

Human Pancreatic α -Cell Heterogeneity and Trajectory Inference Analyses Reveal SMOC1 as a β -Cell Dedifferentiation Gene

Corresponding Author: Dr Adolfo Garcia-Ocana

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

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

In the manuscript, Kang et al. analyzed scRNA-seq and snRNA-seq data from human pancreatic islets and performed velocity analysis. The authors found SMOC1 as a candidate gene involved in beta to alpha transdifferentiation. Further functional study in the primary human islets and in Endoc- β H1 cells indicate that this gene is negatively correlated with beta cell function. Overall, the study centers on important topics related to islet cell heterogeneity and plasticity. However, major flaws in the single-cell analysis and lack of sufficient orthogonal validation undermine the major conclusions of the research.

Major points:

1. One of the major issues is that the authors did not convincingly show that AB cluster represents a real subtype in vivo, rather than doublets/aggregates in the scRNA-seq/snRNA-seq workflow. The high percentage of AB positive cells (especially in the snRNA-seq 12%) exceeds the estimate of bihormonal cells in the control pancreas (for example, Yoneda et al. JCEM, 2013.), suggesting a large fraction of these cells might be an artifact.
2. Another major issue is the velocity analysis. As well recognized, the velocity analysis has potential pitfalls that need to be carefully addressed to ensure that the trajectory reflects true biology. The operation of integrating scRNA and snRNA-seq data and then inferring velocity is problematic because, by its nature, the snRNA-seq data contain transcripts with more intron-retention and, hence will have increased unspliced rates compared with scRNA-seq data where more cytoplasmic RNA are captured. The data presented by the authors fail to provide confidence and raise significant concerns from the reviewer. For example, the velocity stream in 3B does not reflect the pseudotime values in 3C; the Figure 3D portrait plots do not show strong dynamic patterns; Figure 5 velocity and pseudotime from controls do not match the velocity pattern from Figure 3.
3. It is difficult to judge the staining pattern of SMOC1 with glucagon and insulin.
4. SMOC1 was identified because they posit between beta to alpha transition. However, all the functional data the authors shown indicate beta cell dysfunction induced by SMOC1 overexpression. Does SMOC1 induces beta cell to alpha cell conversion?

Minor points:

1. Line 17, the “expression of GCG and ALDH1A1 genes was ... similar in scRNA- and snRNA-seq datasets”? However, it looks like the scRNA-seq had way higher expression levels compared with snRNA-seq.

Reviewer #2

(Remarks to the Author)

In this manuscript, Kang et al. use single-cell and single-nuclei RNAseq and analysis of human islets to interrogate gene signatures and RNA velocity across alpha- and beta-cell groups. They find gene sets of genes that change as cells move toward either an alpha- or beta-cell fate in normal donors. However, they also analyze available dataset that includes T2D

donors and find that the trajectory of cells is weighted towards an alpha-cell fate. The conclusion that beta-cells are more alpha-cell like or more dedifferentiated in T2D is not novel. However, the velocity analysis, gene signatures discovered by RNA velocity analysis, and picking out SMOC1 as a gene whose activation leads to beta-cell dedifferentiation is. While this manuscript has interesting and exciting results, there are still a few questions remaining regarding its methodology and findings.

Major Points.

1. It was unclear to me whether the velocity analysis was performed just on snRNAseq or scRNAseq, or both. Does limiting it to one or the other change velocity outcome? If so, how did you decide which dataset to use?
2. Is it possible to know whether the effective concentrations of SMOC1 in Figure 8 are similar to what beta cells are exposed to in vivo? Please discuss this point.

Minor points.

