EMBO reports

Single-cell transcriptome analysis reveals thyrocyte diversity in the zebrafish thyroid gland

Pierre Gillotay, Meghna Shankar, Benoit Haerlingen, Sema Elif Eski, Macarena Pozo-Morales, Inés Garteizgogeascoa Suñer, Susanne Reinhardt, Annekathrin Kränkel, Juliane Bläsche, Andreas Petzold, Nikolay Ninov, Gokul Kesavan, Christian Lange, Michael Brand, Anne Lefort, Frédéric Libert, Vincent Detours, Sabine Costagliola, and Sumeet Pal Singh
DOI: 10.15252/embr.201050612

Corresponding author(s): Sumeet Singh (sumeet.pal.singh@ulb.ac.be) and Sabine Costagliola (scostag@ulb.ac.be)

Review Timeline:

Transfer from Review Commons:	9th Apr 20
Editorial Decision:	20th Apr 20
Revision Received:	14th Jul 20
Editorial Decision:	2nd Sep 20
Revision Received:	30th Sep 20
Accepted:	6th Oct 20

Editor: Esther Schnapp

Transaction Report: This manuscript was transferred to EMBO reports following peer review at Review Commons

1st Editorial Decision 20th Apr 2020

Dear Sumeet,

Thank you for the submission of your peer-reviewed manuscript and point-by-point response to EMBO reports. I have now read and discussed all with my colleagues here, and we would like to invite you to revise your study for publication here.

Please address all referee concerns in a complete point-by-point response, and please add the missing experiment that you mention and that is important.
Acceptance of the manuscript will depend on a positive outcome of a second round of review. It is EMBO reports policy to allow a single round of major revision only and acceptance or rejection of the manuscript will therefore depend on the completeness of your responses included in the next, final version of the manuscript.

Revised manuscripts should be submitted within three months of a request for revision; they will otherwise be treated as new submissions. Please contact us if a 3-months time frame is not sufficient for the revisions so that we can discuss this further. It is of course fine to extend the revision time because of lab closure due to the corona virus.

You can either publish the study as a short report or as a full article. For short reports, the revised manuscript should not exceed 27,000 characters (including spaces but excluding materials & methods and references) and 5 main plus 5 expanded view figures. The results and discussion sections must further be combined, which will help to shorten the manuscript text by eliminating some redundancy that is inevitable when discussing the same experiments twice. For a normal article there are no length limitations, but it should have more than 5 main figures and the results and discussion sections must be separate. In both cases, the entire materials and methods must be included in the main manuscript file.

Regarding data quantification, please specify the number "n" for how many independent experiments were performed, the bars and error bars (e.g. SEM, SD) and the test used to calculate p-values in the respective figure legends. This information must be provided in the figure legends. Please also include scale bars in all microscopy images.

IMPORTANT NOTE: we perform an initial quality control of all revised manuscripts before re-review. Your manuscript will FAIL this control and the handling will be DELAYED if the following APPLIES:
1) A data availability section providing access to data deposited in public databases is missing. If you have not deposited any data, please add a sentence to the data availability section that explains that.
2) Your manuscript contains statistics and error bars based on n=2 or on technical replicates. Please use scatter blots in these cases. No statistics can be calculated if n=2.

When submitting your revised manuscript, please carefully review the instructions that follow below. Failure to include requested items will delay the evaluation of your revision.

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

2) individual production quality figure files as .eps, .tif, .jpg (one file per figure).
See https://wol-prod-cdn.literatumonline.com/pb-assets/embo-

site/EMBOPress_Figure_Guidelines_061115-1561436025777.pdf for more info on how to prepare your figures.

3) We replaced Supplementary Information with Expanded View (EV) Figures and Tables that are collapsible/expandable online. A maximum of 5 EV Figures can be typeset. EV Figures should be cited as 'Figure EV1, Figure EV2" etc... in the text and their respective legends should be included in the main text after the legends of regular figures.

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

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

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

5) a complete author checklist, which you can download from our author guidelines . Please insert information in the checklist that is also reflected in the manuscript. The completed author checklist will also be part of the RPF.

6) Please note that all corresponding authors are required to supply an ORCID ID for their name upon submission of a revised manuscript (). Please find instructions on how to link your ORCID ID to your account in our manuscript tracking system in our Author guidelines

7) Before submitting your revision, primary datasets produced in this study need to be deposited in an appropriate public database (see https://www.embopress.org/page/journal/14693178/authorguide#datadeposition). Please remember to provide a reviewer password if the datasets are not yet public. The accession numbers and database should be listed in a formal "Data Availability" section placed after Materials & Method (see also https://www.embopress.org/page/journal/14693178/authorguide#datadeposition). Please note that the Data Availability Section is restricted to new primary data that are part of this study. * Note - All links should resolve to a page where the data can be accessed. *
If your study has not produced novel datasets, please mention this fact in the Data Availability Section.

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

9) Our journal also encourages inclusion of *data citations in the reference list* to directly cite datasets that were re-used and obtained from public databases. Data citations in the article text

are distinct from normal bibliographical citations and should directly link to the database records from which the data can be accessed. In the main text, data citations are formatted as follows: "Data ref: Smith et al, 2001" or "Data ref: NCBI Sequence Read Archive PRJNA342805, 2017". In the Reference list, data citations must be labeled with "[DATASET]". A data reference must provide the database name, accession number/identifiers and a resolvable link to the landing page from which the data can be accessed at the end of the reference. Further instructions are available at https://www.embopress.org/page/journal/14693178/authorguide#referencesformat

We would also welcome the submission of cover suggestions, or motifs to be used by our Graphics Illustrator in designing a cover.

As part of the EMBO publication's Transparent Editorial Process, EMBO reports publishes online a Review Process File (RPF) to accompany accepted manuscripts. This File will be published in conjunction with your paper and will include the referee reports, your point-by-point response and all pertinent correspondence relating to the manuscript.

You are able to opt out of this by letting the editorial office know (emboreports@embo.org). If you do opt out, the Review Process File link will point to the following statement: "No Review Process File is available with this article, as the authors have chosen not to make the review process public in this case."

I look forward to seeing a revised version of your manuscript when it is ready. Please let me know if you have questions or comments regarding the revision.

Kind regards,
Esther

Esther Schnapp, PhD
Senior Editor
EMBO reports

Reviews and Authors' Response transferred from Review Commons

Response to Reviewers:

We would like to thank the reviewers for taking the time to provide us with insightful and constructive comments, which helped us in improving the manuscript. We have performed additional experiments and data analysis while also improving the presentation of our analysis.

Please find below our point-by-point response.

Reviewer #1:

Gillotay et al. performed an unbiased profiling of the zebra fish thyroid gland and captured different cell populations by single-cell gene expression analysis. Using bioinformatic tools, they identified seven clusters corresponding to expected, but also poorly characterized, sub-populations, such as non-follicular epithelial cells. They also found two transcriptionally distinct types of thyrocytes and validated this heterogeneity using a new transgenic Pax2a reporter line. Using this tool, they identified and located Pax2a-low and -high thyrocytes within thyroid follicles. Finally, they highlighted a dense intercellular signaling network based on ligands expressed by the diverse sub-populations present in the thyroid and receptors expressed by the thyrocytes. This is a descriptive work calling for more in-depth analyses.

The authors thank the reviewer #1 for acknowledging the strengths of the manuscript and for stating that it could be of interest to a larger audience. We appreciate the reviewer's advice on supportive experiments and for improving the clarity of the analysis and presentation. We have experimentally addressed the concerns of the reviewer and hope this provides in-depth substantiation of our observations.

****Major comments on main conclusions :****

Conclusion 1 : Identification of 7 clusters
- This reviewer was particularly surprised by the relative abundance of the different sub-populations identified. For example, 267 thyrocytes out of 6249 cells from the thyroid gland is less than 5% of the total thyroid cell number. In comparison, authors identified three times more immune cells or non-follicular epithelial cells. Authors should comment on these numbers, on the dissection (contaminants?) and dissociation procedure.
Do these relative abundance reflect the proportion of thyrocytes, immune, stromal cells they normally observe in adult thyroid sections ? It would be interesting to have a better resolution for figure 6A in order to evaluate the number of nuclei stained only with DAPI as compared to nuclei stained with DAPI and surrounded by E-cadherin staining. Based on the image, this reviewer seriously doubts that follicular cells represent less than 5% of the total cell number in this organ, considering that the colloid is a cell-free zone.

We fully agree with the reviewer statement. We have clarified the nature of the dissociated tissue in Results, Methods and Fig. 1 C- G. In this, we have performed 3D confocal imaging of the region utilized for single-cell RNA-Seq. (Fig. 1C). Further, we quantified percentage of thyrocytes in transverse sections across the dissociated region (Fig. 1 D – E). Our results

demonstrate presence of 5.9 ± 1.9 % thyrocytes in the region. Lastly, we provide the FACS plots of the cells utilized for single-cell RNA-Seq. In this, we obtained around 4% thyrocytes among the live cells. Taken together, our quantifications of the thyrocyte proportion in the tissue matches well with the percentage of thyrocytes obtained in the single-cell atlas, suggesting lack of thyrocyte loss during the procedure.

We agree with the reviewer's observation that the follicle lumen represents a cell-free zone. However, it is worth noting that the cells surrounding the follicles, in particular gills and stroma, have a higher cell density. Additionally, the zebrafish thyroid follicles sit loosely on the ventral aorta, thus making it difficult to manually separate the follicles from the surrounding tissue without destroying the organ. We avoided injuring the organ to minimize cell-death associated with manual dissection.

- When using the webtool developed on the thyrocyte population, one can notice that only a small fraction of the thyrocyte population expresses common thyroid-specific genes such as Tpo or duox. Was this expected ? It would be interesting to comment on this observation and confirm using standard localization technique to demonstrate that this is real and not due to the sequencing. Is it specific to the zebrafish ? On the other hand, Tg is expressed in most thyrocytes, but surprisingly also in all the clusters at a fairly good level. This should be commented... Is it normal ? due to the sequencing quality ? or clustering ?

We believe these are technical issues related to single-cell sequencing and thank the reviewer for their insight in seeing this.

The non-uniform expression of thyroid-specific marker genes (*Tpo* and *duox*) likely represents dropout effects, possibly due to the low expression levels of these genes. To address this issue, we performed bulk RNA-Seq. of thyrocytes segregated by *pax2a* expression levels (Fig. 9D, E), which is more sensitive than single-cell RNA-Seq. The gene *tg* (thyroglobulin) was one of the top 15 expressed genes in the two populations (Fig. EV5). Further, genes involved in thyroid functionality: *nkx2.4b* (zebrafish homologue of NKX2.1 / TTF1), slc5a5 (NIS), TPO, tshr, ctsba, ctsk (cathepsin K), were detected in the two populations (Fig. 9E, Table EV4).

The expression of *tg* in non-thyrocyte population likely represents cross-contamination of free RNAs released from ruptured cells. Since *tg* mRNA is highly expressed in the thyroid follicular cells, release of the mRNA from a few injured cells would contaminate the cell suspension, leading to its detection in non-thyroid cells. However, the expression represents background noise signal. To test this, we utilized DecontX, a recently developed approach for background correction (S. Yang et al. 2020). In this, the expression of a gene is modelled as a mixture of expression in the expected population plus background expression. With this, we could robustly reduce *tg* mRNA expression in non-thyrocytes (Appendix Fig S1). This supports our hypothesis that the *tg* expression in non-thyrocytes likely represents cross-contamination of mRNA from ruptured cells.

- In order to validate and locate the different populations identified in the thyroid, this reviewer suggests to perform in situ hybridization or immunostaining, based on the specific marker genes identified in each cluster. This experiment could lead to the precise

identification of the different sub-populations and their respective localization. These experiments would also help in the interpretation of the cellular interaction network.

