nature portfolio

Peer Review File

Molecular logic for cellular specializations that initiate the auditory parallel processing pathways

Corresponding Author: Dr Xiaolong Jiang

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)
Jing et al. report a landmark study - in many respects the first of its kind across any field of neuroscience - that showcases a truly remarkable alignment of genetics, morphology, and intrinsic biophysical properties in the first station of sound processing in the mammalian brainstem. This study will be impactful and highly cited for two reasons: 1) it synthesizes genetic, computational, and electrophysiological approaches with a large sample size and rigorous, careful quantification that will provide a valuable public data resource; 2) In so doing, it reveals categorically different principles of neural organization than have been identified in previous patch-seq efforts from the mammalian forebrain. All in all, this is among the very best manuscripts I have reviewed for any journal in the last few years. The correspondence between morphology, intrinsic electrophysiological properties, and genetics is nothing short of stunning. It is a rare paper these days that leverages cutting-edge technology to simplify and harmonize our understanding of a brain circuit instead of revealing greater depths of complexity (and confusion). It's the kind of paper should appear in undergraduate textbooks because it shows students that in some cases the organization of the brain makes perfect sense.

I have but a few minor suggestions:

1) The authors might consider rewording the phrase on Ln 35, "These CN cell types are perhaps one of the best-defined in the nervous systems". I'm not sure that this is true, how one would back up this claim, nor whether it does them any favors. It puts readers in an adversarial mind frame, as I began thinking of all the cell types that might prove that statement wrong (e.g., cortical inhibitory cell types, hippocampal cell types, certain cell types in the cerebellum etc.). At the very least, "systems" shouldn't be plural. How about a short positive statement about CN cell types that doesn't invoke a comparison to other cell types?
2) Figure 1a could be improved to help orient the uninitiated. Perhaps a recognizable drawing of the mouse brain with a blowout box to highlight the CNC and plane of section used in the cartoon. Perhaps anatomical axes and a scale bar?
3) I'm assuming that the cells in 1a are cartoon drawings but the cells in 2a are actual fills? Can this be clarified?
4) I was confused by Figure 5g vs 5h. What are ME-types? It isn't defined? Does it correspond to morphology (which was defined as M-clusters in Fig. 2 and elsewhere)? More to the point, it wasn't clear why the UMAP visualization was performed independently for 5g and 5h.
5) I think the interpretation and discussion of the data presented in Figure 6 and 7 could be made clearer for the reader. In particular, a sizeable fraction of readers will be poring through the text and associated data looking for candidate genes for future Cre lines that will provide the means to monitor and manipulate individual CN cell types. My interpretation of the data is that, in a nutshell, there aren't many cases where this would be true. Instead, cell identity can only be distinguished through synergies and complementarities of multiple genes encoding transcription factors or ion channel subunits. There may be one or two examples to the contrary (which could be more clearly highlighted) but I think the takeaway message was that there aren't great single gene candidates to isolate CN cell types from these coding sequences (which is what lead Fishell and colleagues to atac-seq to isolate enhancers for various forebrain cell types). For this reason, it seems like the sentences on Lns 465-8 are a little misleading for most readers. As I understand things, the current state of genetics are such that there aren't many candidate genes identified here that would inform the generation of Cre driver lines for cell type manipulations. If I am wrong about that, it only underscores the point that they should distill the discussion of these points for the average reader seeking this information.
6) The Discussion is already pretty long but there was one aspect of their data set that wasn't mentioned that might deserve some consideration. DCN stands apart from a/pVCN in that DCN cell types have a propensity for developing hyperactivity

after cochlear deafferentation whereas a/pVCN cell type activity tends to track afferent input from the auditory nerve, becoming depressed after most forms of peripheral deprivation. It might prove interesting to see if DCN cell types (fusiform cells in particular) are enriched for transcripts involved in homeostatic plasticity (e.g., tnf-alpha) that aren't observed in a/pVCN. This might explain the propensity of some CN cell types to exhibit hyperactivity after hearing loss, which has been suggested to be a root cause of common neurological disorders like tinnitus.

Reviewer #2

(Remarks to the Author)
This manuscript by Jing et al describes a cell type taxonomy of the mouse cochlear nucleus (CN) based on single-nucleus RNA sequencing (snRNA-seq) and multimodal profiling of neuronal transcriptome, morphology, and electrophysiology using patch-seq. Their overall experimental approach is not only appropriate but also cutting-edge and exciting, and their data analysis methods are thorough and rigorous. Moreover, their interpretation of results is clear, logical, and well-founded for the most part, supported by engaging, thoughtful writing and well-organized, nicely illustrated figures. As the first comprehensive molecular atlas of CN cell types, this work will no doubt serve as a valuable reference and springboard for future studies on CN contributions to auditory processing.

Listed below are some points of concern the authors should address.

1) The methodology behind mapping of patch-seq transcriptomic profiles onto UMAP space depicting snRNA-seq clusters (e.g., 2f, 3a and 4a) needs to be elaborated. This visualization is an important aspect of their study given its significance for reconciliation of cell type definitions. In lines 754-755, they wrote: "Each Patch-seq cell was assigned to a cluster to which it had the highest correlation." This however only explains cluster assignment, not 2D embedding. Same applies for mapping of tdTomato cells on E- and M-clusters in Fig. 5l. I can see that the mapping was done using the python code from an earlier Berens Lab (Tübingen) publication, but the underlying logic needs to be explained in plain text in Methods. Additionally, 'rnaseqTools.py' on the Github repository for this study appears to be a copy of the python code from Berens Lab repositories. If this is true, the authors are advised to note in the header that the code was authored by someone else, and/or point out explicitly in README.md that the code was copied in whole or part from [source].

2) In some cases, patch-seq profiles seem to form subclusters when mapped onto snRNA-seq UMAP space even though patch-seq T-clusters do not show such structure (e.g., 2f vs. 2e). What is the significance of these subclusters? Could they represent additional subtypes that the snRNA-seq profiles do not have the resolution to capture? This is important to consider because patch-seq profiles can have higher information content than snRNA-seq profiles. The vastly different cut-offs used in this study for gene detection (4000 vs. 200) makes this distinction even more stark, so it is important to note that the patch-seq dataset, although handicapped somewhat by sample size, may have higher predictive power for subtypes of principal CN neurons than the snRNA-seq dataset in some cases.

3) Their E-, M- and T- clusters are significantly different in sample size (Fig. 2c and 2d). This is likely explained by exclusion of certain cells: "For morphological feature extraction, only cells with a sufficient morphological recovery were included" (Lines 690-691). Please explain what objective criteria were used to determine 'sufficient'. In addition, if the explanation for why cells fail to qualify is incomplete filling of dendrites, please elaborate on why some cells have this issue (e.g., do they correlate with recording duration?) and the possibility that some morphological features that may further distinguish subtypes (e.g., fine dendritic branches that may take much longer to fill) may have been missed.