1. P 4 line 37 . . . you should cite your additional file 2: Table S2 since Figure 2D doesn't alone hold this information.
2. Page 5, lines 18-19. Authors state that PDX1, INSM1 and SMAD1 are transdifferentiation markers, but don't cite papers showing this, or demonstrate it with their own data.
3. P 10, line 41. Please specify that this work was done in zebrafish.
4. P 11, lines 1-5. It is unclear to me what these sentences mean. Does your data contradict this previous study? What was different between the two studies?
5. P 12, line 11—use of "beta" spelled out rather than the Greek symbol, unlike the rest of manuscript.
6. Please acknowledge HPAP funding, as agreed to when using HPAP datasets. From the website: Citation requirement: You agree to cite this database and the corresponding data providers every time you use this data for any publication purpose. Please use this NIH grant number- UC4-DK112217 and UC4-DK112232, an HPAP identifier (RRID:SCR_016202), and the HPAP database URL- <https://hpap.pmacs.upenn.edu/> for citation purposes. Please use the name "Human Pancreas Analysis Program (HPAP) Database, consortia under Human Islet Research Network (RRID:SCR_014393).
7. Several grammatical and idiomatic errors were noted; further editing would benefit the readability of this manuscript.

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

In the revised manuscript by Kang et al., the authors addressed some of my questions. However, major concerns remain:

1. It is still unclear to me whether the AB cluster is a real subtype, doublets/aggregates or some other artifacts such as cells with low quality. What is the logic behind "the AB cluster ... display enhanced chromatin accessibility for INS and GCG genes, strongly suggesting that AB cells are bona fide INS-positive and GCG-positive cells and not cell doublets/aggregates"? If the AB cells are indeed alpha/beta doublets, won't they also show open chromatin for both INS and GCG? The authors did not show enough evidence that AB cluster exhibit distinct expression profiles that cannot be explained by mixture from different islet cell types.
2. Integration of scRNA-seq and snRNA-seq data and then perform RNA velocity analysis is not recommended, with reasons detailed in my previous comment. The authors should at least analyze the two types of data separately and interpret the velocity results in parallel. Furthermore, the lack of distinct direction in pseudotime plots (Figures 5c and 5d), as well as the inconsistency between Figure 5 and Figure 3 velocities are all indicative of potential issues in RNA velocity analysis.

Reviewer #2

(Remarks to the Author)

All of my points have been successfully addressed.

Version 2:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

The author has addressed all my concerns.

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

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

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

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

POINT-BY-POINT RESPONSE TO THE REVIEWERS

We would like to thank the Reviewers for their thorough evaluation and positive comments of our manuscript. They mention that "...the study centers on important topics related to islet cell heterogeneity and plasticity" and "...this manuscript has interesting and exciting results..." We have addressed all their comments and concerns and revised the manuscript accordingly. We believe the comments from the Reviewers have strengthened the data analysis, rigor, novelty, and conclusions of the manuscript.

Of note, changes in the manuscript appear in red font.

Reviewer #1:

We thank the Reviewer for his/her helpful comments. Next, we address each of the issues raised by the Reviewer separately.

1." One of the major issues is that the authors did not convincingly show that AB cluster represents a real subtype in vivo, rather than doublets/aggregates in the scRNA-seq/snRNA-seq workflow. The high percentage of AB positive cells (especially in the snRNA-seq 12%) exceeds the estimate of bihormonal cells in the control pancreas (for example, Yoneda et al. JCEM, 2013.), suggesting a large fraction of these cells might be an artifact."

We believe the Reviewer raises an important point in this comment. Yoneda et al. (*JCEM* 98, 2053–2061, 2013) reported that 0.33% to 0.82% of immunolabeled insulin-positive cells are also glucagon-positive cells (bi-hormonal cells) in human pancreas sections. However, other studies have found that this percentage can be higher. Mezza et al. (*Diabetes* 63, 994-1007, 2014) reported 4.5% bi-hormonal cells in human pancreas from insulin sensitive donors and 10.9% bi-hormonal cells in human pancreas from insulin resistant subjects. More recently, the same group (*JCEM* 101, 470-475, 2016) has described that bi-hormonal cells in human pancreas sections are in the range of 4 to 30% depending on the insulin secretory capacity assessed in hyperglycemic clamps in patients undergoing pancreateoduodenectomy. In our recent publication (Karakose et al. *Cell Reports Medicine* 5, 12101832, 2024), we found that the percentage of immunolabeled bi-hormonal cells in isolated human islets can reach up to 4% of total insulin-positive cells, similar to the values found in human pancreas sections in the reports from Mezza et al. Therefore, in the literature, there is wide range of results regarding bi-hormonal cells in human pancreas with values between 0.33 and 4.5% in non-diabetic donors that can reach up to 10.9-30% depending on the study and whether the organ donor was from a subject with insulin resistance or diabetes.

