA insertion. This way each cell would have methylation data and a LARRY barcode. From inspecting the code, this seems the most likely scenario. However no methods as to how LARRY amplicons may be generated is given. This is a key experiment that validates the whole procedure, but not enough information is provided as to how it was performed.

We apologize for the unclear description of the amplification of the LARRY barcode. As already correctly inferred by the reviewer, the approach was simply to include primers specific to the LARRY construct into the primer pool for scTAM-seq. Since the LARRY construct does not carry a HhaI recognition site, it is not affected by the digest with the enzyme included as part of the scTAM-seq protocol. LARRY barcodes are then read out using standard sequencing of the scTAM-seq library. We now clarified this point in the main text and methods section;

(Results, line 146)

The LARRY barcode was read out directly from the DNA by including a LARRY-specific amplicon in our targeting panel for scTAM-seq.

(Methods, lines 1055-1058)

For amplifying the LARRY barcodes, we spiked in an additional primer into the primer pool targeting the LARRY barcode sequence (forward: GCATCGGTTGCTAGGAGAGA, backward: GGGAGTGAATTAGCCCTCCA). We can thus read-out the LARRY barcode together with information about the DNA methylation state from the same single cell.

b. Figure legends are really minimal, methods are choppy and overall schemes of how each step of the experiments and each step of the analysis are integrated is missing. The authors should also improve consistency in naming of experiments or features throughout the manuscript. For example, they often use barcodes in a general sense, but they should be more precise (cellular barcodes, LARRY barcodes...). Another example is: The authors list all the experiments performed in mice in line 108 but then they do not use consistent naming for each of the experiments later on, so it is confusing. At line 141 they talk about "native hematopoiesis", I guess this is experimental batch iii) "wildtype mouse without LARRY barcodes"?

We thank the reviewer for these suggestions. We now

- Extended and improved figure legends
- Restructured the methods and provide a methods summary
- Always specify if referring to cellular or LARRY barcodes.
- List all experiments upfront (lines 130-134), provide a detailed schematic overview of all mouse experiments (Figure S1A), and a detailed supplementary table with references to the methods section and links to figshare for data availability (Table S1). We also implemented consistent codes for the experiments (mouse: M.1-M.8, human: A.1-A.7 and B.1-B.5, methodological expansions: X.1 & X.2) and used them throughout, in the main text, methods, and figures.

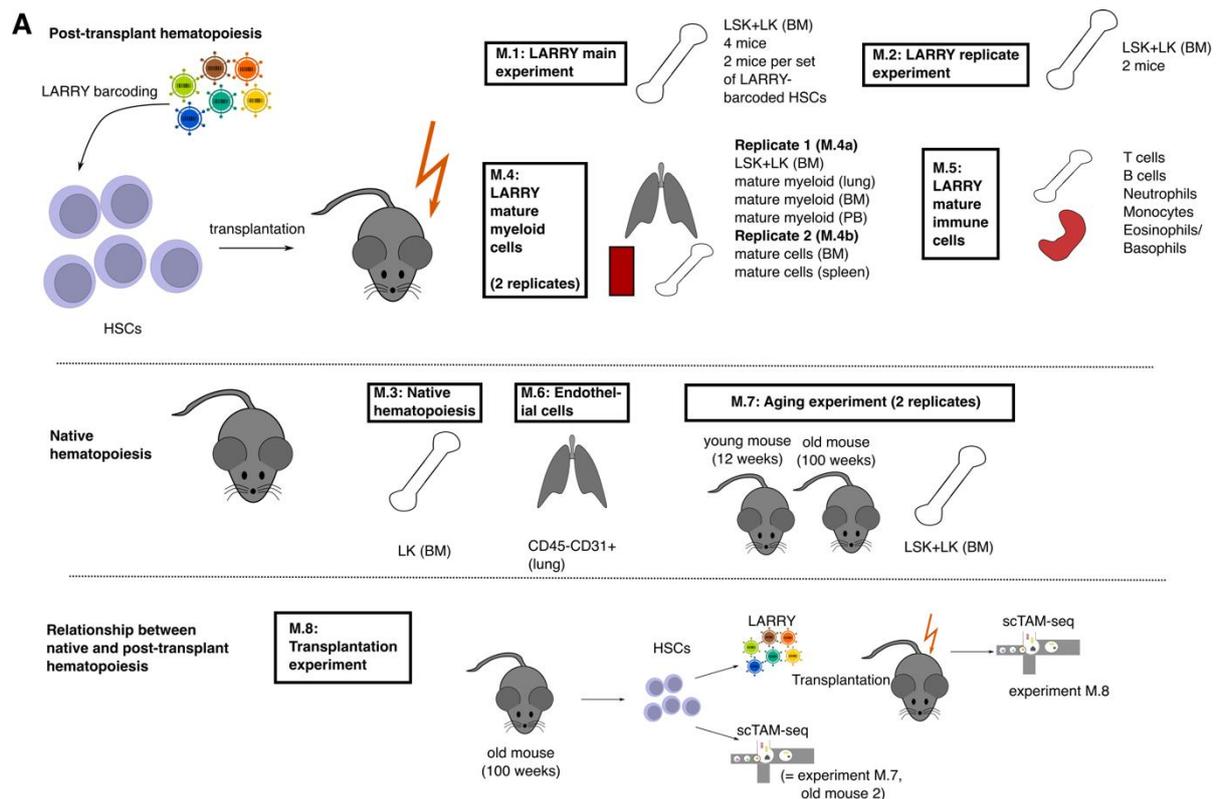


Figure S1A, reproduced here for convenience: Overview of experimental design of the mouse experiments M.1-M.8. See also table S1. LSK: *Lin-Sca1+c-Kit+*, LK: *Lin-c-Kit+*.

2. The method of defining static CpGs as part of EPI-clone is elegant and the overlap with LARRY barcoding convincing. The fact that cellular state can be inferred from the methylation data alone is also convincing. However the biological conclusions from the experiments on ageing (Figure 4 and 5) are somewhat underwhelming in terms of novelty as well as in terms of how the data is presented.

Thank you for pointing out that the validation we used for EPI-Clone are convincing and that there is strong information about the cell state available in the data. We have substantially extended our study of ageing (now figure 4-6) to enhance the biological novelty of our study (see also general reply to this reviewer, above).

a. The main message from the mouse aging study is that with ageing there are more expanded clones (from 40% to 78% of cells) and these tend to be larger. This message is novel to my knowledge. However the authors should improve the robustness of this finding.

We have improved and expanded the analyses on that central point. Specifically, we have included a second biological replicate of young and old animals, which displayed very similar clone size distributions in the different cellular compartments (updated figure 4c,d, also Figure S9, S10). We have added quantifications and statistical analyses, as detailed in response to the next points, and clarified text and figure.

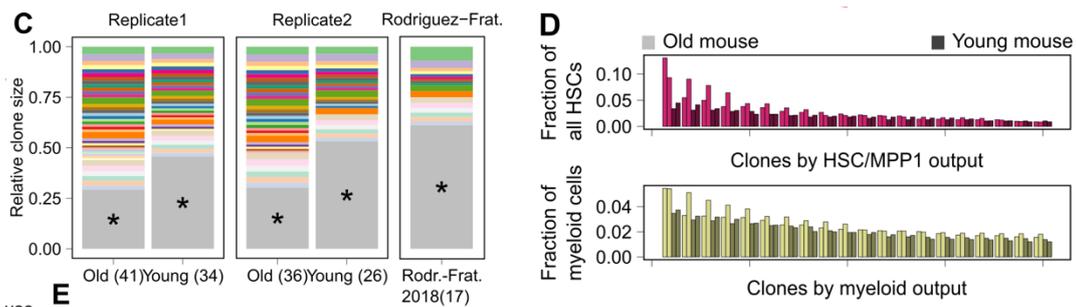


Figure 4c,d, reproduced here for convenience. **C.** Comparison of clone sizes, measured using the fraction of cells in the clone in comparison to all cells, for the old/young mice (2 biological replicates). Clones with relative clone size less than 1% are shown in grey. The last chart reproduces the clone sizes in a young mouse reported in Rodriguez-Frat. 2018¹⁴. **D.** Comparison of HSC/MPP1 output and myeloid output for the 20 clones with the highest HSC/myeloid output between the young and old mouse (2 replicates). The output is defined as the fraction of all HSC/MPP1 or myeloid cells per EPI-Clone cluster compared to all HSC/MPP1 or myeloid cells per experiment.

Finally, we now further characterize these clones by tracking their behavior through transplantation.

(Results, line 291-301)

To determine the long-term stability of the HSC-expanded clonal behavior, we performed a transplantation assay using an aged donor mouse. We used EPI-Clone to compare the clonal composition of the native (pre-transplant) and post-transplant hematopoietic systems, using LARRY barcoding as an additional control during transplantation (experiment M.8, Figure 4h, Figure S11a). We first demonstrated that clonal identities defined through EPI-Clone remained stable during transplantation (Figure S11b-e). HSCs with abundant progeny pre-transplant showed poor engraftment, in line with serial transplantation studies using lentiviral barcoding^{6,14} (Figure 4i, Figure S11f, $p < 0.05$). Intriguingly, HSC-expanded clones also engrafted poorly, identifying non-expanded HSCs as the major drivers of transplantation hematopoiesis (Figure 4j). Clones with quantifiable output pre- and post-transplant showed a stable lineage bias that was inherited post-transplantation (Figure 4k, Figure S11g, $p < 0.05$).

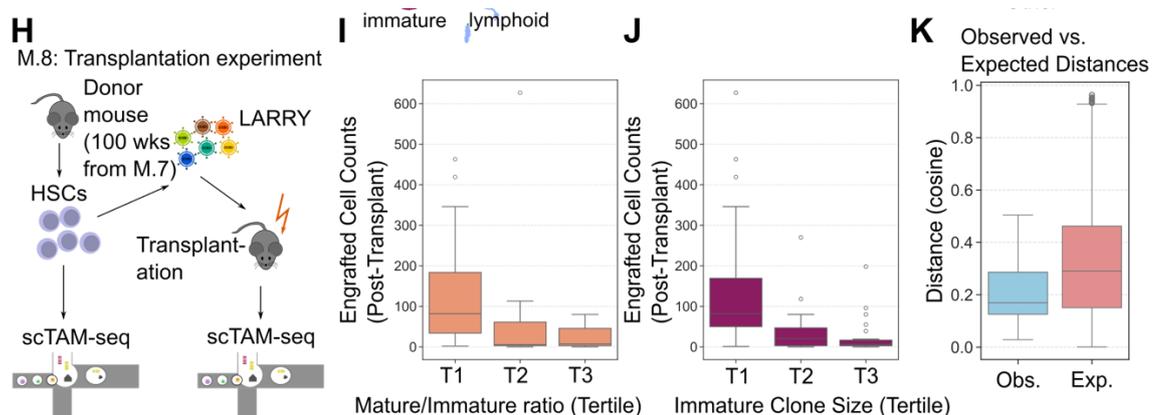


Figure 4h-k, reproduced here for convenience. **H.** Overview of the experimental design for the transplantation experiment. A subset of the HSCs from a donor mouse (old mouse from experiment M.7) was used for scTAMseq processing, while the remaining part was labeled with LARRY barcodes and transplanted into a recipient mouse, which was also profiled with scTAMseq (+LARRY profiling) six months post transplantation. **I.** Boxplot showing the

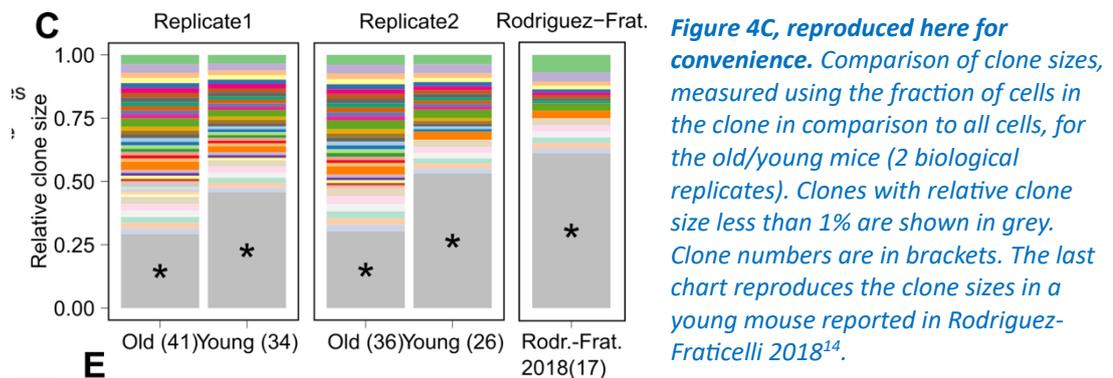
distribution of post-transplant clone sizes, comparing clones with different pre-transplant differentiation bias, calculated as the ratio of mature vs. immature cells per clone (grouped by tertiles). *Wilcoxon $p(T1 \text{ vs } T3) < 0.05$ J. Boxplot showing the distribution of post-transplant clone sizes, comparing clones with different pre-transplant immature clone size (number of cells annotated as HSC/MPPs per clone, grouped by tertiles). *Wilcoxon $p(T1 \text{ vs } T3) < 0.05$ K. Boxplot showing the distribution of pairwise cosine distances (pre and post-transplant) computed using the cell-type distribution of each clone. Observed data is compared with a null model created by randomly shuffling the clonal identities of post-transplant clones (1000 times). *Wilcoxon $p < 0.05$

- Figure 4a: quantification of the total number of clones would be useful. Are these conclusions based on a low number of clones or not? Is this total number also similar to that of other methods (Rodriguez-Fraticelli, 2018)?

We agree with the reviewer that the total number of expanded clones per sample is an important information that was missing in the original version of the manuscript. We now added these numbers to Figure 4D of the manuscript and mention them in the main text. This number is similar to that of Rodriguez-Fraticelli et al., 2018 (see updated figure 4c)

(Results, line 274-276)

When comparing the EPI-clone result, we observed that 70.8%/69.7% of cells were part of 41/36 large clones in the two biological replicates of two old mice, compared to only 54.4%/46.9% in the young (Figure 4c,d, Figure S9, S10, $n=34/26$ expanded clones).



- Figure 4b, d, e and f: could the authors add statistics to each of these observations, given that the main novel conclusion of the paper is drawn from these?

- Line 237: "the large, low-output stem cells clones were mostly myeloid biased": is this qualitative statement only from 4i? Could the authors actually quantify this as again this is presented as a key finding?

We now moved figure 4i to the supplement (figure S9 and S10, for the two replicates) and replaced it by a more quantitative analyses (new figure 4g). We also included relevant statistics as follows:

(Results, line 281-289)

Next, we measured the distribution of cell types for each clone across the various stem and progenitor clusters. In the old mice, we observed several expanded clones containing mostly HSCs across both of our replicates (Figure 4d-f, KS test $p\text{-value} < 0.05$), which were not present in the young mice. These HSC-expanded clones contained large numbers of stem cells

apparently incapable of proceeding with differentiation, and little progeny. Old mice showed a very moderate increase in the number of myeloid-biased clones, in contrast with results from classic transplantation experiments⁵⁵⁻⁵⁸ (Figure 4d, Figure S9, S10). However, the rare HSC-expanded clones were mostly myeloid biased (Figure 4g, Wilcoxon test p-values: 0.01/0.076).

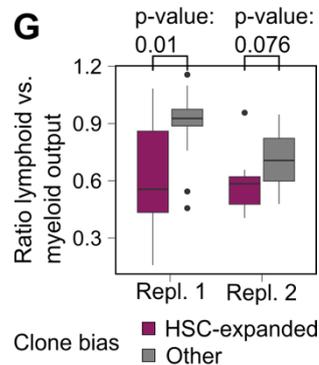


Figure 4G, reproduced here for convenience. Comparison of the ratio between lymphoid and myeloid output (defined as the fraction of all lymphoid/myeloid cells per EPI-Clone cluster compared to all lymphoid/myeloid cells per experiment). The clone bias annotation of the clones that are shown here is defined in Figure S9 (replicate 1) and Figure S10 (replicate 2), respectively. P-values indicate the result of a Wilcoxon test.

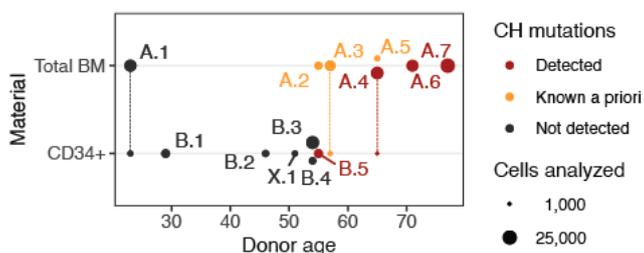
- Importantly, while this reviewer appreciates the complexity of these experiments, the conclusions are drawn from one experiment with one young mouse and one old mouse. Is this really enough to make any conclusions? How variable may be the oligoclonality between old mice?

We agree with the reviewer that this statement requires replicates. We now included a replicate experiment of the mouse aging study and included them in the updated Figure 4. We found that the results were highly reproducible across the two biological replicates (see above, and see figure S10 for full data of the replicate experiment).

B. The authors also apply their method to investigate human aging. With this analysis they largely confirm a previous study reporting oligoclonality in individuals aged over 70 years old (Mitchell et al., Nature, 2022).

The human study has now been substantially expanded to include 13 individuals (Figure 5a), and to focus on a comparison of lineage bias between clones with clonal hematopoiesis (CH) driver mutation, and clones without a known driver (“non-CH clones”). The approach used by Mitchell et al. cannot be used to study stem cell clonal behaviors other than expansion in vitro. Furthermore, colony-based whole-genome sequencing is limited to cells with strong proliferative capacity in culture, and it may also miss expansions of clones with lost regenerative potential.