4) There is no explanation for what 'expert classification' really is: Who is the expert? If there were >1, did they independently assign cell types? How were disagreements reconciled? All of these should be noted in Methods.

5) UMAPs in 3a and 3b appear to be the same but they aren't identical. While cell location in much of the two panels appear to be similar, there are clear differences toward the lower left corner. For example, some cells visible in 3b are missing in 3a.

6) The authors wrote in lines 259-261: "the random forest classifier and clustering-based k-means classifier based on either electrophysiology or morphology distinguished each cell type pair with high accuracy, especially with additional anatomic location supervision." Please explain in Methods what exactly 'additional anatomic location supervision' is and how it was implemented, noting especially whether measures were taken to keep it unbiased.

7) Transcriptomic studies are excellent for generating hypotheses but presence (or level) of transcripts for a particular gene does not guarantee specific neuronal properties. Many inferences have been drawn in this study based on correlation-based approaches, but correlation is merely suggestive—it does not substantiate a working model. Although it is appreciated that the authors are careful to note the suggestive/speculative nature of some of their discussion points, this consideration is lacking in some parts. I note a few examples below, but the authors are advised to amend their wording where necessary throughout the manuscript.

Lines 392-393: "This analysis substantiated Kv1.1 and HCN1 as molecular determinants of single-spike firing property…" In this example, their analysis simply suggests that the named genes are candidate molecular determinants.

Lines 380-381: "To substantiate this contention, and identify these molecular determinants, we took advantage of the bimodal gene expression and physiological data…" In this example, the sparse reduced-rank regression they performed is interesting, but in no way does it substantiate the contention in the preceding sentence, nor identify molecular determinants.

What it does is generate candidate molecular determinants, the caveat of relying on mRNA instead of proteins levels notwithstanding.

Lines 559-560: “We identified three transcripts including Kcna1, Hcn1, and Kcnq4 as molecular determinants for single-spike firing property,…”

8) Synergistic core identification, Lines 914-915: “For each target subpopulation, a set of TFs that are specifically expressed in the target cell type (specific TFs) is identified.” This does not fit with a note elsewhere in the manuscript that “Besides a few genes (e.g., Atoh7), we did not find TFs whose expression correlates strictly with each cell type, indicating the importance of combinatorial TF codes in specifying cell type identity.” As is evident in Fig. 3b, even Atoh7 does not really correlate strictly with a cell type – it is expressed at slightly higher levels in one subtype of BCs than the other. Given the general lack of truly cell type specific-TFs and the significance of this criterion for how the computational approach developed by Okawa and Del Sol (2019) works, it is not clear if the synergistic cores identified in this study tell us anything biologically meaningful.

9) Their discussion of the “terminal selector” concept in describing transcription factors that may control CN cell type identities is highly speculative. In lines 324-326, they wrote: “Each IF core is characterized by the close-knit group of TFs that regulate each other (including themselves, i.e., auto-regulatory) and other downstream targets specific to this cell population, suggesting they are “terminal selector” for each CN cell type.” There are well-established criteria for whether a gene is a terminal selector. See Hobert (2016) (PMID: 26970634) for example. Just because a TF is part of a regulon and is maintained in mature neurons does not make it a candidate terminal selector. Loss-of-function studies and experimental validation of gene regulatory relationships in cell types in question would be necessary to ascertain this.

10) Lines 530-531: “Our use of multiple bioinformatic tools identified high-confidence candidate identity TFs for each CN cell type.” I agree candidate identity TFs were identified but the basis for high confidence is not clear. Given the predictive nature of most bioinformatic tools, even when several of them point to the same working hypothesis, if they utilize the same underlying incomplete model of how certain TFs work and whether their roles are context-dependent, the likelihood of a given TF regulating neuronal identity does not go up.

11) Lines 551-552: “Some of these families are shared with other brain circuits, but many appear to be specific to auditory systems.” Please point here to data in the manuscript that supports this claim.

Reviewer #3

(Remarks to the Author)
The cochlear nuclei contain a variety of neuronal cell types that underly this initial stage of auditory processing, but definitive molecular identifiers for many of types are not known. In this study, the authors use an single-nucleus RNA-seq data set to define transcriptomic cell types, then use the Patch-seq technique to link those types to previously established types by examining their distinguishing electrophysiological and morphological properties. For some neuronal types, the transcriptomic types suggest a more granular distinction than the previously established cell types, which the authors find also correlate to further phenotypic differences.

Overall, the study is well designed and executed, and the results relating the cell types identified through different methods are clear and compelling. The authors also connect different neuronal subtypes to spatial locations in the CN as further validation of their distinctiveness. The transcriptomic characterization of these major cell types in the CN will be useful for producing molecular tools to manipulate and measure the system, and the additional classifications will assist in understanding the neuronal circuits in finer detail. I do have more questions, however, about some of the later analyses presented in the manuscript that are less clearly presented than the earlier, core components of the study.

The description and presentation of the analyses in Figure 6 are much more difficult to follow than the rest of the paper. I can understand that certain gene families are more able to distinguish the identified cell types (Fig. 6a-b), and that certain transcription factors are differentially expressed across the projection neuron types (Fig. 6c), but then the figures and description of results become harder to understand. I don't know why all the neuronal types are used in Fig. 6a-b but then only the five projection types are shown in the rest of the figure (while the other types are presented in less detail in Extended Data Fig. 12). I'm not sure what it means on lines 311-312 where "Besides a few genes (e.g. Atoh7) we did not find TFs whose expression correlates strictly with each cell type" - Atoh7 expression seems quite similar to the other TFs expressed by the Bushy_Atoh7 type, and TFs for other types seem similar respective to their particular type, as well. It is not clear what the referenced SCENIC analysis shows or how that relates to the values shown in Fig. 6d (or why those particular genes are shown in the figure). The description of the "synergistic core analysis" in the Results and Methods is pretty impenetrable - is Fig. 6e just showing TF expression again (as in Fig. 6c), but now organized by some dendrogram on the left? Again, I don't understand what sentences like (lines 323-324) "Their regulons could be mostly predicted in each cell type, except for a few understudied ones (e.g., Tshz2 and Tschz3)" mean - how do Tshz2 and Tshz3 appear to be different in Fig. 6e from neighbors like Rarb and Rere? It is difficult to know what specific conclusions to take away from these analyses.

I also am not as clear on the utility of the spare reduced-rank regression analyses in Figure 7. The authors seem to conclude that particular subsets of genes are responsible for particular sets of electrophysiological features (e.g. Kcna1, Hcn1, and Kcnq4 on lines 389-395). But those genes don't appear to be particular well-correlated with the regression components (their vectors in Fig. 7b are fairly short), and the bi-biplots don't really illustrate the predictive power of those genes for particular

features. For example, Hcn1 seems to be expressed in essentially all the cells (Fig. 7a), but to different mean expression values, which is similar to the expression of several sodium channel genes in Fig. 7a, which are said to not be determinants for electrophysiological variation in these neurons. Perhaps that is because the sodium channels vary less in the five projection neuron types than Hcn1, but the presented analysis only includes those types? (As in Figure 6, it is not stated why the analysis moves to only focus on the excitatory types). Another issue with using the sparse RRR method to identify determinant genes is that because it is a sparse model, if two genes have similar expression patterns, it will likely choose only one (arbitrarily) because it can predict the electrophysiological features equally well with either - but both may be involved mechanistically in determining those properties. Finally, I'm not sure why the model hyperparameters were selected to identify a target number of 25 genes - it seems like it would be of more use to find the most predictive model, even if more genes were identified as being relevant? In general, I believe that many of the identified genes are in fact related to the electrophysiological phenotypes of these neurons (there are strong mechanistic reasons to think so), but the statements in the Results seem overly strong compared with what the analyses can really demonstrate.

Despite those concerns with the latter two figures, this is a well done study with clear, novel results that help better characterize the neuronal components of auditory circuits. It demonstrates the strong utility of the multiple modality Patch-seq data set and should enable future work to more precisely target neuronal populations in the CN.

Minor points:
- Fig. 7c would benefit from some additional labels to help readers more quickly identify the genes associated with the receptor types discussed in the text (AMPARs, NMDARs, kainate receptors)
- Extended Data Fig. 9 could benefit from some illustrations of specific electrophysiological features that differ between the subtypes. For example, the electrophysiology UMAP could be colored by particular electrophysiology feature values, like input resistance or spike delay.

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
No further comments

Reviewer #2

(Remarks to the Author)
I am satisfied with the authors' responses and changes + clarifications in the revised manuscript. A minor suggestion for lines 323-338 and 339-353: although based on different methods, there is significant conceptual overlap in these sections (combinatorial TF code for regulation of cell type identity, continuity of TF expression from development to maturity). They could be tied together better, either by merging or drawing a stronger logical link between them while minimizing redundancy.

Reviewer #3

(Remarks to the Author)
In their revision, Jiang et al. have clarified and improved an already strong manuscript. They have addressed my concerns with the original submission in detail: in particular, they have improved the interpretability and description of Figure 6 and provided substantial justification for the results of Figure 7 (with the revisions to Extended Figure 12). I have no remaining major concerns with the manuscript.

AUTHORS' RESPONSE TO THE REVIEWERS:

We are deeply grateful for the reviewer's thorough and insightful reviews of our manuscript. Your comments and suggestions have greatly improved the quality and clarity of our work, and we appreciate the time and effort you took to provide such detailed feedback.

Reviewer #1 (Remarks to the Author):

Jing et al. report a landmark study - in many respects the first of its kind across any field of neuroscience - that showcases a truly remarkable alignment of genetics, morphology, and intrinsic biophysical properties in the first station of sound processing in the mammalian brainstem. This study will be impactful and highly cited for two reasons: 1) it synthesizes genetic, computational, and electrophysiological approaches with a large sample size and rigorous, careful quantification that will provide a valuable public data resource; 2) In so doing, it reveals categorically different principles of neural organization than have been identified in previous patch-seq efforts from the mammalian forebrain. All in all, this is among the very best manuscripts I have reviewed for any journal in the last few years. The correspondence between morphology, intrinsic electrophysiological properties, and genetics is nothing short of stunning. It is a rare paper these days that leverages cutting-edge technology to simplify and harmonize our understanding of a brain circuit instead of revealing greater depths of complexity (and confusion). It's the kind of paper should appear in undergraduate textbooks because it shows students that in some cases the organization of the brain makes perfect sense.

We thank the reviewer for these laudatory comments.

I have but a few minor suggestions:
1) The authors might consider rewording the phrase on Ln 35, "These CN cell types are perhaps one of the best-defined in the nervous systems". I'm not sure that this is true, how one would back up this claim, nor whether it does them any favors. It puts readers in an adversarial mind frame, as I began thinking of all the cell types that might prove that statement wrong (e.g., cortical inhibitory cell types, hippocampal cell types, certain cell types in the cerebellum etc.). At the very least, "systems" shouldn't be plural. How about a short positive statement about CN cell types that doesn't invoke a comparison to other cell types?

We thank the reviewer for this suggestion. We changed the wording as suggested (lines 35-36).

2) Figure 1a could be improved to help orient the uninitiated. Perhaps a recognizable drawing of the mouse brain with a blowout box to highlight the CNC and plane of section used in the cartoon. Perhaps anatomical axes and a scale bar?

We thank the reviewer for this suggestion. We revised the figure as recommended (See revised Fig1a).

3) I'm assuming that the cells in 1a are cartoon drawings but the cells in 2a are actual fills? Can this be clarified?

Correct. We have revised the figure legends for Figures 1a, 2a, and 5a to clarify this distinction.

4) I was confused by Figure 5g vs 5h. What are ME-types? It isn't defined? Does it correspond to morphology (which was defined as M-clusters in Fig. 2 and elsewhere)? More to the point, it wasn't clear why the UMAP visualization was performed independently for 5g and 5h.

We apologize for the confusion by using the undefined term "ME-types". The term refers to morpho/electrophysiologically defined cell types (i.e., bushy cells, T-stellate cells, et.al). We now replaced it with a more precise description in the revised legend. Fig 5g shows a fully annotated

UMAP plot for excitatory neurons (Fig.2) and, based on data from 5a-f, inhibitory neurons as well. Fig 5h shows UMAP visualization of only inhibitory neuron clusters. We did not include excitatory neurons in Fig 5h as this figure focused on marker genes for only glycinergic cells. We revised the figure legends to make these points clearer.

5) I think the interpretation and discussion of the data presented in Figure 6 and 7 could be made clearer for the reader. In particular, a sizeable fraction of readers will be poring through the text and associated data looking for candidate genes for future Cre lines that will provide the means to monitor and manipulate individual CN cell types. My interpretation of the data is that, in a nutshell, there aren't many cases where this would be true. Instead, cell identity can only be distinguished through synergies and complementarities of multiple genes encoding transcription factors or ion channel subunits. There may be one or two examples to the contrary (which could be more clearly highlighted) but I think the takeaway message was that there aren't great single gene candidates to isolate CN cell types from these coding sequences (which is what lead Fishell and colleagues to atac-seq to isolate enhancers for various forebrain cell types). For this reason, it seems like the sentences on Lns 465-8 are a little misleading for most readers. As I understand things, the current state of genetics are such that there aren't many candidate genes identified here that would inform the generation of Cre driver lines for cell type manipulations. If I am wrong about that, it only underscores the point that they should distill the discussion of these points for the average reader seeking this information.

We agree that clarity in identifying candidate genes for future Cre lines for cell-type-specific manipulations is crucial. As noted, cell identity often depends on a combination of multiple genes, including transcription factors and ion channel subunits. However, we respectfully disagree with the assertion that there are no single gene candidates that could inform Cre driver line generation. There are numerous successful cases in the neocortex and other brain regions where a single gene could be used to label a specific cell type (Tasic et al., 2016). When a single gene alone is insufficient, intersectional approaches utilizing two genes have also proven effective in defining and manipulating specific cell types. In the context of CN cell types, we have identified several promising candidates, including Atoh7, Hhip, Penk, Necab2, Stac, Phgdh, which show restricted expression patterns. We have validated the existing Penk-Cre line as a driver line for vertical cells in this manuscript, providing the first demonstration of a single gene labeling a specific CN cell type. We are also actively generating and validating Cre lines based on these and other candidate genes, with many showing promising results. To enhance clarity for readers, we have revised the Discussion section to highlight these candidate genes and their potential utility for Cre-loxP mediated Cre driver line development (lines 479-485).

6) The Discussion is already pretty long but there was one aspect of their data set that wasn't mentioned that might deserve some consideration. DCN stands apart from a/pVCN in that DCN cell types have a propensity for developing hyperactivity after cochlear deafferentation whereas a/pVCN cell type activity tends to track afferent input from the auditory nerve, becoming depressed after most forms of peripheral deprivation. It might prove interesting to see if DCN cell types (fusiform cells in particular) are enriched for transcripts involved in homeostatic plasticity (e.g., tnf-alpha) that aren't observed in a/pVCN. This might explain the propensity of some CN cell types to exhibit hyperactivity after hearing loss, which has been suggested to be a root cause of common neurological disorders like tinnitus.

We agree that our dataset could be a valuable resource for investigating the molecular basis of the differential homeostatic plasticity observed between DCN and a/pVCN cell types. We also agree that TNF-alpha is a plausible molecule mediating homeostatic plasticity after sensory deprivation. We examined the expression of TNF-alpha (Tnf) and its receptors (Tnfrsf1a, Tnfrsf1b) within our dataset, but these genes were either not detected or expressed at very low

levels in all CN cell types, including those in the DCN. While this doesn't rule out a role for TNF-alpha, it suggests that other mechanisms are likely involved. We have added a brief discussion of this point in the revised manuscript, highlighting the potential for future studies to leverage our dataset to identify the molecular drivers of the distinct responses of DCN and VCN to deafferentation, a phenomenon that may contribute to conditions like tinnitus (lines 488-489).

Reviewer #2 (Remarks to the Author):

This manuscript by Jing et al describes a cell type taxonomy of the mouse cochlear nucleus (CN) based on single-nucleus RNA sequencing (snRNA-seq) and multimodal profiling of neuronal transcriptome, morphology, and electrophysiology using patch-seq. Their overall experimental approach is not only appropriate but also cutting-edge and exciting, and their data analysis methods are thorough and rigorous. Moreover, their interpretation of results is clear, logical, and well-founded for the most part, supported by engaging, thoughtful writing and well-organized, nicely illustrated figures. As the first comprehensive molecular atlas of CN cell types, this work will no doubt serve as a valuable reference and springboard for future studies on CN contributions to auditory processing.

We thank the reviewer for the laudatory and encouraging comments.

Listed below are some points of concern the authors should address.
1) The methodology behind mapping of patch-seq transcriptomic profiles onto UMAP space depicting snRNA-seq clusters (e.g., 2f, 3a and 4a) needs to be elaborated. This visualization is an important aspect of their study given its significance for reconciliation of cell type definitions. In lines 754-755, they wrote: “Each Patch-seq cell was assigned to a cluster to which it had the highest correlation.” This however only explains cluster assignment, not 2D embedding. Same applies for mapping of tdTomato cells on E- and M-clusters in Fig. 5l. I can see that the mapping was done using the python code from an earlier Berens Lab (Tübingen) publication, but the underlying logic needs to be explained in plain text in Methods. Additionally, ‘rnaseqTools.py’ on the Github repository for this study appears to be a copy of the python code from Berens Lab repositories. If this is true, the authors are advised to note in the header that the code was authored by someone else, and/or point out explicitly in README.md that the code was copied in whole or part from [source].

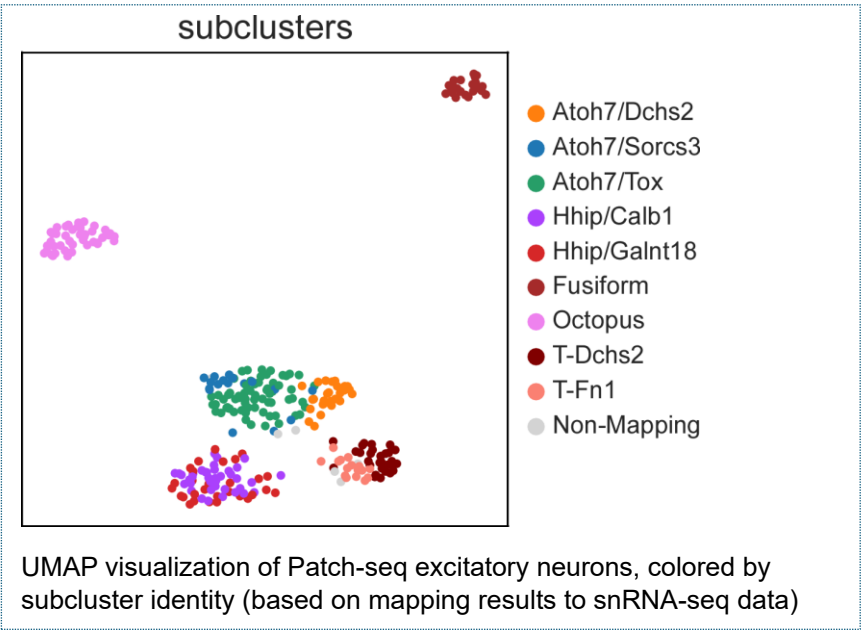
We thank the reviewer for these suggestions. We followed our previously established method (see cited refs. 28, 114) to position Patch-seq cells on snRNA-seq UMAP embeddings. Briefly, each Patch-seq cell was positioned at the median UMAP coordinates of its ten nearest neighbors in snRNA-seq UMAP embeddings, as determined by Pearson correlation. For Fig. 5l., a different approach was used to position tdTomato$^+$ cells on electrophysiological or morphological UMAP embeddings (E- or M-clusters): we first extracted the electrophysiological and morphological features of those tdTomato$^+$ cells, and these features were then used to cluster tdTomato$^+$ cells together with all other recorded inhibitory neurons (as shown in Fig.2c). Visualizing tdTomato$^+$ cells in the resulting UMAP embeddings revealed that they clustered with other recorded vertical cells. We have updated the methods section (lines 808-810; lines 879-884; lines 899-904) to clarify these procedures.

As suggested, we also updated the README.md file in our GitHub repository (https://github.com/bcmjianglab/cn_project/tree/main/patchseq/Mapping) to specify the source of the code.

2) In some cases, patch-seq profiles seem to form subclusters when mapped onto snRNA-seq UMAP space even though patch-seq T-clusters do not show such structure (e.g., 2f vs. 2e).

What is the significance of these subclusters? Could they represent additional subtypes that the snRNA-seq profiles do not have the resolution to capture? This is important to consider because patch-seq profiles can have higher information content than snRNA-seq profiles. The vastly different cut-offs used in this study for gene detection (4000 vs. 200) makes this distinction even more stark, so it is important to note that the patch-seq dataset, although handicapped somewhat by sample size, may have higher predictive power for subtypes of principal CN neurons than the snRNA-seq dataset in some cases.

The reviewer raised an important question regarding scRNA-seq experimental design: balancing library size (number of cells captured) with sequencing depth (number of reads per cell) to optimize cell type classification. Different scRNA-seq protocols, such as droplet-based and Smart-seq-based methods like Patch-seq, offer distinct advantages and disadvantages in this regard. Droplet-based snRNA-seq, despite capturing fewer genes than Patch-seq, demonstrates a similar ability to identify potential subtypes due to its substantially larger sample size, while deeply sequencing fewer libraries with Patch-seq do not enable better classification when the overall number of cells analyzed is limited (Shekhar et al., 2016; Ding et al., 2020). Indeed, Patch-seq profiles mapped onto the snRNA-seq UMAP space reveal subclusters not evident in the Patch-seq UMAP alone (Fig. 2f vs. 2e). This finding aligns with our subsequent analyses (Extended Data Fig. 9 and Fig. 4), where in-depth snRNA-seq analysis confirms multiple subpopulations within each of two bushy cell subtypes and two subtypes of T-stellate cells—also supported by Patch-seq mapping. The lack of distinct subclusters within the Patch-seq UMAP embedding itself is primarily attributed to the limited sample size. We anticipate that increasing Patch-seq sampling will reveal clearer subtype distinctions, mirroring the snRNA-seq resolution. It's worth noting that even within the current Patch-seq T-clusters, cells mapping to different subtypes or subclusters often occupy distinct corners, aligning with their positions in the snRNA-seq UMAP embedding (see attached figure below). We have revised the Result section to reflect these points (lines 186-188).



UMAP visualization of Patch-seq excitatory neurons, colored by subcluster identity (based on mapping results to snRNA-seq data)

3) Their E-, M- and T- clusters are significantly different in sample size (Fig. 2c and 2d). This is likely explained by exclusion of certain cells: “For morphological feature extraction, only cells with a sufficient morphological recovery were included” (Lines 690-691). Please explain what objective criteria were used to determine ‘sufficient’. In addition, if the explanation for why cells

fail to qualify is incomplete filling of dendrites, please elaborate on why some cells have this issue (e.g., do they correlate with recording duration?) and the possibility that some morphological features that may further distinguish subtypes (e.g., fine dendritic branches that may take much longer to fill) may have been missed.

Our criterion for "sufficient morphological recovery" was the presence of intact dendritic structures. Axons were excluded from this assessment as they are typically severed during slice preparation. The dendritic filling is considerably faster than axonal filling, and recording for over 5 minutes (our recording typically lasts 5-20 mins) is sufficient for complete dendritic filling, including that of fine dendrites, consistent with many prior studies (Lee et al., Elife, 2021; Scala et al., Nature, 2020). Cells that did not meet our criteria typically exhibited truncated dendrites (a result of slicing) or poor overall staining due to the aspiration procedure involved in Patch-seq (unclear or invisible soma and dendrite, likely caused by soma collapse during nucleus extraction and subsequent biocytin leakage). Our exclusion criteria do not target specific dendritic features and are not expected to bias potential subcategory resolution. Moreover, our morphological analysis results, in conjunction with electrophysiological properties and anatomical location within the CN, align with established cell type classifications and do not suggest the presence of morphological subcategories. We revised the Methods section (lines 718-723) to explicitly state our inclusion criteria for morphological analysis.

4) There is no explanation for what 'expert classification' really is: Who is the expert? If there were >1, did they independently assign cell types? How were disagreements reconciled? All of these should be noted in Methods.

Cell classification was performed by four individuals (TN, LT, JJ, MM), all with extensive expertise in the CN cell types. Each individual independently assessed each cell based on its location, morphology, input resistance, and responses to a family of depolarizing and hyperpolarizing current steps. The results of these independent evaluations were then compared to established features in the CN literature. Any disagreements were resolved through discussions in group meetings until a consensus identification was reached for each recorded cell. We have revised the Methods section (lines 945-950) to include these details, as suggested.

5) UMAPs in 3a and 3b appear to be the same but they aren't identical. While cell location in much of the two panels appear to be similar, there are clear differences toward the lower left corner. For example, some cells visible in 3b are missing in 3a.

We thank the reviewer for catching this error in Fig. 3a. The missing cells in Fig. 3a were inadvertently excluded due to an error in setting the y-axis limit during the initial figure generation. We have corrected this error and revised the figure (revised Fig.3) to ensure that all data points are accurately represented in the figure.

6) The authors wrote in lines 259-261: "the random forest classifier and clustering-based k-means classifier based on either electrophysiology or morphology distinguished each cell type pair with high accuracy, especially with additional anatomic location supervision." Please explain in Methods what exactly 'additional anatomic location supervision' is and how it was implemented, noting especially whether measures were taken to keep it unbiased.

We thank the reviewer for the opportunity to clarify the "additional anatomic location supervision". This refers to the inclusion of the recording location (either VCN or DCN) as an additional feature for cell classification. Inhibitory neurons recorded in the VCN were arbitrarily assigned a value of 0, while those recorded in the DCN were assigned a value of 1. These values were then incorporated alongside the electrophysiological values in clustering analysis and random forest classification (Fig.5d, Extended Data Fig.13g-h). To avoid bias, we

systematically tested different value assignments and found no differences in classifier performance. We have revised the Methods section to include these details (lines 873-878; lines 953-957). We also rephrased "additional anatomic location supervision" to "anatomic location information" in the revised text to better reflect the classification scheme and avoid confusion (Line 268).

7) Transcriptomic studies are excellent for generating hypotheses but presence (or level) of transcripts for a particular gene does not guarantee specific neuronal properties. Many inferences have been drawn in this study based on correlation-based approaches, but correlation is merely suggestive—it does not substantiate a working model. Although it is appreciated that the authors are careful to note the suggestive/speculative nature of some of their discussion points, this consideration is lacking in some parts. I note a few examples below, but the authors are advised to amend their wording where necessary throughout the manuscript. Lines 392-393: "This analysis substantiated Kv1.1 and HCN1 as molecular determinants of single-spike firing property…" In this example, their analysis simply suggests that the named genes are candidate molecular determinants. Lines 380-381: "To substantiate this contention, and identify these molecular determinants, we took advantage of the bimodal gene expression and physiological data…" In this example, the sparse reduced-rank regression they performed is interesting, but in no way does it substantiate the contention in the preceding sentence, nor identify molecular determinants. What it does is generate candidate molecular determinants, the caveat of relying on mRNA instead of proteins levels notwithstanding. Lines 559-560: "We identified three transcripts including Kcna1, Hcn1, and Kcnq4 as molecular determinants for single-spike firing property,…"

We agree with the reviewer's assessment. We have carefully amended our wording throughout the text to more accurately reflect the nature of our findings and the limitations of transcriptomic data in establishing causation. Specifically, we have: 1) revised language to highlight the hypothesis-generating nature of transcriptomics and emphasized that our study primarily generates hypotheses rather than definitively proving causal relationships; 2) we have replaced terms like "substantiated" and "identified" with more appropriate language such as "suggested," "implicated," or "predicted as candidate."; 3) we have clarified that correlations observed in our analyses are suggestive and require further validation to establish causal mechanisms.

8) Synergistic core identification, Lines 914-915: "For each target subpopulation, a set of TFs that are specifically expressed in the target cell type (specific TFs) is identified." This does not fit with a note elsewhere in the manuscript that "Besides a few genes (e.g., Atoh7), we did not find TFs whose expression correlates strictly with each cell type, indicating the importance of combinatorial TF codes in specifying cell type identity." As is evident in Fig. 3b, even Atoh7 does not really correlate strictly with a cell type – it is expressed at slightly higher levels in one subtype of BCs than the other. Given the general lack of truly cell type specific-TFs and the significance of this criterion for how the computational approach developed by Okawa and Del Sol (2019) works, it is not clear if the synergistic cores identified in this study tell us anything biologically meaningful.

We thank the reviewer for raising this important concern regarding the identification of synergistic cores and the specificity of transcription factors (TFs) in our study. We apologize for the inconsistency between the statements regarding TF specificity. To clarify, our statement that "Besides a few genes (e.g., Atoh7), we did not find TFs whose expression correlates strictly with each cell type..." is indeed a more accurate reflection of our findings. Most TFs, including Atoh7, did not exhibit exclusive expression patterns in specific cell types. We have revised the text to reflect this observation (lines 323-324) and clarify that "specific TFs" refers to TFs that are preferentially expressed in a target cell type compared to others (line 1001). These TFs may not

be exclusive to one cell type, but their higher expression suggests a potential role in shaping its transcriptional profile.

It has been well-documented that a small number of TFs is often sufficient to establish control of the gene expression programs that define cell identity. Consequently, the identification of these core TFs has gained significant momentum, with various methods developed for such an analysis. The computational approach developed by Okawa and Del Sol (2019) has been promising, as the predicted candidate core TFs often overlap well with factors already used for cell type conversion and validated experimentally. This analysis doesn't rely on strict TF specificity but leverages the concept of combinatorial TF action and key principles of core TF expression. We believe the identified putative synergistic cores provide valuable biological insights. These cores represent groups of TFs that, even without strict specificity, collectively contribute to cell type-specific gene regulation and potentially play a role in establishing cell identity.

Nevertheless, we acknowledge this bioinformatic tool generates predictions rather than proving definitive relationships. We thus have relocated this portion of the analysis to the supplemental table and adjusted our language to reflect this accordingly (Extended Data Table 8). In addition, we emphasize that the data will be a valuable resource for researchers to use alternative approaches to validate our candidate TFs and explore additional candidates (lines 550-552, lines 554-555).

9) Their discussion of the “terminal selector” concept in describing transcription factors that may control CN cell type identities is highly speculative. In lines 324-326, they wrote: “Each IF core is characterized by the close-knit group of TFs that regulate each other (including themselves, i.e., auto-regulatory) and other downstream targets specific to this cell population, suggesting they are “terminal selector” for each CN cell type.” There are well-established criteria for whether a gene is a terminal selector. See Hobert (2016) (PMID: 26970634) for example. Just because a TF is part of a regulon and is maintained in mature neurons does not make it a candidate terminal selector. Loss-of-function studies and experimental validation of gene regulatory relationships in cell types in question would be necessary to ascertain this.

We agree with the reviewer that our previous use of the term "terminal selector" was overly speculative and did not meet the rigorous criteria established in the field. We have removed this term in the Result section (previous lines 324-326) and briefly discussed the need for further experimental validation in the Discussion section (lines 554-555).

10) Lines 530-531: “Our use of multiple bioinformatic tools identified high-confidence candidate identity TFs for each CN cell type.” I agree candidate identity TFs were identified but the basis for high confidence is not clear. Given the predictive nature of most bioinformatic tools, even when several of them point to the same working hypothesis, if they utilize the same underlying incomplete model of how certain TFs work and whether their roles are context-dependent, the likelihood of a given TF regulating neuronal identity does not go up.

We agree with the reviewer's concern about the phrasing of our statement regarding "high-confidence" candidate identity TFs. We have revised the sentence to more accurately reflect the nature of our findings and the limitations of bioinformatic predictions (lines 548-550). We have also added a sentence to the Discussion section to explicitly acknowledge the need for experimental validation (line 550-552).

11) Lines 551-552: “Some of these families are shared with other brain circuits, but many appear to be specific to auditory systems.” Please point here to data in the manuscript that supports this claim.

We apologize for the lack of support in the original text. Upon revisiting the literature and conducting a more thorough database search, we found that the most families we identified are shared across different brain circuits. We have revised this sentence and claim accordingly. Nevertheless, a few CAMs in some families seem specific to the auditory system, which are highlighted in the revised Discussion section (including the supporting literature (lines 572-575).

Reviewer #3 (Remarks to the Author):

The cochlear nuclei contain a variety of neuronal cell types that underly this initial stage of auditory processing, but definitive molecular identifiers for many of types are not known. In this study, the authors use a single-nucleus RNA-seq data set to define transcriptomic cell types, then use the Patch-seq technique to link those types to previously established types by examining their distinguishing electrophysiological and morphological properties. For some neuronal types, the transcriptomic types suggest a more granular distinction than the previously established cell types, which the authors find also correlate to further phenotypic differences.

Overall, the study is well designed and executed, and the results relating the cell types identified through different methods are clear and compelling. The authors also connect different neuronal subtypes to spatial locations in the CN as further validation of their distinctiveness. The transcriptomic characterization of these major cell types in the CN will be useful for producing molecular tools to manipulate and measure the system, and the additional classifications will assist in understanding the neuronal circuits in finer detail. I do have more questions, however, about some of the later analyses presented in the manuscript that are less clearly presented than the earlier, core components of the study.

We thank the reviewer for the supportive comments.

1. The description and presentation of the analyses in Figure 6 are much more difficult to follow than the rest of the paper. I can understand that certain gene families are more able to distinguish the identified cell types (Fig. 6a-b), and that certain transcription factors are differentially expressed across the projection neuron types (Fig. 6c), but then the figures and description of results become harder to understand. I don't know why all the neuronal types are used in Fig. 6a-b but then only the five projection types are shown in the rest of the figure (while the other types are presented in less detail in Extended Data Fig. 12). I'm not sure what it means on lines 311-312 where "Besides a few genes (e.g. Atoh7) we did not find TFs whose expression correlates strictly with each cell type" - Atoh7 expression seems quite similar to the other TFs expressed by the Bushy_Atoh7 type, and TFs for other types seem similar respective to their particular type, as well. It is not clear what the referenced SCENIC analysis shows or how that relates to the values shown in Fig. 6d (or why those particular genes are shown in the figure). The description of the "synergistic core analysis" in the Results and Methods is pretty impenetrable - is Fig. 6e just showing TF expression again (as in Fig. 6c), but now organized by some dendrogram on the left? Again, I don't understand what sentences like (lines 323-324) "Their regulons could be mostly predicted in each cell type, except for a few understudied ones (e.g., Tshz2 and Tschz3)" mean - how do Tshz2 and Tshz3 appear to be different in Fig. 6e from neighbors like Rarb and Rere? It is difficult to know what specific conclusions to take away from these analyses.

We thank the reviewer for the feedback on Figure 6 and the related text. We have carefully addressed the concerns raised and made the following revisions to enhance clarity and comprehensiveness:

1) The purpose of the prior focus on the projection neurons was to leave room to provide more detailed explanations in the text. We now have expanded Figure 6 to encompass data from both projection neurons and inhibitory neurons (revised Fig.6d, 6g), aligning it with the full scope of the paper. The corresponding text has been revised to reflect the inclusion of inhibitory neurons and to provide a more comprehensive interpretation of the findings (Lines 334-336, Lines 366-378).

2) We acknowledge that most TFs, including Atoh7, did not exhibit exclusive expression patterns in specific cell types. The text has been revised to state this observation explicitly (Lines 323-324).

3) SCENIC Analysis and Figure 6d: The SCENIC is a bioinformatics tool designed to predict the gene regulatory activity of specific transcription factors (TFs) expressed within particular cell types. However, we recognize that the accuracy of this prediction is significantly constrained by the current understanding of each TF, particularly for those understudied TFs. Our current analysis and interpretation are thus preliminary and inherently speculative, which does not yield additional biological insights. To avoid this over-speculative section clouding the rest of study, we have opted to remove this portion of the analysis (prior Figure 6d) from the manuscript. We thank the reviewer for spurring us to do so.

4) Synergistic Core Analysis and Figure 6e: The description of the synergistic core analysis in both the Results and Methods sections has been revised to enhance clarity. Initially, Figure 6e presented the expression of those TFs identified through Core Analysis, with some overlap but also differences compared to Figure 6c. To reduce redundancy, we have replaced Figure 6e with a table (Extended Data Table 8) that now lists the TFs identified by Core Analysis.

5) Tshz2 and Tshz3 are regarded as "understudied" because there is limited research on these TFs and their downstream targets, making it challenging to predict their regulatory activity using SCENIC analysis (previously mentioned in lines 323-424). We have removed this sentence from the manuscript, as the SCENIC analysis has been excluded.

Additionally, we have carefully adjusted the language throughout the manuscript to emphasize the hypothesis generating nature of the bioinformatic tools employed, ensuring that we do not overstate the conclusions drawn from these analyses. Furthermore, the portion of the analysis related to synergistic core analysis has been relocated to the supplemental table (Extended Data Table 8).

2. I also am not as clear on the utility of the spare reduced-rank regression analyses in Figure 7. The authors seem to conclude that particular subsets of genes are responsible for particular sets of electrophysiological features (e.g. Kcna1, Hcn1, and Kcnq4 on lines 389-395). But those genes don't appear to be particular well-correlated with the regression components (their vectors in Fig. 7b are fairly short), and the bi-biplots don't really illustrate the predictive power of those genes for particular features. For example, Hcn1 seems to be expressed in essentially all the cells (Fig. 7a), but to different mean expression values, which is similar to the expression of several sodium channel genes in Fig. 7a, which are said to not be determinants for electrophysiological variation in these neurons. Perhaps that is because the sodium channels vary less in the five projection neuron types than Hcn1, but the presented analysis only includes those types? (As in Figure 6, it is not stated why the analysis moves to only focus on the excitatory types).

The sRRR model was developed to illustrate the relationship between high-dimensional single-cell transcriptomes and electrophysiological properties in the bimodal Patch-seq dataset. This relationship is illustrated using bi-biplots, where the directions and magnitudes of lines (representing individual genes and electrophysiological features) reflect their correlations to two high variance-explaining components. By comparing these directions and magnitudes between the two biplots, potential associations between specific genes and electrophysiological traits can

be inferred. The key advantage of this approach is its ability to simultaneously visualize multiple candidate genes and their predictions of correlated electrophysiological features. To bolster this approach, we have validated the sRRR model using classical statistical measures (simple linear correlation and regression), including assessing the predictive power of each individual gene selected by the model (originally presented in Extended Data Fig. 13d, now in revised Extended Data Fig. 12c). We acknowledge that the biplots do not directly illustrate the predictive power of each individual gene and that our initial presentation of classical statistical measures lacked clarity. Therefore, we have thoroughly revised the text and the presentation of classical statistical measures, allowing for better interpretation of each gene's potential contribution (see revised Extended Data Figure 12d-f).

In the case of Hcn1 versus sodium channel genes, Hcn1 was selected due to its greater variability across projection neuron types, suggesting it may play a more significant role in the observed electrophysiological differences within this specific subset of neurons (i.e., projection neurons). Our subsequent classical statistical analyses verified the correlation of Hcn1 with features related to single-spiking property (see original Extended data Fig.13d, and revised Extended data Fig.12c-d), while none of sodium channel genes showed such correlations. Our current analysis primarily focuses on excitatory projection neurons due to their larger representation in our dataset. Including a wider range of neuronal types, such as inhibitory neurons, could potentially reveal contributions from sodium channels to electrophysiological properties in those cell types. However, the limited sample size of inhibitory neurons in our current dataset restricts the statistical power and the interpretability of results for this group. As we collect more data, we intend to perform a comprehensive analysis encompassing diverse neuronal populations. We have revised the relevant sections (Results, Discussion, and Methods) to clarify these points (lines 391-393, lines 602-604, lines 908-909).

3. Another issue with using the sparse RRR method to identify determinant genes is that because it is a sparse model, if two genes have similar expression patterns, it will likely choose only one (arbitrarily) because it can predict the electrophysiological features equally well with either - but both may be involved mechanistically in determining those properties.

We thank the reviewer for raising this important point regarding gene selection in sparse regression models like sRRR. It's true that when genes have highly correlated expression patterns, a model with strong sparsity (high penalty on the number of included variables) may select only one from the group. While this is a potential limitation of a sparse model, our analysis suggests it did not significantly impact our results. We systematically tested different sparsity levels (α = 0.5, 0.75, and 1) and found no major differences in gene selection across these values (see attached Figure above, also now included in the revised Extended Data Fig.12). This indicates that our model was not overly sensitive to the choice of sparsity and did not systematically exclude relevant genes due to correlated expression.

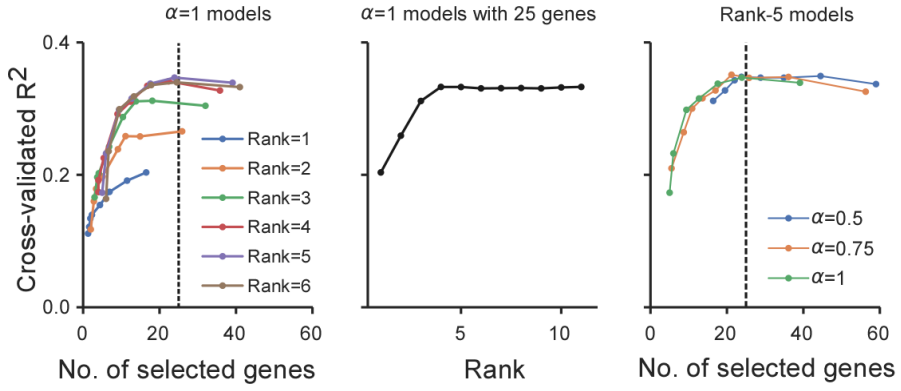


Figure. Cross-validation performance of sRRR model (cross-validated R-squared) with the different hyperparameters (alpha, rank and gene number). The vertical lines indicate the performance at 25 selected genes. With Rank=5 and alpha=1, 25 genes have near-optimal predictive accuracy and increasing more genes would not significantly enhance overall performance.

In addition, we have explored the classical statistical measure which is not prone to the limitations of the sparse model, and confirmed that our sRRR model did not overlook relevant genes due to expression similarity (original Extended data Fig.13d, and revised Extended data Fig.12c).

We have revised the Methods section (lines 909-915, lines 921-929) to provide a more detailed explanation of our gene selection process, including the evaluation of sparsity levels and comparison with alternative simple regression method.

4. Finally, I'm not sure why the model hyperparameters were selected to identify a target number of 25 genes - it seems like it would be of more use to find the most predictive model, even if more genes were identified as being relevant?

We apologize for the initial lack of clarity regarding the number of genes selected. Our primary goal was indeed to identify the most predictive model, rather than restricting ourselves to a specific number of genes. The choice of 25 genes was the result of a hyperparameter optimization process aimed at maximizing the model's predictive performance. As detailed in attached Figure above (now included in the revised Extended Data Fig.12), we used a cross-validated, systematic evaluation of a wide range of hyperparameter combinations, including those controlling the number of selected genes on the model's predictive performance (cross-validated R-squared). Our analysis revealed that a model with 25 genes achieved near-maximal predictive accuracy. While a slightly larger set of genes could potentially offer marginal improvements, we found that the added complexity did not substantially enhance the model's overall performance, and thus chose to use the smaller model.

We have revised the Methods section to provide the rationale behind our final selection of hyperparameters, including the number of selected genes (lines 909-915, Extended Data Fig.12b).

5. In general, I believe that many of the identified genes are in fact related to the electrophysiological phenotypes of these neurons (there are strong mechanistic reasons to think so), but the statements in the Results seem overly strong compared with what the analyses can really demonstrate. Despite those concerns with the latter two figures, this is a well done study with clear, novel results that help better characterize the neuronal components of auditory circuits. It demonstrates the strong utility of the multiple modality Patch-seq data set and should enable future work to more precisely target neuronal populations in the CN.

We agree that our initial wording may have overstated the definitive nature of these connections. We have thoroughly revised the Results section to ensure that our conclusions are presented more cautiously and accurately reflect the limitations of our analyses. We have replaced overly assertive language with more nuanced phrasing that acknowledges the correlative nature of our findings and the need for further investigation to establish direct causal links.

We appreciate the reviewer's constructive feedback, which has helped us refine our manuscript and ensure a more accurate and balanced presentation of our results. We believe these revisions further strengthen our study and highlight the potential for future work to build upon our findings and elucidate the precise mechanisms underlying neuronal function in auditory circuits.

Minor points:

- Fig. 7c would benefit from some additional labels to help readers more quickly identify the genes associated with the receptor types discussed in the text (AMPARs, NMDARs, kainate receptors)

We revised the figure as suggested.

- Extended Data Fig. 9 could benefit from some illustrations of specific electrophysiological features that differ between the subtypes. For example, the electrophysiology UMAP could be colored by particular electrophysiology feature values, like input resistance or spike delay.

We revised the figure as suggested (See revised figure).

References:

Ding, J., et al. Systematic comparison of single-cell and single-nucleus RNA-sequencing methods. *Nature Biotechnology*(2020). DOI:10.1038/s41587-020-0465-8

Lee, B., et al. Scaled, high fidelity electrophysiological, morphological, and transcriptomic cell characterization. *eLife*(2021). DOI:10.7554/eLife.65482

Okawa, S., del Sol, A. A general computational approach to predicting synergistic transcriptional cores that determine cell subpopulation identities. *Nucleic Acids Research*(2019). DOI:10.1093/nar/gkz147

Scala, F., et al. Phenotypic variation of transcriptomic cell types in mouse motor cortex. *Nature*(2021). DOI:10.1038/s41586-020-2907-3

Shekhar, K., et al. Comprehensive Classification of Retinal Bipolar Neurons by Single-Cell Transcriptomics. *Cell*(2016). DOI:10.1016/j.cell.2016.07.054

Tasic, B., et al. Adult mouse cortical cell taxonomy revealed by single cell transcriptomics. *Nature Neuroscience*(2016). DOI:10.1038/nn.4216

AUTHORS' RESPONSE TO THE REVIEWERS:

We sincerely appreciate the time and effort the reviewers dedicated to evaluating our manuscript.

REVIEWERS' COMMENTS

Reviewer #1 (Remarks to the Author):

No further comments

Reviewer #2 (Remarks to the Author):

I am satisfied with the authors' responses and changes + clarifications in the revised manuscript. A minor suggestion for lines 323-338 and 339-353: although based on different methods, there is significant conceptual overlap in these sections (combinatorial TF code for regulation of cell type identity, continuity of TF expression from development to maturity). They could be tied together better, either by merging or drawing a stronger logical link between them while minimizing redundancy.

We thank the reviewer for the additional suggestion and have revised the sections accordingly (lines 318-339).

Reviewer #3 (Remarks to the Author):

In their revision, Jiang et al. have clarified and improved an already strong manuscript. They have addressed my concerns with the original submission in detail: in particular, they have improved the interpretability and description of Figure 6 and provided substantial justification for the results of Figure 7 (with the revisions to Extended Figure 12). I have no remaining major concerns with the manuscript.