THE EMBO JOURNAL

Manuscript EMBO-2017-97105

Single-cell RNA sequencing reveals developmental heterogeneity among early lymphoid progenitors

Llucia Alberti-Servera, Lilly von Muenchow, Panagiotis Tsapogas, Giuseppina Capoferri, Katja Eschbach, Christian Beisel, Rhodri Ceredig, Robert Ivanek, Antonius G. Rolink

Corresponding author: Llucia Alberti-Servera, University of Basel

Review timeline:

Submission date:	13 April 2017
Editorial Decision:	18 May 2017
Revision received:	11 August 2017
Accepted:	13 September 2017

Editor: Karin Dumstrei

Transaction Report:

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

1st Editorial Decision 18 May 2017

Thank you for submitting your manuscript to The EMBO Journal. Your study has now been seen by two referees and their comments are provided below.

As you can see the referees find the analysis interesting and timely, but also find that further work is needed to consider publication here. Should you be able to address the concerns raised then we are interested in considering a revised version. The issues raised are clearly outlined below and the most of them should be fairly easy to sort out. Referee #2 would like to see that the findings are confirmed in a wt setting. I don't know if you have data on hand to address this issue and I happy to discuss this point further.

When preparing your letter of response to the referees' comments, please bear in mind that this will form part of the Review Process File, and will therefore be available online to the community. For more details on our Transparent Editorial Process, please visit our website: http://emboj.embopress.org/about#Transparent_Process

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

REFEREE REPORTS

Referee #1:

Review of EMBOJ-2017-97105. This is a timely and interesting study. The authors have analyzed a cell population in mouse bone marrow that is endowed (based on previous work by this group) with

both lymphoid and myeloid potential (EPLM; early progenitor with lymphoid and myeloid potential). As with any 'common' progenitor population, a key question is whether single cells in this population have indeed more than one developmental potential versus mixed populations of cells with separate potential. Alberti-Servera et al. have first dissected EPLM by cell surface phenotype into four subsets. These were tested in limiting dilution assays for lymphoid and myeloid potential which yielded frequencies for progenitor activities. These populations were next subjected to single cell RNA seq, currently a widely employed technique to study the phenotypic heterogeneity of cell populations. This probably provides the highest possible phenotypic resolution. The authors of this paper are careful and do not fall into this trap (with some possible caveats, see below). Analysis of single cell RNA expression data revealed, interestingly, that gene expression profiles were either myeloid or lymphoid but, in single cells, not shared. This appears to contradict the former conclusion that EPLM are common progenitor for lymphoid and myeloid potential. However, single cell RNA seq does not reveal developmental 'trajectories' or fates because single cell analysis does not contain precursor - product relationship information.

Questions and suggestion:
1. The limiting dilution experiments provide frequency information but they are not true single cell in vitro assays. Do the authors have any information (either published data or own new experiments in which they have tested whether single EPLM have dual potential for lymphoid and myeloid cells? This could be examined, for example, in OP9 cultures supplemented with both IL7, Flk-ligand and myeloid cytokines. If they do, there would be an interested discrepancy between the single cell RNA seq data and in vitro potential. If they don't, both assay are consistent. It is indeed nor clear from reading the paper, including the abstract, whether EPLM at the single cell level are bipotent for lymphoid and myeloid lineage. Collectively, this is important to clarify.
2. My advice would be to check the wording in the paper very carefully against over interpretation of the single cell RNA seq data. Because there is no time scale in these experiments, claims such as gene expression is upregulated are not justified. Moreover, the authors tent to conclude from gene expression on developmental potential which is also (in my view) impossible. The authors seem (at least partially) aware of this problem.

With such information and possibly data provide, I think the paper would be an interesting example of careful use of single cell RNA seq analysis (including the recognition and discussion of its limitations), eventually revealing population heterogeneity.