A



New figure 5A, reproduced here for convenience. Donor characteristics including donor age, presence of CH mutations and number of cells analyzed. See also Methods, Human samples and table S1. Dots connected by dotted lines denote samples that were analyzed as part of the TBM and the CD34+ dataset

We first extensively evaluate EPI-Clone's ability to recapitulate CH clones, and additionally discover clonal expansions with no known driver (Figure 5a-e, and see also one of the reviewer's further points). EPI-Clone enables a comparison between clones with and without known driver mutations (CH and non-CH clones). We describe our results as follows:

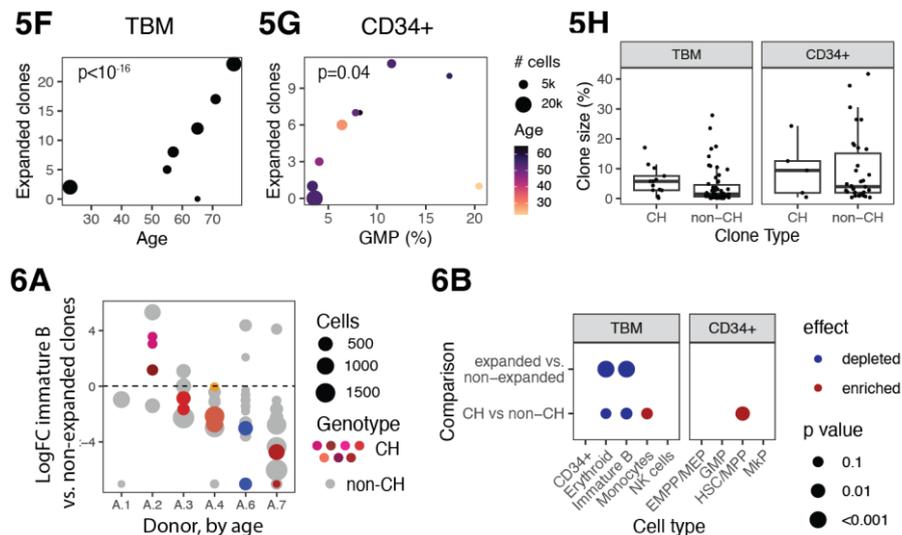
(Results, line 375-382)

CH clones tended to be more expanded than non-CH clones, but were not always among the largest ones (Figure 5h). Expanded clones were significantly depleted (compared to cells from non-expanded clones) from the B and erythroid lineages (Figure 6a,b, S13f), suggesting a link between myelopoiesis and expansion even for non-CH clones. Compared to non-CH clones, CH clones were significantly enriched in HSCs/MPPs, in line with known specific effects of CH mutations, while being depleted from the B and erythroid lineages^{48,62} (Figure 6b, S13f,g). These results highlight a stem-cell bias in age-expanded clones that is conserved across mice and humans, and they support a model where CH clones are part of a spectrum of such age-expanded clones.

Beyond these fundamental new capabilities, our study also adds to Mitchell et al (ref 18) by profiling more individuals in the critical age period of 50-60 years, where clonal expansions begin to accumulate, and thereby provide new insights into processes driving clonal expansion:

(Results, line 365-373)

As expected¹⁸, in the TBM cohort, we observed an age-dependent accumulation of expanded CH and non-CH clones (Figure 5f). Interestingly, in the CD34+ cohort, which mostly was sampled from individuals aged 50-60, we identified a correlation between the fraction of GMPs in the sample and the accumulation of expanded clones (Figure 5g, S13c), suggesting that cues that enhance myelopoiesis also lead to more clonal expansions.



New figure 5f-h, 6a,b, reproduced here for convenience. **5F.** Scatter plot relating donor age to the number of clones identified by EPI-clone, in the TBM cohort. p -value is from a generalized linear model of the poisson family, using number of cells observed as a weight. Dot size denotes number of cells analyzed, see panel B for a scale. **5G.** Scatter plot relating the

presence of GMPs (expressed as % of the CD34+ compartment) to the number of clones identified by EPI-clone, in CD34+ cells from all samples where CD34+ cells were covered with > 1,000 cells (B.1-B.5, X.1, A1, A.3 and A.4). P-value is from a generalized linear model of the poisson family, using number of cells observed as a weight and age as a covariate. **5H.** Box plot depicting clone size stratified into clones carrying CH mutations and clones for which no CH mutation was identified. **6A.** Scatter plot depicting the fraction of immature B cells observed per clone, relative to the fraction of immature B cells observed in non-expanded clones from the same patient. Grey dots correspond to EPI-clones with no known driver mutation. Dots in colors correspond to EPI-clones dominated by a CH mutation, see figure 5c for a color scheme. **6B.** Dot plot depicting p values for enrichments and depletion of cell types in expanded vs. non-expanded and CH vs. non-CH clones. For this analysis, cell type composition of clones (e.g. % of clone CD34+, or % of clone GMP) were transformed via a logit transform and p-values were computed using a mixed effect model, using donor as a random effect and clone type (expanded/non-expanded or CH/non-CH) as a fixed effect. See also figure S13f,g.

- Figure 5D: why were B and T cells removed from the analysis? Does EPI-clone work in these cells as well? Please comment in text or figure legend.

We now evaluate EPI-Clone's performance on various immune cell types in detail, in mouse and human.

In the context of the mouse study, we

(Results, line 223-233)

asked whether EPI-Clone could also determine clonal identity in mature immune cells. To that end, we collected mature immune (lymphoid and myeloid) cells from bone marrow and spleen (experiment M.5, Figure 3g, Figure S5, Table S1) and profiled surface antigen expression as well as DNA methylation at the same CpGs as in experiments M.1-M.3. Using the static CpGs defined from experiment M.1, EPI-Clone again yielded a clonal clustering that recapitulated ground truth clonal labels (Figure 3h). We separately computed ARIs between EPI-Clone results and LARRY barcodes. ARIs were higher than 0.7 for monocytes, neutrophils, other myeloid cells, CD8+ T cells, and one B cell subset, higher than 0.4 for CD4+ T cells, and low for macrophages and a second B cell subset (Figure 3i, j). Most T- and B-cells belonged to lymphoid-dominated (LARRY- and EPI-)clones (Figure 3i, Figure S5d), suggesting an origin in lymphoid-biased or restricted progenitors⁵².

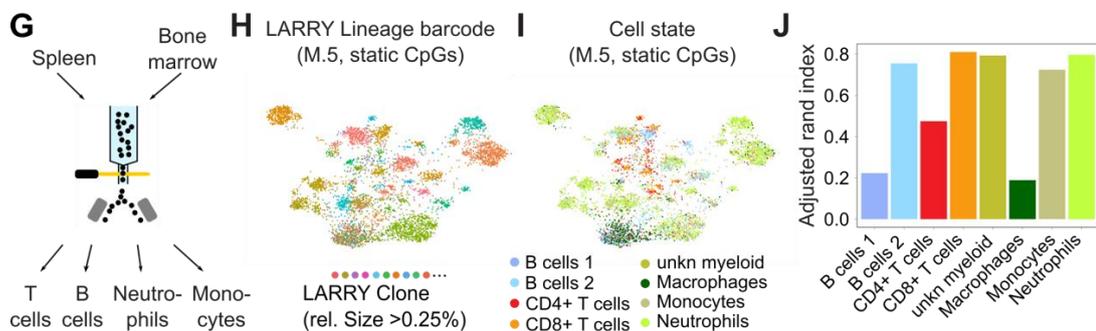


Figure 3g-h, reproduced here for convenience. G. Scheme illustrating experiment M.5: LARRY mature immune cell experiment. **H.** DNA methylation UMAP representation based on the static CpGs for cells from expanded clones in experiment M.5. Cells are colored by LARRY barcode. The static CpGs identified from experiment M.1 were used. **I.** Same UMAP representation as in H highlighting the cell state annotation as defined in Figure S5. Of note,

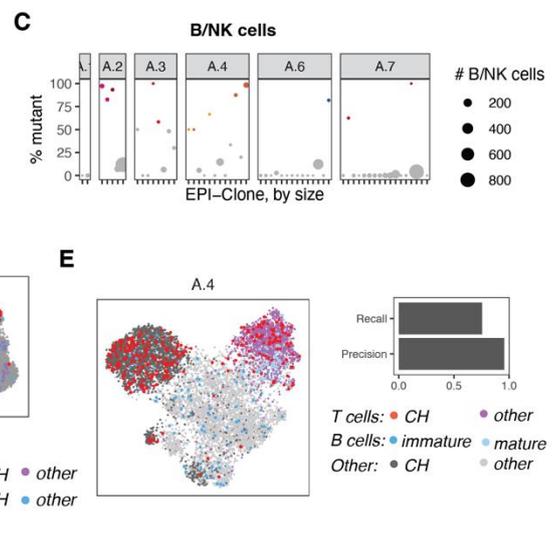
most EPI-clones were specific for T, B or myeloid cells, in line with the result from LARRY (Figure S5d). J. Adjusted rand indices between the ground truth clonal label (LARRY) and the clones identified by EPI-Clone stratified by cell type.

In human, our findings were in line with this observation and are now described as follows

(Results, line 346-353)

We included NK and immature B cells in our analysis and used CH mutations to validate that these cells also clustered by clone (Figure S13c,d). When T and mature B cells were included, they associated with lymphoid-dominant clusters, in line with the result from mouse (Figure S13e, and cf. Figure 3i) and suggesting distinct clonal origins compared to the other cells. In patient A.4, where a large CH clone contributed to T cells, mutant T cells mostly clustered with the remaining CH-derived cells (Figure S13e). Together with the mouse LARRY experiment, this constitutes evidence that EPI-clones remain stable from HSC to myeloid, T, NK, and immature B cells.

Figure S13C-D, reproduced here for convenience. C. Scatter plot displaying the percentage of cells from each EPI-Clone displaying CH mutations, for NK and immature B cells. EPI-Clone was run on all cells except T and mature B cells, but the overlap was computed on NK and immature B cells only. See main figure 5c for color scheme. D. Static CpG UMAPs as in main figure 5c/d, highlighting NK and immature B cells classified according to CH status. E. Static CpG UMAP computed for all cells (including mature B and T cells) for patient A.4, highlighting T cells classified according to CH status. Mature and immature B cells are also highlighted to demonstrate that mature B and T cells mostly cluster in lymphoid clusters. Barchart depicts precision and recall for the task of classifying T cells as CH or non-CH based on EPI-Clone labels.



We further use mitochondrial lineage tracing to validate that EPI-Clone had correctly identified several expanded T cell clones in a PBMC sample (Figure 6e-g).

In conclusion, EPI-clone works well in myeloid cells, NK cell, T cells and immature B cells. The mature B cell compartment in human is highly polyclonal, and even in autoimmune disease, expanded B cell receptor clones make up <<1% of B cells outside of the plasma cell compartment (ref 10.1038/s41586-019-1595-3). We cannot currently explain why EPI-Clone does not work in mouse macrophages, we suspect that they may have “eaten” DNA from other cells. Macrophages, unlike other cell types also often display more than one LARRY barcode (Figure S6M).

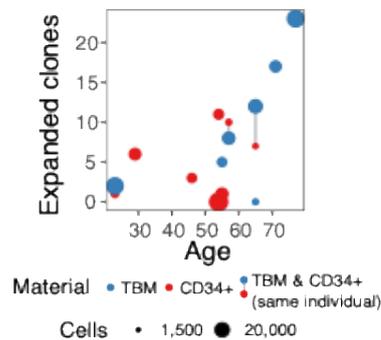
The reason we excluded T- and mature B cells from most EPI-Clone analysis in human was that due to their distinct clonal origin and polyclonal nature, their presence in clustering reduced resolution on the remaining cells.

(Results, line 368)

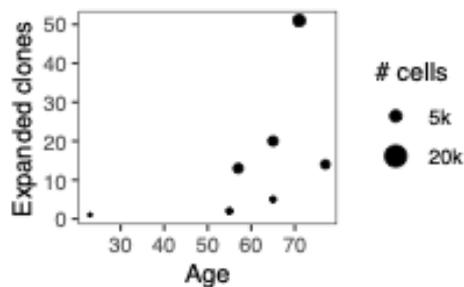
Due to the putatively distinct clonal origins, we excluded T and mature B cell from this analysis.

- Figure 5E: Is oligoclonality more marked at the CD34+ level than at the My level? What about lymphoid cells? This would be a novel finding not demonstrated in previous studies.

We now thoroughly evaluate the ability of EPI-Clone to detect clonal expansions and find that EPI-Clone reliably detects clonal expansions of size >1%, see the reviewer's next point. The overall number of such clonal expansions in the TBM (without T and mature B cell) and CD34+ compartment were, after accounting for age, not different. These results are included in figure 5f,g but can for this concrete comparison be more suitably visualized as follows (Scatter plot depicting number of expanded clones detected by EPI-Clone, as a function of age and compartment).



T cells also displayed a trend towards more clonal expansions with age. However, that trend was not significant, possibly due to an insufficient sampling of cells in that compartment in some patients. We did not include it in the manuscript so as to not dilute the focus on processes in the stem- and myeloid compartment (Scatter plot depicting number of expanded clones detected by EPI-Clone in the T cell compartment, as a function of age).



Functionally, across all patients profiled, individual expanded clones tend to variably contribute to the CD34+ compartment (Figure S13f), but did not follow any specific trend of being enriched or depleted from that compartment. Rather, expanded clones were depleted from the erythroid and immature B compartments, and within CD34+ compartment, tended to be enriched for HSC/MPP (Figure 6a,b, S13f,g).

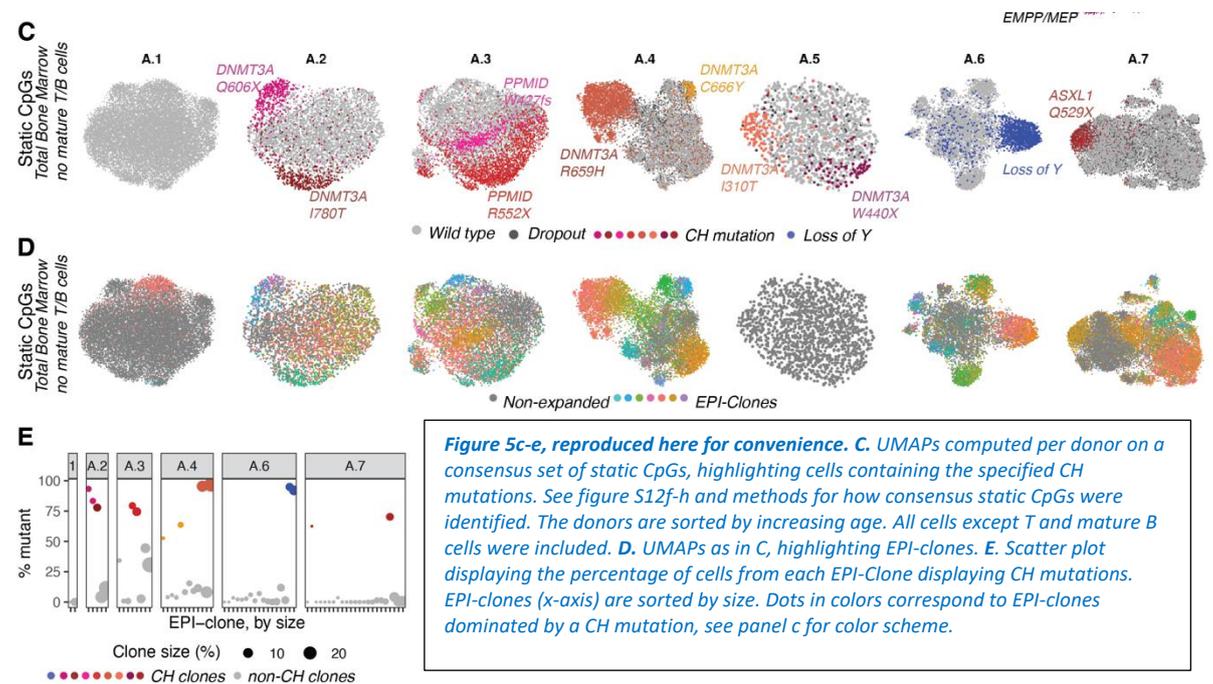
- The authors designed a panel allowing for the assessment of mutations in a quite large number of genomic regions (>140) usually associated with clonal hemopoiesis (CH). However no comment is made about these in the findings. Do expanded clones preferentially carry mutations in CH genes? Can the author address the relationship between individuals mutations and CH

mutations? In other words, is EPI-clone more sensitive at finding clones than just using CH mutations?

The comparison between CH mutations and EPI-clones is now the main focus of the human study (Figure 5&6). We now included several samples from a CH study (Jakobsen et al., Cell Stem Cell 2024, DOI: 10.1016/j.stem.2024.05.010), for which the presence of CH mutations was known a priori. For the remaining samples, we applied SComatic (Muyas et al., Nature Biotech 2023, DOI: 10.1038/s41587-023-01863-z) to identify CH mutations *de novo* from the genotyping regions. In total, eleven CH clones were thereby identified across the thirteen samples. We found that:

(Results, line 331-344)

To assess the fidelity of static CpGs for identifying clones, we exploited the CH mutations/LoY events as a clonal ground truth. CH clones clustered together in static CpG UMAPs in all cases (Figure 5c, S13a). EPI-clone recapitulated the CH clones in all donors except A.5, which was covered with substantially less cells compared to the rest of the TBM cohort (Figure 5d,e, S13b). Quantitatively, EPI-clones dominated by CH mutant cells were on average 78.8% mutant and EPI-clones dominated by wild-type cells were on average 95.4% wild-type (Figure 5e). These numbers likely underestimate the true overlap between EPI-clones and CH clones, due to allelic dropout of CH mutations. We observe a stronger separation of EPI-clones, and better overlap with CH mutations, in older donors, suggesting that EPI-Clone most accurately identifies clones in hematopoietic systems of reduced clonal complexity. Besides the CH clones, EPI-clone also identified a total of 67 other clonal expansions in the seven TBM donors, suggesting the capacity of EPI-clone to recapitulate clonal expansions driven by known and unknown drivers.