We have characterized the different cell types surrounding the thyroid follicles using various reporter lines. The data is presented in Fig. 4.

Conclusion 2 : two distinct types of thyrocytes
- This is an interesting observation. However, from a non-expert it is difficult to understand why the authors propose two populations. Based on the points distribution (Figure 4A), this reviewer would rather identify 3 or 4 clusters but not the two shown in red and blue.... Did the authors impose two populations for the clustering ? Did they perform a permutation test to confirm the pax2a significant fold change seen between clusters is not a false positive generated by the clustering ? Could they show, as a supplementary file, the same graph with points colored based on Pax2a expression ?

We concur with the reviewer that the number of potential clusters in thyrocyte population might be more than two. In-fact, there is no upper limit to the diversity present in the thyrocyte population. However, our message in the manuscript is that the population is not homogenous, and there are at-least two populations based on *pax2a* expression level. In this regard, we do believe that we have only scratched the tip of the iceberg and further investigation is needed to completely answer this issue. Nonetheless, we are the first to demonstrate genetic heterogeneity among the population.

In response to this, and in complete agreement with the reviewer's comment, we have edited the results to showcase the presence of transcriptional heterogeneity using single-cell RNA-Seq. data. For this, we conducted analysis of transcriptional diversity in the population using a recently developed method called ROGUE (Ratio of Global Unshifted Entropy) (Liu et al. 2019). The method provides genes that display transcriptional heterogeneity within the cell population. Assessment is based on expression entropy, a measure of the degree of uncertainty, or promiscuity, in the expression of a gene (Teschendorff and Enver 2017). For this, we utilized raw counts of thyrocytes so as to provide an alternative analysis of our data. The analysis, presented as Fig. 6A, demonstrates significant entropy, or transcriptional heterogeneity, for *pax2a* and cathepsin B (*ctsba*). The full list of genes displaying transcriptional heterogeneity in thyrocytes is provided as Table EV3.

The expression of *pax2a* and *tg* in the single-cell atlas is provided as Fig. 6B, B'.

Using this, we first demonstrate the presence of transcriptional heterogeneity within the thyrocyte population. We do not restrict the division of thyrocytes into two clusters. Using *pax2a* reporter line and RNA-Sequencing, we demonstrate the presence of at-least two sub-populations of thyrocytes with differential *pax2a* expression (Fig. 8, 9). Notably, *pax2a*-Low thyrocytes express *tg* (Fig. 8E, EV5), demonstrating their identity as differentiated thyrocytes.

In the revised text, we talk about transcriptional heterogeneity when discussing the thyrocyte population (Fig. 6), and discuss the thyrocyte sub-populations in the context of *pax2a* diversity that shows a bimodal expression distribution (Fig. 6B', 9C).

- This reviewer was also surprised by the relatively "heavy" approach (generation of the Pax2a reporter line) used to demonstrate the existence of two types of thyrocytes. Knowing that the reporter line was validated with a very good Pax2a antibody... The use of the reporter line is a bit short. The authors could for example validate the two populations of Figure 4A using the Pax2a FACS-sorted cells and RT-qPCR.

We completely agree with the suggestion of the author, and have performed the experiment as suggested. Using FACS, we demonstrate heterogeneity in *pax2a* expression levels with the thyrocytes in Fig. 9A – C. Further, the FACS-enriched sub-populations were transcriptionally profiled, which demonstrated differential gene expression in *pax2a*, mKO2 (the reporter for *pax2a*) and 2060 other genes. Together the differentially expressed genes represent 20% of the detected transcriptome, suggesting molecular differences between *pax2a*-High and *pax2a*-Low thyrocytes.

- In addition, data available in the sequencing dataset could be used to prove that the two populations are really active thyrocytes. This reviewer would suggest to present a table with the expression level of common and thyroid-specific genes such as TshR, Nis, Tpo, Duox, Tg, Pax2, TTF1 and other known transcription factors in the two populations to demonstrate that these two types of cells are indeed thyrocytes. Finally, image quality (Figure 6) could be improved and high-magnification images with several thyrocyte marker could be shown to convince the readers.

We strongly agree with the reviewer that this is a very important concern to address. To address this, we have taken three steps:

1. We have included the expression level of *tg* in the thyrocyte cluster in Fig. 6B, B'.
2. We performed antibody staining against *pax2a* on thin section obtained from *Tg(tg:nls-EGFP)* animals (Fig. 8 E-F). In this, we observed PAX2A-Low cells with *tg* reporter expression, suggesting that they are indeed differentiated thyrocytes.
3. We performed bulk RNA-sequencing of cells from *pax2a*-Low and *pax2a*-High population (Fig. 9D – E). The gene *tg* (thyroglobulin) was one of the top 15 expressed genes in the two populations (Fig. EV5). Further, genes involved in thyroid functionality: *nkx2.4b* (zebrafish homologue of NKX2.1 / TTF1), slc5a5 (NIS), TPO, tshr, ctsba, ctsk (cathepsin K), were detected in the two populations (Fig. 9E, Table EV4).

Conclusion 3 : cellular interaction network

- Most of the interactions revealed by the analysis seem to belong to the extracellular matrix and not to classical ligands such a Wnt, TGFb, FGF, PDGF,..... could the authors comment on this ? Considering that both endothelial cells and epithelial cells assemble their own basement membrane, the analysis will obviously reveal interactions between endothelial cells and epithelial cells....

We appreciate that the reviewer pointed this out. The enrichment of ligands related to extracellular matrix, and not growth factors, likely represents the homeostatic nature of the organ. Growth at 2 and 8 mpf is low (if not absent). Correspondingly, gene expression related to development and cell-cycle might be reduced. As stated in response to the next concern, extending the atlas to juvenile stage (1 mpf) would be beneficial to understand the regulators of cell-cycle.

However, to improve the cellular interaction network, we have incorporated physical information from the characterization of cell-populations surrounding the thyroid follicles (Fig. 4). Our experiments suggested that stromal, gills and NFE do not physically contact the thyrocytes. Thus, interactions based on ligands incorporated into the cell membrane were removed for these cell-populations.

****Minor comments to improve the Ms :****

- Could you explain how from 2 x 12 000 FACS-sorted live-cells (from six animals at each stage; 2 mpf and 8 mpf) you obtain 6249 cells (pooled of 2 mpf and 8 mpf), and why the two stages were first sorted separately and then pooled (?), as no differential analysis is carried out for the two stages.

The number of cells obtained for analysis represents cells that were successfully incorporated into droplets during library preparation and generated high-quality data that passed quality control (Fig EV2). FACS sorted cells were utilized for droplet generation using 10X Chromium that encapsulates cells with single-Poisson distribution (Zheng et al. 2017). This leads to approximately 50% cell capture rate, which is the ratio of the number of cells detected by sequencing and the number of cells loaded. Thus, we obtained 13,106 sequenced cells from 24,000 input cells (54.6 % cell capture rate). Further, the quality control criteria removed 6,857 low-quality cells (52.3 % dropout rate). We chose a stringent cut-off for quality control so as to remove technical artefacts from the analysis. We have added these detail to the Result section.

We pooled the two samples as the stages represent the range of homeostasis in zebrafish. We decided not to include differential expression between the two stages as the number of cells in multiple clusters were too low for individual stages (less than 100), and thus not trustworthy. In future, it would be of interest to extend the analysis for rapidly-growing juvenile (less than 1 mpf) and old-age (greater than 1.5 ypf animals) animals and to perform single-cell or bulk RNA-Seq. with high cell numbers. We have mentioned this drawback in the discussion section.

- Which method was used for the graph-based clustering ? KNN ? Louvain ? Random walk ?

The details have been added to the Method section. Specifically, the clustering was performed using graph-based method, Shared Nearest Neighbour (SNN), which is default for Seurat 2.3 package.

- How did you define the numbers of clusters ?

The number of clusters were defined by using the first five principal components as they displayed significant deviation from uniform distribution as accessed by JackStraw analysis. Further, a resolution of 0.3 was used in Seurat as the clusters generated by this parameter could be annotated using a cell-type specific marker from literature.

- Figure 4B, the color-code for the expression level would help the reader.

The color code has been added (Fig. 6B in revision).

- Figure 4C, violin plot for Pax2a: why do we find cells that do not express this gene in the two populations ? The same is true for tbx2a and ahnak ... is the clustering optimal ?

The detected expression of *pax2a* depends on its biological expression and technical dropout rate. Thus, the *pax2a*-High cluster also contains cells with no detectable expression of *pax2a*. Similar detection dynamics can be expected for other genes.

We have experimentally validated the variability in *pax2a* expression using antibody staining for endogenous *pax2a* protein in *tg:nls-EGFP* transgenic line (Fig. 8 E-F). With this, we can validate the presence of *pax2a* heterogeneity within thyrocytes.

- Figure 4C, blue violin plot for ptp4a3 does not seem to fit with the distribution of the points.

Due to the high number of cells that do not express *ptp4a3*, the cells collapse on each other at the bottom of the graph, thus making the violin plot seem different from the distribution. However, the plots were made with Seurat without changing any parameters.

- what is the function of tbx2a, ahnak, ptp4a3 and dusp5, which are not mentioned in the text.

The genes have been implicated in regulation of cell proliferation. However, we acknowledge that the evidence based on a handful of genes needs to be strengthened. For this, we have removed the figure panels from the revision, and instead identify genetic markers based on bulk RNA-sequencing analysis of *pax2a*-High and *pax2a*-Low population (Fig. 9D, E).

- Line 195: "pax2amKO2 reporter expression perfectly overlapped with PAX2A antibody staining". This reviewer would be more cautious as the images (Fig. 5 C, D and F) do not show a perfect colocalization: one can observe only blue or only red staining.

We have edited the text to "The *pax2a*$^{pax2a\text{-}T2A\text{-}mKO2}$ (abbreviated as *pax2a*^{mKO2}) reporter expression overlapped with PAX2A antibody staining in a majority of regions at 9.5 hours post-fertilization." The regions with single colors in Fig. 5C (Fig. 7C in revision) are due to differences in staining intensity between different regions.

- Line 246: it is proposed to "study the functional and replicative differences among the two sub-populations of thyrocytes". This reviewer completely agrees and the suggestion

made (vide supra) to use the datasets to assemble a table with the expression level of common and thyroid-specific genes such as TshR, Nis, Tpo, Duox, Tg, Pax2, TTF1 and other known transcription factors in the two populations could already give some indications on the functionality of these two types of cells. Expression of genes involved in cell-cycle control and/or apoptosis would be another possibility to better characterize the two populations. Lastly, the authors could perform the comparative analysis of ligand-receptor pairs between these two sub-populations to examine if they differentially interact with their environment.

We agree with the reviewer. Using bulk RNA-Seq. for *pax2a*-High and *pax2a*-Low thyrocytes (Fig. 9D, E), we now demonstrate that *pax2a*-High sub-population is enriched for genes involved in thyroid function (nkx2.4b, slc5a5 / NIS, TPO, tshr). The experiment suggests that *pax2a*-High cells are functionally active. We could not detect major differences in genes related to cell-cycle. However, this could be attributed to the fact that we have performed our analysis at stages where thyroid gland growth is low. In future, it would be of interest to profile the thyroid gland during metamorphosis, a stage of rapid growth in zebrafish.

Further, ligands and receptors differentially expressed in the two sub-populations are marked in Table EV4. In total, we detected 62 ligands and 66 receptors were differentially expressed between the two populations. Notably, we observed enrichment for *cxcl12a* expression in *pax2a*-Low population. The cytokine *cxcl12a* (zebrafish homologue of CXCL12 / SDF1) has been shown to recruit leukocytes to tissues (Ankamreddy et al. 2020; Isles et al. 2019). This gene is highlighted in Fig. 9E.