We initially believed that the comparison of the percentage of bi-hormonal islet cells using immunolabeling and single cell transcriptomics could be misleading. This is because of the difference in the biological abundance of proteins vs. RNAs and the different sensitivity of both methods of analysis (immunolabeling vs. transcriptomics). However, we surprisingly found that the percentage of bi-hormonal cells in both cases is not that different. In our studies, we find that AB cells (double positive for *INS* and *GCG*) are in the range of 3.8-11.9% of alpha cells, that is 2.2% and 6.9% of beta cells (*INS*⁺ cells) depending on whether scRNA-seq or snRNA-seq was performed. These numbers are similar to the range of immunolabeled bi-hormonal cells per total insulin-positive cells shown in the studies reported by Mezza et al and Karakose et al (see above).

It is important to mention that we used a stringent hyperparameter in the SoupX (20% contamination estimation) and employed the Doubletfinder package (20% doublet ratio) to algorithmically eliminate ambient RNA contamination and doublets/aggregates in our dataset,

increasing the reliability of our findings. Finally, to further address that the AB cluster could represent a human islet cell subtype, we analyzed the datasets from single nucleus multi-omics of non-diabetic adult human islets performed by Millman's group in *Nat Cell Biol*, 2023 (ref. 34). We found that the AB cluster (GCG^+INS^+ cells) display enhanced chromatin accessibility for *INS* and *GCG* genes (see new Suppl. Fig. 2), strongly suggesting that AB cells are *bona fide* *INS*⁺ and *GCG*⁺ cells and not cell doublets/aggregates. Based on these new data and all the information indicated above, we have modified the Methods (page 16), Results (pages 4 and 5) and Discussion (pages 10 and 11) sections in the manuscript to address this point.

2. Another major issue is the velocity analysis. As well recognized, the velocity analysis has potential pitfalls that need to be carefully addressed to ensure that the trajectory reflects true biology. The operation of integrating scRNA and snRNA-seq data and then inferring velocity is problematic because, by its nature, the snRNA-seq data contain transcripts with more intron-retention and, hence will have increased unspliced rates compared with scRNA-seq data where more cytoplasmic RNA are captured. The data presented by the authors fail to provide confidence and raise significant concerns from the reviewer. For example, the velocity stream in 3B does not reflect the pseudotime values in 3C; the Figure 3D portrait plots do not show strong dynamic patterns; Figure 5 velocity and pseudotime from controls do not match the velocity pattern from Figure 3.

The Reviewer raises here another interesting point. Analysis of RNA velocity has been traditionally performed with scRNA-seq and not with snRNA-seq because snRNA-seq data contain transcripts with more intron-retention and, hence will have increased unspliced rates compared with scRNA-seq data where less intron-retention cytoplasmic mature transcripts are captured. However, we reasoned that integrating these two modalities in the same samples will allow to capture both the early transcriptional dynamics (from snRNA-seq) and the downstream more mature expression profiles (from scRNA-seq), thereby offering a more complete temporal trajectory of cell state transitions and enhancing overall velocity confidence. Indeed, RNA velocity confidence score using integrated scRNA-seq and snRNA-seq was significantly improved (mean 0.81; median 0.93, Mann-Whitney-Wilcoxon test, $p=1.55 \times 10^{-135}$) compared with scRNA-seq alone (mean 0.71; median 0.80). This information is now included in the Methods (page 15) and Results (pages 5 and 6) sections and in the new Suppl. Fig. 4

The Reviewer indicates that “velocity stream in Fig. 3B does not reflect pseudotime values in Fig. 3C”. Pseudotime assumes a smooth continuum of cell state derived from a static snapshot of the transcriptome (Trapnell et al. *Nature Biotechnology* 32, 381–386, 2014; Haghverdi et al. *Nature Methods* 13, 845–848, 2016), in other words, it averages gene expression over longer periods and may smooth over abrupt transitions. RNA velocity, on the other hand, captures rapid transcriptional dynamics because it directly measures the unspliced RNA that represents the immediate “future” of the transcriptome. This difference in temporal resolution can lead to divergent patterns. The discrepancies between pseudotime and RNA velocity do not indicate that one of the analyses is incorrect. Instead, they provide complementary views of cellular dynamics. For example, if RNA velocity suggests a directional shift that is not reflected in the pseudotime trajectory, it might be revealing a transient state, or an impending cell fate change that the static pseudotime order has not captured yet. Divergent patterns should prompt further validation in the future (e.g., via lineage tracing or additional time-series experiments) to resolve complex cell-state transitions. This complementary use can enrich the biological interpretation rather than compromise the analysis.