Importantly, we also confirm EPI-Clone results in human with an other form of orthogonal genetic label, namely, mitochondrial variants, thereby demonstrating that the overlap of EPI-clones and CH mutation does not stem from the effect of CH mutations on the methylome:

(Results, line 398-418)

In the field, there is some controversy regarding the potential of other somatic events, in particular low-heteroplasmy mtDNA variants, for lineage tracing²²⁻²⁴. To perform a direct experimental comparison, we analyzed peripheral blood from a 38 year old healthy donor (X.2) that had previously been characterized by a state-of-the-art single cell mitochondrial lineage tracing method, mt-scATAC-seq^{20,65}. This donor had 44 EPI-Clones, and in particular displayed prominent clonal expansions of NK- and T-cells (Figure S15a). By including a mitochondrial targeting panel into scTAM-seq, we achieved a median coverage of 176 reads per cell on the mitochondrial genome (Figure 6e, S15b,c). Of the 23 mtDNA variants previously identified⁶⁵ in this donor, five had clear phylogenetic relationships with EPI-Clones, i.e. were either subclones to single EPI-Clones, or were parental to several EPI-Clones (Figure 6f), and one variant was observed in two EPI-Clones. A highly abundant variant (mt:7076A>G) was strongly enriched or depleted in 17 T- or NK-cell EPI-clones, but was observed in a subset of cells of the remaining, mostly multilineage- or B-cell, EPI-clones (Figure 6g). This variant was likely present before epimutational patterns were established and repeatedly underwent selection at late stages of cell differentiation. Finally, the remaining 16 low-heteroplasmy mitochondrial variants did not segregate with EPI-clones (Figure S15d,e). These findings are in line with a recent preprint²³ observing that only some observed mitochondrial variants carry phylogenetic information, and illustrate the complexity of mitochondrial genetics, where selection of variants can happen during differentiation⁶⁵. They further provide additional orthogonal validation of EPI-Clone outside of the setting of CH.

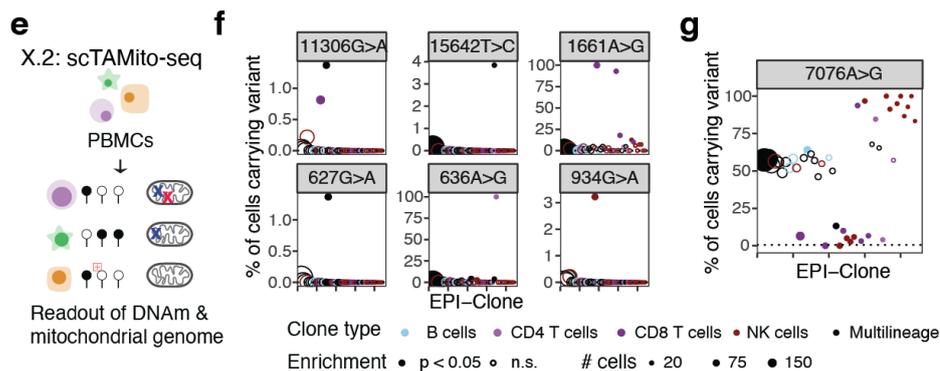


Figure 6E-F, reproduced here for convenience. E. Scheme illustrating experiment X.2. Mitochondrial mutations were read out together with DNA methylation states with scTAMito-seq. A PBMC sample characterized previously for mitochondrial mutations⁶⁵ was analyzed. See also figure S15. **F.** Scatter plot depicting the presence of six mitochondrial variants in the different EPI-clones from X.2. Cells were scored as positive for the variant, if at least 5% of reads supported the variant. The enrichment of variants in EPI-Clones was determined by a binomial test. EPI-Clones were classified as B, T or NK cell clones if at least 80% of cells were from a single lineage, or as multilineage clones otherwise. **G.** Like F, for the mt:7076A>G variant.

3. EPI-clone seems to track clonal somatic epimutations that are stable over several months in mice, and likely many years in humans. The discussion has an interesting paragraph comparing EPI-clone with mtDNA tracking. Could the authors estimate the "timescale" of tracking of the two methods?

We now investigate this point by experimentally comparing mtDNA lineage tracing and epimutations (see previous point). We found a mitochondrial mutation (mt:7076) that is present in 39% of cells (it possibly occurred in early development or was present in the zygote) and appears to repeatedly have undergone selection. We also observe other mitochondrial variants that appear to have formed later.

The timescale and emergence of epimutations is now a focus in the revised discussion:

(Discussion, line 431-446)

This raises the question of where and how clonal epimutations arise. We found that they (i) randomly occur, but remain stable over many cell divisions, (ii) their number does not increase during differentiation (Figure S16), and (iii) they are enriched for heterochromatic and late-replicating domains. We propose that some developmental events that are characterized by rapid cellular proliferation and/or a remodeling of the methylome, such as the specification of HSCs⁶⁷ essentially randomize the methylation state of CpGs in heterochromatin and late replicating regions. As a potential explanation, in fast-dividing cells DNMT1 may act insufficiently to copy the DNA methylation state to the nascent DNA strand. Consistent with this, a recent study of bulk methylome profiles from blood cells in monozygotic twins suggests that clone-associated variation of the methylome may be established during embryonic development²⁶. In the case of some large CH clones, we observed additional diversification of epimutational patterns.

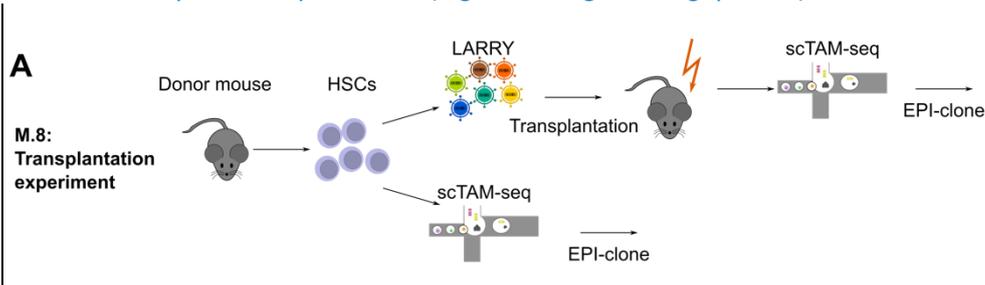
We therefore propose that variably methylated CpGs in non-regulatory genomic regions can act as a digital barcode of clonal origin. The digital and stochastic nature of epimutations makes single-cell methods that are capable of mapping methylation state of single CpGs at high confidence, such as scTAM-seq, a powerful tool for lineage tracing.

Another unresolved question is how diverse is the functional behaviour of the clones expanded with age on shorter time scale (for example during differentiation). This is something that could be addressed by combining EPI-clone with shorter LARRY barcoding follow-ups.

Our new aged BM transplantation experiment provides interesting insights regarding the stability of these different clonal behaviors (Figure 4h-k, S11, and see reviewer's point 2a):

(Results, line 300-301)

Clones with quantifiable output pre- and post-transplant showed a stable lineage bias that was inherited post-transplantation (Figure 4k, Figure S11g, $p < 0.05$).



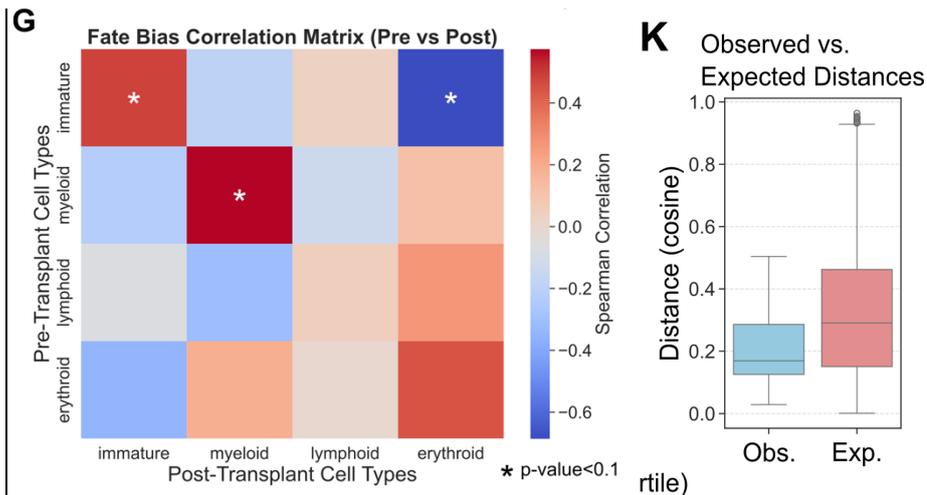


Figure S11a,g and main figure 4k, included here for convenience. A. Overview of the experimental design for experiment M.8: transplantation experiment. HSCs from an old donor mouse (100 weeks) were either LARRY-barcoded and transplanted into a recipient mouse or directly used for processing with scTAM-seq/EPI-clone. **G.** Spearman correlation between the clonal output of each clone towards the three main blood lineages compared between the donor mouse and the transplanted mouse. The asterisk indicated p-values below 0.1 from a correlation test. **K.** Boxplot showing the distribution of pairwise cosine distances (pre and post-transplant) computed using the cell-type distribution of each clone. Observed data is compared with a null model created by randomly shuffling the clonal identities of post-transplant clones (1000 times). *Wilcoxon $p < 0.05$

We believe that a more comprehensive functional characterization of mouse HSC-expanded clones is out of scope for the present manuscript, and we have instead focused our revision efforts on the new and extensive examination of human aged hematopoiesis.

Minor comments:

Line 141: the statement that clonal information was removed when three different experiments were integrated should be backed up by a figure.

We thank the reviewer for pointing out the missing justification for this statement. We now included the corresponding UMAP highlighting the LARRY barcodes in the integrated UMAP as Supplementary Figure S3A. We cannot find an association between the LARRY barcodes and the differentiation trajectory, indicating that the UMAP indeed captures only differentiation state.

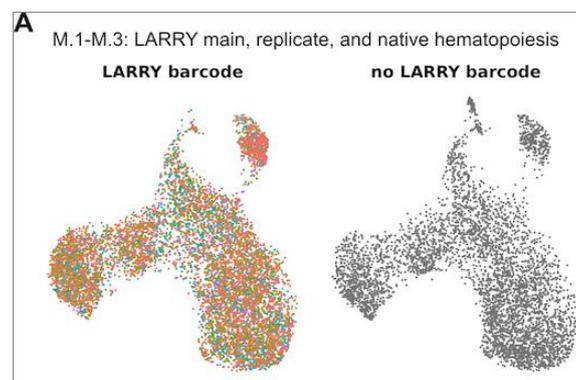


Figure S3A, reproduced here for convenience. Integrated UMAP of the LARRY main experiment, replicate, and native hematopoiesis (experiments M.1-M.3) as in Figure 1c, highlighting the LARRY barcodes.

Line 152: shouldn't this come earlier? This is how the clusters in Figure 2A have been annotated, correct? Very confusing

Thank you for pointing out the missing clarity in cell state annotation. To improve the flow of the overall manuscript, we now first introduce the cell state annotation in Figure 1. We believe that with this modified structure, the overall message of the manuscript is easier to understand.

Line 217: it is not clear how was this done. Please also repeat which populations were sorted.

We appreciate the reviewer's concern and understand that the experimental design of the aging experiment was insufficient. To that end, we create an outline of the experiment as the new Figure S8a and also updated the text accordingly. We now worked on explaining our experimental design in an easier way and decided to use more graphical outlines of the experiments in the Supplementary Figures (see also new Figure S1a, above).

Figure 4H: it would be really useful to have the annotated map next to these clone examples for orientation

Thank you for this suggestion. We now included the cell-state UMAP for better orientation in Figure 4F.

How are low-output and high-output defined? Is there a numeric criterium? Even if not, these should be explained in the main text, currently this is only relatively clear if one has read previous studies from the authors

Figure 4I: how do we distinguish low-output from high-output in these graphs? In the old mouse, is it the first split in the tree on the left that distinguishes them (low output on top, high output at the bottom).

To avoid confusion, we do not use the terms low-output and high-output in the revised manuscript. Rather we use the term "stem cell expanded clones", which we now define by at least 30% of cells from the clone being HSCs. The heatmap visualizations (formerly figure 4i, now figure S9 and S10) now only serve to confirm that these clones indeed mostly consist of HSCs without displaying further patterns of lineage bias.

Referee #1 (Remarks on code availability):

The authors have provided code in GitHub which is relatively clear and concise. While I did not try to run the code, it is useful that the authors provide the Seurat objects and overall it should be possible to reproduce the code. However there are no links provided to the raw data, which is not ideal. Most importantly the manuscript and/or the GitHub page both lack a schematic explanation of how the different pieces of data generated with which libraries are all integrated together. Consistency in terminology used in GitHub and in the methods of the manuscript could also be improved.

We would like to thank the reviewer for pointing out some missing clarity in the code documentation. We address the points as follow:

Raw data has been submitted to SRA/GEO (accession numbers GSE282971, SRA : assignment of accession number pending).

Table S1 and Figure S1 were substantially extended to better explain the experimental conditions.

The code was more adapted to the main text.

Referee #2 (Remarks to the Author):

The authors present EPI-Clone, which uses the scTAM-seq technology for lineage tracing purposes in addition to cell typing. The approach is based on the Mission Bio Tapestry system and provides amplicon readouts and methylation status of a pre-defined set of CpG sites which enable lineage assessment due to the stochastic changes at some of these sites, as well as general cell type due to the known status of the specific sites selected. While the technology itself is previously-described, the analysis and application is novel and serves as a powerful example of how DNA methylation can be leveraged for lineage tracing without the need for genetic manipulation. One notable weakness of the method is that a specific panel of CpG sites must be developed for any given target issue, limiting generalizability. The authors apply EPI-Clone to several systems, first validating alongside a LARRY genetic lineage tracing model, which is a key proof of principle followed by assessment in ageing mice and in human. The assessment of clonal lineages reveals the landscape within these systems which generally match with what is previously known, though approaching it from a new angle, and arguably providing improved power and resolution. That being said, there are not any big surprises or entirely novel findings in the analysis, which is largely descriptive. That being said, the method likely has the power to learn some really powerful biology, such as a large-scale analysis across many human samples as opposed to a vignette example.

We would like to thank the reviewer for the detailed evaluation of our work, and the encouragement to apply EPI-clone to a large-scale analysis of human samples. We now compiled a cohort of 13 bone marrow donors, including seven total bone marrow samples with CD34+ enrichment, and six pure CD34+ samples (new Figure 5a). In seven of these donors, we identified CH mutations or LoY events, which mapped to specific EPI-clones (new Figure 5c-e, S13a-e). Alongside these CH-driven clonal expansion, we also identified tens of EPI-clones not associated with mutations (new figure 5d). This cohort therefore enabled us to systematically compare “driver-free” and CH-driven clonal expansions in human ageing. In brief, while the existence of clonal expansions without known driver in bone marrow from elderly has been described previously, these clones have never been functionally characterized, due to a lack of suitable methods. We describe our novel biological findings in detail in the main text:

(Results, line 365-382)

We leveraged EPI-Clone’s ability to trace both CH clones, which are well characterized in human^{48,62}, and clones without known driver mutations (“non-CH clones”) to functionally compare these two types of clonal expansions in our total bone marrow and CD34+ cohorts. Due to their putatively distinct clonal origins, we excluded T and mature B cell from this analysis. As expected¹⁸, in the TBM cohort, we observed an age-dependent accumulation of expanded CH and non-CH clones (Figure 5f). Interestingly, in the CD34+ cohort, which mostly was sampled from individuals aged 50-60, we identified a correlation between the fraction of GMPs in the sample and the accumulation of expanded clones (Figure 5g, S13c), suggesting that cues that enhance myelopoiesis also lead to more clonal expansions.

CH clones tended to be more expanded than non-CH clones, but were not always among the largest ones (Figure 5h). Expanded clones were significantly depleted (compared to cells from

non-expanded clones) from the B and erythroid lineages (Figure 6a, S13f), suggesting a link between myelopoiesis and expansion even for non-CH clones. Compared to non-CH clones, CH clones were significantly enriched in HSCs/MPPs, while being depleted from the B and erythroid lineages (Figure 6b, S13f,g). These results highlight a stem-cell bias in age-expanded clones that is conserved across mice and humans, and they support a model where CH clones are part of a spectrum of such age-expanded clones.

We discuss the biological importance of these findings:

(Discussion, lines 459-468)

Our data further put CH mutations into a perspective with clonal expansions without known driver: CH clones are more strongly biased towards the myeloid lineage and towards an expansion of stem cells, but together with “non-CH” clones form part of a spectrum of age-related clonal expansions that display similar functional properties. In aged mice, we similarly detected large HSC-expanded clones that had reduced regenerative capacity. Together with recent transplantation studies of human HSCs⁶⁸ this suggests conservation of the processes that drive hematopoietic ageing and decline in clonal complexity, and highlights that CH mutations might not be the main driver of this process. Epidemiological studies found an increased mortality risk in carriers of driver-free expanded clones⁴⁵. These results call for a broader investigation of age-related decline in clonality instead of a strict focus on CH.

We would like to briefly comment on the perceived weakness of EPI-clone, that a specific panel of CpG sites must be developed for any given target tissue. We now included data demonstrating the successful application of the original mouse panel (designed for hematopoietic stem and progenitor cells) to endothelia (new figure S7). Surprisingly, a similar set of static and dynamic CpGs defines clones and differentiation states, respectively, in endothelia and hematopoiesis:

(Results, line 242-253)

Finally, we asked whether EPI-clone would also be applicable to tissues other than blood. We therefore applied the same CpG panel to sorted endothelial cells (ECs) from lung, which share a common developmental origin with blood (experiment M.6, Figure S8A). Using the dynamic (differentiation-associated) CpGs defined in hematopoiesis, we identified three main clusters (Figure S8B). Based on the surface expression of CD31, Sca-1 and Podoplanin these were identified as highly (CD31+Sca1-Podoplanin-) or lowly (CD31+Sca1+) abundant Capillaries (labelled as capillary 1 and 2, according to scRNA-seq references), and more rare Lymphatic (CD31+Podoplanin+) endothelial cells (Figure S8C-F). Using the same set of static CpGs as in hematopoiesis, EPI-Clone revealed cell state independent, yet statistically supported clusters (Figure S8G, H), which we interpret as clones. Notably, most clones behaved in a similar manner, majorly contributing to generate the most abundant capillaries rather than other endothelial cells, independently of the clone size (Figure S8H). These observations suggest that a similar set of static and dynamic CpGs defines clones and differentiation states, respectively, in endothelia and hematopoiesis (Figure S8I).