Referee #2:

B cell development and commitment is a process well characterized in the mouse system. However the stage just before the cell commits and become CD19+ B cell is less defined. Rolink et al have in an earlier study characterized a fraction of B220+ cells, (B220+CD117intCD19-NK1.1-, related to the preproB, Fraction A) which appeared multi-potent. In this study the heterogeneity of this population is resolved by sub dividing the population based on three different surface markers, Ly6D and two dendritic cell markers; SiglecH and CD11c. The authors found four subpopulations with distinct developmental potentials. The Ly6D+ cell fraction turned out to be lymphoid restricted and formed mainly B cell. It appeared to be the direct precursors of the first CD19+ committed B cells. SC RNA seq could further identify two different clusters within this population. The B220+ sub-population that lacked expression of the three surface markers formed B/T and myeloid cells in vitro. SC RNA seq could further separate these cells into three clusters, one with myeloid, one with lymphoid and one with cDC signature. The myeloid and lymphoid genetic signatures did not appear to be co-expressed at the single cell level.
The paper is well written and easy to read and follow, and the single cell RNA seq technology used in a potent way to resolve the heterogeneity within the cell population investigated. The step preceding B cell commitment and the expression of CD19 has so far not been fully characterized at the single cell level and here the study adds knowledge to the field. The surface markers used to sub-fractionate the B220+ cells, have been used in an earlier study to purify related preproB cells, though they were not investigated at the single cell level (Medina et al Plos one 2013). Some issues need to be resolved before publication and are listed below.

Major points

1. RNA seq is performed on cells from Flt3Ltg mice. As mentioned in the manuscript important B cell transcription factors like Ebf1 and Pax5 are reduced in these mice. I therefore think it is important to confirm the conclusions in a WT setting.

2. Pax5 is upregulated in the transition from CLPLy6D- to Ly6D+ (Inlay et al, Genes and Dev. 2009). How is expression of Pax5 in the B220+ populations investigated? It does not seem to come up as an upregulated gene in G1 Ly6D+ cells, although these cells are proposed to be direct precursors of the CD19+ committed B cells (Figure 4C).

3. In the stroma co-culture assays the lineage output (B/T/Myeloid) should be confirmed with for instance FACS. What kind of myeloid cells were formed from the TN fraction?

Minor points

4. CLP has in an earlier study been divided based on the surface marker Ly6D and the Ly6D+ fraction shown to have mainly B cell potential (Inlay et al, Genes and Dev. 2009). What is the relationship between the B220+Ly6D+ and the TN populations identified here and the CLP Ly6D+ and Ly6D- cells?

5. Ly6D and surface markers for dendritic cells have previously been used to subdivide and purify preproB cells (Medina et al, Plos one 2013). How are these populations and findings related to the populations described herein?

6. How is FLT3 expressed in the different subpopulations?

7. The T cell potential readout in Figure 1E is low for all populations investigated. Was a positive control for the assay included?

8. The surface markers used to phenotypically divide the B220+ fraction should preferably be explained and referenced when introduced.

1st Revision - authors' response 11 August 2017

Here follows a detailed point-by-point response to the Reviewers' comments:

Reviewer #1:

1. **The limiting dilution experiments provide frequency information but they are not true single cell in vitro assays. Do the authors have any information (either published data or own new experiments in which they have tested whether single EPLM have dual potential for lymphoid and myeloid cells? This could be examined, for example, in OP9 cultures supplemented with both IL7, Flk-ligand and myeloid cytokines. If they do, there would be an interested discrepancy between the single cell RNA seq data and in vitro potential. If they don't, both assay are consistent.**

As suggested, we have assessed whether single EPLM are bipotent for lymphoid and myeloid lineage by sorting single Ly6D$^+$ or TN cells into 96-well plates and co-culturing them with OP9-stromal cells supplemented with IL-7 and MCSF cytokines, the most potent cytokines supporting lymphoid and myeloid differentiation, respectively. This experiment has led to a similar conclusion as the single-cell RNAseq data. Ly6D$^+$ cells mainly differentiated into B-cells whereas TN into myeloid cells. No mixed lymphoid-myeloid clones were found in the Ly6D$^+$ clones, while very few were found in the TN cultures. The frequency of mixed clones was even lower than that of the observed single-cells expressing both lymphoid and myeloid genes. The results are shown in Fig. 5E and mentioned in the results section (page 16).

It is indeed not clear from reading the paper, including the abstract, whether EPLM at the single cell level are bipotent for lymphoid and myeloid lineage. Collectively, this is important to clarify.

EPLM have been identified and characterized as bipotent cells at the population level. It was one of the goals of the present study to assess whether they are bipotent at the single cell level or not. Based on our single cell RNAseq analysis and the new experiment performed, we conclude that EPLM are, with few exceptions, not bipotent for lymphoid and myeloid lineage at the single-cell level. We have now stated this more clearly in the text and the abstract.