The Reviewer also indicates that *“the Figure 3D portrait plots do not show strong dynamic patterns”*. In our study, we have applied stochastic modeling rather than dynamical modeling. Stochastic modeling in scVelo does not provide the fitted dynamical model (dynamical fit) and does not attempt to estimate the full dynamical splicing kinetics (not parameterizing rates of transcription, splicing and degradation in a time resolved way). Thus, it does not fit the parametric curve relating spliced and un-spliced counts.

The Reviewer also indicates that *“Figure 5 velocity and pseudotime from controls do not match the velocity pattern from Figure 3”*. As indicated above, integrated scRNA-seq and snRNA-seq provides a better and more significant RNA velocity confidence score than scRNA-seq, so it is not surprising that pseudotime and velocity in islets from nondiabetic donors in both datasets do not perfectly match, ours done with integrated scRNA- and snRNA-seq and HPAP done with scRNA-seq only. However, in both non-diabetic human islets datasets bifurcate trajectories are observed. Although HPAP datasets from islets of non-diabetic and T2D donors only contain scRNA-seq data, we were able to compare trajectories between both conditions showing different pattern of trajectories, from the bifurcate more complex trajectories in non-diabetic islets to the unique trajectories from beta to alpha cells observed in islets from T2D donors.

3. It is difficult to judge the staining pattern of SMOC1 with glucagon and insulin.

As requested by the Reviewer, we now include new confocal microscopy images at higher magnification of SMOC1 and C-peptide, and SMOC1 and glucagon immunolabeling of pancreas sections from non-diabetic and T2D human organ donors. These new images substitute the images in Fig. 6E.

4. SMOC1 was identified because they posit between beta to alpha transition. However, all the functional data the authors shown indicate beta cell dysfunction induced by SMOC1 overexpression. Does SMOC1 induces beta cell to alpha cell conversion?

To address the Reviewer's question, we overexpressed hSMOC1 in EndoC-BH1 human beta cells and analyzed expression of 8 gene markers that define alpha cells (see Fig. 2A-B) to determine whether there was any “induction” in the beta to alpha cell conversion. As shown in the new Supp. Fig. 9, several genes were upregulated including *MAFB*, *TTR* and *CRYBA2*, (and also *ALDH1A1* as shown Suppl. Fig. 8) while others *GCG*, *ARX*, *SCG2*, *FAP*, *PDK4* and *LOX4* were not altered indicating incomplete induction of conversion of beta cells into bihormonal cells. This indicates that in addition to *SMOC1*, the expression of other genes is required to induce the complete appearance of an alpha cell phenotype in beta cells. Methods (page 17) and Results (page 9) have been modified to include this new information. It is important to note that the title of the manuscript mentions *SMOC1* as a beta cell dedifferentiation gene and does not mention that *SMOC1* induces the conversion of beta cells into alpha cells. Although *SMOC1* expression in beta cells alone induces changes in the expression of some gene markers that define alpha cells, it is not enough by itself to cause the conversion of beta cells into alpha cells.

Minor points:

1. Line 17, the “expression of GCG an ALDH1A1 genes was ... similar in scRNA- and snRNA-seq datasets”? However, it looks like the scRNA-seq had way higher expression levels compared with snRNA-seq.

Thank you for noticing this inaccuracy. This phrase has been modified.

Reviewer #2

We thank the Reviewer for his/her helpful comments. Next, we address individually each of the issues raised by the Reviewer.

1. “It was unclear to me whether the velocity analysis was performed just on snRNAseq or scRNAseq, or both. Does limiting it to one or the other change velocity outcome? If so, how did you decide which dataset to use?”