While we still cannot exclude that different CpGs will be required to profile very different tissues (e.g. neurons), the design of the scTAM-seq panel follows a standard procedure that makes use of broadly available bulk DNA methylation data.

Finally, we would like to briefly summarise other additions to the manuscript during these revisions.

- We have strengthened the mouse ageing study through inclusion of replicates and an experiment where we used EPI-Clone to compare an aged native hematopoietic system to the same system post-transplant (Figure 4)
- We demonstrate that EPI-Clone can be combined with targeted RNA-seq from the same single cell (Figure 6c,d and S14)
- We experimentally compare EPI-Clone and mitochondrial lineage tracing (Figure 6e-g and S15).

On the technical side, the experiments and analysis are sound. Additional detail on the specific readout and how those data are leveraged would be helpful in the main text. The github vignette documentation is transparent and clear, detailing exactly what analysis was performed which is important and should be commended. The repository was very helpful in deciphering exactly what analysis was performed and aided in review of the manuscript. An archived version should accompany any final release of the work.

We thank the reviewer for pointing out the potential improvements in the main text and addressed these points in the revised version of the manuscript. We also appreciate the positive feedback on the documentation in the GitHub repository, which we expanded, and will publish an archived version together with the manuscript. An archived version of the code at initial submission is available from GitHub (<https://github.com/veltenlab/EPI-clone/releases/tag/v1.0>).

The flow of the paper itself could use some work – there are multiple instances where later sections are referenced, forcing the reader to jump around, sometimes even having to jump to a later section / figure then jump to another later one, to then get back to the earliest point. Structuring the work in a way that does not require flipping back and forth would be a huge improvement.

We appreciate this suggestion, which helped us to improve the presentation of our results: We have now re-written parts of the manuscript to avoid cross-referencing earlier or later parts of the manuscript. Most notably, we now introduce the cells state annotation already in Figure 1.

Minor: The authors state that single-cell DNA methylation is limited to a few hundred cells; however, datasets from the BRAIN Initiative are in the 100's of thousands. There is also a commercially-available kit for single-cell methylation that has a throughput in the 10's of thousands per experiment. Though notably, coverage sparsity is an issue for all of those methods.

We agree and now stress the sparsity issue clearly in the introduction and discussion. We removed the statement on cell number limitations.

(Discussion, lines 444-446)

The digital and stochastic nature of epimutations makes single cell methods that are capable of mapping methylation state of single CpGs at high confidence, such as scTAM-seq, a powerful tool for lineage tracing.

It is a bit unclear in the main text results what the specific readout is from the assay. 453 CpG sites are targeted and assessed using a methyl-sensitive restriction enzyme as opposed to bisulfite methods. Is the readout presence/absence? Or is there data on the actual methylation state? If it is presence/absence then dropout would contribute to the absence readout. The analysis suggests this

is not the case, so clarification without having to read the scTAM paper would benefit the manuscript greatly.

We fully agree with the reviewer that the paper should be self-sustained and the reader should not be required to read the original scTAM-seq publication to follow the manuscript. We thus now extended the description of the method in the introduction and in the methods part as follows

Introduction, line 116-118:

Briefly, scTAM-seq uses a methylation-sensitive restriction enzyme to selectively cut unmethylated CpGs and thus allows for differentiating methylated from unmethylated CpGs through the presence of sequencing reads (Methods).

Methods, lines 1138-1146:

At the single-cell level, we differentiated methylated from unmethylated CpGs through the presence of at least one sequencing read for the corresponding amplicon as in the original publication of scTAM-seq. Sequencing reads can uniquely originate from amplicons harboring methylated CpGs, while the lack of sequencing reads from an amplicon originates either from an unmethylated CpG or from a dropout. To minimize the effect of dropout, we determine the primer combinations that reliably amplify in our panel using a single experiment without the restriction enzyme. For this experiment, Lin-cKIT⁺ cells from a young, wildtype mouse (12 weeks) were used and we determined that 453 of the 573 non-control amplicons (79%) amplified in more than 90% of the cells. These amplicons were used for subsequent analysis.

While we agree with the reviewer that the lack of sequencing read can come either from an unmethylated CpGs or from a technical dropout, we found that in the original publication of scTAM-seq technical dropout can be controlled for using an undigested control experiment. We also performed this control experiment in this study and found that 79% if the amplicons have a dropout rate of less than 10% (Table S4, S5). We exclusively included the amplicons with a dropout rate lower than 10% during all analyses presented in the paper.

Is “uMAP” different from “UMAP”? “UMAP” is generally the standard term.

Thank you for pointing out that we did not follow the standard naming scheme of the low dimensional representation. Indeed, uMAP also refers to Uniform Manifold Approximation and Projection and we now replaced all appearances of uMAP by UMAP.

Figure 1C – label it “LARRY Lineage Barcode” to be more descriptive

Thank you for this suggestion. We replaced the label in the plots accordingly.

Figure 4 – the vast majority of cells profiled fall into clones with <75 cells. Is this due to an inability to properly assign? A limitation of the cell numbers profiled? Or expected and the focus is just on expanded clones? It is hard to know if one could even ascertain this.

We thank the reviewer for this comment. In the revised version of the manuscript, we now revisited the identification of expanded clones by using the relative clone size instead of the absolute clone size. Using this approach, we can now identify more than 50% of all of the cells being as part of expanded clones with a size >1% in the young and the old mice across two biological replicates. We would like to point out that we cannot capture non-expanded

clones with our technology, since EPI-clone relies on the epimutational signatures being present across various cells. We mention this limitation in the discussion part of the manuscript:

Discussion, lines 478-481

Cells belonging to very small clones (<0.25% of cells in post-transplant, and <1% in native hematopoiesis) can be identified as not belonging to expanded clones, but their clonal identity cannot be inferred with the cell numbers used in this study.

Referee #2 (Remarks on code availability):

The code was provided and is detailed.

Referee #3 (Remarks to the Author):

The paper "Somatic Epimutations Enable Single-Cell Lineage Tracing in Native Haematopoiesis Across the Murine and Human Lifespan" by Michael Scherer et al. introduces a method called EPI-clone to identify both clonal identities and cell differentiation states in haematopoiesis in humans and mice. EPI-clone is based on the scTAM-seq method (Agostina Bianchi, et al., Genome Biol 23, 229, 2022), which allows for the simultaneous measurement of single-cell DNA methylation profiling and surface-protein expression. In this study, the authors use a specially designed set of primers to amplify selected CpG sites that are differentially methylated during different stages of haematopoiesis. They also use a set of oligo-tagged antibodies to identify cell surface protein markers associated with cell differentiation.

By studying murine haematopoiesis, the authors demonstrated that EPI-Clone can identify CpGs correlated with differentiation and clonal identity, referred to as Dynamic and Static CpGs, respectively. The accuracy of Static CpGs is compared to the ground truth orthogonal LARRY lentiviral barcodes, and the accuracy of Dynamic CpGs is assessed using cell surface protein expression. Static CpGs show high correlations with the ground truth LARRY lentiviral barcodes in blood stem and progenitor cell fate in transplantation experiments and slightly weaker correlations with mature immune cells. The applicability of this method is also demonstrated in aged vs young mice, and aged human male donors supporting EPI-clone's utility in important areas of biology (ageing and clonal expansion). The authors discuss this method as an alternative to synthetic barcoding systems (which are not always feasible, such as in human research) and argue its superiority compared to other approaches based on mitochondrial DNA mutations and WGS.

While this study presents fascinating findings with the potential to advance synthetic barcode-free lineage tracing in human and murine models, there are some limitations that currently hinder its broader applicability. Were they to be addressed, one envisages this technique could be widely adopted by the scientific community, progressing the analysis of lineage relationships of tissue development and homeostasis, cancer, inflammation and numerous other fields.

We would like to thank the reviewer for their detailed evaluation of our work and encouragement to demonstrate broader applicability. We now

- demonstrate application of EPI-Clone to a human cohort of 13 donors of different age. These analyses are focused on a comparison between clonal hematopoiesis (CH) and clonal expansions without known drivers, and we carefully evaluate and validate the ability of EPI-Clone to track clonal expansions in that setting (Figure 5+6).
- demonstrate that EPI-Clone works in most mature immune cells in mouse and human (Figure 3g-j).

- demonstrate application of EPI-Clone to endothelia. Surprisingly we found that the same set of dynamic and static CpGs as in hematopoiesis capture cell state and, potentially, clone (Figure S7).
- combine EPI-Clone with mitochondrial lineage tracing from the same cell, to further validate and compare the two approaches (Figure 6e-g and S15)
- combine EPI-Clone with targeted RNA-seq from the same cell, to enable a transcriptomic characterization of clones, and to refine cell state analyses (Figure 6c,d and S14).

We have also addressed all point regarding bioinformatic analyses and readability.

1. Applicability Beyond HSPCs: The study is mostly restricted to HSPCs, and its applicability for analysis of downstream haematopoietic cell types and other tissues/ cell types is not clear. Below is a list of several suggestions but not all need to be addressed. For example, the authors could either focus on EPI-clone's utility beyond haematopoiesis (i.e. other tissue types), or deeper within haematopoiesis (including capturing clonality within more mature lineages and young human donors)

a) The authors could demonstrate the applicability of EPI-Clone for at least one non-hematopoietic tissue in humans and mice. This would demonstrate the versatility and robustness of the method across different tissue types. This might also help determine if there are overlapping (universal?) static CpGs.

We now demonstrate that the same panel developed for mouse hematopoietic tissue can be used to profile cell state and, potentially, clone in endothelia:

*Finally, we asked whether EPI-clone would also be applicable to tissues other than blood. We therefore applied the same CpG panel to sorted endothelial cells (ECs) from lung, which share a common developmental origin with blood (experiment M.6, Figure S7A). Using the dynamic (differentiation-associated) CpGs defined in hematopoiesis, we identified three main clusters (Figure S7B). Based on the surface expression of CD31, Sca-1 and Podoplanin these were identified as highly (CD31+Sca1-Podoplanin-) or lowly (CD31+Sca1+) abundant Capillaries (labelled as capillary 1 and 2, according to scRNA-seq references), and more rare Lymphatic (CD31+Podoplanin+) endothelial cells (Figure S7C-F). Using the same set of static CpGs as in hematopoiesis, EPI-Clone revealed cell state independent, yet statistically supported clusters, which we interpret as clones (Figure S7G, H, see also methods section EPI-clone of endothelial data). Notably, most clones behaved in a similar manner, majorly contributing to generate the most abundant capillaries rather than other endothelial cells, independently of the clone size (Figure S7H). **These observations suggest that a similar set of static and dynamic CpGs defines clones and differentiation states, respectively, in endothelia and hematopoiesis (Figure S7I).***

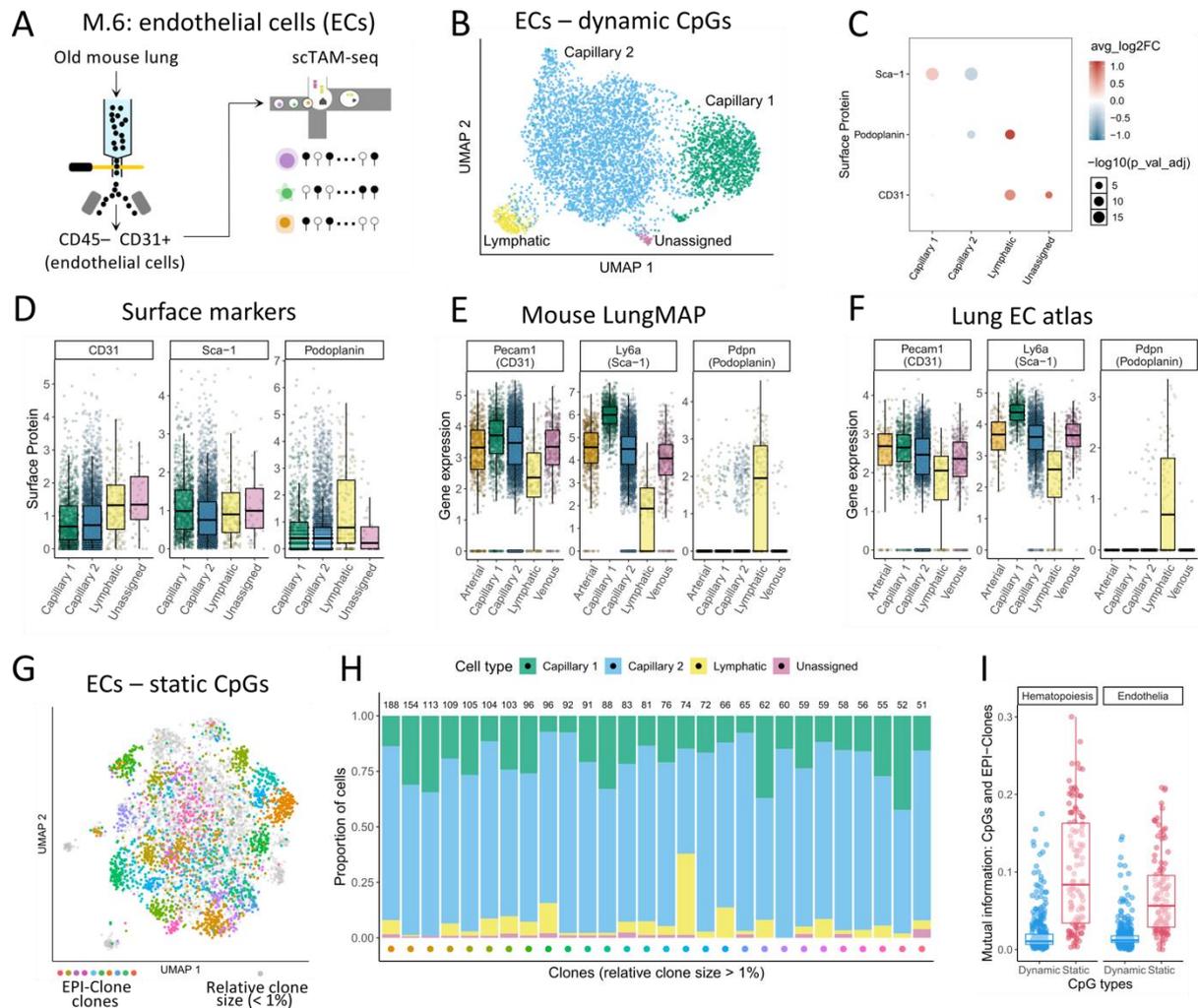


Figure S7, reproduced here for convenience: Cell type mapping and clonality of lung endothelial cells by scTAM-seq and EPI-clone. **A.** Lung cells were isolated from an old mouse, then purified and sorted to filter out CD45+ cells and enrich for CD31+, before profiling with scTAM-seq. **B.** UMAP embedding and low-resolution clustering of endothelial cells using the dynamic CpGs identified in experiment M.1. **C.** Differential expression analysis of surface markers in the different clusters from panel B. **D.** CLR-normalized expression values of surface markers across the different clusters. **E.** Normalized expression of the corresponding genes (scRNA-seq) for endothelial cells from the Mouse LungMAP, only for adult samples⁵³. **F.** Normalized expression of the corresponding genes (scRNA-seq) for endothelial cells from the lung EC atlas⁸⁰. **G.** UMAP computed on static CpGs (identified in experiment M.1). Colors highlight clones identified by EPI-Clone with a relative clone size greater than 1%. **H.** Barplot of endothelial cell types contributions across clones; again, only EPI-clones with a relative clone size greater than 1% are visualized; numbers in the top of the bars represent the absolute clone size, i.e. number of cells. **I.** Mutual information between methylation status of all CpGs and the EPI-clones for endothelial and hematopoietic cells.

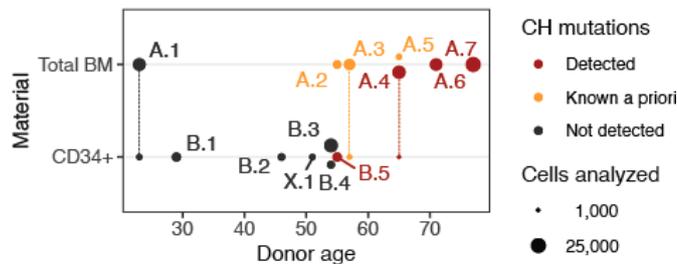
A limitation here is that we do not have clonal ground truth information for endothelia. Nonetheless, the appearance of clusters in static CpGs that are not associated with differentiation makes it likely that epimutations can also be used for lineage tracing in other tissues.

b) The authors could also test EPI-Clone in a young human donor/s, as it is currently tested only in two aged donors where clonal haematopoiesis/oligoclonality is a factor (i.e., the number of clones is

very limited while the size of clones is quite large). This would provide a more comprehensive understanding of the method's applicability across different age groups, not only aged humans.

Thank you for this suggestion. We have now substantially expanded the human study to include a total of 13 BM donors of different age. This dataset is described in detail in the revised manuscript, figures 5 and 6.

A



New figure 5A, reproduced here for convenience. Donor characteristics including donor age, presence of CH mutations and number of cells analyzed. See also Methods, Human samples and table S1. Dots connected by dotted lines denote samples that were analyzed as part of the TBM and the CD34+ dataset

To concisely address this point of the reviewer: The availability of ground truth CH and mitochondrial mutations allows us to estimate that EPI-Clone detects clones with a relative size of at least 1%.

(Results, line 355-357)

To establish a conservative estimate for a minimum clone size of EPI-clone in human, we determined the smallest CH clone identified by EPI-clone. This clone, DNMT3A C666Y in donor A.4, had 145 cells or a relative size of 1% in the myeloid compartment.