2. **My advice would be to check the wording in the paper very carefully against over interpretation of the single cell RNA seq data. Because there is no time scale in these experiments, claims such as gene expression is upregulated are not justified. Moreover, the authors tent to conclude from gene expression on developmental potential which is also (in my view) impossible. The authors seem (at least partially) aware of this problem.**

We have tried to adjust our wording on the interpretation of single-cell RNAseq data according to the reviewer's comment, with which we agree. We have rephrased our relevant statements. The words up- and down-regulated have been substituted by "higher level of expression", "higher expressed" etc (pages 10, 13, 20 and, legends of figures 3 and 4). We have been more careful about concluding from gene expression on developmental potential (pages 10, 14, 19, 21). We believe that other statements are now more justified after the single-cell cultures results (page 18, 20, 21).

Reviewer #2:

Major points

1. **RNA seq is performed on cells from Flt3Ltg mice. As mentioned in the manuscript important B cell transcription factors like Ebf1 and Pax5 are reduced in these mice. I therefore think it is important to confirm the conclusions in a WT setting.**

The genomics single-cell facility and the system used (C1 platform) for the single-cell RNAseq experiment imposed a restriction in cell numbers. We had to provide 30K cells in 100µl per run (mentioned in M&M page 25) and, therefore, it was not feasible on a WT setting where the EPLM subpopulations are very rare. Recent technological advances, such as the 10X Genomic Platform (mentioned in discussion page 18), would allow single-cell RNAseq on EPLM from WT. We would be very much interested in performing the single-cell RNAseq experiment on WT EPLM cells but a new experiment would take months to complete (taking into consideration the availability of the platform and the time required for analysis of the results) and therefore it would be impossible to include in the present manuscript within the revision time, or after a reasonable extension.

2. **Pax5 is upregulated in the transition from CLPLy6D- to Ly6D+ (Inlay et al, Genes and Dev. 2009). How is expression of Pax5 in the B220+ populations investigated? It does not seem to come up as an upregulated gene in G1 Ly6D+ cells, although these cells are proposed to be direct precursors of the CD19+ committed B cells (Figure 4C).**

In the first version of the manuscript we did not include *Pax5* expression from the single-cell RNAseq analysis since we found only few cells expressing Pax5. We have now added that (violin plot in Fig 4C). Few cells are found positive for *Pax5* mRNA (8 out of 56 cells) but all of them are within the G1 Ly6D$^+$ group, while no *Pax5*$^+$ cells were found in the G2 Ly6D$^+$ group. We have recently gained access to Pax5-reporter mice, which we crossed to our *Flt3L*tg mice in order to assess the expression of *Pax5* within these EPLM populations at the protein level. We found no expression of the Pax5-reporter in TN EPLMs (data not shown) while 37.7% of WT Ly6D$^+$ EPLMs were positive for Pax5, a percentage that dropped to 9.2% in the *Flt3L*tg Ly6D$^+$ EPLM (Appendix Fig S2C). It is possible that sustained Flt3 signalling might delay and/or downregulate Pax5 expression. *In vitro* evidence for such a hypothesis has been published by Holmes et al. 2006 (added reference page 9).
In addition, *Pax5* is differentially expressed in the bulk RNAseq experiment when comparing Ly6D$^+$ and TN (see Table EV1). Therefore, the low detection of *Pax5* transcript in the single-cell RNA seq

experiment might be due to sensitivity reasons. *Pax5* gene is low-mid expressed and for those genes the drop-out rate is quite high in single cell RNA-seq.

We have also analyzed Ebf1-reporter mice. Preliminary analysis of Ly6D$^+$ cells from WT and *Flt3L*tg Ebf1-reporter mice has shown that the percentage of Ebf1$^+$Ly6D$^+$ does not seem to change significantly between the two genotypes. Therefore, we now only mention reduction of Pax5 in the manuscript (page 9). Moreover, this would indicate that the B-cell specification signature (G1 Ly6D$^+$, exemplified by Ebf1 expression) observed in the sc-RNAseq experiment using our mice model, might not be significantly different than the WT, with the exception of Pax5.