Analysis of RNA velocity has been traditionally performed with scRNA-seq and not with snRNA-seq because snRNA-seq data contain transcripts with more intron-retention and, hence will have increased unspliced rates compared with scRNA-seq data where less intron-retention cytoplasmic mature transcripts are captured. However, we reasoned that integrating these two modalities in the same samples will allow to capture both the early transcriptional dynamics (from snRNA-seq) and the downstream more mature expression profiles (from scRNA-seq), thereby offering a more complete temporal trajectory of cell state transitions, enhancing overall velocity confidence. Indeed, RNA velocity confidence score using integrated scRNA-seq and snRNA-seq was significantly improved (mean 0.81; median 0.93, Mann-Whitney-Wilcoxon test, $p=1.55e^{-135}$) compared with scRNA-seq alone (mean 0.71; median 0.80). This information is now included in the Methods (page 15) and Results sections (pages 5 and 6) and in the new Suppl. Fig. 4. Of note, in the case of the HPAP dataset, there is only scRNA-seq available, so we performed RNA velocity comparison between ND and T2D samples only from the data repository.

2. Is it possible to know whether the effective concentrations of SMOC1 in Figure 8 are similar to what beta cells are exposed to in vivo? Please discuss this point.

SMOC1 concentrations in the circulation of non-diabetic subjects are in the range of 50 to 500ng/ml as previously reported by Montgomery et al. (ref. 28). However, the level of SMOC1 in the circulation of T2D subjects has not been reported. Therefore, we measured the circulating levels of SMOC1 in non-diabetic and T2D subjects and found that SMOC1 levels are significantly increased in T2D subjects compared with non-diabetic individuals, reaching levels up to 800 ng/ml. These results are shown in the new Fig. 8A. These SMOC1 levels are close to the ones used in the *in vitro* experiments in Figs. 8B-E. The Methods (page 17), Results (page 10) and Discussion (page 12) sections have been modified to reflect these new data.

Minor

1. P 4 line 37 . . . you should cite your additional file 2: Table S2 since Figure 2D doesn't alone hold this information.

This has been cited. Thank you.

2. Page 5, lines 18-19. Authors state that PDX1, INSM1 and SMAD1 are transdifferentiation markers, but don't cite papers showing this, or demonstrate it with their own data.

These genes are included in the trans-differentiation pathway in the Molecular Signatures Database (MSigDB) of the gene set enrichment analysis https://www.gsea-msigdb.org/gsea/msigdb/human/geneset/GOBP_TRANSDIFFERENTIATION.html. However, since we do not have evidence that these are transdifferentiation markers, we have removed the mention of these genes from the text.

3. P 10, line 41. Please specify that this work was done in zebrafish.

As requested, this has been included in the text.

4. P 11, lines 1-5. It is unclear to me what these sentences mean. Does your data contradict this previous study? What was different between the two studies?

Thank you for noticing that these sentences were unclear. We have changed the text to indicate that “The orphan nuclear receptor NR4A1 is important for β -cell mitochondrial function and insulin secretion (59). In the current study, we find that NR4A1 is expressed in α -cells and β -cells in non-diabetic islets, and it is upregulated in β -cells of T2D subjects, perhaps as a compensatory mechanism for insulin resistance in T2D. This result is in contrast with a previous report indicating that NR4A1 expression assessed by qPCR is not altered in whole islets from T2D subjects (60). The different method employed for gene expression analysis and the use of whole islets vs. single cells could explain the difference in the results.”

5. P 12, line 11—use of “beta” spelled out rather than the Greek symbol, unlike the rest of manuscript.

Thank you. This has been corrected in the text.

6. Please acknowledge HPAP funding, as agreed to when using HPAP datasets. From the website: Citation requirement: You agree to cite this database and the corresponding data providers every time you use this data for any publication purpose. Please use this NIH grant number- UC4-DK112217 and UC4-DK112232, an HPAP identifier (RRID:SCR_016202), and the HPAP database URL- <https://hpap.pmacs.upenn.edu/> for citation purposes. Please use the name “Human Pancreas Analysis Program (HPAP) Database, consortia under Human Islet Research Network (RRID:SCR_014393).

We apologize for not properly acknowledging HPAP throughout the manuscript. This has been corrected as indicated in pages 8, 13, 16, and 19.

7. Several grammatical and idiomatic errors were noted; further editing would benefit the readability of this manuscript.

The manuscript has been revised, and readability has been improved. Thank you.

ANSWERS TO THE REVIEWERS-R2

We would like to thank the Reviewers for their comments on our revised manuscript. Next, we address point-by-point all your comments.

Reviewer #1 (Remarks to the Author):

Thank you for your comments which we address point-by-point below.

1. It is still unclear to me whether the AB cluster is a real subtype, doublets/aggregates or some other artifacts such as cells with low quality. What is the logic behind “the AB cluster ... display enhanced chromatin accessibility for INS and GCG genes, strongly suggesting that AB cells are bona fide INS-positive and GCG-positive cells and not cell doublets/aggregates”? If the AB cells are indeed alpha/beta doublets, won't they also show open chromatin for both INS and GCG? The authors did not show enough evidence that AB cluster exhibit distinct expression profiles that cannot explained by mixture from different islet cell types.

As indicated by the Reviewer in the first round of comments, the existence of human islet bi-hormonal cells (positive for insulin and glucagon immunostaining) has been already reported. In our previous rebuttal, we included further evidence supporting similar percentages of reported human islet bi-hormonal cells (immunostaining) and AB cells in our study. Importantly, if bihormonal cells exist at the protein level, as shown in the literature and indicated by the Reviewer, it is likely that they exist at the transcriptional level as well. Indeed, several publications have already reported the existence of *GCG+//INS+* islet cells/clusters using scRNA-seq:

a. Karakose, E., Wang, X., Wang, P., Carcamo, S., Demircioglu, D., Lambertini, L., Wood, O., Kang, R., Lu, G., Scott, D. K., Garcia-Ocaña, A., Argmann, C., Sebra, R. P., Hasson, D., & Stewart, A. F. (2024). Cycling alpha cells in regenerative drug-treated human pancreatic islets may serve as key beta cell progenitors. *Cell reports. Medicine*, 5(12), 101832. (ref. 33)

b. Elgamal, R. M., Kudtarkar, P., Melton, R. L., Mummey, H. M., Benaglio, P., Okino, M. L., & Gaulton, K. J. (2023). An Integrated Map of Cell Type-Specific Gene Expression in Pancreatic Islets. *Diabetes*, 72(11), 1719–1728. (ref. 47)

c. Saikia, M., Holter, M. M., Donahue, L. R., Lee, I. S., Zheng, Q. C., Wise, J. L., Todero, J. E., Phuong, D. J., Garibay, D., Coch, R., Sloop, K. W., Garcia-Ocana, A., Danko, C. G., & Cummings, B. P. (2021). GLP-1 receptor signaling increases PCSK1 and β cell features in human α cells. *JCI insight*, 6(3), e141851. (ref. 49)

d. Segerstolpe, Å., Palasantza, A., Eliasson, P., Andersson, E. M., Andréasson, A. C., Sun, X., Picelli, S., Sabirsh, A., Clausen, M., Bjursell, M. K., Smith, D. M., Kasper, M., Ämmälä, C., & Sandberg, R. (2016). Single-Cell Transcriptome Profiling of Human Pancreatic Islets in Health and Type 2 Diabetes. *Cell metabolism*, 24(4), 593–607. (ref. 12)

e. Marquina-Sanchez B, Fortelny N, Farlik M, Vieira A, Collombat P, Bock C, Kubicek S. (2020). Single-cell RNA-seq with spike-in cells enables accurate quantification of cell-specific drug effects in pancreatic islets. *Genome Biol.* May 6;21(1):106.

Therefore, the identification of the *GCG+//INS+* (named AB in this manuscript) cell subtype in scRNA-seq studies has already been reported.

In the current study, we used stringent quality controls to eliminate doublets/aggregates and cells with low quality as reported in Kang et al. *Genome Medicine* 2023. Specifically, the AB alpha cell cluster passed all the standard cell quality controls: they have 1) more than 500 UMI count, 2) more than 250 gene varieties, 3) more than 0.8 log10 genes per UMI (gene density), and 4) 20% or less mitochondrial genes. The median nFeatures_RNA for AB cells is 692 which is within the range expected for cells and not background/noise or potential doublets/aggregates as indicated by the two most prominent community analysis packages recommended by 10X Genomics [Seurat: nFeature_RNA>200 for real cells, and scanpy: n_genes_by_counts (synonym for nFeature_RNA) >100 for real cells]. This information is now included in the Results section, page 4.

To eliminate doublets/aggregates, we used the DoubletFinder Algorithm pANN (proportion of artificial nearest neighbors), which provides a score as a decision factor to distinguish between doublets and singlets. The score of the AB cluster is 0.23, very similar to the score of the rest of the islet cell types (see new Supplementary Figure 1c). Normally, the droplet is more likely to contain a doublet/aggregate if the score is above 0.5 and closer to 1, and it is more likely to contain a singlet if the score is below 0.5 and closer to 0. This information is now included in the Results (pages 4 and 5) and Discussion (page 11).

It has been reported that cellular stress can promote non- β cells in the islet to exhibit β cell-like properties (Cigliola V, et al. Stress-induced adaptive islet cell identity changes. *Diabetes Obes Metab.* 2016;18-Suppl 1:87–96; Saikia M, et al. GLP-1 receptor signaling increases PCSK1 and β cell features in human α cells. *JCI insight*, 2021; 6, e141851). To determine whether AB cells could be endocrine cells under high levels of stress that could lead to expression of other hormone genes, we performed GSEA of stress pathways in these cells. We found that stress pathways were reduced in the AB cell cluster compared with the other alpha cell clusters (Supplementary Fig. 4a). The stress module score was lower in AB cells compared with other alpha cells and similar to beta-1 and beta-3 cell clusters (Supplementary Fig. 4b). This information is now included in Results (pages 5-6), Discussion (page 11) and the new Supplementary Fig. 4a-b.

Finally, we now include a new Supplementary Table 3 with the exclusive genes present in AB cells compared with beta and alpha cells suggesting that a specific transcriptional profile might define AB α cells. This information is included in the Results section, page 5.

In summary, (1) the previously reported existence of bihormonal cells at a protein and transcriptional level, (2) the stringent quality controls in our studies, (3) the number of n_Features and the DoubletFinder pANN Score within the expected range for cells (not background/noise nor potential doublets/aggregates), (4) the enhanced chromatin accessibility for *INS* and *GCG* genes in the ATAC-seq analysis, (5) the low levels of cellular stress, and (6) the exclusive genes present in AB cells strongly suggest the existence of these cells in human islets. Based on these points, we have modified the Results section (pages 4-6), Discussion (page 11) and added new Supplementary Figures 1c and 4a-b and new Supplementary Table 3.

2. Integration of scRNA-seq and snRNA-seq data and then perform RNA velocity analysis is not recommended, with reasons detailed in my previous comment. The authors should at least analyze the two types of data separately and interpret the velocity results in parallel. Furthermore, the lack of distinct direction in pseudotime plots (Figures 5c and 5d), as well as the inconsistency between Figure 5 and Figure 3 velocities are all indicative of potential issues in RNA velocity analysis.

As requested by the Reviewer, we now provide a new Supplementary Fig. 5b-f with RNA velocity and trajectory analysis of scRNA-seq data only of non-diabetic human islets. We compare these results with the results obtained with integrated scRNA-seq and snRNA-seq data already in Fig. 3. We did not perform RNA velocity and trajectory analysis of the snRNA-seq data alone since, as indicated by the Reviewer, this is not recommended. As shown in the new Supplementary Fig. 5f, we observe a similar trajectory pattern to the one obtained with integrated scRNA-seq and snRNA-seq (Fig. 3e) depicting a bifurcating trajectory pattern from the AB cluster both towards the beta and alpha cell subclusters. Additional trajectories using only the scRNA-seq data are 1) from less differentiated beta cells (beta 3) to more differentiated beta cells (beta 1), and 2) from alpha 3 to AB which further reinforces the bifurcating trajectory potential of AB cells. This suggests that both references (integrated sc/snRNA-seq or scRNA-seq alone) provide a bifurcating trajectory from AB to both alpha cells and beta cells in our dataset. We describe this information in the Results section (pages 6 and 7).

As indicated in our answer to the previous Reviewer's comments, the RNA Velocity confidence score was superior using integrated scRNA-seq and snRNA-seq than with scRNA-seq only (Suppl. Fig. 4, now Suppl. Fig. 5a). For the cell trajectory studies, we have used the packages Velocity and scVelo. In Scvelo, we can calculate the RNA Velocity confidence score which is a quality metric for how well the predicted RNA velocity aligns with each gene's observed splicing dynamic. When we do RNA velocity analysis, the package fits the model to the spliced/unspliced RNA, and this fit yields a measurement that indicates how reliable each velocity estimate is. The fit is based on how close the observed spliced/unspliced RNA matches the model's predicted splicing kinetic. The algorithm makes the prediction for the next cell's splicing dynamic and compares it with the real splicing ratio. It also considers the consistency of the direction, whether the predicted progression (for example increasing or decreasing spliced mRNA gradient) corresponds to the transcriptional states along a cell's trajectory. In short, the higher prediction score implies the splicing model fits the data more robustly, resulting in more trustworthy prediction. This has already been described in ref. 39 (Bergen, V., Lange, M., Peidli, S. et al. Generalizing RNA velocity to transient cell states through dynamical modeling. *Nat Biotechnol* 38, 1408–1414, 2020). Based on all this information, we believe the analysis of RNA Velocity and PAGA trajectories using the integrated scRNA-seq and snRNA-seq data is appropriate.

Regarding your comment “...the lack of distinct direction in pseudotime plots (Figures 5c and 5d)”, it is important to note that pseudotime can suggest useful abstraction of the cellular progression, but they generally have limited ability for capturing the complexity of multi-branching trajectories as suggested by Saelens et al. A comparison of single-cell trajectory inference methods. *Nat Biotechnol* 37, 547–554 (2019), and Li, S. et al. A relay velocity model infers cell-dependent RNA velocity. *Nat Biotechnol* 42, 99–108 (2024). Pseudo-temporal algorithm is based on transcriptional similarity rather than suggesting trajectory vectors, and it does not clearly provide directional information unless anchored by specific root cells or guided by prior knowledge. In the case of Figs. 5c (Fig. 5d is the PAGA analysis), we used un-biased, algorithm defined root from RNA velocity for pseudotime trajectory, and the visual representation in pseudotime may not explicitly reflect the directionality since we have multi-branching transition as we can observe from RNA velocity. The aim of pseudotime analysis was to capture the relative progression of cellular states, rather than suggesting the strict temporal directionality. This is an assumed limitation in trajectory inference. Zheng, S.C. et al. Pumping the brakes on RNA velocity by understanding and interpreting RNA velocity estimates. *Genome Biol* 24, 246 (2023). Nevertheless, pseudotime plots with a pre-defined root by the PAGA shows distinct directions between non-diabetic and T2D

human islet cell clusters in the HPAP data sets. This is now shown in the new Supplementary Fig. 7 and in the Results (page 9).

Regarding your comment “...the inconsistency between Figure 5 and Figure 3 velocities...”, we agree that the velocities and trajectories in the HPAP non-diabetic dataset only partially resemble the trajectory pattern in the integrated scRNA-seq and snRNA-seq data, and now also in the scRNA-seq data alone of our islet cells from non-diabetic human donors (new Supplementary Fig. 5f). We believe this can be related to: 1) the smaller number of samples in our dataset compared with HPAP; 2) the wide range in age (range 4-55 years) and BMI (20.6-38.1) of the HPAP non-diabetic human islet donors compared with our samples (age 24-42, BMI 20-26.3) that can influence the transcriptional gene profile of islet cells; and, 3) the inclusion of 3 donors with HbA1c in the range of 5.7-6.0 in the HPAP non-diabetic human islet samples suggesting potential undiagnosed pre-diabetes that could affect the results. Results (page 8) have been modified including all these comments. However, it is important to emphasize that Figure 5 makes the important point that the cell trajectories between non-diabetic and T2D human islet clusters in the HPAP dataset are very different since T2D islet cell clusters show unidirectional trajectories from β -cells to α -cells suggesting potential pressure on β -cells to become less differentiated or converted to α -like cells. Introduction (page 3) and Discussion (page 12) have been modified to emphasize this point.

Reviewer #2 (Remarks to the Author):

1. ***All of my points have been successfully addressed.***

Thank you.