With this in mind, we observe that the number of clones identified by EPI-Clone in our human cohort steeply increased at an age of around 60 years (new Figure 5f), in line with an increase in oligoclonality at that age. Bone marrow of young donors is highly polyclonal. We now included two young donors, of age 23 (TBM) and 29 (CD34+ BM), for whom EPI-clone identified one and six expanded clones. We also include PBMCs from a 38 year old donor. Here, we observed 10s of expanded T- and NK-cell clones, validated by mitochondrial mutations (see also below). In sum, we conclude that EPI-clone is applicable to different age groups, and that young humans in bone marrow only have few clones that are sufficiently expanded for EPI-Clone to detect them.

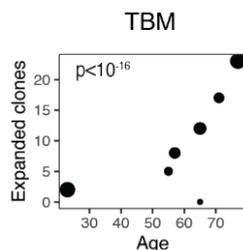


Figure 5F, reproduced here for convenience: Scatter plot relating donor age to the number of clones identified by EPI-clone, in the TBM cohort. p-value is from a generalized linear model of the poisson family, using number of cells observed as a weight.

c) As mentioned, the study does not cover many mature blood cell types. The authors could consider the performance of EPI-Clone for clonal identification (i.e., static CpGs) for lymphocytes, where T or B Cell receptor (TCR/BCR) repertoire could additionally provide a ground truth barcode after activation/expansion. Also, a broader range of myeloid subtypes in addition to monocytes (e.g. neutrophils, eosinophils, dendritic cells etc).

We now evaluate EPI-Clone's performance on various immune cell types in detail, in mouse and human.

In the context of the mouse study, we

(Results, line 223-233)

asked whether EPI-Clone could also determine clonal identity in mature immune cells. To that end, we collected mature immune (lymphoid and myeloid) cells from bone marrow and spleen (experiment M.5, Figure 3g, Figure S5, Table S1) and profiled surface antigen expression as well as DNA methylation at the same CpGs as in experiments M.1-M.3. Using the static CpGs defined from experiment M.1, EPI-Clone again yielded a clonal clustering that recapitulated ground truth clonal labels (Figure 3h). We separately computed ARIs between EPI-Clone results and LARRY barcodes. ARIs were higher than 0.7 for monocytes, neutrophils, other myeloid cells, CD8+ T cells, and one B cell subset, higher than 0.4 for CD4+ T cells, and low for macrophages and a second B cell subset (Figure 3i, j). Most T- and B-cells belonged to lymphoid-dominated (LARRY- and EPI-)clones (Figure 3i, Figure S5d), suggesting an origin in lymphoid-biased or restricted progenitors⁵².

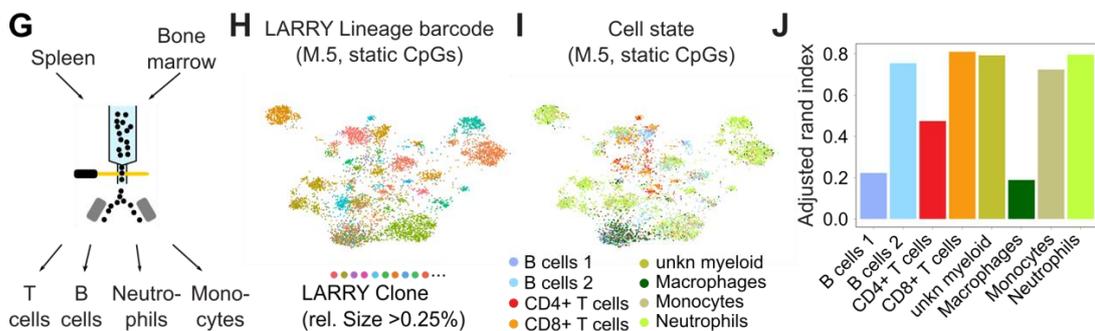
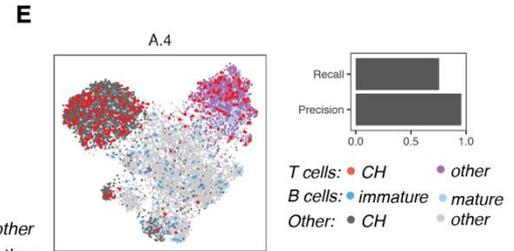
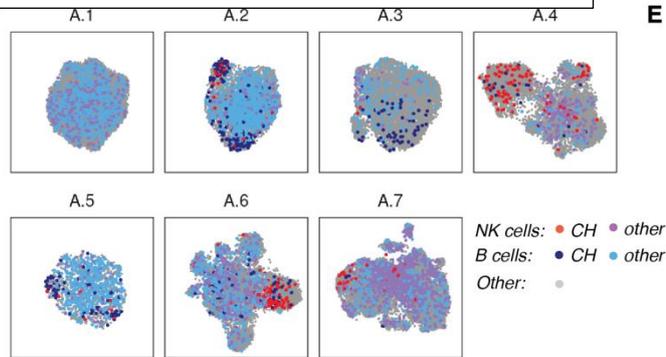
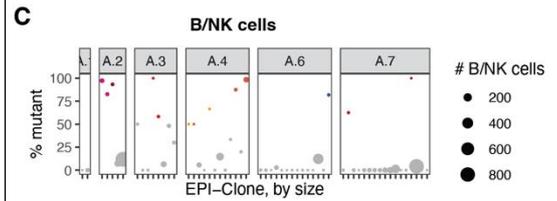


Figure 3g-h, reproduced here for convenience. **G.** Scheme illustrating experiment M.5: LARRY mature immune cell experiment. **H.** DNA methylation UMAP representation based on the static CpGs for cells from expanded clones in experiment M.5. Cells are colored by LARRY barcode. The static CpGs identified from experiment M.1 were used. **I.** Same UMAP representation as in H highlighting the cell state annotation as defined in Figure S5. Of note, most EPI-clones were specific for T, B or myeloid cells, in line with the result from LARRY (Figure S5d). **J.** Adjusted rand indices between the ground truth clonal label (LARRY) and the clones identified by EPI-Clone stratified by cell type.

In human, our findings were in line with this observation and are now described as follows (line 346-353):

We included NK and immature B cells in our analysis and used CH mutations to validate that these cells also clustered by clone (Figure S13c,d). When T and mature B cells were included, they associated with lymphoid-dominant clusters, in line with the result from mouse (Figure S13e, and cf. Figure 3i) and suggesting distinct clonal origins compared to the other cells. In patient A.4, where a large CH clone contributed to T cells, mutant T cells mostly clustered with the remaining CH-derived cells (Figure S13e). Together with the mouse LARRY experiment, this constitutes evidence that EPI-clones remain stable from HSC to myeloid, T, NK, and immature B cells.

Figure S13C-D, reproduced here for convenience. C. Scatter plot displaying the percentage of cells from each EPI-Clone displaying CH mutations, for NK and immature B cells. EPI-Clone was run on all cells except T and mature B cells, but the overlap was computed on NK and immature B cells only. See main figure 5c for color scheme. D. Static CpG UMAPs as in main figure 5c/d, highlighting NK and immature B cells classified according to CH status. E. Static CpG UMAP computed for all cells (including mature B and T cells) for patient A.4, highlighting T cells classified according to CH status. Mature and immature B cells are also highlighted to demonstrate that mature B and T cells mostly cluster in lymphoid clusters. Barchart depicts precision and recall for the task of classifying T cells as CH or non-CH based on EPI-Clone labels.



We further use mitochondrial lineage tracing to validate that EPI-Clone had correctly identified several expanded T cell clones in a PBMC sample (Figure 6e-g and see also below).

In conclusion, EPI-clone works well in myeloid cells, NK cell, T cells and immature B cells. The mature B cell compartment in human is highly polyclonal, and even in autoimmune disease, expanded B cell receptor clones make up <<1% of B cells outside of the plasma cell compartment (ref 10.1038/s41586-019-1595-3). We cannot currently explain why EPI-Clone does not work in mouse macrophages, we suspect that they may have “eaten” DNA from other cells. Figure S6m provides some evidence for that hypothesis:

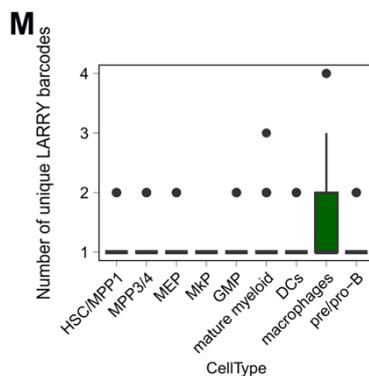


Figure S6m, reproduced here for convenience. Number of unique LARRY barcodes per cell type cluster. The elevated number of LARRY barcodes per cell in the macrophage cluster suggests the presence of contaminant DNA from doublets or phagocytosis in this cluster.

d) Although not essential, for human studies, an analysis of EPI-clone’s utility using barcoded human hematopoietic stem and progenitor cells (HSPCs) could be conducted to compare its performance with the ground truth barcodes (e.g., LARRY). We are also concerned about the quality/resolution of data afforded by their antibody panel (see 3a)

While we considered xenotransplant experiment unrealistic in the time frame of these revisions, our expanded human cohort displayed a total of 11 CH mutations and LoY events, which served as a partial ground truth and allowed us to determine that EPI-clone identified these clones well, down to a relative size of 1% (Figure 5c-e). We now also include mitochondrial somatic variants as an additional validation (see the reviewer’s next point).

We do not share the concerns on antibody data quality. In the context of the mouse study, antibody data enables accurate separation of progenitor populations (HSC, MPP1-4: Figure 1g). In the context of the human CD34+ data, we now identified CD34+CD38- HSC/MPPs, CD34+CD38+/-CD71+ EMPP/MEPs and CD34+CD38+CD45RA+ GMPs and CD34+CD38+CD41+CD61+ MkPs (Figure 5b, S12d,e). We further validate these cell state annotations by integrating RNA-seq into EPI-clone (see the reviewer's point 3a and minor point 6).

2. Benchmarking EPI-Clone's Performance:

a) The paper argues its superiority over other approaches based on mitochondrial DNA mutations. A systematic comparison with mtDNA-based approaches such as Weng, C., et al. (Nature 627, 389–398, 2024) would provide a clearer picture of the relative strengths and weaknesses of EPI-Clone.

We now include an experimental comparison.

(Results, line 398-418)

In the field, there is some controversy regarding the potential of other somatic events, in particular low-heteroplasmy mtDNA variants, for lineage tracing²²⁻²⁴. To perform a direct experimental comparison, we analyzed peripheral blood from a 38-year-old healthy donor (X.2) that had previously been characterized by a state-of-the-art single cell mitochondrial lineage tracing method, mt-scATAC-seq^{20,65}. This donor had 44 EPI-Clones, and in particular displayed prominent clonal expansions of NK- and T-cells (Figure S15a). By including a mitochondrial targeting panel into scTAM-seq, we achieved a median coverage of 176 reads per cell on the mitochondrial genome (Figure 6e, S15b,c). Of the 23 mtDNA variants previously identified⁶⁵ in this donor, five had clear phylogenetic relationships with EPI-Clones, i.e. were either subclones to single EPI-Clones, or were parental to several EPI-Clones (Figure 6f), and one variant was observed in two EPI-Clones. A highly abundant variant (mt:7076A>G) was strongly enriched or depleted in 17 T- or NK-cell EPI-clones, but was observed in a subset of cells of the remaining, mostly multilineage- or B-cell, EPI-clones (Figure 6g). This variant was likely present before epimutational patterns were established and repeatedly underwent selection at late stages of cell differentiation. Finally, the remaining 16 low-heteroplasmy mitochondrial variants did not segregate with EPI-clones (Figure S15d,e). These findings are in line with a recent preprint²³ observing that only some observed mitochondrial variants carry phylogenetic information, and illustrate the complexity of mitochondrial genetics, where selection of variants can happen during differentiation⁶⁵. They further provide additional orthogonal validation of EPI-Clone outside of the setting of CH.

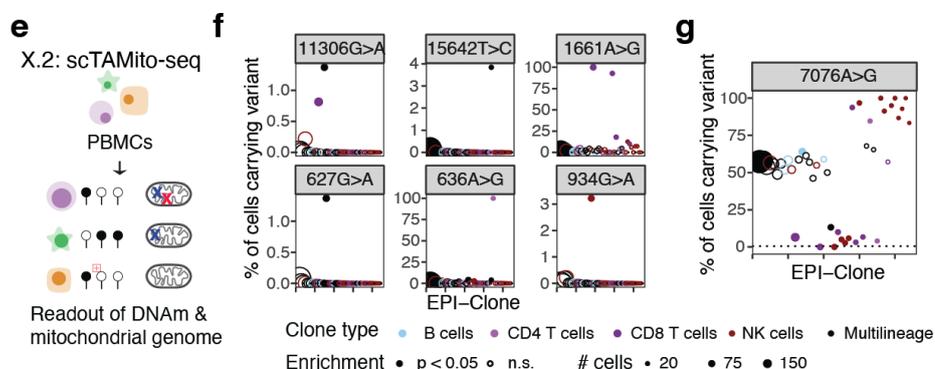


Figure 6e-g, reproduced here for convenience. E. Scheme illustrating experiment X.2. Mitochondrial mutations were read out together with DNA methylation states with scTAMito-seq. A PBMC sample characterized previously for mitochondrial mutations⁶⁵ was

analyzed. See also figure S15. **F.** Scatter plot depicting the presence of six mitochondrial variants in the different EPI-clones from X.2. Cells were scored as positive for the variant, if at least 5% of reads supported the variant. The enrichment of variants in EPI-Clones was determined by a binomial test. EPI-Clones were classified as B, T or NK cell clones if at least 80% of cells were from a single lineage, or as multilineage clones otherwise. **G.** Like F, for the mt:7076A>G variant.

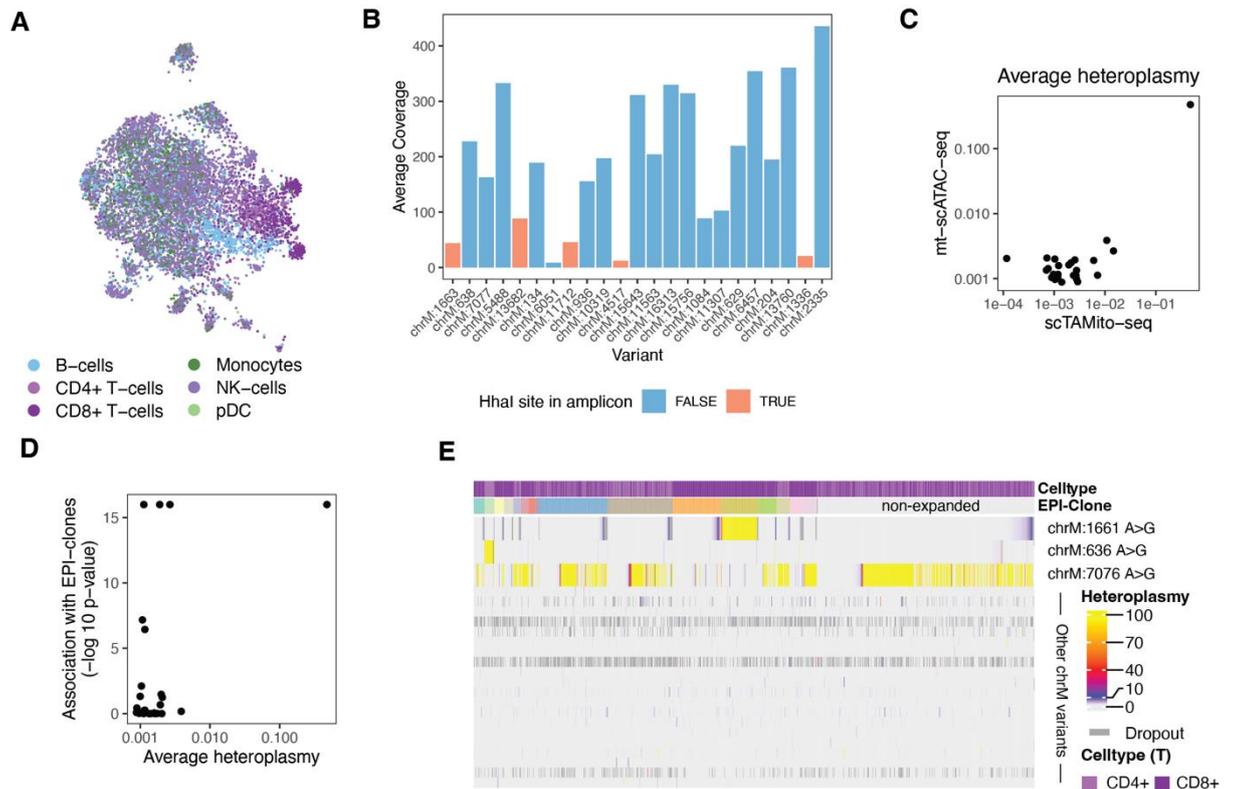


Figure S15, reproduced here for convenience. Comparison of EPI-Clone and mitochondrial lineage tracing by scTAMito-seq. A. Static CpG UMAP computed on all cells from the patient, highlighting cell types identified using surface antigen expression levels. **B.** Average coverage in reads per cell for the mitochondrial variants previously described for donor X.2⁶⁵. **C.** Scatter plot comparing average heteroplasmy for these mutations, as determined by mt-scATAC-seq (reference⁶⁵) or scTAMito-seq (this study). **D.** Scatter plot depicting, for all mitochondrial variants, the average heteroplasmy and the statistical association with EPI-Clone. Specifically, a linear model was trained on EPI-Clone clusters to predict heteroplasmy at the single cell level, and the *p* value from an *F*-test is shown. **E.** Heatmap relating the single-cell heteroplasmy of mitochondrial variants to EPI-Clones, for T cells only. The columns correspond to different T cells and the rows comprise mitochondrial mutations measured by scTAMito-seq.