3. **In the stroma co-culture assays the lineage output (B/T/Myeloid) should be confirmed with for instance FACS. What kind of myeloid cells were formed from the TN fraction?**

In all our *in vitro* cultures the B/T/myeloid lineage output can be clearly identified based on the distinct morphology of the corresponding clones under the inverted microscope: B and T cells are small, with B cell colonies round shaped and tightly packed while T cell colonies more spread. Myeloid cells were always significantly larger than lymphoid. However, we have initially analysed by FACS the lineage output of several wells (we provide representative analyses in Appendix Fig S1B, mentioned both in results (page 7) and Appendix Supplementary Methods (page 9)) and have found 100% agreement between the resulting FACS phenotype and the observed microscopic morphology. Therefore, we have continued to score the colonies by microscopy for the rest of the experiments.
We have performed a more detailed analysis of the particular type of myeloid cells by FACS. This analysis showed that the vast majority of myeloid colonies consisted mainly of macrophages (F4/80$^+$CD11b$^+$) with few cells being F4/80$^-$CD11b$^-$ and very few showing a CD11c$^+$ phenotype. Representative FACS analysis of such a colony is shown in Sup. Fig. 1B right. We have not gone into a detailed analysis of the types of myeloid cells resulting in our ST2 cultures, as it was not within the scope of our study.

Minor points

4. **CLP has in an earlier study been divided based on the surface marker Ly6D and the Ly6D+ fraction shown to have mainly B cell potential (Inlay et al, Genes and Dev. 2009). What is the relationship between the B220+Ly6D+ and the TN populations identified here and the CLP Ly6D+ and Ly6D- cells?**

Ly6D$^+$ EPLM are Flt3$^+$CD127$^+$ and therefore phenotypically overlap to a large extent with B220$^+$CD19$^-$ pre-proB and partly also with Ly6D$^+$ CLP (to the extent that some Ly6D$^+$ CLP can be B220^{low}). A significant fraction of TN EPLM also expresses Flt3 and CD127, therefore representing a population highly similar to Ly6D$^-$ CLP in phenotype. However, TN EPLM also contains cells that would not be placed in a "Ly6D$^-$ CLP gate". We postulate that these cells would be the G4 and G5 TN groups identified in our present single-cell RNAseq analysis. Accordingly, we would consider the G3 TN group as phenotypically overlapping with Ly6D$^-$ CLP.

5. **Ly6D and surface markers for dendritic cells have previously been used to subdivide and purify preproB cells (Medina et al, Plos one 2013). How are these populations and findings related to the populations described herein?**

As mentioned above in our response to point 4, Ly6D$^+$ EPLM partially overlap with Ly6D$^+$ CLP and pre-proB cells, which are the populations Medina et al. identify to contain PDCA1$^+$ cells. Therefore, the SiglecH$^+$ subpopulation of EPLM would represent the PDCA1$^+$ fraction of the cells identified by Medina et al, as PDCA1 and SiglecH are both specific plasmacytoid dendritic cell markers. We now mention in the discussion (page 20) that the G1 Ly6D$^+$ subset is phenotypically closely related to PDCA-1$^-$ BLP and PDCA-1$^-$ Pre-pro B cells and included the corresponding reference (Medina et al. 2013) (pages 6 and 20). The PDCA-1$^+$ BLP and PDCA-1$^+$ Pre-pro resemble to our SiglecH$^+$ EPLM.

6. **How is FLT3 expressed in the different subpopulations?**

In *Flt3L*tg mice, presumably due to continuous engagement and internalization of the receptor, the Flt3 receptor cannot be detected on the cell surface by flow cytometry (Tsapogas et al, 2014). We

can however assess the transcript levels. In the bulk RNAseq experiment, Flt3 is not differentially expressed between Ly6D and TN (see Table EV1 and now mentioned in the results, page 10). Moreover, in our single-cell RNAseq analysis 90% and 70% of Ly6D$^+$ and TN, respectively, are positive for FLt3 mRNA expression (data not shown).

7. **The T cell potential readout in Figure 1E is low for all populations investigated. Was a positive control for the assay included?**

We have performed the T cell differentiation assay in Fig. 1E (WT) in parallel with the one shown in Fig 2E (*Flt3L*tg), which shows a much more robust T cell potential, therefore we are confident for the efficiency of the assay.

8. **The surface markers used to phenotypically divide the B220+ fraction should preferably be explained and referenced when introduced.**

We have now introduced the Ly6D$^+$, SiglecH and CD11c markers in the results section (Page 6) instead of the discussion and stressed the reasoning behind their use.

Accepted 13 September 2017

Thank you for submitting your revised manuscript.

The study has now been seen by the original referees and as you can see below both referees appreciate the introduced changes and support publication in The EMBO Journal.

I am therefore very pleased to accept the manuscript for publication here - congratulations on a very nice paper.

REFEREE REPORTS

Referee #1:

The authors have thoughtfully considered the points I raised and revised the paper accordingly. I now support publication of this paper.

Referee #2:

It is with regret I received the news of the decease of Professor Antonius Rolink, a most distinguished colleague and scientist.

The revised manuscript by Alberti-Servera et al has improved. New data has been added that addresses my issues and comments and the new functional data strengthens the paper. I think this is an interesting study that shows how SC RNA seq can resolve heterogeneity within a defined population.

EMBO PRESS

YOU MUST COMPLETE ALL CELLS WITH A PINK BACKGROUND ↓

PLEASE NOTE THAT THIS CHECKLIST WILL BE PUBLISHED ALONGSIDE YOUR PAPER

Corresponding Author Name: Llucia Alberti-Servera
Journal Submitted to: The EMBO Journal
Manuscript Number: EMBOJ-2017-97105

Reporting Checklist For Life Sciences Articles (Rev. July 2015)

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

A- Figures

1. Data

The data shown in figures should satisfy the following conditions:

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

2. Captions

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

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

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

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

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

B- Statistics and general methods

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

1.a. How was the sample size chosen to ensure adequate power to detect a pre-specified effect size?	For the single-cell RNAseq experiment we decided to perform three runs per population in order to profile at least 100 single-cells per cell type with good quality. When deciding the experimental set up, we also took into account the number of mice required and the prize of the experiment. For the in vitro and in vivo experiments, experiments were reproduced with completely independent biological replicates.
1.b. For animal studies, include a statement about sample size estimate even if no statistical methods were used.	The animal size was estimated based on the number of cells needed to perform the ex vivo experiment. We mainly used a transgenic model that has increased number of cells of interest in order to reduce the use of mice. Normally, 5 independent mice were analyzed per experiment. When pooling mice was required, we used 2 and a maximum of 8 for WT settings.
2. Describe inclusion/exclusion criteria if samples or animals were excluded from the analysis. Were the criteria pre-established?	Samples (cells) that did not pass quality control durig single-cell RNAseq analysis were excluded. Established criteria: at least 60% of mapped reads, 200K counts, and 800 detected genes.
3. Were any steps taken to minimize the effects of subjective bias when allocating animals/samples to treatment (e.g. randomization procedure)? If yes, please describe.	all mice used were 6 to 11 weeks old and matched by age and sex for each experiment
For animal studies, include a statement about randomization even if no randomization was used.	randomization was not taken into account but each experiment was independently repeated at least three times
4.a. Were any steps taken to minimize the effects of subjective bias during group allocation or/and when assessing results (e.g. blinding of the investigator)? If yes please describe.	each experiment was independently repeated at least three times and performed by different authors
4.b. For animal studies, include a statement about blinding even if no blinding was done	NA
5. For every figure, are statistical tests justified as appropriate?	statistical tests are indicated in materials and methods and figure legends
Do the data meet the assumptions of the tests (e.g., normal distribution)? Describe any methods used to assess it.	statistical methods for normalisation of transcriptomics data are indicated in materials and methods

USEFUL LINKS FOR COMPLETING THIS FORM

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

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

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

http://datadryad.org

http://figshare.com

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

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

http://biomodels.net/

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

Is there an estimate of variation within each group of data?	standard error of the mean (sem). For expression data, the dispersion was calculated with edgeR R package.
Is the variance similar between the groups that are being statistically compared?	Not assessed

C- Reagents

6. To show that antibodies were profiled for use in the system under study (assay and species), provide a citation, catalog number and/or clone number, supplementary information or reference to an antibody validation profile. e.g., Antibodypedia (see link list at top right), 1DegreeBio (see link list at top right).	Antibodies were used (from BD Pharmingen, eBioscience, BioLegend, or produced in house): anti-B220 (RA3-6B2), anti-CD117 (2B8), anti-CD19 (1D3), anti-NK1.1 (PK136), anti-SiglecH (551), anti-CD11c (HL3), anti-Ly6D (49-H4), anti-Thy1.2 (53-2.1), anti-F4/80 (F4/80) conjugated with FITC, PE, PE/Cy7, APC, BV421 or Biotin. Biotin-labelled antibodies were revealed using streptavidin-BV650.
7. Identify the source of cell lines and report if they were recently authenticated (e.g., by STR profiling) and tested for mycoplasma contamination.	All cultures were performed from ex vivo mice-derived cells. ST2 (Ogawa et al, 1988), OP9 (Nakano et al, 1994) and OP9-DL1 (Schmitt & Zuniga-Pflucker, 2002)

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

D- Animal Models

8. Report species, strain, gender, age of animals and genetic modification status where applicable. Please detail housing and husbandry conditions and the source of animals.	C57BL/6 (B6), B6 Rag2-deficient (Shinkai et al, 1992), B6 Flt3L transgenic (Flt3Ltg, (Tsapogas et al, 2014)), and Pax5-reporter (Fuxa & Busslinger, 2007) mice used herein were 6 to 11 weeks old and matched by age and sex for each experiment. All mice were bred and maintained in our animal facility under specific pathogen-free conditions
9. For experiments involving live vertebrates, include a statement of compliance with ethical regulations and identify the committee(s) approving the experiments.	All animal experiments were carried out according to institutional guidelines (authorization numbers 1886 and 1888 from Kantonales Veterinäramt, Basel)
10. We recommend consulting the ARRIVE guidelines (see link list at top right) (PLoS Biol. 8(6), e1000412, 2010) to ensure that other relevant aspects of animal studies are adequately reported. See author guidelines, under 'Reporting Guidelines'. See also: NIH (see link list at top right) and MRC (see link list at top right) recommendations. Please confirm compliance.	We ensure that animal studies are adequately reported

E- Human Subjects

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

F- Data Accessibility

18: Provide a "Data Availability" section at the end of the Materials & Methods, listing the accession codes for data generated in this study and deposited in a public database (e.g. RNA-Seq data: Gene Expression Omnibus GSE39462, Proteomics data: PRIDE PXD000208 etc.) Please refer to our author guidelines for 'Data Deposition'. Data deposition in a public repository is mandatory for: a. Protein, DNA and RNA sequences b. Macromolecular structures c. Crystallographic data for small molecules d. Functional genomics data e. Proteomics and molecular interactions	The bulk RNAseq as well as the single-cell RNAseq data from this publication have been deposited to NCBI's Gene Expression Omnibus database (Barrett et al, 2013; Edgar et al, 2002) (https://www.ncbi.nlm.nih.gov/geo/) and assigned the GEO Series accession number GSE102456.
19. Deposition is strongly recommended for any datasets that are central and integral to the study; please consider the journal's data policy. If no structured public repository exists for a given data type, we encourage the provision of datasets in the manuscript as a Supplementary Document (see author guidelines under 'Expanded View' or in unstructured repositories such as Dryad (see link list at top right) or Figshare (see link list at top right).	NA
20. Access to human clinical and genomic datasets should be provided with as few restrictions as possible while respecting ethical obligations to the patients and relevant medical and legal issues. If practically possible and compatible with the individual consent agreement used in the study, such data should be deposited in one of the major public access-controlled repositories such as dbGAP (see link list at top right) or EGA (see link list at top right).	NA
21. Computational models that are central and integral to a study should be shared without restrictions and provided in a machine-readable form. The relevant accession numbers or links should be provided. When possible, standardized format (SBML, CellML) should be used instead of scripts (e.g. MATLAB). Authors are strongly encouraged to follow the MIRIAM guidelines (see link list at top right) and deposit their model in a public database such as Biomodels (see link list at top right) or JWS Online (see link list at top right). If computer source code is provided with the paper, it should be deposited in a public repository or included in supplementary information.	NA

G- Dual use research of concern

22. Could your study fall under dual use research restrictions? Please check biosecurity documents (see link list at top right) and list of select agents and toxins (APHIS/CDC) (see link list at top right). According to our biosecurity guidelines, provide a statement only if it could.	No, it does not.