Text improvement:
Intro: thyroglobulin (TG) appears twice (line 46 and 49)
Results: Fig. 5 (not 8) (line 203 and 205)
Figure 3: stromal? (not skeletal)
Figure 4: fold change scale is missing
Figure 5: Thyroid gland (Gland)
Figure sup 2: Fluorescence-activated cell sorting (FACS) of zebrafish thyroid gland
Figure sup 3: number of unique molecular identifier
Figure sup 4: "are present in the zebrafish"

We thank the reviewer for pointing these errors. We have edited them.

Reviewer #1 (Significance (Required)):

The work performed by Gillotay et al. is clearly novel but descriptive. It provides a useful database to propose hypotheses to further study the thyroid gland. The single-cell RNA-seq analysis brings a molecular appreciation of the various thyroid cell populations, thyrocyte heterogeneity and intercellular signaling network. Although focused on the thyroid, results will be of interest to a larger audience than the thyroid community, especially the demonstration (if further and better validated) of thyrocyte heterogeneity and the intercellular communication possibilities.

In response to comments by Reviewer 1, we performed immunofluorescence imaging of pax2a (Fig. 8E, F) and bulk RNA-Sequencing of *pax2a*-Low and *pax2a*-High thyrocytes (Fig. 9D, E). We hope that this addresses the concerns of the reviewer.

Reviewer #2:

****Summary****

In this manuscript the authors present a single-cell transcriptome atlas of the zebrafish thyroid gland (possibly also including some adjacent tissues depending on how the dissection was performed, see comments below). The atlas includes cell clusters that are expected to be found in the thyroid of any higher vertebrate species (thyrocytes, blood vessels, lymphatic vessels, immune cells and fibroblasts), but also musculature/gills and a less well-defined population of non-follicular epithelium. The data will be made publicly available as a resource, by what seems to be a user-friendly web-interface (more accessible to a broader audience than customary raw sequencing data deposition, that I suppose will also be provided). The results are used to describe putative autocrine or paracrine signaling networks. The authors are able to further subdivide the thyrocyte cell cluster into two sub-populations with different transcriptomic features. Interestingly, these populations differ in their expression of for instance the key transcription factor pax2a, which is further demonstrated by the use of a novel zebrafish reporter strain.
In general, this is a clearly interesting, novel, nicely structured and well written manuscript and the data presented seems to be of high quality.

We would like to thank reviewer 2 for the constructive comments. We are glad that the reviewer finds our work interesting, novel and of high quality. We appreciate the reviewer's advice on additional experiments, analysis and on improving the clarity of the text. We have addressed all the concerns raised, and hope that our revised manuscript satisfies the reviewer.

****Major comments****

Key conclusions are convincing and performed with scientific rigor. As will be further discussed below the seemingly superficial description of the extent of tissue that was subjected to transcriptome analysis makes it a bit difficult to assess reproducibility outside the authors' lab.

We acknowledge the lack of clarity in the description of the tissue utilized for single-cell analysis. We have corrected this by providing a detailed step-by-step protocol for dissecting the organ in Methods Section, titled 'Dissection of the zebrafish thyroid gland'. Additionally, using immunofluorescence based imaging of the region and FACS, we estimate the proportion of thyroid follicular cells within the region. The results are presented in Fig. 1 C – G.

As far as I can see very little methodological detail or information is provided about how the dissection of the thyroid region was performed, more than that "the thyroid gland was collected" or that "we dissected out the entire thyroid gland". This is essential to

understand the significance of the cell populations that are described based on the transcriptomic data. The section "Tissue collection" describes dissection of the thyroid for whole-mount imaging. From Fig. 6A it seems that substantial amounts of non-thyroid tissue are included in this dissection. Is it the same kind of dissection that was performed for transcriptomics? Was the string of thyroid follicles shelled out from their surroundings or was some kind of en bloc dissection, including other neighboring tissues, performed (as suggested from the transcriptome cell populations data including e.g. gill transcripts)? In the latter case it would be good if the authors discuss in more detail what other tissues or structures that are expected to also be included in the dissected tissue and transcriptomic data.

In response to this concern of the reviewer, we have improved the clarity of the text in Results and Methods section. We have added the following text to the Result section, "The zebrafish thyroid gland is composed of follicles scattered in the soft tissue surrounding the ventral aorta (Fig. 1 A, B). Ventral aorta extends from the outflow tract of the zebrafish heart and carries blood from the ventricle to the gills. Dissection of the ventral aorta associated region (detailed in Methods section) provided us with tissue that included the thyroid follicles and parts of zebrafish gills (Fig. 1C). Using *Tg(tg:nls-EGFP)* transgenic line, which labels thyrocytes with nuclear green fluorescence (Fig. 1D), we estimated presence of 5.9 ± 1.9 % thyrocytes within the dissociated region (Fig. 1E)."

Further, the Methods section now defines a step-by-step protocol for dissociation ('Dissection of the zebrafish thyroid gland').

In addition, we have improved the characterization of the dissected region, as stated in response to the next comment.

It would facilitate understanding if the thyroid is outlined in Fig. 1A as well as the region that was dissected for downstream single cell sequencing.

A whole-mount 3D reconstruction of the region is presented in Fig. 1C. A transverse section from the region is presented as Fig. 1D, while quantification of the percentage of thyrocytes in the transverse sections is provided in Fig. 1E.

The clusters seem logical given what cell types that could be expected in the region (but depending on how dissection was performed). The only exception is cluster 7 (non-follicular epithelium; NFE). I do understand that relative sizes of the clusters do not necessarily reflect the endogenous relative abundance of different cell types, as I guess they may be more or less prone to enzymatic dissociation, survival etc. Nevertheless, the number of cells in the NFE cluster (831 cells) seems sizeable relative to the number of thyrocytes (267 cells). In my opinion, a major weakness of the current manuscript is that little detail is provided about this cell population and that no attempt to at least spatially localize these cells is presented.

Detailed characterization of the cell-populations surrounding the thyroid follicles is now presented in Fig. 4. In addition, we have quantified the percentage of thyrocytes in the region (Fig. 1 E), which demonstrates that thyrocytes represent 5.9 ± 1.9 % of the cell-

population. Additionally, we have presented FACS analysis of the dissociated region as Fig. 1 F - G, which corroborates the imaging based quantification. Both quantifications are in the same range as the proportion of thyrocytes identified in the single-cell analysis (4.27 % - 267 out of 6249 cells). Thus, we do not believe that cell-loss had a big impact on the relative abundance of thyrocytes in the single-cell atlas.

A detailed characterization of NFE cells is provided in response to the next three comments, which includes visualization of the population using TP63 / *p63* antibody staining in Fig. 4D. Particularly, Fig. 4D contains 72 thyrocytes and 302 TP63+ nuclei, thereby pointing to the higher relative abundance of NFE in the region.

The NFE cells are characterized by TP63 expression and the authors speculate that they might show homology to main cells of solid cell nests. From previous zebrafish literature it seems like what is supposedly ultimobranchial bodies (or ultimobranchial glands more similar to those in avian species) are located rather distant from the thyroid follicles (Alt et al 2006). Is it possible that these structures are included in the dissection that has been performed? As solid cell nests are supposed to be ultimobranchial body remnants (with calcitonin positive and calcitonin negative epithelial cells) it would be good if the authors discuss in more detail what is known about the ultimobranchial bodies in zebrafish, if they are located inside the zebrafish thyroid, in an anatomical region that is included in the dissected tissue of this study or in a region that is likely not included.

As stated by the reviewer, the ultimobranchial bodies lie distant to the thyroid gland. They lie as a pair of follicles on top of the sinus venous (Alt et al. 2006), which is a blood vessel that delivers blood to the atrium. In contrast the thyroid follicles sit loosely on top of ventral aorta that connects to the ventricle via the outflow tract (Fig. 1B). Thus, the collection of the ultimobranchial bodies is not expected. This is also corroborated by the absence of *calcitonin* (*calca*) expression in the NFE (Table EV1). This has been added to Discussion section.

In higher vertebrates, P63 expression is typically seen in basal cells of stratified epithelia (as for instance in the esophagus), in myoepithelial cells, in the urothelium and in the thymus. Is it possible that the TP63 expressing NFE population corresponds to cells of the zebrafish thymus (that might perhaps explain the seemingly large immune cell population in cluster 4)? Could TP63 expressing NFE cells represent the esophagus (if included in the dissection)? As so little detail is provided about the dissection procedure this opens up for speculation and it would be good if the authors discuss these possibilities, as some of them might perhaps be unlikely or even impossible given regional anatomy of the zebrafish and how the dissection was performed.

The dissection region is now characterised in detail in Fig. 1 C – E. The presence of immune cells (macrophages) in the proximity of thyroid follicles is validated in Fig. 4B. The presence of NFE is presented as Fig. 4D, and explained in detail in response to the next comment.

To gain better understanding of the sizeable TP63 expressing NFE population the authors briefly mention the possibility of future studies utilizing a TP63 reporter. If a reporter line is not available, immunofluorescent detection of P63 (as presented for PAX2A in Fig. 5 and

E-cadherin in Fig. 6) would be desirable to provide more insight into the location of the NFE population. Given the proficiency the authors demonstrate in this manuscript when it comes to zebrafish imaging, at least whole-mount immunostaining of P63 in the region that was dissected for transcriptome analysis seems clearly feasible, both with respect to resources and time needed (perhaps in the range of 1-3 months).

To clarify the presence of NFE cells, we have followed the suggestion of the reviewer and performed immunostaining against TP63. The result is presented as Fig. 4D. The staining was performed on thin (8 µm) sections, allowing us to ensure uniform antibody penetration. As depicted in the image, TP63+ cells are part of the gills. This population possibly represents a progenitor population, similar to the TP63+ basal layer in the zebrafish (Guzman et al. 2013) and mammalian (A. Yang et al. 1999) epithelium. Additionally, a sub-set of TP63+ cells were observed in the region between the thyroid follicles and gills. Our data provides an exciting opportunity for an in-depth study of these cells in future, particularly using *tp63*-regulatory region driven transgenic reporter and Cre lines.

****Minor comments****

It is a little bit confusing that different color coding for the various cell populations is used in Fig. 3B as compared to Figs. 1 and 2.

The color coding for Fig. 3B (Fig. 5B in revised manuscript) has been modified in accordance to the once used in Fig. 1 and 2.

In the discussion of the intercellular interaction network (Fig. 3B) the authors clearly point out that anatomical barriers are not modelled. Nevertheless, it would be more informative if this description was able to sort out ligands that are secreted, from ligands that are not secreted and would require physical interaction between thyrocytes and a different cell population for signaling to occur.

We have now improved the intercellular network to resemble the putative physical interactions. By characterizing the different cell-populations present in the atlas (Fig. 4), we recognized that gills, stromal and NFE are not in physical proximity of the thyrocytes. Thus, these three cell-populations would not be able to communicate via ligands attached to the cell membrane. Hence, we have pruned the network to remove cell-membrane attached ligands for these three cell-populations. Only secreted ligands were considered for the mentioned cell-populations. In accordance, the figure and Table EV2 has been updated.

The authors describe thyroid solid cell nests as "... lumen containing irregular structures located between the thyroid lobes in mammals...". Solid cell nests of the thyroid in higher vertebrates (e.g. humans and dogs) are located within the thyroid lobes and not between the lobes (i.e. the right and left thyroid lobes). Moreover, the authors write that "Recently, epithelial cells have been reported in a structure called the Solid Cell Nests (SCN) of the thyroid..." and give reference to a paper from 1988. If that is recent or not might be a matter of opinion, but to the best of my knowledge, solid cell nests were describe already by Getzowa in 1907 and I suppose that the epithelial identity (or at least

epithelioid morphology) has been appreciated for long.

We thank the reviewer for pointing this out. We have added the reference to the original study by Dr. Sophia Getzowa identifying SCN (Getzowa 1907). As the original study is in German, we have also added a reference to a recent article attributing the discovery of SCN to Dr. Getzowa and performing immunohistochemistry analysis of the tissue (Ríos Moreno et al. 2011). Notably, the authors note the presence of TP63 staining, along with the absence of TG and Calcitonin staining, in SCN main cells – similar to the expression profile of NFE in our atlas. Finally, we have edited the text to accurately describe their location in the mammalian thyroid gland.

Reviewer #2 (Significance (Required)):

The authors provide a single-cell transcriptomic atlas of the zebrafish thyroid gland. To the best of my knowledge this is certainly a unique resource. In that sense it provides novel and significant information that will surely facilitate our further understanding of thyroid biology. It will surely be of great interest and value to the thyroid community, but probably also to a wider audience interested in e.g. zebrafish biology, endodermal biology and the biology of endocrine glands. Their finding and direct demonstration of transcriptional heterogeneity within the thyrocyte population is very interesting, also in different contexts of thyroid disease. However, I leave it to other referees to comment on the conceptual uniqueness of the current manuscript (i.e. a single-cell transcriptomic atlas of a zebrafish organ). Does it provide conceptually unique information, or does it add to an expanding collection of single-cell transcriptomic atlases of zebrafish organs?

References:

Alt, Burkhard, Saskia Reibe, Natalia M. Feitosa, Osama A. Elsalini, Thomas Wendl, and Klaus B. Rohr. 2006. “Analysis of Origin and Growth of the Thyroid Gland in Zebrafish.” *Developmental Dynamics* 235 (7): 1872–83. https://doi.org/10.1002/dvdy.20831.

Ankamreddy, Harinarayana, Heiyeun Koo, Young Jae Lee, and Jinwoong Bok. 2020. “CXCL12 Is Required for Stirrup-shaped Stapes Formation during Mammalian Middle Ear Development.” *Developmental Dynamics*, May, dvdy.180. https://doi.org/10.1002/dvdy.180.

Getzowa, Sophia. 1907. “Über Die Glandula Parathyreodeaa, Intrathyreoideale Zellhaufen Derselben Und Reste Des Postbranchialen Körpers.” *Virchows Archiv Für Pathologische Anatomie Und Physiologie Und Für Klinische Medizin* 188 (2): 181–235. https://doi.org/10.1007/BF01945893.

Guzman, A., J. L. Ramos-Balderas, S. Carrillo-Rosas, and E. Maldonado. 2013. “A Stem Cell Proliferation Burst Forms New Layers of P63 Expressing Suprabasal Cells during Zebrafish Postembryonic Epidermal Development.” *Biology Open* 2 (11): 1179–86. https://doi.org/10.1242/bio.20136023.

Isles, Hannah M., Kimberly D. Herman, Anne L. Robertson, Catherine A. Loynes, Lynne R. Prince, Philip M. Elks, and Stephen A. Renshaw. 2019. “The CXCL12/CXCR4 Signaling Axis Retains Neutrophils at Inflammatory Sites in Zebrafish.” *Frontiers in Immunology* 10 (July). https://doi.org/10.3389/fimmu.2019.01784.

Liu, Baolin, Chenwei Li, Ziyi Li, Xianwen Ren, and Zemin Zhang. 2019. “ROGUE: An Entropy-Based Universal Metric for Assessing the Purity of Single Cell Population.” *BioRxiv*,

January, 819581. https://doi.org/10.1101/819581.
Ríos Moreno, María José, Hugo Galera-Ruiz, Manuel De Miguel, María Inés Carmona López, Matilde Illanes, and Hugo Galera-Davidson. 2011. “Inmunohistochemical Profile of Solid Cell Nest of Thyroid Gland.” *Endocrine Pathology* 22 (1): 35–39. https://doi.org/10.1007/s12022-010-9145-4.
Teschendorff, Andrew E., and Tariq Enver. 2017. “Single-Cell Entropy for Accurate Estimation of Differentiation Potency from a Cell’s Transcriptome.” *Nature Communications* 8 (1): 15599. https://doi.org/10.1038/ncomms15599.
Yang, A, R Schweitzer, D Sun, M Kaghad, N Walker, R T Bronson, C Tabin, et al. 1999. “P63 Is Essential for Regenerative Proliferation in Limb, Craniofacial and Epithelial Development.” *Nature* 398 (6729): 714–18. https://doi.org/10.1038/19539.
Yang, Shiyi, Sean E. Corbett, Yusuke Koga, Zhe Wang, W Evan Johnson, Masanao Yajima, and Joshua D. Campbell. 2020. “Decontamination of Ambient RNA in Single-Cell RNA-Seq with DecontX.” *Genome Biology* 21 (1): 57. https://doi.org/10.1186/s13059-020-1950-6.
Zheng, Grace X. Y., Jessica M. Terry, Phillip Belgrader, Paul Ryvkin, Zachary W. Bent, Ryan Wilson, Solongo B. Ziraldo, et al. 2017. “Massively Parallel Digital Transcriptional Profiling of Single Cells.” *Nature Communications* 8 (1): 14049. https://doi.org/10.1038/ncomms14049.

ULB UNIVERSITÉ LIBRE DE BRUXELLES

16 July, 2020

To,
Dr. Esther Schnapp,
Editor, EMBO Reports,
Heidelberg, Germany

Dear Dr. Schnapp,

We are pleased to submit our revised manuscript entitled "**Atlas of cell-cell communication and thyrocyte diversity for the zebrafish thyroid gland**". The manuscript was reviewed by Review Commons (RC-2020-00157R) and invited for Revision at EMBO Reports (EMBOR-2020-50612V2). We would like to thank the reviewers for taking the time to provide us with insightful and constructive comments, which helped us in improving the manuscript.

This study provides the first single-cell atlas of the thyroid gland, along with the first documentation of cellular heterogeneity within the thyroid follicular cells. As mentioned by Reviewer #2, "this is a clearly interesting, novel, nicely structured and well written manuscript and the data presented seems to be of high quality." Our manuscript will contribute significantly to the basic understanding of the thyroid gland and the different cell-types within the organ, including blood vessels, lymphatic vessels, immune cells, fibroblasts and zebrafish gills. As stated by Reviewer #1, the "results will be of interest to a larger audience than the thyroid community."

Our revision addresses all the concerns of the reviewers. During invitation to submit to EMBO Reports, we were asked to perform additional analysis, namely validation of transcriptional heterogeneity using bulk RNA-Seq. analysis. We have performed the experiment, which does indeed validate our findings. The revision is formatted according to EMBO guidelines and accompanied by point-by-point response to reviewers and a suggestion for Cover Art.

We thank you for your time and consideration. Please address all correspondence concerning this manuscript to me at sumeet.pal.singh@ulb.ac.be.

Sincerely,

Sumeet Pal Singh, PhD

IRIBHM
Campus Erasme ULB
808 route de Lennik
1070 Brussels
Belgium

https://iribhm.org/

IRIBHM

1st Revision - Editorial Decision 2nd Sep 2020

Dear Sumeet,

Thank you for the submission of your revised manuscript. We have now received the enclosed reports from the two referees.

As you will see, while referee 1 is more critical, referee 2 supports the publication of your study following minor revisions. Given the disagreement, I contacted an expert advisor for arbitration, and the advisor concurs with referee 2 that the findings are interesting and the data strong.

I would therefore like to ask you to address the remaining referee concerns (also in a point by point response) and the following editorial comments, so that we can proceed with the official acceptance of your newly revised, final manuscript.

Please address the following points:

- Remove cell communication from the title, as no data is provided in the manuscript to support this claim.

- not more than 5 keywords can be added, please correct.

- Our REFERENCE FORMAT has changed to Harvard style, please correct. The EMBO reports style is also in EndNote.

- Please delete the DATASET/TABLE EV LEGENDS from the article file and add them to the first tab of the EV excel files.

- The movie needs to be ZIPed with its legend and uploaded as one ZIPed file.

- The EV figure legends need to be moved to after the main figure legends.

- I attach to this email a related manuscript file with comments by our data editors. Please address all comments in the final manuscript.

- The synopsis image looks good, but the text is too small at the final image size of 550 pixels x 200 pixels. Can you please send us an image at the correct size with larger text?

- Please also send us a short (1-2 sentences) summary of your findings and their significance and 3-4 bullet points highlighting key results for our website.

I would like to suggest a few changes to the abstract that needs to be written in present tense. Please address the following comments and let me know whether you agree with the changes:

The thyroid gland regulates growth and metabolism via the production of thyroid hormone in follicles composed of thyrocytes. So far, thyrocytes have been assumed to be a homogenous population. To uncover heterogeneity in the thyrocyte population, and molecularly characterize the non-thyrocyte cells surrounding the follicle, we developed a single-cell transcriptome atlas of the zebrafish thyroid gland [please re-write, see referee comment]. The 6249-cell atlas includes profiles

of thyrocytes, blood vessels, lymphatic vessels, immune cells and fibroblasts. Further, the thyrocytes show expression heterogeneity, including bimodal expression of the transcription factor pax2a. To validate thyrocyte heterogeneity, we generated a CRISPR/Cas9-based pax2a knock-in line that monitors pax2a expression in the thyrocytes. A population of pax2a-low mature thyrocytes interspersed in individual follicles can be distinguished. We corroborate heterogeneity within the thyrocyte population using RNA-Sequencing of pax2a-high and pax2a-low thyrocytes, which demonstrates 20% differential expression between the two sub-populations. Our results identify and validate transcriptional differences within the presumed homogenous thyrocyte population.

I look forward to seeing a final version of your manuscript as soon as possible.

Best regards,
Esther

Esther Schnapp, PhD
Senior Editor
EMBO reports

Referee #1:

Gillotay et al. propose a descriptive molecular study of the thyroid gland in Zebrafish. They first present transcriptomic profiling at the single cell level of a large dissociated thyroid gland region. Obviously, single-cell sequencing revealed the presence of different cell populations (7 clusters) and the presence of these populations were logically confirmed by immunolocalization studies. One of the identified cluster, termed non-follicular epithelial cells, clearly deserve further studies. Another cluster, which was the main focus of the Ms correspond to thyrocytes. As single-cell sequencing provides a higher resolution of cellular differences, it is not surprising that in-depth analysis of the thyrocyte cell cluster revealed transcriptional heterogeneity. Using immunolocalization, a newly created transgenic reporter line (under the control of the Pax2a locus) and bulk sequencing, they validated the existence of two transcriptionally distinct types of "thyrocytes". All in all, this manuscript contribute to the academic knowledge, and further analyses should reveal the potential impact of this work and the functional significance of these observations, in particular of the most interesting data : the two thyrocyte populations.

According to this reviewer, the manuscript by Gillotay could be published as it is in another journal (PlosONE or Biology Open), but it lacks conceptual advances, functional or physiological significance for EMBO reports.

Minor comments :

- Is the term "Atlas of cell-communication » in the first part of the title fully justified ? It is solely based on sequence analysis of ligands and receptors and is currently not supported by biological data. It could help if some of the ligands and receptors would be characterized by ISH or immunolocalization on sections.

- It is clearly mentioned that transcriptomic data were obtained at 2 mpf and 8 mpf. « The process provided single-cell transcriptomic profiles for 2986 and 3263 individual cells for 2 mpf and 8 mpf, respectively (Fig 1F and G). In the rest of the paper, stages are not indicated anymore (except in

images) and it seems that the sequencing analysis results come from the pooling of the two stages. If this is correct, part of the transcriptional heterogeneity could be attributed to the pooling of the data from the two stages. This pooling should be better justified and discussed by the authors.

- It would have been nice to show/color the cells from the two stages in the t-SNE of figure 2. TPO intensity clearly show a graded repartition....

- Scale bars should be added in Figure 1A-C.

- Transcriptional heterogeneity is exemplified by comparing normalized gene expression of Tg (unimodal) and Pax2a (bimodal). This reviewer would have compared Pax2a expression level with another transcription factor (generally expressed at a lower level) and not with one of the most highly expressed thyroid gene, thyroglobulin !!

- Generation of Pax2a knock-in reporter line. The knock-in line displayed mKO2 fluorescence in the thyroid gland (Fig 7C-F). In Figure 7C, the thyroid is clearly green (Tg-EGFP) but there is no red (pax2a-mKO2) or blue (pax2a-Ab) fluorescence. The authors nevertheless conclude : "overlay in a majority of regions... », « Faithfull recapitulation..." . Higher magnification of the thyroid should be shown. Data are convincing in the pronephros (Fig 7D; Nuclear green+ Blue and cytoplasmic red) but not for the thyroid (Fig 7E) !

- Quality of Figure 8 is considered as low for publication by this reviewer.

- In panels B,C and D, of Figure 8, 1 to 4 nuclei per follilcular section do not express Pax2A reporter gene (mKO2), while in panel F, more than 50% of the thyrocytes do not express Pax2A. Is it due to the half life of the mKO2 as compared to Pax2A ? Is it due to the sensibility of the antibodies used ? This should be discussed as it could help proposing working hypotheses.

- In panel 8F, Tg-NLS-EGFP is highly expressed in all the thyrocytes (DAPI+), thus in Pax2A high thyrocytes and in Pax2A low thyrocytes.... « Notably, both PAX2A-low and -High cells display tg promoter-driven EGFP expression, thereby confirming their differential status ».
However, in the bulk sequencing (Fig 9E), the Pax2A-low population reveal that Tg is weakly expressed in this population (blue or light yellow as compared to orange)..... This contradicts with their previous conclusion : « Notably, both PAX2A-low and -High cells display tg promoter-driven EGFP expression, thereby confirming their differential status » and « The high expression of tg in both populations confirms the identity of the cells as thyrocytes ». Authors should explain or discuss this discrepancy.

- In addition, analysis of the bulk sequencing of the Pax2A-low population (Fig 9E) suggests that this population expresses very low level of most thyroid-specific genes (Tg, nkx2.4, slc5a5, TPO, TshR, ctsba, ctsk). One can thus wonder whether these cells are active thyrocytes ? or could correspond to a kind of inactive cellular state comparable to the hot and cold follicles described in mammals.

Referee #2:

I think the authors have carefully addressed comments from both referees. This has led to

important clarifications (that have raised a few further questions, see below) as well as substantial rewriting of the manuscript. I usually refrain from raising what might be perceived as new issues at a revision stage. When reading this revised version, my comments mainly revolve around two themes. The first one is the result of the new, more detailed information about the thyroid dissection provided in the revision, that was not clear before. The second one is something that passed unnoticed by me when reading the prior version, but that became evident when reading the revised manuscript. In my opinion, even though I elaborate on these two themes at some length, these comments do not require additional experiments, but potentially some textual revision of the current manuscript.

Major comments:

1. On one hand I find it commendable that the authors in the revised manuscript have described dissection of the zebrafish "thyroid gland" in much more detail. On the other hand, when I thus now get a better picture of how the dissection was performed it raises some additional comments and questions.

I do not doubt for a second that the zebrafish has thyroid follicular tissue and that this is an extremely valuable model organism for thyroid research. It is not readily apparent to me though if the thyroid tissue of the zebrafish can be described as "a thyroid gland". As far as I understand, zebrafish thyroid tissue is more or less a number of follicles loosely distributed along the aorta. To me, a gland is an anatomically and histologically coherent and often encapsulated organ, even though it might be argued that for instance colonic crypts are sometimes referred to as colonic glands. When the dissection procedure is now nicely detailed in the revised manuscript, this issue becomes not only a matter of semantics or definitions in my opinion.
If I understand correctly, the dissection encompasses the ventral aorta, parts of the gills, the thyroid follicles and presumably surrounding connective tissue, which to me seems like kind of a regional en bloc dissection. This is obviously the result of very skilled and careful experimental work. Still, can the resulting tissue that is then sequenced be equaled to that of a "zebrafish thyroid gland", an expression that is repeatedly used in the manuscript, even though "the zebrafish thyroid gland" is certainly present as a constituent part of the dissection? To me it seems that analysis has been performed of the "zebrafish thyroid region"/"the zebrafish thyroid and its immediate surroundings" rather than "the zebrafish thyroid gland". This seems to be reflected by the fact that follicular cells constitute only approximately 6% of the cells in the dissected region. In thyroid glands (as encapsulated coherent structures) of higher vertebrates, follicular cells are certainly in majority with only minor contributions/fractions of other cell types.
The authors seem to more or less equal or define "the zebrafish thyroid gland" as the anatomical region that they have dissected. This region obviously includes for instance part of the gills. Do the authors consider these to be part of "the zebrafish thyroid gland"? For instance, to me it seems likely that a major part of the tp63-positive NFE cell population are constituent cells of the gills. The size of the "zebrafish thyroid gland" and its proportion of NFE cells then largely becomes a matter of how much gill tissue is included in the dissection. As a comparison, even though the trachea, parathyroids and thymus are immediately adjacent to the thyroid gland in higher vertebrates, these structures would not be described as parts of the thyroid gland, even if they were included in an en bloc dissection.
When reading some parts of the manuscript, this becomes a matter of some concern to me, e.g:

p. 4, line 76: "... we develop the first atlas of the thyroid gland..."
p. 4, line 80: "... comprehensively represent the cells present in the zebrafish thyroid gland"
p. 7, line 132: "Identification of cell-types present in the zebrafish thyroid gland."

p. 9, line 184: "Having defined the cell types of the thyroid gland, we quantified potential cell-cell interactions between thyrocytes and all cell types present in the organ..."
p. 9, line 198: "Further, the ligand cyr61 is broadly expressed in the thyroid gland..." (note that for instance expression is seen also in the gills, that makes me come back to my question if the authors consider these as a part of the zebrafish thyroid gland).
p. 15, line 327: "This allowed us to capture yet poorly characterized cell-populations within the thyroid gland"

Can really the identified cell populations be considered to be "within the thyroid gland"?

I think the authors need to elaborate on this and possibly modify some parts of the text. Also, the anatomical differences between thyroid in zebrafish and higher vertebrates may need to be detailed even more clearly than presently at the beginning of the results section. In my opinion the authors have characterized the zebrafish thyroid region, rather than the zebrafish thyroid gland (even though the thyroid follicles are a part of the dissection). I do not think this makes the findings less interesting at all, as signaling cues may be provided by adjacent cell populations, regardless if they are considered constituents of a "zebrafish thyroid gland" or not. Moreover, the clustering certainly sorts out the follicular cells and provides a catalogue of their transcriptomes as well as that of other cell populations of the region. Still, I think this issue may warrant some rewriting.

I furthermore think the authors may overstate the NFE population as part of the zebrafish thyroid gland and I have some comments about this population. I would not really consider tp63 as a "...known marker of ... epithelial tissues", but rather as a protein that may be expressed in some epithelia. Indeed, TP63 is expressed in human stratified epithelia and urothelium, but absent in most other epithelial tissues. Moreover, TP63 is characteristically expressed also in myoepithelial cells. The latter cell type seems to be well represented in gills (PMID: 20143320). Indeed, the authors find tp63 positive "... NFE celler scattered throughout the gills and in the region adjacent to the follicles". Perhaps this should be further commented as a major part of the NFE population thus perhaps cannot really be considered a constituent of the "zebrafish thyroid gland" (see above).
What about more prototypic epithelial markers in the NFE population (such as cytokeratins and E-cadherin)?

2. Even though I am aware that a distinction may be made between molecular and functional homogeneity, when rereading the manuscript I think that the authors incorrectly describe the prevailing view of thyroid follicular cells as a homogenous population. That is not the case. It has for long been appreciated that thyroid follicular cells are functionally heterogeneous, both in normal conditions as well as in disease, which is perhaps most apparent in multinodular goiter. This is evident already from the histology of the normal human thyroid gland, where inactive follicles outlined by a flat epithelium are mixed with active follicles (colloid scalloping etc) outlined by taller epithelial cells. I think it is fair to say that this heterogeneity has generally been assumed to be reflected also at the molecular level. Just to mention a few references:

Naturally occurring clones of cells with high intrinsic proliferation potential within the follicular epithelium of mouse thyroids
Smeds et al. Cancer Res. 1987. 47(6):1646-51. PMID: 3815361

Natural heterogeneity of thyroid cells: the basis for understanding thyroid function and nodular goiter growth
Studer et al. Endocr Rev. 1989. 10(2):125-35. PMID: 2666115

Various facets of the intercellular heterogeneity in thyroid primary culture
Baptist et al. Thyroidology. 1991. 3(3):109-13. PMID: 1726924

Intercellular heterogeneity of early mitogenic events: cAMP generalizes the EGF effect on c-Fos protein appearance but not on MAP kinase phosphorylation and nuclear translocation in dog thyroid epithelial cells
Baptist et al. Exp Cell Res. 1995. 221(1):160-71. PMID: 7589241

I think this is in stark contrast to some statements of the manuscript, e.g.

p. 1, line 20: "So far, the thyrocytes have been assumed to be a homogenous population."
p. 1, line 34: "... nominally homogenous thyroid population."
p. 2 line 39: "Single-cell analysis uncovers latent heterogeneity in thyroid follicular cells." (what does "latent" mean in this context?)
p. 3, line 59: "... it remains unknown if all the thyrocytes resident in the thyroid gland are equally capable of generating thyroid hormones."
p. 15, line 319: "Our result suggest that the gland may be divided into functional and resting thyrocytes." (I think it is a prevailing view that this is the case, but that the current study demonstrates a genetic/transcriptomic correlate to this notion, see below).
p. 18, line 401: "... beyond a nominally homogenous endocrine cell population".

I do not mention this to discredit the present manuscript and I do not think it makes it less interesting. On the contrary I think it provides a valuable step towards a genetic/transcriptomic correlate to the existing notion of thyrocyte heterogeneity. I think that this needs to be addressed in the sense that previous relevant literature is acknowledged, sections that convey the impression that thyrocyte has previously not been appreciated be modified and putting the current findings into this context.

Minor comment:

The introduction makes some statements about thyroid biology and pathology that strike me as a bit odd.

p. 3, line 46: The authors write that:
- "Thyroid dysfunction affects 100 million people worldwide and is treatable by hormone replacement." Do the authors refer to hypo- as well as hyperthyroidism in this description? The latter condition is not primarily treatable by hormone replacement (but most often requires it after e.g. radioidine ablation or surgical treatment.
- if left untreated may result in "profound adverse effects of the human body, including mental retardation, goiter or dwarfism". Whereas it can be debated if goiter can really be considered a "profound adverse effect", the other two effects ("mental retardation" and "dwarfism") only relate to untreated congenital hypothyroidism, that constitutes a very small fraction of these 100 million people. Profound adverse effects of no treatment, such as mental retardation and dwarfism, are only relevant when considering this comparatively very small group and is not generally applicable on untreated thyroid disease. I think this needs to be clarified.

On balance, even though it might seem that I have quite a lot of comments, these are possible to address by rephrasing some parts of the manuscript and do not require new experiments being performed. Even though I am aware that EMBO Reports usually allows only a single revision I think the present manuscript is a nice contribution and a valuable resource that should be offered a

possibility to address these issues of textual modification. I consider this manuscript to be of high quality, novel and important not only to the thyroid field, but to a broader audience.

Response to Reviewers

We thank the reviewers for insightful feedback that helped improve the manuscript considerably. In response, we have made multiple edits in the manuscript that we hope satisfy the issues raised by the referrers. Edit made in the manuscript are marked in 'blue' to allow easy visualization.

Please find our point-by-point response below.

Referee #1:

Gillotay et al. propose a descriptive molecular study of the thyroid gland in Zebrafish. They first present transcriptomic profiling at the single cell level of a large dissociated thyroid gland region. Obviously, single-cell sequencing revealed the presence of different cell populations (7 clusters) and the presence of these populations were logically confirmed by immunolocalization studies. One of the identified cluster, termed non-follicular epithelial cells, clearly deserve further studies. Another cluster, which was the main focus of the Ms correspond to thyrocytes. As single-cell sequencing provides a higher resolution of cellular differences, it is not surprising that in-depth analysis of the thyrocyte cell cluster revealed transcriptional heterogeneity. Using immunolocalization, a newly created transgenic reporter line (under the control of the Pax2a locus) and bulk sequencing, they validated the existence of two transcriptionally distinct types of "thyrocytes". All in all, this manuscript contribute to the academic knowledge, and further analyses should reveal the potential impact of this work and the functional significance of these observations, in particular of the most interesting data : the two thyrocyte populations.

According to this reviewer, the manuscript by Gillotay could be published as it is in another journal (PlosONE or Biology Open), but it lacks conceptual advances, functional or physiological significance for EMBO reports.

We thank the reviewer for acknowledging that the study would positively contribute to the current knowledge in the field. We agree that more work is needed to completely understand the complex biology of the thyroid gland. However, we do believe that we have shown, for the first time, molecular heterogeneity in the thyrocyte population and as such provide a pioneering study on genetic diversity in the organ.

Minor comments :

- Is the term "Atlas of cell-communication » in the first part of the title fully justified ? It is solely based on sequence analysis of ligands and receptors and is currently not supported by biological data. It could help if some of the ligands and receptors would be characterized by ISH or immunolocalization on sections.

The *in silico* connectome provided in the study would act as a reference for generating hypothesis related to cell-cell communication in the organ. A majority of thyrocytes ligands identified in single-cell atlas were also detected in the bulk RNA-Seq. of thyrocytes (marked in Table EV4). However, we do agree that we have not validated the functional significance

of the expression of ligands or receptors in this study (we are currently investigating the significance of cxcl12 – cxcr4 axis between thyrocytes – macrophages in a follow-up study, an axis identified using the atlas). Thus, we have removed the focus on 'cell-communication' in the title of the study and focussed on heterogeneity in the thyrocyte population, which is validated using multiple follow-up experiments.

- It is clearly mentioned that transcriptomic data were obtained at 2 mpf and 8 mpf. « The process provided single-cell transcriptomic profiles for 2986 and 3263 individual cells for 2 mpf and 8 mpf, respectively (Fig 1F and G). In the rest of the paper, stages are not indicated anymore (except in images) and it seems that the sequencing analysis results come from the pooling of the two stages. If this is correct, part of the transcriptional heterogeneity could be attributed to the pooling of the data from the two stages. This pooling should be better justified and discussed by the authors.

- It would have been nice to show/color the cells from the two stages in the t-SNE of figure 2. TPO intensity clearly show a graded repartition....

The data from the two stages has been added as Appendix Figure S4 and shows that the cells from two ages do not separate in t-SNE plot. The cells from two ages mix with each other suggesting that the sampling age does not play a major role in defining the gene expression. Further, the heterogeneity in pax2a expression is observed at both stages. This is to be expected as both ages correspond to healthy, young adult. Old age in zebrafish would be closer to two years of age, which would be a fantastic time-point to profile in future.

- Scale bars should be added in Figure 1A-C.

We have added this.

- Transcriptional heterogeneity is exemplified by comparing normalized gene expression of Tg (unimodal) and Pax2a (bimodal). This reviewer would have compared Pax2a expression level with another transcription factor (generally expressed at a lower level) and not with one of the most highly expressed thyroid gene, thyroglobulin !!

The t-SNE plot and histogram for *pax2a* serves to demonstrate the expression distribution, as was requested by the reviewer in the previous review. It is not used to identify genes that display transcriptional heterogeneity, for which we used gene entropy (Fig. 6A).

- Generation of Pax2a knock-in reporter line. The knock-in line displayed mKO2 fluorescence in the thyroid gland (Fig 7C-F). In Figure 7C, the thyroid is clearly green (Tg-EGFP) but there is no red (pax2a-mKO2) or blue (pax2a-Ab) fluorescence. The authors nevertheless conclude : "overlay in a majority of regions... », « Faithfull recapitulation..." . Higher magnification of the thyroid should be shown. Data are convincing in the pronephros (Fig 7D; Nuclear green+ Blue and cytoplasmic red) but not for the thyroid (Fig 7E) !

The figure panel 7 C – F displays expression from *pax2a*^{mKO2} (red); *Tg(tg:nls-EGFP)* (green) animals immunostained with PAX2A antibody (cyan). The expression of pax2a reporter (red)

and PAX2A antibody (cyan) is clearly distinguishable in the pronephros and brain. However, in the thyroid the green colour from the tg reporter driven nls-EGFP masks the cyan from PAX2A antibody. The image panel with thyroid (7E) is repeated as Appendix Figure S5 with each colour separated. Appendix Figure S5 clearly shows specific expression of *pax2a* reporter and PAX2A antibody stain specifically in the thyrocytes.

- Quality of Figure 8 is considered as low for publication by this reviewer.

We are disheartened by the fact that the Figure 8 is not of publication standard as deemed by the reviewer. We had put a lot of effort into generating images where the entire thyroid gland is captured in 3D without destroying the tissue, while at the same time providing images with single-cell resolution. This took months of optimization, including testing of various embedding methods and techniques for image capture (including tissue clarification). Finally, we are able to reasonably capture the entire organ at single-cell resolution (Fig. 1C and Fig 8). We put this effort to obtain unbiased data, rather than displaying sections with a couple of follicles that could be cherry-picked depending on the requirement. Notably, our presentation of the entire adult thyroid gland organ at cellular resolution is unique in zebrafish literature. In addition, we do believe that it demonstrates the heterogeneity in *pax2a* expression. We would appreciate any specific feedback on improving the quality of presentation as this would be utilized in future publications as well.

- In panels B,C and D, of Figure 8, 1 to 4 nuclei per follilcular section do not express Pax2A reporter gene (mKO2), while in panel F, more than 50% of the thyrocytes do not express Pax2A. Is it due to the half life of the mKO2 as compared to Pax2A ? Is it due to the sensibility of the antibodies used ? This should be discussed as it could help proposing working hypotheses.

We thank the reviewer for pointing this out. We agree with reviewer that the mKO2 driven by *pax2a* regulatory sequences and PAX2A antibody staining mark distinct proteins that would have different half-lives. As such, we have noted this in the revised manuscript.

In the result section, we have added,
p. 12, line 237: “The mKO2 coding sequence was inserted in-frame with the C-terminus of *pax2a*, separated by a T2A-cleavable linker. This leads to generation of two proteins, PAX2A and mKO2, from the *pax2a* regulatory sequence. The two proteins could display differential protein turnover characteristics.”, and
p.12, line 246: “The difference in levels of PAX2A antibody staining and mKO2 fluorescence could be attributed to different half-lives of the two proteins.”

- In panel 8F, Tg-NLS-EGFP is highly expressed in all the thyrocytes (DAPI+), thus in Pax2A high thyrocytes and in Pax2A low thyrocytes.... « Notably, both PAX2A-low and -High cells display tg promoter-driven EGFP expression, thereby confirming their differential status ». However, in the bulk sequencing (Fig 9E), the Pax2A-low population reveal that Tg is weakly expressed in this population (blue or light yellow as compared to orange)..... This contradicts with their previous conclusion : « Notably, both PAX2A-low and -High cells display tg promoter-driven EGFP expression, thereby confirming their differential status » and « The

high expression of tg in both populations confirms the identity of the cells as thyrocytes ». Authors should explain or discuss this discrepancy.

Figure 9E presents the relative levels of gene expression between pax2a-Low and pax2a-High population. Thus, in comparison to the expression in pax2a-High population, *tg* is weakly expressed in pax2a-Low population.

As mentioned in the Result Section, p. 15, line 293: "Differential expression of tg was observed in-spite of the fact that tg was detected as one of the top 15 highly expressed gene in both pax2a^{mKO2}-High and pax2a^{mKO2}-Low populations (Fig EV5)."

Tg is the gene with the highest expression in pax2a^{mKO2}-High population, while it is the 15 highest expression gene in the pax2a^{mKO2}-Low population (Fig EV5, Table EV4). As 10,062 genes were detected in each population by bulk-sequencing, this puts *tg* expression in the top 0.15 % of genes by absolute expression levels in each population, suggesting its robust and strong expression in both populations.

- In addition, analysis of the bulk sequencing of the Pax2A-low population (Fig 9E) suggests that this population expresses very low level of most thyroid-specific genes (Tg, nkx2.4, slc5a5, TPO, TshR, ctsba, ctsk). One can thus wonder whether these cells are active thyrocytes ? or could correspond to a kind of inactive cellular state comparable to the hot and cold follicles described in mammals.

Similar to the previous remark, the thyroid specific genes are differentially expressed between the two population. Thus, they are expressed at lower level in pax2a-Low population in relation to the pax2a-High population. However, they are robustly expressed in both the populations in absolute levels (Table EV4).

We completely agree with the reviewer that this could be a representation of non-functional, yet mature, thyrocytes. As mentioned in the Discussion, p. 17, line 340: "Our result suggests that the gland maybe divided into functional and resting thyrocytes. It would be of interest to build on this study and investigate the functional and replicative differences among the two subpopulations of thyrocytes. Moreover, it would be important to elucidate the dynamics and plasticity between the two subpopulations as thyrocytes might shuffle between the two states."

The connection between the hot and cold follicles observed in mammals is completely valid here. We thank the reviewer for pointing this out. However, those observations classified complete follicles in functional (hot) or non-functional (cold) categories. Our study, on the other hand, provides evidence that function and resting thyrocytes are present within the same follicle. It is tempting to speculate that hot follicles are follicles with a higher proportion of functional thyrocytes than cold follicles. Our study, performed at single-cell resolution, provides evidence that a follicle is not a homogenous unit, but rather a mix of thyrocytes in at-least two distinct states.
We have added this in the Discussion, p. 16, line 313: "Previous studies on mammalian thyroid gland have identified functional and morphological heterogeneity between the thyroid follicles (Baptist *et al*, 1995; Smeds *et al*, 1987; Struder *et al*, 1989). Using

histological analysis, functionally active follicles were identified as being outlined with tall columnar thyrocytes, while inactive follicles were marked by an outline of low cuboidal to almost squamous thyrocytes (Studer *et al*, 1978; Gerber *et al*, 1987). However, two open questions remain unanswered in the field: 1. What are the genes responsible for the functional heterogeneity; 2. Do all thyrocytes in a follicle display uniform functional capacity. Here, by applying unbiased single-cell gene expression analysis to the thyroid gland for the first time, we have identified transcriptional heterogeneity within the thyrocyte population. Further, we demonstrate that transcriptionally diverse thyrocytes are present in the same follicle."

Referee #2:

I think the authors have carefully addressed comments from both referees. This has led to important clarifications (that have raised a few further questions, see below) as well as substantial rewriting of the manuscript. I usually refrain from raising what might be perceived as new issues at a revision stage. When reading this revised version, my comments mainly revolve around two themes. The first one is the result of the new, more detailed information about the thyroid dissection provided in the revision, that was not clear before. The second one is something that passed unnoticed by me when reading the prior version, but that became evident when reading the revised manuscript. In my opinion, even though I elaborate on these two themes at some length, these comments do not require additional experiments, but potentially some textual revision of the current manuscript.

We are pleased to know that the reviewer found our revision satisfactory. Further, we thank the reviewer for pointing out textual changes that would improve the accuracy of the reporting of our observations. We agree with most of the edits raised by the reviewer. Our response and edits are mentioned below.

Major comments:

1. On one hand I find it commendable that the authors in the revised manuscript have described dissection of the zebrafish "thyroid gland" in much more detail. On the other hand, when I thus now get a better picture of how the dissection was performed it raises some additional comments and questions.

I do not doubt for a second that the zebrafish has thyroid follicular tissue and that this is an extremely valuable model organism for thyroid research. It is not readily apparent to me though if the thyroid tissue of the zebrafish can be described as "a thyroid gland". As far as I understand, zebrafish thyroid tissue is more or less a number of follicles loosely distributed along the aorta. To me, a gland is an anatomically and histologically coherent and often encapsulated organ, even though it might be argued that for instance colonic crypts are sometimes referred to as colonic glands. When the dissection procedure is now nicely detailed in the revised manuscript, this issue becomes not only a matter of semantics or definitions in my opinion.
If I understand correctly, the dissection encompasses the ventral aorta, parts of the gills, the thyroid follicles and presumably surrounding connective tissue, which to me seems like kind

of a regional en bloc dissection. This is obviously the result of very skilled and careful experimental work. Still, can the resulting tissue that is then sequenced be equaled to that of a "zebrafish thyroid gland", an expression that is repeatedly used in the manuscript, even though "the zebrafish thyroid gland" is certainly present as a constituent part of the dissection? To me it seems that analysis has been performed of the "zebrafish thyroid region"/"the zebrafish thyroid and its immediate surroundings" rather than "the zebrafish thyroid gland". This seems to be reflected by the fact that follicular cells constitute only approximately 6% of the cells in the dissected region. In thyroid glands (as encapsulated coherent structures) of higher vertebrates, follicular cells are certainly in majority with only minor contributions/fractions of other cell types.
The authors seem to more or less equal or define "the zebrafish thyroid gland" as the anatomical region that they have dissected. This region obviously includes for instance part of the gills. Do the authors consider these to be part of "the zebrafish thyroid gland"? For instance, to me it seems likely that a major part of the tp63-positive NFE cell population are constituent cells of the gills. The size of the "zebrafish thyroid gland" and its proportion of NFE cells then largely becomes a matter of how much gill tissue is included in the dissection. As a comparison, even though the trachea, parathyroids and thymus are immediately adjacent to the thyroid gland in higher vertebrates, these structures would not be described as parts of the thyroid gland, even if they were included in an en bloc dissection.
When reading some parts of the manuscript, this becomes a matter of some concern to me, e.g:

p. 4, line 76: "... we develop the first atlas of the thyroid gland..."
p. 4, line 80: "... comprehensively represent the cells present in the zebrafish thyroid gland"
p. 7, line 132: "Identification of cell-types present in the zebrafish thyroid gland."
p. 9, line 184: "Having defined the cell types of the thyroid gland, we quantified potential cell-cell interactions between thyrocytes and all cell types present in the organ..."
p. 9, line 198: "Further, the ligand cyr61 is broadly expressed in the thyroid gland..." (note that for instance expression is seen also in the gills, that makes me come back to my question if the authors consider these as a part of the zebrafish thyroid gland).
p. 15, line 327: "This allowed us to capture yet poorly characterized cell-populations within the thyroid gland"

Can really the identified cell populations be considered to be "within the thyroid gland"?

I think the authors need to elaborate on this and possibly modify some parts of the text. Also, the anatomical differences between thyroid in zebrafish and higher vertebrates may need to be detailed even more clearly than presently at the beginning of the results section. In my opinion the authors have characterized the zebrafish thyroid region, rather than the zebrafish thyroid gland (even though the thyroid follicles are a part of the dissection). I do not think this makes the findings less interesting at all, as signaling cues may be provided by adjacent cell populations, regardless if they are considered constituents of a "zebrafish thyroid gland" or not. Moreover, the clustering certainly sorts out the follicular cells and provides a catalogue of their transcriptomes as well as that of other cell populations of the region. Still, I think this issue may warrant some rewriting.

We thank the author for pointing this out, and we do agree that the atlas includes cells adjoining the zebrafish thyroid gland. However, we are not sure of the interpretation of the 'gland' as stated by the reviewer. By definition, a gland is a collection of cells with secretory responsibilities. Their anatomical isolation is not necessary. The prime example being the sweat gland, millions of which are scattered throughout the body and not anatomically isolated. Thus, we do believe that the collection of thyroid follicles, vasculature, lymphatic vessels, fibroblasts and immune cells is the zebrafish thyroid gland, which constitutes more than 80% of the atlas. Cells belonging to the gills and NFE constitute less than 20% of the atlas.

This is also in line with the reporting of thyroid gland in zebrafish and fish in general. A study published in 1973 states, "The presence of a thyroid gland and its ability to produce the thyroid hormones thyroxine and triiodothyronine are striking characteristics of vertebrate organization. **The thyroid gland is composed of hollow follicles of cells surrounding a liquid colloid.** This colloid forms a reserve of potential hormone. The thyroid gland is the only vertebrate endocrine organ with such an extracellular store. In most agnathans and teleosts, thyroid follicles are scattered around the ventral aorta while in other vertebrates the follicles are aggregated into one or two discrete thyroid glands. Such an evolution of scattered elements into a compact gland is a common feature of endocrine evolution. Similiar changes have occurred in the evolution of the adrenal medulla and cortex and to a lesser extent the pancreatic islets." (Sage, 1973)
Thus, it is clear that the anatomical isolation of the thyroid gland is an evolutionary adaptation, but not a criteria for the function or classification of the gland.

Further, during the review for the current publication a Human Cell Atlas was published that included single-cell RNA-Seq. of the human thyroid gland (Han *et al*, 2020). The human thyroid single-cell atlas contains two donors and can be viewed here: https://db.cngb.org/HCL/dpline.html?tissue=Thyroid.
In the human thyroid single-cell atlas, thyrocytes constitute 22 % and 23 % of the total cells in the two donors. Thus, thyrocytes might not be the majority population in mammalian thyroid gland. Certainly, the thyroid follicles occupy the largest area in the gland, but they are hollow spheres and thus rather empty in terms of cell occupancy. Immune cells, fibroblasts and endothelium could be major contributors to the cellular composition of the thyroid gland.

However, as stated earlier, we do agree that cells adjoining the zebrafish thyroid gland were profiled and are part of the atlas. Thus, we have made the following edits in the text:

p. 1, line 23: 'we developed a single-cell transcriptome atlas of **the region containing** the zebrafish thyroid gland'.

p. 11, line 204: 'Further, the ligand cyr61 is broadly expressed in the thyroid gland **and adjoining gills**, with one of its receptors, itgb5, an integrin isoform, expressed specifically by the thyrocytes.'

References:
Han X, Zhou Z, Fei L, Sun H, Wang R, Chen Y, Chen H, Wang J, Tang H, Ge W, *et al* (2020)

Construction of a human cell landscape at single-cell level. *Nature* 581: 303–309
Sage M (1973) The Evolution of Thyroidal Function in Fishes. *Am Zool* 13: 899–905

I furthermore think the authors may overstate the NFE population as part of the zebrafish thyroid gland and I have some comments about this population. I would not really consider tp63 as a "...known marker of ... epithelial tissues", but rather as a protein that may be expressed in some epithelia. Indeed, TP63 is expressed in human stratified epithelia and urothelium, but absent in most other epithelial tissues. Moreover, TP63 is characteristically expressed also in myoepithelial cells. The latter cell type seems to be well represented in gills (PMID: 20143320). Indeed, the authors find tp63 positive "... NFE celler scattered throughout the gills and in the region adjacent to the follicles". Perhaps this should be further commented as a major part of the NFE population thus perhaps cannot really be considered a constituent of the "zebrafish thyroid gland" (see above).
What about more prototypic epithelial markers in the NFE population (such as cytokeratins and E-cadherin)?

The labelling of Cluster 7 as Non-Follicular Epithelium (NFE) was based on Gene Ontology (GO) analysis of marker genes specific for the cluster. As mentioned in the text, this identified categories such as 'epithelium development' (Appendix Fig S2). The utilization of *tp63* as a marker for this cluster was based on its specificity and on the availability of antibody for immunostaining. However, other marker genes expressed in the cluster include cytokeratins (*krt5*, *krt14*) and E-cadherin (*cdh1*) (Table EV1) (thank you for pointing these critical genes that are representative of epithelial tissue). We have listed these critical genes in the Results Section:
p. 9, line 168: 'Notably, cluster seven displayed expression of cytokeratins (*krt5*, *krt14*) and E-cadherin (*cdh1*) (Table EV1).'

2. Even though I am aware that a distinction may be made between molecular and functional homogeneity, when rereading the manuscript I think that the authors incorrectly describe the prevailing view of thyroid follicular cells as a homogenous population. That is not the case. It has for long been appreciated that thyroid follicular cells are functionally heterogeneous, both in normal conditions as well as in disease, which is perhaps most apparent in multinodular goiter. This is evident already from the histology of the normal human thyroid gland, where inactive follicles outlined by a flat epithelium are mixed with active follicles (colloid scalloping etc) outlined by taller epithelial cells. I think it is fair to say that this heterogeneity has generally been assumed to be reflected also at the molecular level. Just to mention a few references:

Naturally occurring clones of cells with high intrinsic proliferation potential within the follicular epithelium of mouse thyroids
Smeds et al. Cancer Res. 1987. 47(6):1646-51. PMID: 3815361

Natural heterogeneity of thyroid cells: the basis for understanding thyroid function and nodular goiter growth
Studer et al. Endocr Rev. 1989. 10(2):125-35. PMID: 2666115

Various facets of the intercellular heterogeneity in thyroid primary culture
Baptist et al. Thyroidology. 1991. 3(3):109-13. PMID: 1726924

Intercellular heterogeneity of early mitogenic events: cAMP generalizes the EGF effect on c-Fos protein appearance but not on MAP kinase phosphorylation and nuclear translocation in dog thyroid epithelial cells
Baptist et al. Exp Cell Res. 1995. 221(1):160-71. PMID: 7589241

I think this is in stark contrast to some statements of the manuscript, e.g.

p. 1, line 20: "So far, the thyrocytes have been assumed to be a homogenous population."
p. 1, line 34: "... nominally homogenous thyroid population."
p. 2 line 39: "Single-cell analysis uncovers latent heterogeneity in thyroid follicular cells." (what does "latent" mean in this context?)
p. 3, line 59: "... it remains unknown if all the thyrocytes resident in the thyroid gland are equally capable of generating thyroid hormones."
p. 15, line 319: "Our result suggest that the gland may be divided into functional and resting thyrocytes." (I think it is a prevailing view that this is the case, but that the current study demonstrates a genetic/transcriptomic correlate to this notion, see below).
p. 18, line 401: "... beyond a nominally homogenous endocrine cell population".

I do not mention this to discredit the present manuscript and I do not think it makes it less interesting. On the contrary I think it provides a valuable step towards a genetic/transcriptomic correlate to the existing notion of thyrocyte heterogeneity. I think that this needs to be addressed in the sense that previous relevant literature is acknowledged, sections that convey the impression that thyrocyte has previously not been appreciated be modified and putting the current findings into this context.

We thank the reviewer for bringing up studies on potential heterogeneity in the thyroid follicles observed in mammals. This is similar to the last comment made by referee #1. The notion of hot (functional) or cold (non-functional / resting) follicles is well documented in mammalian thyroid gland. As mentioned by the referee, “This is evident already from the histology of the normal human thyroid gland, where inactive follicles outlined by a flat epithelium are mixed with active follicles (colloid scalloping etc) outlined by taller epithelial cells.”

However, the literature considers thyroid follicles as a homogenous unit of thyroid follicular cells. Our study not only improves on the observation by providing, as stated the referee, “a genetic/transcriptomic correlate to the existing notion of thyrocyte heterogeneity”, but further provides evidence that functional and resting thyrocytes are present together in the same follicle. Thus, a follicle is a mix of functional and resting cells.

However, we agree that it is important to acknowledge the studies. Hence, the abstract has been edited to state,
p. 1, line 33: “Our results identify and validate transcriptional differences within the **presumed** homogenous thyrocyte population.”

p. 2, line 37: The summary has been made more precise, “Single-cell analysis uncovers **transcriptional** heterogeneity in thyroid follicular cells.”

p. 3, line 58: The results have been made more precise, “Though the machinery responsible for the production of thyroid hormones by thyrocytes is well established, it remains unknown if all the thyrocytes resident **in a follicle** are equally capable of generating thyroid hormones.”

p. 15, line 313: And in the discussion we have added the following text:
“Previous studies on mammalian thyroid gland have identified functional and morphological heterogeneity between the thyroid follicles (Baptist et al, 1995; Smeds et al, 1987; Struder et al, 1989). Using histological analysis, functionally active follicles were identified as being outlined with tall columnar thyrocytes, while inactive follicles were marked by an outline of low cuboidal to almost squamous thyrocytes (Studer et al, 1978; Gerber et al, 1987). However, two open questions remain unanswered in the field: 1. What are the genes responsible for the functional heterogeneity; 2. Do all thyrocytes in a follicle display uniform functional capacity. Here, by applying unbiased single-cell gene expression analysis to the thyroid gland for the first time, we have identified transcriptional heterogeneity within the thyrocyte population. Further, we demonstrate that transcriptionally diverse thyrocytes are present in the same follicle.”

And edited this part:
p. 20, line 424: “We hope that our efforts will expand the understanding of thyrocytes as a **collection of heterogenous** endocrine cell population; providing a complex picture of the diversity in thyrocyte identity and function.”

Minor comment:

The introduction makes some statements about thyroid biology and pathology that strike me as a bit odd.

p. 3, line 46: The authors write that:
- "Thyroid dysfunction affects 100 million people worldwide and is treatable by hormone replacement." Do the authors refer to hypo- as well as hyperthyroidism in this description? The latter condition is not primarily treatable by hormone replacement (but most often requires it after e.g. radioidine ablation or surgical treatment.
- if left untreated may result in "profound adverse effects of the human body, including mental retardation, goiter or dwarfism". Whereas it can be debated if goiter can really be considered a "profound adverse effect", the other two effects ("mental retardation" and "dwarfism") only relate to untreated congenital hypothyroidism, that constitutes a very small fraction of these 100 million people. Profound adverse effects of no treatment, such as mental retardation and dwarfism, are only relevant when considering this comparatively very small group and is not generally applicable on untreated thyroid disease. I think this needs to be clarified.

We have clarified the text on p. 3, line 42 as: “Thyroid dysfunction afflicts almost 100 million people worldwide (Taylor et al, 2018). Hypothyroidism can be efficiently managed by lifelong hormone replacement therapy, while hyperthyroidism is treated with antithyroid medication, surgery or ablation, depending on the underlying disorder. Congenital hypothyroidism if left untreated may result in profound adverse effects on development, including mental retardation, goiter or dwarfism.“

On balance, even though it might seem that I have quite a lot of comments, these are possible to address by rephrasing some parts of the manuscript and do not require new experiments being performed. Even though I am aware that EMBO Reports usually allows only a single revision I think the present manuscript is a nice contribution and a valuable resource that should be offered a possibility to address these issues of textual modification. I consider this manuscript to be of high quality, novel and important not only to the thyroid field, but to a broader audience.

2nd Revision - Editorial Decision 6th Oct 2020

Dr. Sumeet Singh
IRIBHM, ULB
808 route de Lennik
Campus Erasme, Building C
Brussels 1070
Belgium

Dear Sumeet,

I am very pleased to accept your manuscript for publication in the next available issue of EMBO reports. Thank you for your contribution to our journal.

At the end of this email I include important information about how to proceed. Please ensure that you take the time to read the information and complete and return the necessary forms to allow us to publish your manuscript as quickly as possible.

As part of the EMBO publication's Transparent Editorial Process, EMBO reports publishes online a Review Process File to accompany accepted manuscripts. As you are aware, this File will be published in conjunction with your paper and will include the referee reports, your point-by-point response and all pertinent correspondence relating to the manuscript.

If you do NOT want this File to be published, please inform the editorial office within 2 days, if you have not done so already, otherwise the File will be published by default [contact: emboreports@embo.org]. If you do opt out, the Review Process File link will point to the following statement: "No Review Process File is available with this article, as the authors have chosen not to make the review process public in this case."

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

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

Best regards,
Esther

Esther Schnapp, PhD
Senior Editor
EMBO reports

****

THINGS TO DO NOW:

You will receive proofs by e-mail approximately 2-3 weeks after all relevant files have been sent to

our Production Office; you should return your corrections within 2 days of receiving the proofs.

Please inform us if there is likely to be any difficulty in reaching you at the above address at that time. Failure to meet our deadlines may result in a delay of publication, or publication without your corrections.

All further communications concerning your paper should quote reference number EMBOR-2020-50612V3 and be addressed to emboreports@wiley.com.

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

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: Sumeet Pal Singh
Journal Submitted to: EMBO Reports
Manuscript Number: EMBOR-2020-50612V2

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

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

A- Figures

1. Data

The data shown in figures should satisfy the following conditions:

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

2. Captions

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

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

- the exact sample size (n) for each experimental group/condition, given as a number, not a range;
- a description of the sample collection allowing the reader to understand whether the samples represent technical or biological replicates (including how many animals, litters, cultures, etc.).
- a statement of how many times the experiment shown was independently replicated in the laboratory.
- definitions of statistical methods and measures:
  - common tests, such as t-test (please specify whether paired vs. unpaired), simple χ2 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?	NA
1.b. For animal studies, include a statement about sample size estimate even if no statistical methods were used.	Statement in Statistical Analysis (Methods): "Power analysis was not performed to estimate sample sizes."
2. Describe inclusion/exclusion criteria if samples or animals were excluded from the analysis. Were the criteria pre-established?	Statement in Statistical Analysis (Methods): "No data was excluded from analysis."
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.	Yes
For animal studies, include a statement about randomization even if no randomization was used.	Statement in Statistical Analysis (Methods): "Animals were randomly chosen for experimentation"
4.a. Were any steps taken to minimize the effects of subjective bias during group allocation or/and when assessing results (e.g. blinding of the investigator)? If yes please describe.	No
4.b. For animal studies, include a statement about blinding even if no blinding was done	Statement in Statistical Analysis (Methods): "Blinding was not performed during analysis."
5. For every figure, are statistical tests justified as appropriate?	Yes
Do the data meet the assumptions of the tests (e.g., normal distribution)? Describe any methods used to assess it.	Statement in Statistical Analysis (Methods): "Analysis of normal distribution was performed using Shapiro–Wilk test."
Is there an estimate of variation within each group of data?	Yes
Is the variance similar between the groups that are being statistically compared?	Comparison for RNA-Seq. data was performed using existing tool that adjust for unequal variance.

C- Reagents

USEFUL LINKS FOR COMPLETING THIS FORM

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

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

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

http://datadryad.org

http://figshare.com

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

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

http://biomodels.net/

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

6. To show that antibodies were profiled for use in the system under study (assay and species), provide a citation, catalog number and/or clone number, supplementary information or reference to an antibody validation profile. e.g., Antibodypedia (see link list at top right), 1DegreeBio (see link list at top right).	Statement in Immunofluorescence and image acquisition (Methods): "Primary antibodies used in this study were anti-PAX2A (rabbit, Genetex GTX128127) at 1:250, anti-EGFP (chicken, Abcam ab13970) at 1:1000, anti-E-Cadherin (mouse, BD bioscience cat 610181) at 1:200, anti-monomeric Kusabira-Orange 2 (mouse, MBL amalgaam M-168-3M) at 1:200, anti-monomeric Kusabira-Orange 2 (rabbit, MBL amalgam PM051M) at 1:250 and anti-p63 (mouse, Santa Cruz Biotechnology 4A4) at 1:200. Secondary antibodies at 1:250 dilutions used in this study were Alexa Fluor 488 anti-chicken (Jackson ImmunoResearch laboratories 703-545-155), Alexa Fluor 647 anti-rabbit (Jackson ImmunoResearch laboratories 711-605-152), Alexa Fluor 647 anti-mouse (Jackson ImmunoResearch laboratories 715-605-150), CyTM3-conjugated anti-rabbit (Jackson ImmunoResearch laboratories 711-165-152) and CyTM3-conjugated anti-mouse (Jackson ImmunoResearch laboratories 715-165-150)."
7. Identify the source of cell lines and report if they were recently authenticated (e.g., by STR profiling) and tested for mycoplasma contamination.	NA

* 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.	Statement in Zebrafish strains and husbandry (Methods): "Wild-type or transgenic zebrafish of the outbred AB, WIK, or a hybrid WIK/AB strain were used in all experiments. Zebrafish were raised under standard conditions at 28 °C. Animals were chosen at random for all experiments. Published transgenic strains used in this study were Tg(tg:nls-mVenus-T2A-NTR) (7), Tg(tg:nls-EGFP) (27), Tg(kdrl:GFP)la116 (44), Tg(mpeg1.1:mCherry)gl23 (45) and Tg(col1a2:LOXP-mCherry-NTR)cn11 (referred as Tg(Col1a2:mCherry)) (14)."
9. For experiments involving live vertebrates, include a statement of compliance with ethical regulations and identify the committee(s) approving the experiments.	Statement in Zebrafish strains and husbandry (Methods): "Experiments with Tg(tg:nls-mVenus-T2A-NTR) were conducted in accordance with the Animal Welfare Act and with permission of the Landesdirektion Sachsen, Germany (DD24-5131/346/11, DD24-5131/346/12, DD24.1-5131/476/2, TVV21/2018 and all corresponding amendments). Zebrafish husbandry and experiments with all transgenic lines will be performed under standard conditions as per the Federation of European Laboratory Animal Science Associations (FELASA) guidelines (46), and in accordance with Institutional (Université Libre de Bruxelles (ULB)) and national ethical and animal welfare guidelines and regulation, which were approved by the ethical committee for animal welfare (CEBEA) from the Université Libre de Bruxelles (protocols 578N-579N)."
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 confirm compliance.

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	Data Availability Section (Methods): "The raw 10X data, along with tabulated count data are available publicly from GEO under accession number GSE133466. The atlas for online browsing is available at https://sumeet.shinyapps.io/zfthyroid/. The raw bulk RNA-Seq. data, along with tabulated count data are available publicly from GEO under accession number GSE153197."
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).	Raw RNA-Seq. data deposited to GEO and made public.
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.	Not applicable.