Regarding the ReDeeM paper (ref. 22: Weng, C., et al. Nature 627, 389–398, 2024): ReDeeM performs poorly on ground truth genetic labels (ARI 0.2-0.7, typically around 0.4, this study: 0.7-0.88 in different biological settings), was not validated using CH mutations (this study: extensive validation of EPI-Clone’s ability to detect clonal expansions >1%), and produced interesting results on loss of Y (a weakly significant enrichment of LoY in one “clade”, with an enrichment *p* value of 0.009, which essentially would mean that LoY is mostly not heritable:

Figure 5f of Weng et al.). A recent preprints highlights several issues with the computational analyses of Weng et al (our ref 24: DOI: 10.1101/2024.07.28.605517).

b) Although not essential, studying EPI-clone in a native setting in mice using one of the available in-vivo barcoding mouse models instead of transplanting barcoded cells into irradiated mice would provide a better benchmark in case transplantation is amplifying the reliability of the method and would help validate the method in a more physiologically relevant setting. To this point, when we visually assessed EPI-clone signatures in the authors' data set, we noted that the Transplantation experiments give visually 'stronger' clonal information compared to cells from a native setting (attached).

We would like to thank the reviewer for the detailed investigation of our method. Indeed, the transplantation setting, as opposed to native hematopoiesis show more structure in the clonal UMAP as expected, since fewer clones contribute to blood production, which are then also larger in size. Using in-vivo barcode mouse models is beyond the scope of the current manuscript, since it would require the design of such models being compatible with scTAM-seq readout (i.e., not harboring the HhaI-cutsequence). On the other hand, we now have three strong validations for EPI-Clones performance: (i) Association with the LARRY barcode in mice and (ii) Association with CH mutations and loss of Y in humans and (iii) mitochondrial variants. Taken together, we are convinced that these validations are sufficient to benchmark EPI-Clone's performance also in a native setting.

3. Technological Improvements for parallel cell type/state identification:

The current resolution of cell state, based on dynamic methylation sites and/or antibodies is quite restrictive for broad take up.

a) For antibodies, is this due to sequencing depth and/or too small a panel? Could the authors generate uMAPs only using cell-surface protein expression to demonstrate the efficiency of the panel used for cell identification, or use a more comprehensive panel (e.g. TOTAL-seq)? This would ensure comprehensive haematopoietic cell type coverage and improve the accuracy of cell state identification (related to Point 1).

In the context of the mouse data, we used a very small panel of antibodies and noted that the cell state resolution obtained by scTAM-seq closely recapitulated a widely used gating scheme used to identify HSCs and MPP subsets based on SLAM family markers (Figure 1g).

We have now added a UMAP using cell surface protein expression (Figure S3c,d).

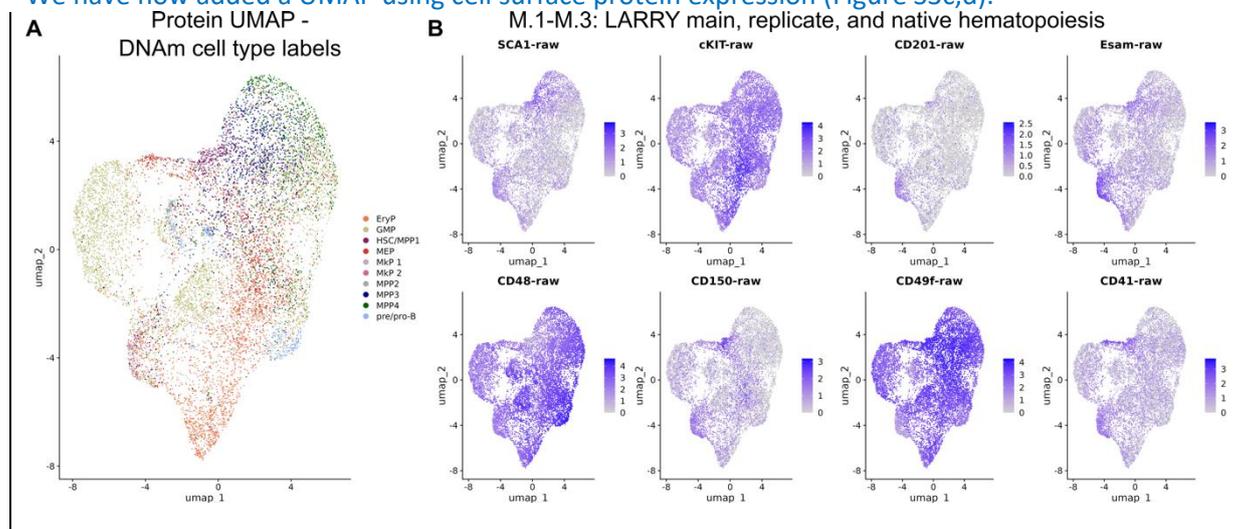


Figure S3C+D, reproduced here for convenience: **C.** Surface protein UMAP of experiment M.1 (13,885 cells) with the cell type labels obtained from the DNA methylation UMAP as shown in Figure 1c. Protein data was normalized using SCTransform prior to generating a low-dimensional representation with PCA and UMAP. **D.** Expression of selected surface proteins in the protein UMAP

In the context of the human experiment, we used the TotalSeq-D HemeOncology cocktail comprising 45 antibodies (<https://www.biolegend.com/de-de/products/totalseq-d-human-heme-oncology-cocktail-v10-20465>). In the human experiments, we observed that the antibodies alone achieve a good cell state resolution across immune cell types, but displayed stronger technical variation related to staining quality. The DNA methylation amplicons achieved a somewhat weaker cell state resolution, but displayed less technical variation. Defining cell states on both modalities gave an excellent cell state resolution robust to inter-individual and technical variation (S12c-e).

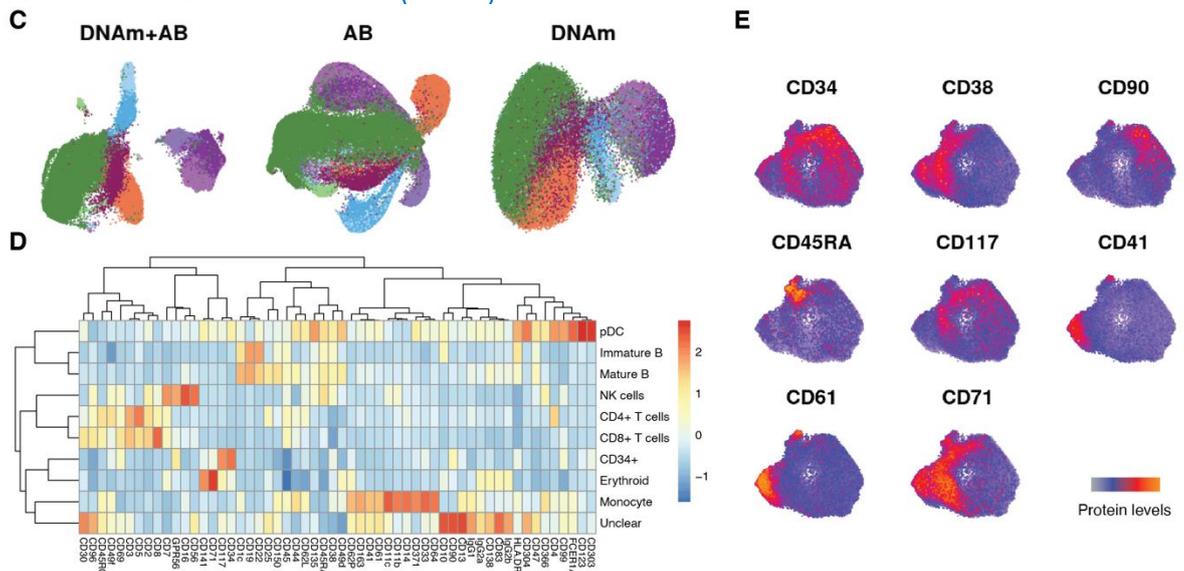


Figure S12C-E, reproduced here for convenience. **C.** Cell state clustering for the TBM cohort using antibodies, DNA methylation or both modalities. Colors correspond to clustering on the DNA methylation (DNAm)+AB data, see main figure 5b for color scheme. UMAPs were generated using data integration by scanorama across donors from the TBM cohort. **D.** Average protein expression levels in the different clusters, for the TBM cohort. **E.** UMAPs of the CD34+ cohort highlighting the surface expression of various antigens. See also main figure 5b.

b) Regardless of 3a) capturing transcriptomes and static CpGs in the same single cells would massively increase EPI-clone's utility for the scientific community. Although the authors have performed scRNA-seq on the same samples, correlating these profiles demonstrates a similarity in the uMAP landscapes, but not for the same cells. Including scRNA-seq on the same cells would be highly advantageous, providing a more comprehensive view of cellular states and enhancing the interpretability of the results. The authors could consider practical ways to achieve this, perhaps the recent preprint from the Landau lab, or other approaches.

We agree that inclusion of RNA-seq can be beneficial in some scenarios (e.g. if populations with no known surface markers need to be characterized, to increase confidence in antibody-based cell type identification, or to characterize transcriptional differences between clones). In the manuscript, we focus on the latter two of these use cases:

(Results, lines 384-395)

To resolve transcriptional differences between clones within the HSC/MPP compartment, we added targeted RNA-seq to the scTAM-seq protocol (single cell targeted Analysis of the Methylome And RnA, scTAMARA-seq, Figure 6f, S15a). To that end, we combined SDR-seq⁶⁶, a recently described targeted RNA-seq protocol for Mission Bio Tapestry, with scTAMseq. We profiled one of the CD34+ BM samples (X.1), obtaining high-quality DNA methylation and RNA-seq data from 2,745 cells (Figure S15b-e). scRNAseq data confirmed the DNA- methylation-based cell state annotation and showed a higher resolution at the level of erythromyeloid progenitors (Figure S15f,g). We then investigated the gene expression pattern of distinct clones. HSC/MPP-biased clones expressed less TAL1/SLC40A1/CDC45 at the HSC/MPP level and more CEBPA, suggesting that clonal fate biases are correlated with gene expression changes at early stem and progenitor states (Figure 6g). Our results further demonstrate the compatibility of EPI-Clone with targeted RNaseq from the same cell.

The SDR-seq manuscript is currently undergoing a 2nd round of review. The most recent version of that manuscript is included in this submission as a “related manuscript file”.

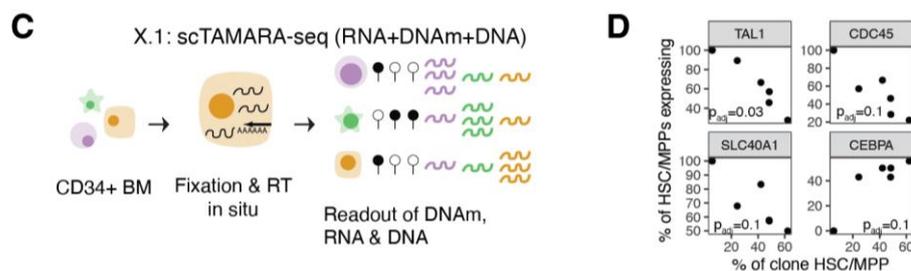


Figure 6c,d, reproduced here for convenience: **C.** Scheme illustrating the scTAMARA-seq protocol, see also Figure S15 and Methods. **D.** Clones were identified on CD34+ cells from donor X.1 using the DNA methylation data from scTAMARA-seq. Subsequently, genes with differential expression between clones were identified. Scatter plots depict genes identified as significantly DE between clones whose expression was additionally correlated with the fraction of HSC/MPP cells of the clone. P values are from Pearson correlation, adjusted for multiple testing.

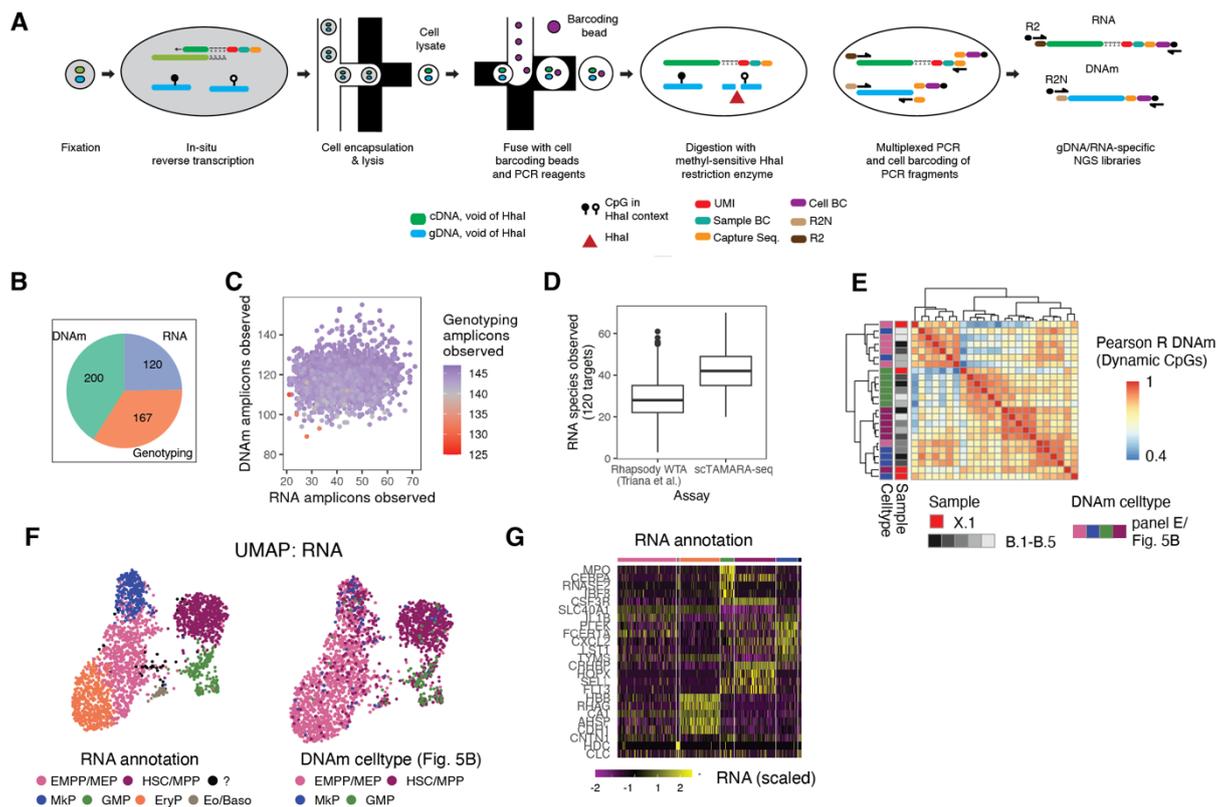


Figure S14, reproduced here for convenience: scTAMARA-seq enables multiplexed readout out RNA, DNA methylation and genotyping amplicons from the same single cell. A. Scheme of the method, adapted from⁶⁶. **B.** Composition of the panel used, see table S5. RNA-seq amplicons were selected using a scRNA-seq reference⁷⁴ to identify the set of 120 genes with highest information on cell states in the CD34+ compartment by LASSO regression. **C.** Scatter plot depicting the number of RNA, DNA methylation (DNAm) and genotyping amplicons observed per cell. **D.** Boxplot comparing the number of features (RNA species) observed per cell in scTAMARA-seq to the number of features observed in whole transcriptome analysis (WTA) on CD34+ cells for the same 120 genes⁷⁴. **E.** Heatmap depicting correlation in DNA methylation profiles between sample X.1 and the other CD34+ BM donors. **F.** UMAPs computed on the RNA information from scTAMARA-seq highlighting cell state annotation based on RNA (left) and based on DNAm (right). **G.** Heatmap depicting scaled expression of marker genes for the different RNA-based cell states.

4. Algorithm Improvements:

a) Statistical Power: In the EPI-Clone algorithm for computing the minimum p-value for association with any protein, a minimum sample size of 3 cells per group can limit the statistical power of the Kolmogorov-Smirnov Test (ks.test). Given the noisy nature of protein expression data, the authors could consider increasing the sample size. This might enhance the reliability of the statistical associations identified.

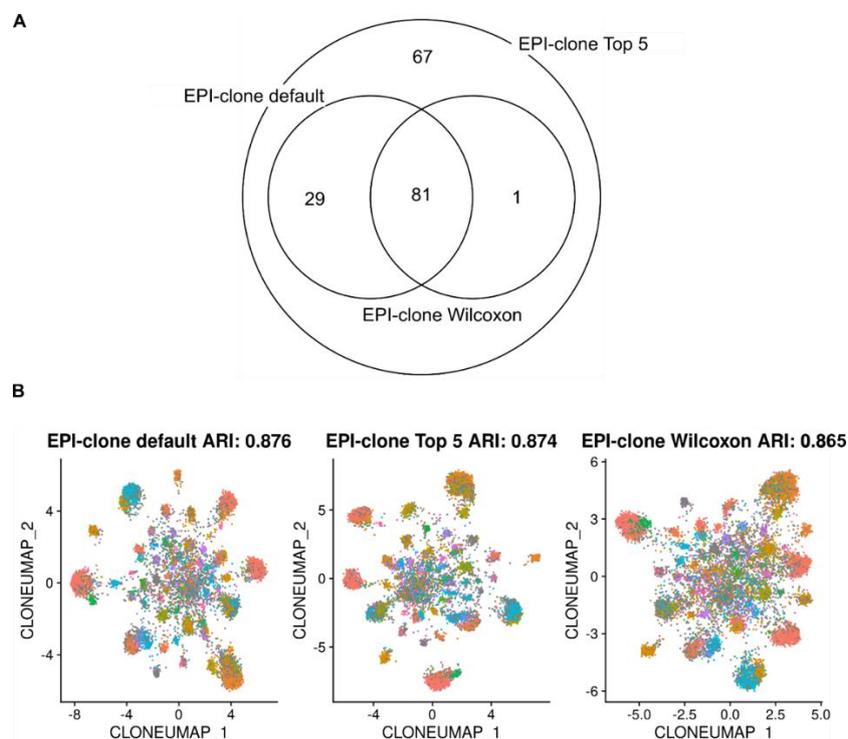
b) Minimum p-values: When analysing the associations between methylation status and protein expression, the algorithm runs a Ks.test for each protein to obtain p-values for the significance of the association. For each methylation site, the minimum p-value across all proteins is selected to identify the most significant association. However, it is important to note that when p-values are extremely small, the exact differences between them may not be practically significant, as they all indicate a strong association. Therefore, while the minimum p-value is used to highlight the most significant association, and authors used this strategy to summarise the data, analysing top-n p-values can also be considered for a more comprehensive analysis.

c) Tied-values: In the context of the KS test, having many tied values can affect the test's performance. The frequency of tied values could change by applying different normalization methods to the raw data. The authors need to inspect this to ensure the test's performance is not affected. If necessary, they should consider alternative statistical tests. This would ensure the robustness of the statistical analyses.

d) Data Normalization: The normalization method provided in the code (ScaleData from Seurat) is robust and widely used for single-cell data. However, alternatives like SCTransform from Seurat might offer better performance for dealing with noisy protein data (Hafemeister, C., Satija, R. Genome Biol 20, 296, 2019). This might improve the reliability of the results.

These points a-d regard the selection of CpGs as static or dynamic, based on statistical association with surface protein expression. We appreciate the reviewer's detailed thoughts and suggestions. The exact set of static CpGs used does not affect the methods performance; we have now increased the minimum sample size (a), also selected CpGs using top-n p-values (b), and we tried a different statistical test (Wilcoxon test instead of KS test) (c). These changes do not affect the performance of EPI-Clone (Response figure below). In response to the reviewer 7th point, below, we have also randomly subsampled the static CpGs and determined robust performance over a wide range. Together these analyses demonstrate the robustness of our algorithm to specific choices in the selection of static CpGs.

Regarding suggestion (d), use of SCTransform, while this did not noticeably impact the choice of static CpGs, it improved the cell state resolution of UMAPs obtained from surface protein data, which we have now included as figure S3a (see also the reviewer's 3rd point)



Response figure: Comparison between different CpG selection methods. A: Venn Diagram describing the overlap between three CpG selection methods: EPI-clone default setting as described in the manuscript; Top 5 p-values, which takes the mean of the 5 lowest association-p-values with proteins; Wilcoxon uses the Wilcoxon test instead of the KS-test. B: Results for EPI-clone using the different sets of static CpGs described in A. The color indicates

LARRY lineage tracing barcodes. ARI: Adjusted rand index. EPI-clones performance is not influenced by the set of CpGs selected.

e) Subclonal structure: Did the authors attempt to resolve clonal/subclonal relationships within a data set? This would be advantageous for the study of phylogeny and e.g. cancer evolution in the future (especially where CNVs are not prevalent).

We observe interesting instances of “sub-clones”, for example, in the context of the human CH clones (Figure 5c,d):

(Results, line 357-359)

We observed that several large CH clones (e.g. DNMT3A R659H in donor A.4) had diversified into two EPI-clones with a similar, but distinguishable static CpG profile, suggesting that over decades epimutations can continue to accrue phylogenetic information.

We also observe mitochondrial clones that had diversified into several EPI-clones (Figure 6f,g).

We currently want to avoid explicit statements on phylogenies between EPI-Clones as we lack a strong ground truth for these statements. It is easy to draw phylogenetic trees and very hard to validate them. Most validation efforts in this manuscript can confirm that clones are tracked correctly, but cannot be used to make statements on their phylogenetic relationships.

We now synthesize our observations into the following discussion:

(Discussion, line 431-446)

This raises the question of where and how clonal epimutations arise. We found that they (i) randomly occur, but remain stable over many cell divisions, (ii) their number does not increase during differentiation (Figure S16), and (iii) they are enriched for heterochromatic and late-replicating domains. We propose that some developmental events that are characterized by rapid cellular proliferation and/or a remodeling of the methylome, such as the specification of HSCs⁶⁷ essentially randomize the methylation state of CpGs in heterochromatin and late replicating regions. As a potential explanation, in fast-dividing cells DNMT1 may act insufficiently to copy the DNA methylation state to the nascent DNA strand. Consistent with this, a recent study of bulk methylome profiles from blood cells in monozygotic twins suggests that clone-associated variation of the methylome may be established during embryonic development²⁶. In the case of some large CH clones, we observed additional diversification of epimutational patterns.

We therefore propose that variably methylated CpGs in non-regulatory genomic regions can act as a digital barcode of clonal origin. The digital and stochastic nature of epimutations makes single-cell methods that are capable of mapping methylation state of single CpGs at high confidence, such as scTAM-seq, a powerful tool for lineage tracing.

5. Blind Design: Since LARRY barcodes are used as the ground truth, the authors could conduct the analysis in a manner where EPI-Clones are blind to the LARRY barcodes. If this approach was used in the current study, please discuss it in detail. Otherwise, for future experiments, we suggest implementing a blind design (train and test) to prevent potential biases and circular logic in the analysis. This would enhance the robustness and credibility of the study's findings.

EPI-clone parameters were determined only for the initial LARRY experiment (M.1: LARRY main experiment). For all subsequent experiments (e.g., M.2: LARRY replicate, figure S4d,e; M.4: LARRY mature myeloid cells, Figure S5 and M.5: LARRY mature immune cell experiments, figure 3g-j), the established parameters were used without regarding their association with the LARRY barcodes, effectively being a “blind” design for the LARRY barcodes. We now include the following sentence in the methods part of the manuscript:

(Methods, line 1234-1236)

The parameters of step (ii) and (iii) were established on the original LARRY experiment (M.1: LARRY main experiment) and employed for all subsequent analyses of the mouse hematopoietic system (M.2-M.5, M.7, M.8) without further adjustments. Static CpGs were defined in experiment M.1 and used for all remaining experiments. In particular, the performance on a replicate LARRY ground truth experiment (M.2) is analyzed in figure S4.

We now also performed clustering using a parameter-free method (CHOIR, 10.1101/2024.01.18.576317v1) which displayed a performance similar to the optimal EPI-clone parameters (Figure S4g). For the analysis of human samples and endothelia, where no detailed ground truth is available to establish EPI-Clone parameters, we used CHOIR and observed that CH clones (and some mitochondrial variants) are recapitulated well:

Quantitatively, EPI-clones dominated by CH mutant cells were on average 78.8% mutant and EPI-clones dominated by wild-type cells were on average 95.4% wild-type (Figure 5e). These numbers likely underestimate the true overlap between EPI-clones and CH clones, due to allelic dropout of CH mutations.

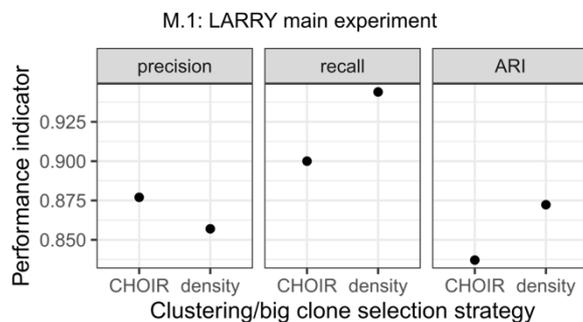


Figure S4g, reproduced here for convenience: Comparison between the performance of the original, density-based clustering of EPI-Clone with the performance of a parameter free method, CHOIR. The results are shown for experiment M.1: LARRY main experiment.

6. LARRY Barcoding: In our exploration of the LARRY barcodes across four mice, we found that mouse 1 and mouse 2 exhibit extremely similar LARRY barcodes, as do mouse 3 and mouse 4,. This similarity is quite unexpected considering barcodes should be random. The authors should provide detailed information on the viral transduction process, transduction efficiency, in vitro expansion, transplantation procedures, and the preprocessing steps used for the LARRY barcodes to understand the observed similarities and ensuring the reliability of the barcoding results.

We apologize that this was not explained in the original submission. We have now added a methods section to clarify what was done:

(Methods, line 841-854)

For validating EPI-Clone using a ground truth genetic lineage tracing experiment, we performed two experiments: the main LARRY experiment (M.1) and the LARRY replicate experiment (M.2). For M.1, two donor mice were sacrificed and HSCs were labeled with LARRY

constructs containing a GFP label in one case, and LARRY constructs containing a Sapphire label in the other case. Subsequently, labeled cells from each donor were transplanted into two recipient mice each. Accordingly, the data set contains cells from four mice that contain two sets of clones, labeled with GFP and Sapphire, respectively. GFP and Sapphire clones did not mix on EPI-Clone UMAPs (Figure S4f), further demonstrating that EPI-Clones are individual-specific. We profiled all four recipient mice after allowing full blood reconstitution over five months. We also repeated this experiment again for validating the computational method (experiment M.2), using only one donor mouse. Both for experiment M.1 and for experiment M.2, we collected LSK and LK cells from the bone marrow and mixed them at 60,000 (LK) plus 50,000 (LSK) before subjecting the cells to the Tapestry technology (Table S1).

7. QC and methylation properties: A useful addition would be to include some QC metrics about the molecular and computational aspects of EPI-clone. In particular, what fraction of all static DNAm sites is detected per cell using scTAM-seq, and in terms of performance in the LARRY experiment, what fraction of the total must be captured to gain ‘meaningful’ clonal information using EPI-clone. And on that note, are some sites more informative than others? Is there a minimal set of sites that captures most information? What is the proportion of cells with Epiclone marks? Perhaps not all cells make random errors during methylation (a rough estimate of probability per cell division or something like that would be useful, if possible). And are some static CpG sites more likely to mutate?

We thank the reviewer for pointing out this very important issue. To follow up on this point, we conducted different analyses. First, we investigated the number of epimutations per cell type in the HSPC experiments and found that epimutations are stably maintained until terminal differentiation, without an apparent increase or decrease with differentiation (new Figure S16b). We also found that, on average, each clone is marked by around 10 epimutations.

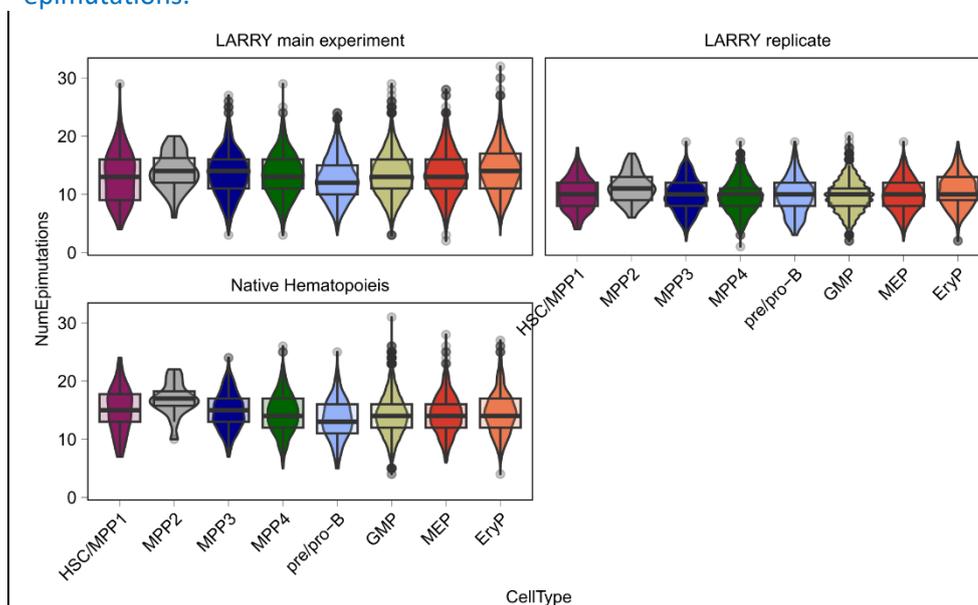
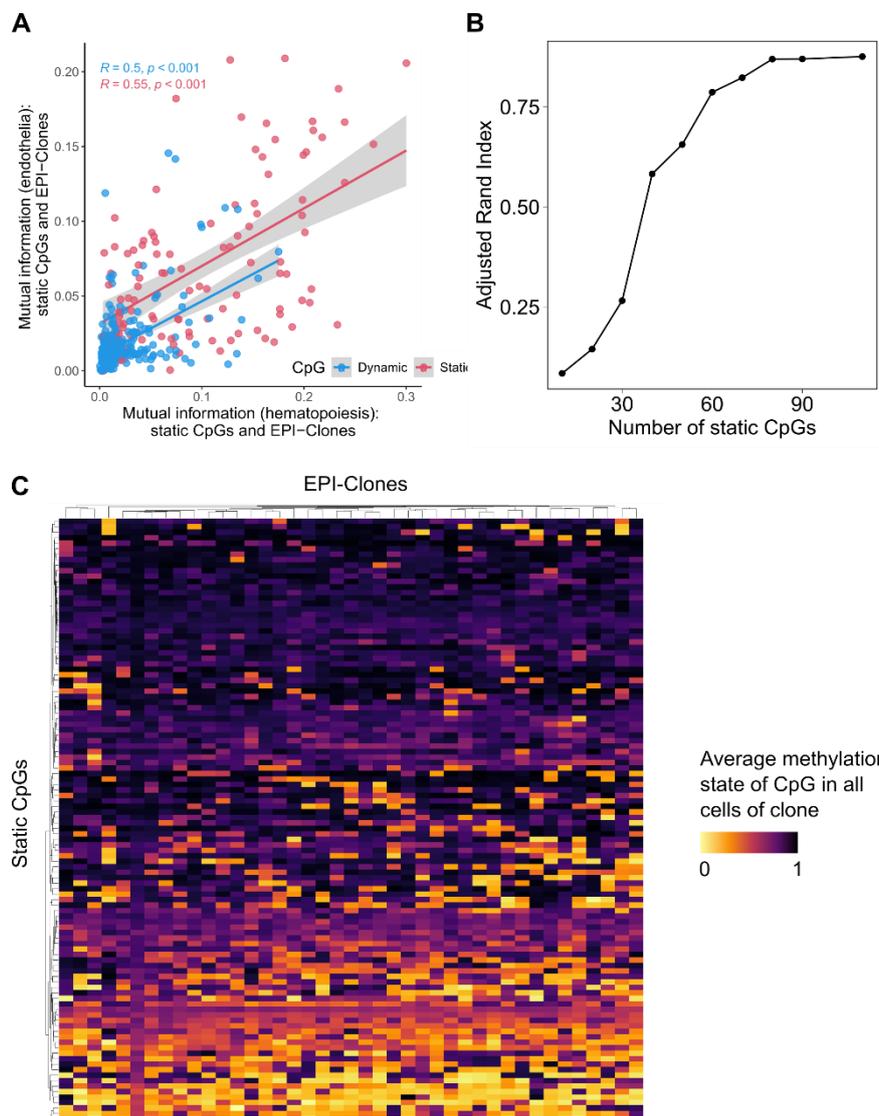


Figure S16B, reproduced here for convenience.: Number of epimutations per cell type for the experiment M.1-M.3. An epimutation is defined as a CpG that shows a DNA methylation state different from the default (i.e. most prevalent across all cells) DNA methylation state of this CpGs according to all profiled cells.

To further investigate whether some CpGs carry more information in marking clonal identity, we computed the mutual information between clonal identity and methylation state of each CpG, both in hematopoiesis and in lung endothelia. Reassuringly we found that the same CpGs that carry a high amount of clonal information in one tissue, also carry a high amount of clonal information in the other tissue (Reviewer figure a, included in a simplified version as figure S7i).

To further investigate the number of CpGs required for successful clustering, we now downsampled the static CpGs to random subsets of size 10-90 and computed the resulting Adjusted Rand Indices. We found that starting from 70 static CpGs, clonal clusters can reliably be identified with ARIs above 0.8 (Reviewer Figure b). Additionally, we found that EPI-Clones are marked by multiple epimutations with variable methylation states across the clones (Reviewer Figure c).



Reviewer figure: A. Comparison between the mutual information between the static/dynamic CpGs and clonal identity defined through EPI-Clone. The y-axis shows data from experiment M.6 (endothelial cells) and the x-axis from experiment M.1 (LARRY main experiment). B. Comparing the number of static CpGs used as input to EPI-clone versus the performance quantified with the adjusted rand index. A random subset of the specified size

(x-axis) of all 110 static CpGs was taken. C. Average DNA methylation state of the static CpGs across all the cells in an EPI-Clone cluster (columns) for experiment M.1.

Minor Comments:

1. Clarify Method Development: The abstract's description, "We develop EPI-clone, a droplet-based method for transgene-free lineage tracing, ..." is ambiguous and might lead readers to assume that scTAM-seq was developed in this study. Please clarify in the abstract and throughout the article where EPI-Clone is introduced, specifying that EPI-Clone builds upon the existing scTAM-seq technology.

We have now changed the wording in the abstract to "We develop EPI-clone, a method for transgene-free lineage tracing based on somatic epimutations of individual CpGs."

2. Same Sample for scRNA-seq: In line 146, "Since we also performed scRNA-seq on the same samples ..."). This would benefit from further disambiguation throughout the article to avoid confusion whether it was the same single cells or 'different cells from the same sample'.

We appreciate the reviewer's suggestion for improving clarity of the experimental setup. In the context of the mouse experiments, scRNA-seq was performed on a different set of cells obtained from the same sample. We reformulated the text as follows:

Since we also performed scRNA-seq on different cells obtained from the same samples, we could compare the DNAm UMAP with a transcriptomic UMAP.

Additionally, we updated Table S1 and Figure S1 to clarify the the experimental setup. Finally, we now also include a human bone marrow sample where RNA and DNA methylation were profiled from the same single cells (see the reviewer's point 3b, above)

3. Literature Citations: Line 46: Perhaps refer to VanHorn S, Morris SA. Next-Generation Lineage Tracing and Fate Mapping to Interrogate Development. Dev Cell. 2021, or similar, and consider a broader range of relevant studies rather than a focus on those of the authors only.

Thank you for pointing towards additional useful citations to further set our work into the context of the scientific community. We now included the following additional references for the first sentence of our manuscript:

"Lineage tracing using genetic or physical labels has been an important tool in developmental and stem cell biology for more than a century¹⁻⁵."

1. Kretzschmar, K. & Watt, F. M. Lineage Tracing. Cell 148, 33–45 (2012).
2. VanHorn, S. & Morris, S. A. Next-Generation Lineage Tracing and Fate Mapping to Interrogate Development. Dev. Cell 56, 7–21 (2021).
3. Baron, C. S. & van Oudenaarden, A. Unravelling cellular relationships during development and regeneration using genetic lineage tracing. Nat. Rev. Mol. Cell Biol. 20, 753–765 (2019).
4. Sankaran, V. G., Weissman, J. S. & Zon, L. I. Cellular barcoding to decipher clonal dynamics in disease. Science (80-.). 378, (2022).
5. Wagner, D. E. & Klein, A. M. Lineage tracing meets single-cell omics: opportunities and challenges. Nat. Rev. Genet. 21, 410–427 (2020).

4. Figure 1C Clarification: Clarify how the uMAP of DNA methylation is highlighted by LARRY barcodes, especially given that these analyses were not performed on the same single cells.

The LARRY barcode was directly read out from the same single cell through a specific primer being spiked-in into the Tapestry platform primer panel. We now clarified this point in the main text and methods section:

(Results, line 146)

The LARRY barcode was read out directly from the DNA by including a LARRY-specific amplicon in our targeting panel for scTAM-seq.

(Methods, lines 1055-1058)

For amplifying the LARRY barcodes, we spiked in an additional primer into the primer pool targeting the LARRY barcode sequence (forward: GCATCGGTTGCTAGGAGAGA, backward: GGGAGTGAATTAGCCCTCCA). We can thus read-out the LARRY barcode together with information about the DNA methylation state from the same single cell.

5. Figure 1H Clarification: Line 130: The description of Static CpGs stochastically gaining or losing methylation (preferentially losing, as shown in Figure 1h) requires more clarification. The current explanation seems counterintuitive and confusing. Please provide a detailed explanation of the methodology and findings to ensure clarity.

We restructured that section and removed, partly speculative, statements on whether methylation has been gained or lost relative to some baseline that we do not know. We had originally alluded to the observation that static CpGs are in general highly methylated in the majority of cells and specifically de-methylated in some of the clones. However, this is a minor point that distracts from the main message of the section *EPI-Clone detects clonal identities from DNA methylation data only*.

6. Quantitative Comparisons: The reliance on visual inspection of uMAPs and identifying a “similar landscape” is prevalent in this article to demonstrate the efficiency of EPI-Clone in identifying cell lineage and state. Whenever possible, use quantitative comparisons, such as a confusion matrix, to show the rates of true positive and true negative classifications. This would provide a more rigorous validation of the method.

Our original manuscript did visually inspect UMAPs to compare the scRNA-seq and the scTAM-seq experiment (Figure 1a vs 1h). To make this comparison more quantitative, we now trained a classifier to predict cell type labels using only surface antigen information from the scRNA-seq experiment. We then applied the same classifier to the surface antigen data from the scTAM-seq experiment. In both scRNA-seq and scTAM-seq, the same Totalseq-B cocktail was used to measure surface antigen expression.

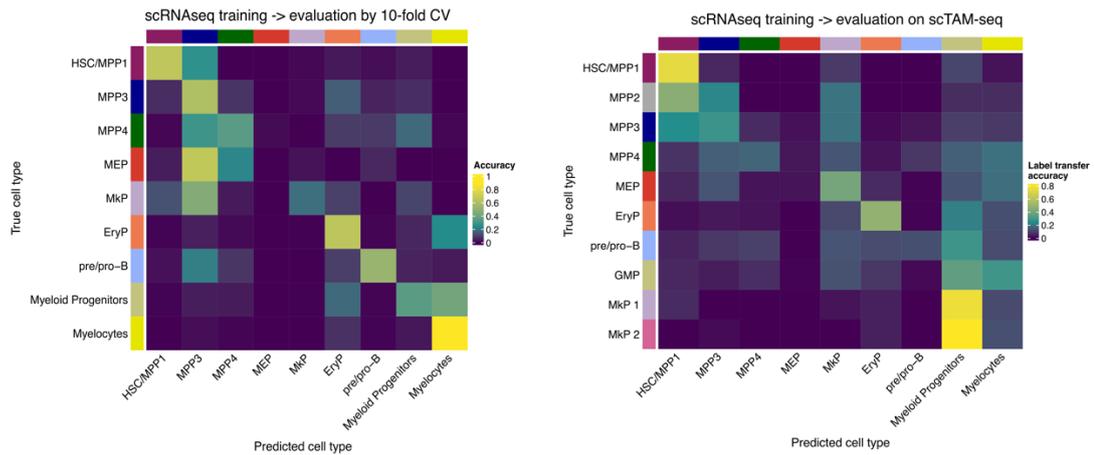


Figure S3E, reproduced here for convenience. Confusion matrices between scRNA-seq celltypes and scTAM-seq celltypes (Figure 1c vs. Figure 1g). To compute the confusion matrix, a random forest classifier was trained to predict cell type from surface antigen expression data, using the scRNA-seq modality. The confusion matrix for that classifier during 10-fold cross validation is shown in the plot on the left. The same classifier was then applied to predict cell type in the scTAM-seq experiment, where the same surface antigens were measured using the same TotalSeq-B cocktail. Label transfer accuracy is shown.

Inspections of clonal UMAPs were accompanied by statistical measures, such as ROC curves in figure 3e, confusion matrices in figure 3f, counts of expanded in clones in figure 4 and 5, overlaps of clones and CHIP mutations in figure 5e, etc.

The use of a UMAP in figure 2a&b, in our view, constitutes a good visualization of the point “both clone and cell state shape the methylome”. Both these points are the followed up with quantitative comparisons in figure 2 &3.

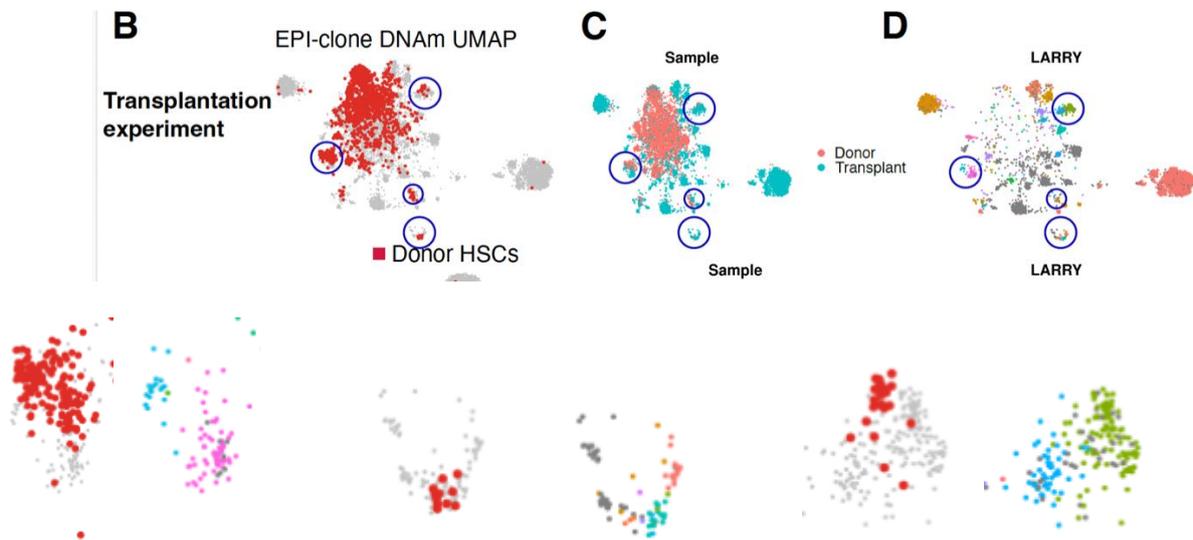
7. Figure 3E Axis Labels: For Figure 3E, why the Y-axis is not labelled as TPR (True Positive Rate)? In a standard ROC plot (TPR vs. FPR), a higher AUC indicates better performance.

We thank the reviewer for spotting this mistake. Indeed, the y-axis legend was incorrect and should be TPR instead of FNR. We corrected this issue in all of the ROC plots show throughout the manuscript and supplement.

8. Figure 3F Overlaps: Figure 3F shows some CpG clusters overlapping with more than one LARRY barcode. Considering this analysis was run only on large clones ($n > 30$ cells), it is worth investigating why EPI-Clone could not classify correctly in these scenarios. Examine whether this issue arises from the panel design, the algorithm, or other factors.

We believe that this plot describes true biology, and is not related to mistakes by the algorithm. These are likely expanded clones that existed in the donor mouse before LARRY infection, and that were infected by several LARRY barcodes. To demonstrate this, we now extracted bone marrow from an aged donor mouse, profiled a part with EPI-clone, LARRY-barcode another part, transplanted, and profiled again by EPI-clone. These data are now included as figure S11. Panel B (included in the response figure below) shows an oligoclonal hematopoietic system post-transplant (grey dots) that derives from a polyclonal system pre-transplant (red dots). Several post-transplant clones (circled) derive from donor cells that were already clonally expanded (several red dots per circle). These clones contain several LARRY barcodes (colors in panel D). It is interesting to see that the different LARRY barcodes

still separate in UMAP-space, this might suggest that further phylogenetic information has been accrued between the expansion of the clone and the LARRY labeling.



9. Minimal Panel Size: The authors could conduct additional analysis using the current data to determine the minimal size of the primer panel (for CpG sites) required to achieve similar resolution in cell lineage and cell state identification with EPI-Clone. This analysis will help identify the smallest set of CpG sites necessary for effective lineage tracing and state identification, potentially making the method easier to apply in various research settings. Demonstrating that EPI-Clone can maintain high accuracy with a reduced panel would significantly enhance its utility and accessibility.

We thank the reviewer for suggesting this additional analysis for showing the minimal panel size required for successfully clustering cells into clones. Using our the data from the M.1: LARRY main experiment, we now investigated how randomly subsetting the number of static CpGs affects the performance of EPI-clone. We found that we require 70 static CpGs for achieving an adjusted rand index higher than 0.8 (see the reviewer's 7th main point, response figure panel b). However, including additional static CpGs still improved the accuracy of the model until reaching the final EPI-clone performance of 0.88. These numbers might be different if additional clonal resolution is required, e.g. for a more polyclonal system.

In the context of the human panel we practically had to make our initial panel smaller, to combine it with RNA-seq from the same cell and mitochondrial lineage tracing. We therefore removed CpGs that did not carry information (because they were only classified as static or dynamic CpGs in few individuals and were mostly fully methylated or fully demethylated, Figure S12f,g), establishing a core panel of 200 CpGs that are maximally informative of cell state and lineage in human hematopoiesis (Table S5).

10. uMAP Annotations: Wherever uMAP plots are used, please specify the type of data used to generate the plot (e.g., DNA methylation, cell surface protein expression, etc.). This will provide clarity and help readers understand the context and significance of each plot.

Thank you for pointing this out. We now indicated which data modality was used to generate the UMAPs throughout the manuscript.

Referee reports

We thank all reviewers for their constructive comments during both rounds of review, which have helped us to realize the full potential of this manuscript.

Referee #1 (Remarks to the Author):

The manuscript by Scherer et al. has been revised with the addition of a large amount of data, which, in my opinion, cover all areas of improvement raised by the reviewers. The authors have made the manuscript substantially stronger and their new experiments now contribute to provide interesting novel insights into haematopoiesis, not reported before with other techniques.

Specifically, the authors:

- Expanded the study on primary human hematopoietic samples from 2 to 13 samples. This part brings very interesting new data that 1) further characterises clonal expansions without driver mutations from what previous reported; 2) links clones traced by EPI-clones to clonal hematopoiesis (CH) mutations.
- Included a replicate of their studies of mouse aging, performed transplantations and improved the quantification and statistical analysis of the mouse data.
- Provided a direct experimental comparison of EPI-Clones to mtDNA tracing, which is most useful to the field given recent controversies on the mtDNA method.
- Combined EPI-Clones with single cell targeted RNAseq in the same cell, being able to provide transcriptional information on specific clones.
- Demonstrated the broader utility of their method by performing lineage tracing of endothelial cells.
- Substantially improved the clarity in which the method is reported in the main text and with much helpful graphical representations of their experiments (Fig.S1A).

Overall, I believe this method will be of great use to the lineage tracing community. In addition, they report novel findings of high relevance to the hematology field, namely: 1) in mice aged HSC-expanded clones contribute little to transplantation, with the bulk of hematopoietic reconstitution provided by non expanded clones; 2) in humans expanded clones can be observed from age 50 onwards, do correspond to clones carrying CH mutations but also include driver-less clones. CH clone tend to be larger, more enriched in HSC and more lineage biased than driver-less clones.

Minor:

Figures 4I-J: please add the number of observations per tertile and clarify in the figure legend that T1 is the tertile with lowest mature output (I) or with smaller immature clone size (J).

We have added this information to figure 3i,j (formerly 4i,j) and its legend.

Referee #1 (Remarks on code availability):

I have reviewed the GitHub page, which seems complete and well laid out. In addition, I am pleased that the authors have deposited all datasets.

Referee #2 (Remarks to the Author):

The authors have addressed all of the minor comments, primarily pertaining to clarifications and additional detail in the manuscript. The authors also apply the method to an additional set of samples - 13 bone marrow samples.

The authors make the case for novel biological discovery and there is some there, but still completely descriptive. The characterization of these populations is a bit of stretch when assessing <500 CpG sites.

Overall, the manuscript is very well put together, has appropriate statistics and detail, and the method appears to be of high quality.

Referee #2 (Remarks on code availability):

Code is provided and well annotated.

Referee #3 (Remarks to the Author):

In this revision the authors have sufficiently addressed our experimental, computational and biological queries. In the spectrum of methods available that balance cost/throughput vs information content, epi-clone is demonstrably powerful in both uncoupling and separately reporting on clonal origin (using static CpGs) and lineage state (using dynamic CpGs, CITE-seq, and targeted mRNA expression). They have added both depth (more patients, CH mutation analysis, more mice, lymphoid cells) to the study of haematopoiesis and breadth (assessment of endothelial lineage clonal origins). The comparison with mtDNA variants shows the superiority of EPI-Clones in tracking human progenitors for in vivo lineage tracing. Furthermore, the presence of EPI-Clones in multiple tissues highlight the biological relevance and wide applicability of this methods. This reviewer thinks that EPI-Clone, and more in general the use of DNA methylation in human lineage tracing, represent a crucial advancement in the field and will be very useful to answer many research questions in the future. Below are minor suggestions/questions that, if addressed, we believe would enhance remaining queries about the study.

- Since submission of the revision, the MethylTree paper in Nature Methods proposes a new method to provide clonal and even subclonal resolution of cell relationships, including unbiased transcriptional state. Please briefly compare the method, for throughput, cost, pros vs cons in discussion.

We have added this comparison in the discussion, as follows:

While this manuscript was under review, a method termed MethylTree³⁹ demonstrated identification of clonal identity from sparse whole-genome, single-cell DNA methylation data. Compared with MethylTree, our approach is more scalable, less expensive and less computationally intense. Conversely, scTAM-seq requires design of a species-specific targeting panel.

- Please provide a version of heatmap (Reviewer Fig C: CpG sites vs single-cells matrix) sorted but not averaged per clone, in addition. This can be in SI. It would allow a

visualisation of the raw data underlying single cell epi-clone features, which is missing to date – even if it doesn't look pretty.

We agree that such a description was missing in our manuscript and thus now included a new Supplementary Fig. 3. This figure describes the DNA methylation states of the static CpGs per cell in each of the EPI-clones.

• Fig 5C: Dnmt3a, the most frequently mutated gene in CH is also involved in CpG methylation. Is it conceivable that mutations in Dnmt3a lead therefore to particularly distinguishable Epiclone signatures (as suggested by UMAP clustering). Conversely could Epiclone signatures allow for biological insights into the epigenomic effects of Dnmt3a mutations at the clonal level? As a potential analysis to explore this, using the existing data, could the authors pool all cells from all Dnmt3a clones (in Fig 5), and create a clustered heatmap of Epiclone signatures (rows: cells, columns: CpG sites), where the rows are annotated/coloured by Dnmt3a mutation, and the columns are annotated/coloured whether the CpG site is a known Dnmt3a target site. ?

We now included the suggested analysis as Supplementary Fig. 8. However, we could not investigate whether some CpGs are more prone to become methylated by DNMT3A, since we are not aware of a study on DNMT3A's site preferences. In any case, this heatmap provides a useful additional visualization for the point "Furthermore, we observed that several large CH clones (e.g. DNMT3A R659H in donor A.4, Fig. 4d and Supplementary Fig. 8) had diversified into two EPI-clones with a similar, but distinguishable static CpG profile"

• Fig 6G: proportions of cells per clone positive for 7076A>G appear tri-modal, i.e., close to 0%, 50% and 100%. Especially 50% is intriguing. Do the authors have an hypothesis to explain this observation?

We have now described this observations more accurately and provide a hypothesis to explain it in line 524-529 (edits in bold):

A highly abundant variant (mt:7076A>G) was strongly enriched or depleted in 17 T- or NK-cell EPI-clones, but was observed in **approximately 50% of** cells of the remaining, mostly multilineage- or B-cell, EPI-clones (Fig. 5g). This variant was likely present before epimutational patterns were established and repeatedly underwent selection **throughout development and adulthood. Therefore, T-cell clones with a recent history of expansion may or may not carry the variant, whereas multilineage clones that expanded before selection of the variant contain a mix of mutant and wild-type cells.**

• Regarding Figure S12C-E:

The authors stated, "The DNA methylation amplicons achieved a somewhat weaker cell state resolution, but displayed less technical variation." Instead of this qualitative description, it would be more useful to include a quantitative metric, such as the Adjusted Rand Index (ARI), to compare resolution more clearly.

The ARI for DNA methylation vs. both modalities is 0.29. For surface protein vs. both modalities, it is 0.4. These numbers indicate a weaker cell state resolution from DNA methylation.

Since this statement was from the response letter to the reviewers, we have not made changes to the manuscript.

Shalin Naik, Tom Weber, Esmaeel Azadian, Sara Tomei

Referee #3 (Remarks on code availability):

We were able to reproduce all the figures, the code is well-written, and the instructions are clear enough for users to follow without issues.

Referee 3 Reviewer attachment 1:
Version 0:

