nature portfolio

Peer Review File

Lifespan single-cell transcriptomic atlas of the human prefrontal cortex

Corresponding Author: Dr Kiran Girdhar

Any redactions in this file are there to maintain patient confidentiality, the confidentiality of unpublished data, or to remove third-party material.

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

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

Version 0:

Reviewer comments:

Referee #1

(Remarks to the Author)
In this article, Yang et al. present a single-cell transcriptomic atlas of the human dorsolateral prefrontal cortex, generated from over 1.3 million nuclei across 284 postmortem samples, spanning birth to 97 years of age. The cohort analyzed includes the following time-points: neonatal, childhood, adolescence, and young, middle, and late adulthood. The authors identified the largest number of significant gene expression changes during the developmental period, encompassing the neonatal, childhood, and adolescent time-points. In contrast, very few differentially expressed genes were observed during young and middle adulthood, while widespread transcriptional changes were identified in late adulthood, though not to the same extent as during the developmental period. Notably, the study found that gene expression changes during development were distinct between excitatory neuron subtypes, whereas in late adulthood, a greater similarity of expression changes between these subtypes was observed. Finally, the authors uncovered significant circadian rhythm reprogramming in late adulthood.
While the study offers valuable insights and a broad dataset, there are a few concerns and areas that could benefit from further clarification and additional analysis, as outlined in the comments below.

Comments

1) Originality and significance
This study represents the largest cohort analyzed by single-cell RNA-seq with a focus on the entire human age spectrum, from birth to 97 years. As such, it provides a valuable resource for the scientific community. However, while the cohort size and age range are impressive, similar studies have been published previously, notably Ling et al. (Nature volume 627, pages604–611 (2024)), which used single-nucleus RNA sequencing to analyze the prefrontal cortex of 191 human donors aged 22–97 years, including both healthy individuals and those with schizophrenia. Although the current study surpasses Ling et al. in terms of the number of individuals analyzed and the broader age range, there is a sense that parts of the approach are not entirely novel. The current study's contribution lies in its scale and the expanded scope of the age range, which should be highlighted as key strengths.

2) Conclusions: robustness, validity, reliability
The analyses in this study are conducted with rigor; however, the majority of the conclusions are based on a single data modality - snRNA-seq. While this provides valuable insights, the robustness of the findings would be significantly strengthened by an independent validation using orthogonal approaches. For example, confirming the age-related changes in cell type composition through complementary techniques would increase confidence in the observed patterns. Additionally, validating some of the widespread transcriptional changes in late adulthood, as well as the differentially expressed genes identified along the pseudotime trajectories, would further support the reliability of the conclusions.

3) If I understood correctly, the differential expression analysis (results presented in Extended Data Fig. 3a-b) was performed within individual age groups (e.g., within the late adulthood group). While this is a valid approach, it might be useful to

explicitly clarify in the text that the analysis was conducted this way, as it could be confusing to readers who might expect a comparison between age groups.

4) The authors state: "Interestingly, down-aDEGs in subclasses such as EN_L3_5_IT_1/3, that were significantly associated with brain related traits (Extended Data Fig. 3c), had significantly higher tolerance for loss-of-function mutations compared to up-aDEGs indicating indeed stronger selection of these genes (Fig. S4c)."
The conclusion that down-aDEGs are under "stronger genetic selection" despite having higher tolerance for loss-of-function mutations compared to up-aDEGs seems counterintuitive. Typically, genes under strong selection are expected to be less tolerant to loss-of-function mutations. If down-aDEGs show higher tolerance for LoF mutations, this would suggest they are under weaker selective constraint, not stronger.
The authors should clarify whether they meant that down-aDEGs have a higher probability of intolerance to loss-of-function mutations. If these down-aDEGs indeed show higher tolerance for LoF mutations, further explanation is needed to clarify how this is congruent with these genes being under stronger selective pressure.

5) Related to the previous comment. The figure legend of Supplementary Figure 4c states: “Heatmap to show differences in loss-of-function mutation scores (pLI) of down and up-aDEGs during development: * and # indicate nominal p-value < 0.05 and FDR < 0.05 from Wilcoxon test.”
Based on the phrasing "of down and up-aDEGs," it is unclear whether the difference in pLI scores was assessed within each group or between the two groups. If the intended comparison is between the two groups, I would suggest adjusting the phrasing to explicitly state this, such as: "differences in pLI scores between down- and up-aDEGs."

6) Technical aspects: Demultiplexing validation and data quality
It appears that the authors used multiplexing for library generation and then demultiplexed the samples using vireoSNP. Given the potential for sample mixing, I suggest the authors provide some validation of their demultiplexing approach. For instance, if male and female cells were mixed, they could validate the demultiplexing procedure by examining the expression of XIST and Y chromosome-associated genes in their demultiplexed samples to ensure correct assignment.
Additionally, I did not find any information regarding the overall quality of the snRNA-seq data. It would be helpful if the authors could provide basic metrics such as the number of genes and transcripts detected per individual cell and compare these metrics to those reported in similar studies.

(Remarks on code availability)

Referee #2

(Remarks to the Author)
Yang et al. present a comprehensive single-cell transcriptomic atlas and analysis of the dorsolateral prefrontal cortex (DLPFC) in a substantial dataset of samples ranging from aged 0 to 97. The study investigates gene expression changes, cell composition, and cell identity as they change with age, through several regression-based analyses. From these analyses, they identify three phases of transcriptional changes, developmental, midlife stabilization, and late adulthood. The changes among these phases are highlighted by neuronal and glial abundance and transcriptional state changes, a plateau of development/regression from 20-59 years, and disruption of circadian rhythms and increased stress/inflammation later in life.

Overall, the dataset is powerful and unique and provides some intriguing insights into how this part of the brain changes through development and the remainder of one’s life. The potential relationship between disease risk loci and gene expression patterns during development potentially clarifies some mechanisms of these diseases and which cells are central to pathogenesis.

On the contrary, there is still much to be desired from this work. Many of the results are either bespoke regression models or are so directly corroborated through cited literature that the early figures of this work do not appear that novel. Primarily, some of the most interesting parts of the manuscript are relegated to the extended data figures or later figures. While the authors make use of some complex regression-based setups, some opportunities to provide in-depth transcriptional evaluation on such a large dataset were missed. Given the frequent use of disease ontology gene sets and the results from most of the gene set analyses, the results leave the reader with little mechanistic interpretation to take away from this work. Despite this, it is clear with some more in-depth, validated computational tools, there is an opportunity for the authors to mine much more information from this dataset and I look forward to seeing such analyses.

MAJOR COMMENTS

Experimental Design
1. There is a vast difference between the number of males and females in this dataset. Given the number of multi-variate models utilized in this work, the authors need to incorporate sex as a blocking variable in these analyses, where available. I can see this was incorporated at least once but has even more chance to be relevant in gene expression modeling.
2. Given that part of the intent of this study is to curate a sufficiently large dataset to address questions at an “atlas” level, there is no mention of statistical power in the manuscript. Considering the number of sample-level multi-variate models and consideration for 12,000+ genes, rather than meta programs, this study is likely underpowered, but the authors could provide some indication as to how underpowered (if it is) the results may be.

Data Analysis and Computation

1. Each sample only averages ~4,600 cells in this dataset. Counting cell abundances when the starting material is >10M cells is quite susceptible to random noise. Although the Crumblr modeling is “robust” as the author says, the example cell types in B do not appear to be logarithmic and the curves appear to be underfit. Considering it is widely known that brain development and cellular changes occur in the first 10-20 years of life, it would be more reasonable to uncouple the first 15 years of these models from the final 85, as you are really capturing changes according to development rather. It is clear that the curve in the later years is heavily influenced by the first 15. Likewise, all the trends can be summarized to just what cellular changes happen in the first 20 years of life anyway. It would, therefore, be more interesting for these more fine-grained populations to see what happens from year 20 to year 100. Essentially, it would be best to consider development and aging as different processes rather than putting them on the same continuum.
2. This manuscript makes use of many tools that lack justification (STARsolo, Crumblr, dreamlet, scDRS, MAGMA, Pegasus, mashr). Great care should be taken to balance unique tools for certain situations and well-established tools.
3. The time trajectory component at least does account for biological sex as a confounding variable, but there is no justification for all the variables included in equation (1). For instance, there are single-cell variables n_genes, mito_genes, mito_ribo, percent_mito, robo_genes, that are somehow accounted for in a pseudobulk measurement, but these variables will have far less effect in a pseudobulk setting. Likewise, these variables, if found to truly impact results, are most often regressed out of the data during normalization. The authors should also clearly note when results are coming from pseudobulk, in the main text. Lack thereof, especially in a single-cell paper, can give the wrong impression about whether the results are from single-cell patterns or not.
4. The authors should avoid overcorrecting for covariates as every covariate included introduces noise to the measurements. It would be interesting to know if Sex, PMI, or Source is a factor in nuclear counts or gene expression. However, if there does not appear to be an effect of these variables on the measurement (especially cell counts), then there is no need to correct it. It is suggested that the author measure the influence of these covariates first before correcting so as not to regress out a real pattern.

Biological Interpretation

1. The authors subcluster inhibitory and excitatory neurons into different groups, but it is unclear how and why this subclustering was performed. The biological rationale for distinguishing these subclusters is not adequately explained, and the manuscript lacks clarity on what these subclusters represent biologically. Providing more detail on the methods and reasoning behind the subclustering, as well as their biological implications, would improve the readers' understanding of the findings. Additionally, the author should provide a feature plot or a dot plot for genes that are specific for each subcluster. Recommendation: include the rationale and parameter selection for subclustering and the meaning of each subcluster.
2. Some findings, such as the logarithmic trend of cellular composition (Fig. 1c and S2), are insufficiently explained. The authors should provide a biological rationale for why a logarithmic model better describes the data compared to a linear one and clarify the biological significance of this trend. The authors should include a discussion on the biological relevance of these trends, particularly whether these findings align with known brain development and aging mechanisms. This interpretation would help readers link the observed patterns to real biological processes. Additionally, the authors conclude that the DLPFC follows a logarithmic trend with age and reaches maturity after undergoing proliferation, differentiation, or apoptosis. However, the current section lacks supporting data to substantiate this claim, and further evidence should be provided to justify this conclusion. For instance, the authors could show marker genes for proliferation/differentiation or apoptosis.
3. The persistent use of the disease ontology gene sets is somewhat confounding to the “healthy aging” nature of this dataset. Since the genes either go up or go down with age, the coordinated enrichment of the GWAS loci with these patterns suggests Schizophrenia (among others) is manifesting in your cohort. However, this is highly subject to the massive changes that occur among the first 20 years of life, in this dataset, so the meaning of these enrichments is puzzling. The authors should reconsider where they use these disease ontology gene set enrichments and give a clearer discussion as to their interpretation and meaning.
4. It is known that nuclei do not always represent the whole cell transcriptome. There is evidence to suggest that microglia, specifically, give differing results depending on the library preparation (cell vs. nuclei). https://www.cell.com/cell-reports/fulltext/S2211-1247(20)31178-5 Given the size of this dataset, it may be a good idea to have a few samples of single-cell RNA-Seq to corroborate the patterns found in single nuclear RNA-Seq. Understandably, this will not work for neurons, but you will still capture most of the non-neuronal populations.
5. The authors conducted a gene set pathway analysis, but they did not identify key transcription factors or regulatory networks driving the gene expression changes across life stages, especially in development and aging. Although this is somewhat captured in Figures 2 and 3, the results are mainly coming from pseudobulked analyses. This represents a missed opportunity to identify regulatory niches that dictate the normal and possibly aberrant aging programs in the PFC. TFs involved in glial and neuronal trajectories should be uncovered to supplement the work shown in Figures 4 and 5.
6. In Figure 5, the analysis primarily centers on psychiatric diseases and pathway analysis. However, a more in-depth investigation is warranted. Specifically, it would be beneficial to explore which genes are involved in neuron maturation, and which genes are significantly associated with or contribute to psychiatric diseases. Additionally, identifying the genes that regulate the branching into upper-IT, deep-IT, and deep-non-IT neurons would provide valuable insights and enhance the overall analysis.
7. A lot of the results, especially in Figure 3 are not new. As the author even cites, a lot of work to “corroborate” their results. However, some focus should be paid to the results which have not been reported previously. This makes it so that the discussion almost reads as a different study than the results.

MINOR COMMENTS

1. Why are down DEGs with age in the extended data figures but up DEGs with age are in the supplemental?
2. Much of this article is relegated to extended data figures. So much so, that the primary (or at least some of the most

interesting) results are almost entirely contained in the extended data figures. Perhaps the authors might consider re-organizing as there are some results in the extended data figures that are more interesting than some of those in the main figures.
3. The authors state they optimize the model of lifespan trajectories from 334,689 genes. Is this coming from ~12-13k genes per class? This is a strange way to word it if so. It is suggested to edit the text to emphasize the per-class comparison. There are apparently not 334,689 genes in the genome.
4. I am not sure why the model incorporates a logarithmic age, since age (as it is cataloged in this dataset) is count data and linear. Also, gene expression is not a polynomial function. It has been shown many times that non-logged normalized gene expression data is best modeled on a linear function.
5. In Fig. 1e, the classification switches from the finer age-group distinctions (e.g., neonatal, childhood, etc.) to broader classifications (development, young, middle, and late adulthood). This lack of consistency might confuse readers. Ensure consistent classification schemes throughout the manuscript or provide a clear explanation for the switch.
6. Some methods, like variancePartition and crumblr, are not explained in detail, in the main text. These methods are essential for interpreting age-related cellular composition changes but may not be familiar to all readers. Provide brief descriptions of these tools in the main text or supplement, explaining how they were applied and what insights they provide.
7. The manuscript does not specify the developmental stage from which the spatial data were prepared. Providing this information is crucial for understanding the context of the data and its relevance to the study. Please clarify the developmental stage used for the spatial data.

(Remarks on code availability)

Referee #3

(Remarks to the Author)
In this manuscript, Yang et al., studied the gene expression patterns associated with aging at single cell resolution by generating snRNA-seq libraries from the dorsolateral prefrontal cortex (DLPFC) of 284 neurotypical male and female controls spanning from infancy to late adulthood. By doing this, they were able to observe that the largest aging-dependent effect on gene expression occurred in childhood followed by late adulthood. Interestingly, neuronal differences were most prominent in the differences observed early in development, while non-neuronal cells contributed the majority of gene expression changes observed in late adulthood. Beyond performing differential expression gene (DEG) analysis across the age spectrum, the researchers also performed several secondary analyses on this generated dataset. First, they conducted trajectory analysis which revealed several non-linear gene expression trends that, to varying degrees were specific or shared to certain cell-types and life periods. They then sought to understand the degree of sharing of age-related changes and found that there was a significant degree of overlap in late adulthood of gene expression changes compared to development across cell types. Third, the researchers performed pseudotime analysis to characterize the dynamics of glial and neuronal cell types. In the glial domain, they highlight two distinct astrocyte pseudotime trajectories with protoplasmic astrocytes maturing later than fibrous astrocytes. For neuronal cells, they describe maturation of deep layer excitatory neurons being observed earlier compared to upper excitatory neurons. Lastly, leveraging the time of death the authors claim to have uncovered a disruption in the rhythmicity of circadian gene expression in late adulthood.

Overall, this work is an impressive resource to have generated, but the paper is lacking in two major ways. First, the authors do not perform validation of any of their age-related findings. While they perform spatial transcriptomic analysis of astrocytes to further work-up their astrocyte pseudotime trajectory, this finding only confirmed a fact that has long been known: that Gfap expression is high in white matter astrocytes and absent in gray matter astrocytes, and they did not leverage this novel technology to shed insights or validate their pseudotime differences. Second, there is very little if any novel biology that has been elucidated by their findings. All of the above points have already been established in the field, and in fact, cited by the authors as evidence to the external validation of their discoveries. Together these two limitations result in a lack of significant enthusiasm for this work, outside of a resource that others in the field may use to probe age-related gene expression changes in the human DLPFC.

Other points:

- Midway through the paper, the researchers choose to expand their dataset to include a published dataset from the DLPFC spanning from mid-gestation to adulthood, without a clear rationale of why this was done, and if so, why it wasn't done for all of their earlier work. Additionally, if this additional dataset spans adulthood, a powerful way to have used it would be to validate their findings in this established dataset. Do their gene expression changes observed in the dynamic developmental period recapitulate what is found in this already published work?

- There is a very skewed representation of adulthood in this dataset with nearly 10-fold greater cells analyzed in late/middle adulthood vs. childhood/neonatal periods. How does this disparity in representation affect the claims and findings of the researchers? This was not discussed at all.

(Remarks on code availability)

Referee #4

(Remarks to the Author)
In this manuscript, Yang and colleagues constructed a comprehensive single-cell transcriptomic atlas of the human dorsolateral prefrontal cortex, encompassing over 1.3 million nuclei from 284 postmortem samples across the human lifespan (0-97 years). They identified distinct phases of transcriptomic activity, revealed ten different trajectories of the transcriptome across all cell types, unveiled crucial gene clusters with dynamic expression changes linked to development, maturation, aging, and brain diseases using pseudotime analysis. One interesting finding of this study is to reveal significant reprogramming of circadian rhythms in late adulthood, marked by disrupted rhythmicity of core clock genes and the emergence of new patterns, especially in microglia and oligodendrocytes. Overall, this extensive atlas provides a crucial foundation for understanding the molecular changes occurring from development to successful aging in the human dorsolateral prefrontal cortex. However, there is a lack of experimental validation for key findings, which is essential to substantiate the conclusions. Below are the questions that the authors need to address.

Major points:
1. This comprehensive study involves a substantial sample size and an extensive number of cells, allowing for a robust analysis of the data. The intensive and thorough data analysis underscores the meticulous approach taken in this research. However, we believe that validation experiments, such as RNA scope and immunostaining, are necessary to enhance the credibility and ensure a more accurate interpretation of the data.
2. The observation of circadian reprogramming in the late adulthood is very interesting. The authors should perform a randomization analysis across time-of-day (TOD) and age groups to validate the robustness of the findings. This will help confirm that the observed patterns are not due to chance.
3. Generate a quality control (QC) matrix comparing metrics before and after filtering to demonstrate sample variation and data quality. Key metrics should include total cell count, mean reads per cell, mean genes detected per cell. This will help assess both the technical consistency across samples and identify any outliers requiring attention.
4. The authors need to explain why the shape of UMAP in Figure 1b and Extended Data Figure 1b is different.

(Remarks on code availability)

Version 1:

Reviewer comments:

Referee #1

(Remarks to the Author)
The authors have adequately addressed all of my comments.

(Remarks on code availability)

Referee #2

(Remarks to the Author)
The authors have addressed our previous comments to a satisfactory degree, and the revisions have improved the clarity and statistical robustness of the manuscript. Upon careful review, substantial edits have been made with additional robust statistical tests supporting their decisions. We appreciate that the authors made a committed effort to not just address all comments with a rebuttal but also performed new tests for most comments.
We have a few minor points that still need to be considered:
1.Regarding logarithmic age, while I appreciate the author's rebuttal, the truth of the matter is that testing multiple models and exponential functions with age modeling is still "move it until it fits". I appreciate that variables may not linearly relate to age, but polynomial fitting is likely an overfitting problem, as increasing polynomial functions have a predisposition to fit with higher power values because they "wiggle" more. I believe the trend that is presented truly exists but likely looks "cleaner" with a polynomial function as opposed to a more logically defined kernel functions (linear, exponential, or logarithmic). This is not a major issue because multiple kernel functions were tested, and the selected model is only a second-order polynomial, despite testing a third-order function as well.
2.In the figure legends, the authors report p-values, but the statistical test used to compute these values is not specified. Please clarify the test applied.
3.For Fig. 1g, showing only a representative image is not sufficient. A quantitative analysis (e.g., boxplot or violin plot) with appropriate statistical testing should be provided to support the findings.
4.In Fig. 2i, the bars representing lifespan discovery (R^2 at trajectories 5 and 10) extend to the very top of the figure panel. Please adjust the figure layout by adding some spacing above the tallest bars to improve readability.
5.In Supplementary Fig. 2d, there is a typographical error ("deveopmenr" instead of "development").

(Remarks on code availability)

Referee #3

(Remarks to the Author)
For their revised submission, Yang et al. should be commended for the significant efforts that went into responding to reviewer's critiques regarding statistical methodology and biological validation.

First, the methodology and statistical analyses of their dataset is significantly improved and more thoroughly detailed in the Methods section. The inclusion of a power analysis, and careful justification of additional datasets to support the author's claims significantly strengthened the manuscript.

Second, we also appreciate the validation of some of their findings via orthogonal methods. Using RNAscope, the authors now verify cell-type proportion changes identified in their data, specifically a decrease in SST interneurons and increase in oligodendrocytes across the lifespan. It should be noted however that age-dependent decreases in SST interneurons (Chen et al., PMID: 36841202) and age-dependent increases in oligodendrocytes (Sams, PMID: 34290887) have been well-described in the literature. Additionally, they now provide a more detailed analysis of their spatial transcriptomic data showing distinct astrocytic programs (beyond GFAP expression) that distinguish protoplasmic from fibrous astrocytes. They also mined their spatial transcriptomic data to validate additional pseudotime trajectories in microglia and oligodendrocytes.

At a high level, the paper is undoubtedly an impressive resource for the field, but remains limited in terms of the novel biological conclusions made with this rich dataset. We appreciate the detailed response with four key biological findings uncovered by this manuscript. And while each of those discoveries is supported by the paper, and at single cell resolution, experimental observations using more traditional methods for each of the points abound in the literature. For example, age-related circadian changes were first described almost 10 years ago by Chen et al. (PMID: 26699485) and these findings have been further validated in both human and mouse models. Glial programs specific to aging that are enriched for immune and unfolded protein response are the topic of a rich body of literature; for review see Latham et al. (37649972) and Wodrich et al. (35283733).

(Remarks on code availability)

Referee #4

(Remarks to the Author)
The authors have satisfactorily addressed all of my concerns in this revised manuscript, and I do not have additional technical questions at this stage. However, I remain concerned about the overall novelty and impact of the study. While the sample size is big, the analyses are carefully conducted, there are already several comprehensive datasets available for human the prefrontal cortex, and the current work does not appear to provide substantial advances beyond what has been published. As such, the incremental contribution of this study may be limited.

(Remarks on code availability)

General Comments

We thank the editor and reviewers for their thorough and constructive feedback on our manuscript. We have carefully considered each comment and have undertaken substantial revisions to address the key concerns raised. In particular, we have improved the clarity of our analytical framework, strengthened biological interpretation, and added orthogonal validation where possible. Major revisions include:

1. Enhanced validation of findings: Added orthogonal confirmation through RNAscope in situ hybridization, replication in an independent lifespan-matched snRNA-seq dataset, and spatial transcriptomics, supporting both transcriptional and nuclei compositional results respectively.

2. Clarified analytical framework: Provided explicit rationale for all computational tools, clearly distinguished pseudobulk from single-nucleus analyses, and detailed covariate usage.

3. Improved figure organization and narrative flow in manuscript: We restructured our findings from Extended Data into main figures to communicate the main message with clarity.

4. Statistical robustness: Performed theoretical and empirical power analyses to support the age associate changes in lifespan and age group analysis and performed rigorous randomization testing to validate circadian rhythmicity associated genes.

We believe these revisions have significantly enhanced the robustness of our findings. Figures labeled as responses to reviewers are not shown in the manuscript. Reviewer comments are marked in black, whereas authors answers are marked in blue. Below, we provide detailed point-by-point responses to all reviewer comments.

Referee #1

Referee #1 (Remarks to the Author):

In this article, Yang et al. present a single-cell transcriptomic atlas of the human dorsolateral prefrontal cortex, generated from over 1.3 million nuclei across 284 postmortem samples, spanning birth to 97 years of age. The cohort analyzed includes the following time-points: neonatal, childhood, adolescence, and young, middle, and late adulthood. The authors identified the largest number of significant gene expression changes during the developmental period, encompassing the neonatal, childhood, and adolescent time-points. In contrast, very few differentially expressed genes were observed during young and middle adulthood, while widespread transcriptional changes were identified in late adulthood, though not to the same extent as during the developmental period. Notably, the study found that gene expression changes during development were distinct between excitatory neuron subtypes, whereas in late adulthood, a greater similarity of expression changes between these subtypes was observed. Finally, the authors uncovered significant circadian rhythm reprogramming in late adulthood. While the study offers valuable insights and a broad dataset, there are a few concerns and areas that could benefit from further clarification and additional analysis, as outlined in the comments below.

Comments

1.1 Originality and significance

This study represents the largest cohort analyzed by single-cell RNA-seq with a focus on the entire human age spectrum, from birth to 97 years. As such, it provides a valuable resource for the scientific community. However, while the cohort size and age range are impressive, similar studies have been published previously, notably Ling et al. (Nature volume 627, pages 604–611 (2024)), which used single-nucleus RNA sequencing to analyze the prefrontal cortex of 191 human donors aged 22–97 years, including both healthy individuals and those with schizophrenia. Although the current study surpasses Ling et al. in terms of the number of individuals analyzed and the broader age range, there is a sense that parts of the approach are

not entirely novel. The current study's contribution lies in its scale and the expanded scope of the age range, which should be highlighted as key strengths.

Response: We thank the reviewer for recognizing the scale and value of our study. While Ling et al. (2024) analyzed 191 donors aged 22-97 years, including both neurotypical individuals and schizophrenia cases, our work uniquely focuses on neurotypical controls spanning the full human lifespan (0-97 years) with a substantially larger cohort (n = 284) and nuclei count (>1.3 million). By concentrating exclusively on healthy aging, we establish a normative transcriptional reference for the DLPFC across the lifespan, providing a critical foundation for future studies investigating disease-associated deviations.

In the revised manuscript, we now emphasize several key innovations and contributions:

1) Largest lifespan-resolved single-cell atlas of human DLPFC: By spanning birth to 97 years, our dataset provides the first cell-type-resolved framework of normative transcriptional trajectories across the human lifespan (**Fig. 1a**).

2) Modeling continuous transcriptomic dynamics: Beyond age-grouped snapshots, we applied non-linear modeling and uncovered 10 distinct transcriptomic trajectories across all major DLPFC cell types. These include groups of genes primarily from neuronal subclasses which are essential for both brain maturation and psychiatric vulnerability (Trajectory 2), and glial aging program marked by an immune and unfolded protein response pathway that is activated during late adulthood (Trajectory 10). Notably, this trajectory also includes genes from microglia subclass implicated in Alzheimer's disease risk, highlighting its potential relevance to neurodegenerative processes (**Fig. 2**).

3) Lifespan disease risk mapping: We integrated single-nucleus disease risk scores with pseudotime trajectories to generate the first lineage-resolved map of genetic vulnerability, linking dynamic maturation of specific cell lineages to neuropsychiatric and neurodegenerative risk (**Figs. 3-5**).

4) Spatial validation of lifespan gene programs: Spatial transcriptomics validated traDEG modules across all major lineages. This included layer-specific localization of upper- and deep-layer excitatory neuron modules, gray versus white matter enrichment of protoplasmic and fibrous astrocyte modules, and the enrichment of aged microglia and oligodendrocyte modules to white matter while their developmental counterparts localized to gray matter (**Figs. 3-5**).

5) Circadian reprogramming in late adulthood: We reveal loss of rhythmicity in core clock genes within neurons and a gain of rhythmicity in microglia and oligodendrocytes, enriched for stress-adaptive pathways. This represents the first profiling of circadian gene expression dynamics across the adult stages of the human DLPFC (**Fig. 6**).

In summary, our study provides a comprehensive, cell-type-resolved atlas of the human DLPFC across the lifespan, uncovering dynamic transcriptional transitions in neurons and glia and revealing how aging reshapes gene regulation, disease risk, and homeostatic systems such as circadian rhythms.

1.2 Conclusions: robustness, validity, reliability

The analyses in this study are conducted with rigor; however, the majority of the conclusions are based on a single data modality - snRNA-seq. While this provides valuable insights, the robustness of the findings would be significantly strengthened by an independent validation using orthogonal approaches. For example, confirming the age-related changes in cell type composition through complementary techniques would increase confidence in the observed patterns. Additionally, validating some of the widespread transcriptional changes in late adulthood, as well as the differentially expressed genes identified along the pseudotime trajectories, would further support the reliability of the conclusions.

Response: We thank the reviewer for highlighting the importance of independent validation. In response, we performed orthogonal analyses using RNAscope in situ hybridization, external snRNA-seq datasets, and spatial transcriptomics.

1) RNAscope in situ hybridization: Consistent with our cell composition estimates from snRNA-seq (**Extended Data Fig. 1a**), RNAscope in situ hybridization performed on eight human DLPFC samples spanning infancy to late adulthood (4 months to 86 years) confirmed a non-linear decline in SST+ inhibitory nuclei and increase in GPR37+ marked oligodendrocyte nuclei (**Fig. 1g** and **Extended Data Fig. 1e**).

2) Replication in an external snRNA-seq dataset: A lifespan-matched, neurotypical-only cohort (n = 303 donors; >2 million nuclei) harmonized from three recent studies[1–3] independently reproduced key findings, including developmental and late-adulthood aDEGs (**Supplementary Fig. 10)** and nonlinear trajectory-based transcriptional changes (**Fig. 2i, j** and **Supplementary Fig. 15**).

3) Spatial transcriptomics validation: Using 10x Visium data from 10 adult donors (ages 33–62 years)[4], we confirmed layer-specific localization of excitatory neuron modules (**Fig. 5d**), gray-vs-white matter enrichment of protoplasmic and fibrous astrocyte modules (**Fig. 3h**), and distinct life stage-specific spatial patterns in microglia and oligodendrocyte , with developmental modules enriched in gray matter and aging modules enriched in white matter (**Fig. 4j**), consistent with pseudotime findings and prior mouse models[5].

Together, these orthogonal validations substantially strengthen the reproducibility and biological reliability of our conclusions. We have incorporated summaries of these analyses in the revised Results and Methods.

1.3 Clarify Within-Group Differential Expression Approach

If I understood correctly, the differential expression analysis (results presented in Extended Data Fig. 3a-b) was performed within individual age groups (e.g., within the late adulthood group). While this is a valid approach, it might be useful to explicitly clarify in the text that the analysis was conducted this way, as it could be confusing to readers who might expect a comparison between age groups.

Response: We thank the reviewer for this helpful observation. To avoid confusion and improve clarity, we have revised the main text to explicitly state the design of the analysis. The updated text is as follows:

*"Using dreamlet[28] at the pseudobulk level (**Methods**), we identified three distinct transcriptional phases in the DLPFC within each age group (**Supplementary Fig. 7a-b** and **Supplemental Data 4**)"*

1.4 Clarify Interpretation of Loss-of-Function Tolerance and Selective Constraint

The authors state: "Interestingly, down-aDEGs in subclasses such as EN L3 5 IT 1/3, that were significantly associated with brain related traits (Extended Data Fig. 3c), had significantly higher tolerance for loss-of-function mutations compared to up-aDEGs indicating indeed stronger selection of these genes (Fig. S4c)." The conclusion that down-aDEGs are under "stronger genetic selection" despite having higher tolerance for loss-of-function mutations compared to up-aDEGs seems counterintuitive. Typically, genes under strong selection are expected to be less tolerant to loss-of-function mutations. If down-aDEGs show higher tolerance for LoF mutations, this would suggest they are under weaker selective constraint, not stronger. The authors should clarify whether they meant that down-aDEGs have a higher probability of intolerance to loss-of-function mutations. If these down-aDEGs indeed show higher tolerance for LoF mutations, further explanation is needed to clarify how this is congruent with these genes being under stronger selective pressure.

Response: We thank the reviewer for this important observation and for pointing out the inconsistency in our original phrasing. The reviewer is absolutely correct, genes under stronger purifying selection are expected to show lower tolerance to loss-of-function mutations, typically reflected by higher pLI scores. To clarify this for readers and accurately reflect the findings, we have carefully revised the text in the main manuscript to correct the interpretation as follows:

*"Down-aDEGs, particularly in EN_L3_5_IT_1/3, also exhibited significantly higher intolerance to loss-of-function mutations (pLI scores[33]) compared to Up-aDEGs, suggesting stronger purifying selection (**Supplementary Fig. 7d**)."*

1.5 Clarify pLI Comparison in Figure Legend

Related to the previous comment. The figure legend of Supplementary Figure 4c states: "Heatmap to show differences in loss-of-function mutation scores (pLI) of down and up-aDEGs during development: * and # indicate nominal p-value < 0.05 and FDR < 0.05 from Wilcoxon test." Based on the phrasing "of down and up-aDEGs," it is unclear whether the difference in pLI scores was assessed within each group or between the two groups. If the intended comparison is between the two groups, I would suggest adjusting the phrasing to explicitly state this, such as: "differences in pLI scores between down- and up-aDEGs."

Response: We thank the reviewer for highlighting this ambiguity. We agree that the original phrasing was unclear and have revised it for clarity. The heatmap in **Supplementary Fig. 7d** (formerly S4c) shows the mean difference in pLI scores between down-aDEGs and up-aDEGs during development for each subclass, with color indicating both the direction and magnitude of the difference. Asterisks (*) denote nominal p-values < 0.05 from Wilcoxon rank-sum tests comparing pLI distributions between down- and up-aDEGs within each subclass, while hash symbol (#) indicate FDR-corrected significance across subclasses. The revised figure legend now reads:

*"Heatmap showing the mean difference in pLI scores between down-aDEGs and up-aDEGs during development. Colors represent the direction and magnitude of mean pLI differences for each subclass. * indicates nominal p-value < 0.05 (Wilcoxon rank-sum test) and # indicates FDR-adjusted p-values < 0.05 across all subclasses."*

1.6 Technical aspects: Demultiplexing validation and data quality

It appears that the authors used multiplexing for library generation and then demultiplexed the samples using vireoSNP. Given the potential for sample mixing, I suggest the authors provide some validation of their demultiplexing approach. For instance, if male and female cells were mixed, they could validate the demultiplexing procedure by examining the expression of XIST and Y chromosome-associated genes in their demultiplexed samples to ensure correct assignment **.**

Response: We agree with the reviewer on the importance of validating our demultiplexing approach, as this step is critical for all downstream analyses. While detailed preprocessing and quality control pipeline are described in our companion data descriptor manuscript[6], we have now included a concise summary in the Methods section ("Preprocessing of snRNA-seq dataset") of the current manuscript:

"Computational Processing and Demultiplexing
Sequencing reads from all pools of multiplexed samples were aligned to the hg38 reference genome using STARsolo[82]. To assign cells from each pool to their respective donors, a genotype-based demultiplexing pipeline was employed, followed by genotype concordance checks. Specifically:

1. *Pile-up of Alleles: Using cellSNP[83], alleles were aggregated from polymorphic sites overlapping snRNA-seq reads within expressed genes (genes expressed by at least 10 cells were included). Polymorphic sites required a minimum minor allele frequency of 0.1 and a minimum aggregated UMI count of 20.*
2. *Donor Assignment: Vireo[84] was used to cluster cells into six distinct donor groups per pool based on allele pile-ups. Each cluster's identity was assigned through genotype concordance analysis, comparing cell clusters to reference genotyping data using QTLtools-mbv[85].*
3. *Baseline QC for Genotyping: Only cells meeting baseline QC thresholds (expressing at least 1,000 genes and with a mitochondrial read fraction below 5%) were included in this analysis."*

1.7 Report snRNA-seq Data Quality Metrics and Benchmarking

Additionally, I did not find any information regarding the overall quality of the snRNA-seq data. It would be helpful if the authors could provide basic metrics such as the number of genes and transcripts detected per individual cell and compare these metrics to those reported in similar studies.

Response: We thank the reviewer for highlighting the importance of reporting data quality metrics. To address this, we now include a supplementary figure (**Supplementary Fig. 2a**) summarizing key QC metrics for this study, including the number of genes and transcripts (UMIs) detected per nucleus, mitochondrial read fractions, and other filtering thresholds. These metrics demonstrate high technical quality and are comparable to or exceed benchmarks reported in similar large-scale snRNA-seq studies.

In addition, we have expanded the Methods section to provide a concise summary of our preprocessing and QC pipeline, including genotype-based demultiplexing, filtering criteria for nuclei and donors, ambient RNA correction, and doublet removal. While full details of the broader PsychAD dataset QC are provided in companion manuscripts (Lee et al., 2024; Fullard et al., 2024)[6,7], the updated Methods clarify the steps applied to this study here.

*"**Preprocessing of snRNA-seq dataset***

The DLPFC lifespan cohort analyzed in this study is a subset of the broader "PsychAD dataset," which comprises over 6 million nuclei and was defined in a separate manuscript (Lee et al., 2024, Fullard et al., 2024).[12,13] Below, we summarize the computational processing and initial quality control (QC) applied to the original PsychAD dataset.

Computational Processing and Demultiplexing

Sequencing reads from all pools of multiplexed samples were aligned to the hg38 reference genome using STARsolo[82]. To assign cells from each pool to their respective donors, a genotype-based demultiplexing pipeline was employed, followed by genotype concordance checks. Specifically:

1. *Pile-up of Alleles: Using cellSNP[83], alleles were aggregated from polymorphic sites overlapping snRNA-seq reads within expressed genes (genes expressed by at least 10 cells were included). Polymorphic sites required a minimum minor allele frequency of 0.1 and a minimum aggregated UMI count of 20.*
2. *Donor Assignment: Vireo[84] was used to cluster cells into six distinct donor groups per pool based on allele pile-ups. Each cluster's identity was assigned through genotype concordance analysis, comparing cell clusters to reference genotyping data using QTLtools-mbv[85].*
3. *Baseline QC for Genotyping: Only cells meeting baseline QC thresholds (expressing at least 1,000 genes and with a mitochondrial read fraction below 5%) were included in this analysis.*

This pipeline detected and corrected occasional sample swaps and mislabeling, ensuring accurate donor assignment for the majority of pools.

Rigorous QC Pipeline

After genome alignment and demultiplexing, downstream QC and processing were conducted using Pegasus (v1.7.0)[86] and Scanpy (v1.9.1)[87], with additional steps to ensure robust data integrity:

1. *Individual Cell QC: Cells were excluded if their UMI counts were outside the range of 1,500 ≤ UMI ≤ 110,000, gene counts were outside 1,100 ≤ genes ≤ 12,500, or if their mitochondrial content exceeded 5% (mitochondrial fraction > 0.05). Ambient RNA contamination was mitigated using CellBender[88], as well as the proportion of reads mapped to non-mRNA categories like rRNA, sRNA, and pseudogenes, in addition to examining confounding factors such as the lncRNA MALAT1. Doublets were identified and removed using Scrublet[89].*
2. *Feature-Level QC: Genes not expressed in at least 0.05% of nuclei were removed.*
3. *Donor-Level QC: Donors with fewer than 50 nuclei were excluded to avoid introducing noise into downstream analyses.*

Data Integration

To correct for (non-biological) variance, including tissue dissection biases or brain bank resources, we employed Canonical Correlation Analysis using the Harmony (v0.1)[90]. After identifying highly variable features based on mean-variance constructed a k-nearest-neighbor (kNN) graph using Harmony-corrected PCA embeddings. The Leiden algorithm was applied to cluster cells by type."

Referee #2

Referee #2 (Remarks to the Author):

Yang et al. present a comprehensive single-cell transcriptomic atlas and analysis of the dorsolateral prefrontal cortex (DLPFC) in a substantial dataset of samples ranging from aged 0 to 97. The study investigates gene expression changes, cell composition, and cell identity as they change with age, through several regression-based analyses. From these analyses, they identify three phases of transcriptional changes, developmental, midlife stabilization, and late adulthood. The changes among these phases are highlighted by neuronal and glial abundance and transcriptional state changes, a plateau of development/regression from 20-59 years, and disruption of circadian rhythms and increased stress/inflammation later in life.

Overall, the dataset is powerful and unique and provides some intriguing insights into how this part of the brain changes through development and the remainder of one's life. The potential relationship between disease risk loci and gene expression patterns during development potentially clarifies some mechanisms of these diseases and which cells are central to pathogenesis.

On the contrary, there is still much to be desired from this work. Many of the results are either bespoke regression models or are so directly corroborated through cited literature that the early figures of this work do not appear that novel. Primarily, some of the most interesting parts of the manuscript are relegated to the extended data figures or later figures. While the authors make use of some complex regression-based setups, some opportunities to provide in-depth transcriptional evaluation on such a large dataset were missed. Given the frequent use of disease ontology gene sets and the results from most of the gene set analyses, the results leave the reader with little mechanistic interpretation to take away from this work. Despite this, it is clear with some more in-depth, validated computational tools, there is an opportunity for the authors to mine much more information from this dataset and I look forward to seeing such analyses.

MAJOR COMMENTS

Experimental Design

2.1 There is a vast difference between the number of males and females in this dataset. Given the number of multi-variate models utilized in this work, the authors need to incorporate sex as a blocking variable in these analyses, where available. I can see this was incorporated at least once but has even more chance to be relevant in gene expression modeling

Response: We appreciate the reviewer's comment on the importance of accounting for sex in our analyses, especially given the imbalance in male and female donors. We confirm that sex was included as a covariate in all major multivariate analyses to mitigate potential confounding effects:

1) Nuclei composition analysis: Sex, postmortem interval (PMI), and source effects were regressed out when quantifying subclass-level nuclei dynamics across age groups and the lifespan. The proportion of variance explained by sex and other covariates is summarized in **Supplementary Fig. 5a**.

2) Pseudobulk gene expression modeling: Sex and other technical/biological covariates were incorporated using a stepwise regression framework for both age-grouped and lifespan lifespan analyses (see **Methods**). Median variance explained by sex across subclasses is shown in **Supplementary Fig. 2b**.

3) Pseudotime trajectory analysis: Sex was included as a covariate when modeling gene expression across pseudotime. As noted in the Methods:

2.2. Given that part of the intent of this study is to curate a sufficiently large dataset to address questions at an "atlas" level, there is no mention of statistical power in the manuscript. Considering the number of sample-level multi-variate models and consideration for 12,000+ genes, rather than meta programs, this study is likely underpowered, but the authors could provide some indication as to how underpowered (if it is) the results may be.

Response: We appreciate the reviewer's concern regarding statistical power, particularly given the atlas-scale scope of our study. In response, we performed comprehensive theoretical and empirical power analyses, now detailed in the revised Methods ("Statistical power analysis") and visualized in **Supplementary Fig. 8**.

For the theoretical evaluation, we estimated statistical power using Cohen's *d*, observed sample sizes per age group, and a Bonferroni-corrected significance threshold accounting for 26 subclasses and ~11,000 genes (minimum number of expressed genes per subclass across age groups and whole lifespan). Across biologically plausible moderated R^2 values (0.5–1), our calculations indicate that even with a sample size of 50, we achieve >80% power to detect moderate-to-large effect sizes ($d \geq 1.2$). Given that each of our age-defined groups includes 50–100 donors (e.g., development $n = 53$; middle adulthood $n = 95$), this analysis supports the adequacy of our study design.

In the empirical analysis, we examined how the number of aDEGs varies across subclasses and found that this variation was influenced by both biological and technical factors - such as the number of nuclei per donor and the number of expressed genes per subclass. Importantly, the robustness of key gene expression trends, particularly those observed during development and aging, was supported by replication in an independent lifespan dataset (**Supplementary Fig. 10**).

The relevant main text and methods snippet is provided below:

*"**Transcriptional remodeling of DLPFC across the lifespan***

*Next, to ensure adequate sensitivity for detecting aDEGs across subclasses, we performed both theoretical and empirical power assessments (**Supplementary Fig. 8a-g**). Power modeling indicated that our design with ≥50 donors per age group and a minimum of ~11,700 expressed genes per subclass, is sufficiently powered to detect age-associated effects across the age groups, confirming that the observed triphasic transcriptional architecture reflects true biological differences rather than technical or power-related artifacts.*

Non-linear gene expression trajectories across the lifespan of the DLPFC

*To model non-linear age-dependent expression patterns, we used subclass-level pseudobulk matrices encompassing 334,689 gene-subclass combinations (median=13,227 genes per subclass). and evaluated twelve alternative aging models. These included linear, log-transformed, and polynomial forms of age. As illustrated in **Fig. 2a**, we selected the optimal model based on median BIC across subclasses and extracted gene-wise model coefficients for downstream clustering and trajectory inference. The model using a second-degree polynomial of log-transformed age consistently provided the best fit, particularly in neuronal subclasses (**Supplementary Fig. 11a** and **Methods**). Importantly, power analysis confirmed that this model is sufficiently powered to detect non-linear age-associated effects across subclasses (**Supplementary Fig. 8h-j**).*

Methods

Statistical power analysis

To evaluate the sensitivity of our dataset to detect age-associated changes for each age group and across lifespan, we conducted both theoretical and empirical power assessments.

Theoretical power estimation

*We first estimated statistical power using Cohen's d, sample size, and a range of moderated R^2 values (0.5 to 1), which reflect the proportion of gene expression variance explained by age. Power calculations were Bonferroni-corrected for multiple testing, accounting for 26 subclasses and a minimum of ~11,700 expressed genes per subclass across age groups and whole lifespan. As shown in **Supplementary Fig. 8a,** with sample size ≥ 50, we achieve >80% power to detect effects of $d \geq 1.2$ across a range of biologically plausible R^2 values. Given that all four of our age-defined groups include 50–100 samples: development (n = 53), young adulthood (n = 54), middle adulthood (n = 95), and late adulthood (n = 82), our design is sufficiently powered to support robust group-level comparisons.*

Empirical power evaluation

*To assess empirical variation in power across cell subclasses, we considered both technical and biological factors that influence sensitivity to detect age-associated differential expression. First, we examined the relationship between the number of expressed genes and the number of donors per subclass. **Supplementary Fig. 8b, e** shows that subclasses with greater donor coverage tend to exhibit higher numbers of expressed genes, reflecting improved transcriptome detection with increased sample size, particularly in the late adulthood group. In contrast, the development group displays substantial variability across subclasses, especially among excitatory neurons, likely due to underlying biological heterogeneity during early brain development.We further evaluated how these factors impact aDEGs detection sensitivity. As shown in **Supplementary Fig. 8c**, **f** there is a positive association between the number of aDEGs and the number of expressed genes per subclass. Finally, we assessed the relationship between aDEG count and the magnitude of age-associated effect sizes. **Supplementary Fig. 8d, g** reveals an inverse relationship: low-powered subclasses (e.g., SMC, PC, VLMC) with fewer expressed genes and low aDEG counts show inflated effect sizes, consistent with the winner's curse, while well-powered subclasses such as neurons and glia exhibit smaller, more stable effect estimates.*

*Finally, we evaluated power using lifespan aDEGs (defined by F-statistics across both linear and non-linear age terms) across subclasses. **Supplementary Fig. 8h** displays the overall distribution of donor coverage and gene expression for each*

subclass using this lifespan non-linear optimal model. ***Supplementary Fig. 8i–j*** *extend our analysis by correlating lifespan aDEG counts with (f) the number of expressed genes and (g) the mean absolute age effect size."*

Data Analysis and Computation

2.3 Each sample only averages ~4,600 cells in this dataset. Counting cell abundances when the starting material is >10M cells is quite susceptible to random noise. Although the Crumblr modeling is "robust" as the author says, the example cell types in B do not appear to be logarithmic and the curves appear to be underfit. Considering it is widely known that brain development and cellular changes occur in the first 10-20 years of life, it would be more reasonable to uncouple the first 15 years of these models from the final 85, as you are really capturing changes according to development rather. It is clear that the curve in the later years is heavily influenced by the first 15. Likewise, all the trends can be summarized to just what cellular changes happen in the first 20 years of life anyway. It would, therefore, be more interesting for these more fine-grained populations to see what happens from year 20 to year 100. Essentially, it would be best to consider development and aging as different processes rather than putting them on the same continuum.

Response: We agree with the reviewer that developmental changes dominate in early life and may bias lifespan modeling. To address this, we performed a new analysis in which we systematically varied the age cutoff (20–30 years) and modeled changes in nuclei abundance separately for samples below and above each threshold (**Supplementary Fig. 5c-d**). This revealed that the majority of significant changes occur before age 23, after which age-related shifts stabilize. We selected age 24 as a representative split and visualized $\log_2$FC trends in each subclass separately during development and adulthood. In the revised Results section "Age-related changes in nuclei composition across the lifespan", we now state:

*"To further interpret this shift, we performed breakpoint analysis (**Methods**), which revealed a sharp decline in the number of subclasses showing significant changes after age 23 years, indicating a transition to relative compositional stability in DLPFC (**Supplementary Fig. 5c**). Based on this inflection point, we selected age 24 as the cutoff and modeled subclass-specific compositional changes across two groups: development (<24 years) and adulthood (≥24 years) (**Supplementary Fig. 5d**)."*

2.4. This manuscript makes use of many tools that lack justification (STARsolo, Crumblr, dreamlet, scDRS, MAGMA, Pegasus, mashr). Great care should be taken to balance unique tools for certain situations and well-established tools.

Response: We appreciate the reviewer's concern regarding the use of numerous computational tools. Given the complexity and scale of our dataset, we carefully selected each tool to address a specific analytical need—balancing well-established community standards with specialized methods optimized for particular tasks. This multifaceted approach allowed us to robustly process, model, and interpret data spanning over 1.3 million nuclei and 284 donors. To improve transparency, we now include a concise rationale for tool selection in the Methods.

STARsolo[8] is the single-cell RNA-seq mode of the STAR aligner, producing fast and accurate gene-count matrices from droplet-based data. It matches CellRanger in accuracy, handles multi-mapped reads, supports multiple protocols, and quantifies beyond gene-level expression (e.g., splicing).
Rationale: We used STARsolo[8] to process raw sequencing reads into expression matrices because it is a community-standard aligner with a proven single-cell module. This allowed us to avoid proprietary software and take advantage of STAR's alignment accuracy and speed. Given our study's scale, STARsolo's efficiency (outperforming pseudo alignment tools and matching CellRanger's accuracy) was critical. In summary, STARsolo provided a robust foundation for downstream analysis by producing high-quality quantification of transcripts per nucleus.

Crumblr[9–11] is a newly developed statistical tool for analyzing cell-type composition changes across samples using precision-weighted linear mixed models. It performs multivariate testing across the cell lineage hierarchy, improving power over single-cell-type tests. Simulations show it enhances detection of true differences while controlling false positives.
***Rationale*:** We included Crumblr to rigorously assess how cell-type proportions in the prefrontal cortex change with age. Traditional methods (e.g. per-cell-type proportion tests) could miss subtle but coordinated shifts among related cell subtypes. Developed in-house by our team, Crumblr was specifically designed for such scenarios, allowing us to leverage

its multivariate framework to find aging-associated composition changes with greater sensitivity. We cited the preprint describing Crumblr to validate its methodology.

Dreamlet[12] is an open-source R/Bioconductor package (Hoffman et al., 2023) for differential expression analysis in multi-sample single-cell datasets. It uses a pseudobulk approach with precision-weighted linear (mixed) models (https://diseaseneurogenomics.github.io/), accommodating batch effects, repeated measures, and continuous covariates. Optimized for large cohorts, it is scalable, memory-efficient, and controls false discovery rates.
***Rationale*:** Our study included many donors and technical variables, requiring a robust framework to detect age-related gene expression per cell type. Dreamlet allowed us to model donor-level variation and batch effects using precision-weighted mixed models, outperforming standard methods. Its efficiency and design-aware approach ensured statistically rigorous identification of differentially expressed genes. Dreamlet provided a *validated, efficient, and design-aware* approach to differential expression that standard single-cell DE methods (like MAST or simple pseudobulk+limma) would struggle to achieve as cleanly.

Single-cell Disease Relevance Score (scDRS)[13] links single-cell expression profiles to polygenic risk by scoring each cell for enrichment of disease-associated genes from GWAS. It identifies subpopulations with elevated expression of risk genes, independent of cell-type labels, enabling discovery of trait-relevant cells not evident from clustering alone.
***Rationale*:** We used scDRS to integrate genetic risk into our aging analysis by identifying individual cells enriched for genes implicated in age-related diseases (e.g., Alzheimer's). Unlike MAGMA, which assesses gene-level GWAS enrichment across predefined cell types, scDRS provides per-nucleus scores by weighting gene expression with GWAS-derived effect sizes. This allowed us to pinpoint disease-relevant subpopulations that may not align with standard cell-type annotations, offering finer resolution into how aging intersects with genetic risk.

MAGMA (Multi-marker Analysis of GenoMic Annotation)[14] is a widely used tool for testing whether GWAS signals are enriched in predefined gene sets or cell type–specific expression profiles. It models gene-level associations using multiple regression while accounting for linkage disequilibrium, improving power and specificity. We used MAGMA to assess whether genes differentially expressed with aging were enriched for genetic risk of age-related traits at the cell-type level. This complements scDRS, which scores individual cells, by providing population-level enrichment aligned to our DE results.

Rationale: To complement scDRS's single-cell analysis, we used MAGMA to test whether aging-related GWAS signals (e.g., for lifespan or Alzheimer's) are enriched in genes expressed by specific cell types. This allowed us to link DE gene profiles to genetic risk at the cell-type level using a well-established framework. MAGMA provided robust gene-set enrichment statistics, reinforcing the biological relevance of our findings in the context of aging and disease.

Pegasus (https://pegasus.readthedocs.io/en/stable/) is a scalable toolkit for processing and analyzing large single-cell datasets, comparable to Seurat or Scanpy. It supports key steps such as QC, normalization, dimensionality reduction, clustering, batch correction, and annotation. Available as both a command-line and Python package, Pegasus integrates with cloud workflows (e.g., Cumulus), making it well-suited for handling millions of cells efficiently.
Rationale: Our snRNA-seq dataset, spanning many donors and hundreds of thousands of nuclei, required a scalable analysis platform. Pegasus enabled efficient processing, from QC to clustering, with better speed and memory performance than other tools. Its use in a cloud workflow allowed full-dataset operations like batch correction to run smoothly. This was essential for generating high-quality clusters and embeddings in aging brain data. Pegasus also offered reproducibility and cloud-based integration, making our pipeline shareable and transparent. Overall, it was a practical and reliable choice for large-scale single-cell analysis.

mashr ("multivariate adaptive shrinkage in R")[15] is a statistical framework that improves effect size estimates by borrowing strength across multiple conditions using Empirical Bayes. Originally developed for multi-tissue eQTLs, it increases power

to detect shared and specific effects that might be missed in separate analyses. In our case, mashr modeled age-related expression across cell types, identifying genes with consistent versus cell-type-specific changes.

Rationale: We applied mashr in order to synthesize and interpret differential expression results across many cell types. After using dreamlet to find age-regulated genes within each cell type, we had a matrix of effect sizes (gene logFCs) across multiple cell populations. mashr allowed us to jointly analyze these results: it improved our ability to detect subtle age effects by leveraging shared patterns across cell types, and conversely, to pinpoint genes with truly cell-type-specific aging responses. This helped address a key biological question – do different cell types age in concert (sharing many common gene expression changes) or in distinct ways? By using mashr (a peer-reviewed empirical Bayes method), we added an extra layer of rigor: rather than just listing overlaps between DE gene sets, we formally modeled the joint distribution of effects. The result was a more nuanced understanding of aging signatures, with increased confidence in shared vs. unique gene changes. We cite mashr's foundational paper Urbut et al. (2019)[15] to underscore the reliability of this approach. In essence, mashr was chosen to ensure our multi-cell-type findings are supported by integrative statistics, maximizing power and insight in our complex dataset.

In conclusion, each tool was chosen to address specific analytical needs—from raw data processing (STARsolo, Pegasus) to detecting composition (Crumblr), expression (dreamlet), cross-cell-type patterns (mashr), and genetic risk integration (scDRS, MAGMA). This multifaceted strategy, combining established and novel methods, was essential for a robust analysis. Each tool contributed uniquely yet complementarily, enabling well-supported insights into how aging shapes the human prefrontal cortex.

2.5. The time trajectory component at least does account for biological sex as a confounding variable, but there is no justification for all the variables included in equation (1). For instance, there are single-cell variables n_genes, mito_genes, mito ribo, percent mito, robo genes, that are somehow accounted for in a pseudobulk measurement, but these variables will have far less effect in a pseudobulk setting. Likewise, these variables, if found to truly impact results, are most often regressed out of the data during normalization. The authors should also clearly note when results are coming from pseudobulk, in the main text. Lack thereof, especially in a single-cell paper, can give the wrong impression about whether the results are from single-cell patterns or not.

Response: We thank the reviewer for highlighting this point. To clarify, all nuclei abundance modeling, gene expression, and circadian analyses were performed on pseudobulk data (aggregated by donor and subclass), while pseudotime trajectory analyses used single-nucleus measurements. This distinction is now made explicit throughout the revised manuscript, including in the Methods, Results, and figure legends.

The QC-derived covariates (n genes, percent mito, mito genes, mito ribo, ribo genes) were aggregated at the donor-by-subclass level and included in the model to account for technical variation in RNA quality and library complexity. To empirically assess their relevance, we performed variance partitioning across four age groups, quantifying the median variance explained by each covariate. As shown in **Supplementary Fig. 2b** , technical covariates typically explain less than 10% of the variance across most subclasses. The correlation heatmap in **Supplementary Fig. 2c** further shows that technical covariates do not colinearize with age, supporting their use as orthogonal controls. To identify age-associated genes, we used a joint modeling approach on normalized pseudobulk counts for each subclass, applying linear mixed models using *dreamlet* tool that include age and other covariates simultaneously. This strategy preserves age-related signals and allows direct estimation of the variance explained by each covariate.

2.6. The authors should avoid overcorrecting for covariates as every covariate included introduces noise to the measurements. It would be interesting to know if Sex, PMI, or Source is a factor in nuclear counts or gene expression. However, if there does not appear to be an effect of these variables on the measurement (especially cell counts), then there is no need to correct it. It is suggested that the author measure the influence of these covariates first before correcting so as not to regress out a real pattern.

Response: We thank the reviewer for raising this important point. To avoid over correcting and to evaluate the influence of covariates such as Sex, PMI, and Source on nuclei abundance, we used the variancePartition[16] framework to systematically quantify the variance in subclass-level nuclei abundance explained by each covariate across the lifespan and within defined age groups. Among all modeled variables, age ($\log_2(\text{Age} + 1)$) consistently explained the largest proportion of variance,

particularly in developmental (median = 11.6%), Late Adulthood (median = 1.3%)(**Fig.1e)**, and across the entire lifespan (median = 8.1%) (**Supplementary Fig. 5b)**. While the variance explained by Sex was low (<1%), we retained it to control for potential sex-specific biological effects that could confound age-related trends. PMI was included due to its known impact on RNA quality and nuclear integrity in postmortem tissue as shown in **Supplementary Fig. 5a**. Source, which is absent in development (as all samples are from a single brain bank), captures cohort-level technical variability related to tissue collection, preservation, and processing in other age groups and showed median variance ranging (1-2%) across all models. Taken together, this modeling strategy balances biological interpretability with statistical rigor, ensuring that the primary signals in nuclei abundance changes are attributed to aging, especially those during development and late adulthood. To make this more clear, we have now added a new **Supplementary Fig. 5a, b** in the revised manuscript.

2.7. The authors subcluster inhibitory and excitatory neurons into different groups, but it is unclear how and why this subclustering was performed. The biological rationale for distinguishing these subclusters is not adequately explained, and the manuscript lacks clarity on what these subclusters represent biologically. Providing more detail on the methods and reasoning behind the subclustering, as well as their biological implications, would improve the readers' understanding of the findings. Additionally, the author should provide a feature plot or a dot plot for genes that are specific for each subcluster. Recommendation: include the rationale and parameter selection for subclustering and the meaning of each subcluster.

Response: We thank the Reviewer for their insightful comment. As mentioned previously, the taxonomy was derived from the broader "PsychAD dataset" as described in Lee et al. (2024)[7] and Fullard et al. (2024)[17], and we have now revised the Methods section to better clarify how the subclustering was performed and the biological rationale behind it. Specifically, subclusters (subclasses and subtypes) were defined through iterative clustering, marker validation, and label transfer using scANVI[18], with additional subclass-level validation using Xenium spatial transcriptomics. This hierarchical taxonomy (class → subclass → subtype) was adapted from previously published classification schemes[19] to capture transcriptional heterogeneity within excitatory and inhibitory neurons, reflecting their distinct molecular identities and functional specializations.

To address the reviewer's request, we now include: a feature plot of key marker genes for each subcluster and a subclass-level correlation with the taxonomy of Mathys et al. (2023) as shown in (**Supplementary Fig. 3**). We believe these additions better contextualize the biological significance of the subclusters and clarify the rationale behind their definition. Furthermore, we include additional explanation of the taxonomy annotation below:

"The taxonomy for the "PsychAD dataset" was established using a modular and iterative clustering approach, as described in Lee et al., 2024 and Fullard et al. 2024. This process involved marker validation, label transfer using scANVI[91] based on the reference dataset from Mathys et al., 2023[14], and subclass-level validation through Xenium spatial experiments. The hierarchy of taxonomy labels - class, subclass, and subtype - was derived from classifications outlined in Mathys et al., 2023[14]. For the current study, the DLPFC lifespan cohort was isolated as a subset of this annotated dataset for further analysis. The subclass level correlation of taxonomy with Mathys et al., 2023[14] as well as visualization of key gene markers is displayed in ***Supplementary Fig. 3****, respectively."*

2.8. Some findings, such as the logarithmic trend of cellular composition (Fig. 1c and S2), are insufficiently explained. The authors should provide a biological rationale for why a logarithmic model better describes the data compared to a linear one and clarify the biological significance of this trend. The authors should include a discussion on the biological relevance of these trends, particularly whether these findings align with known brain development and aging mechanisms. This interpretation would help readers link the observed patterns to real biological processes. Additionally, the authors conclude that the DLPFC follows a logarithmic trend with age and reaches maturity after undergoing proliferation, differentiation, or apoptosis. However, the current section lacks supporting data to substantiate this claim, and further evidence should be provided to justify this conclusion. For instance, the authors could show marker genes for proliferation/differentiation or apoptosis.

Response: We thank the reviewer for this insightful comment. We have clarified the rationale and biological relevance of modeling cellular composition changes using a logarithmic model in the *Results* section of the manuscript. To strengthen our interpretation, we have added new analyses, including: 1) a pathway enrichment analysis of marker genes in subclasses with increasing and decreasing trajectories, and 2) experimental validation using RNAscope to confirm subclass-specific abundance trends.

The logarithmic trajectory reflects nonlinear changes in neuronal and glial nuclei populations, followed by stabilization in young adulthood. To further explore the biological basis of these trends, we performed overrepresentation pathway analysis (ORA) to compare marker gene programs between subclasses with log-increasing versus log-decreasing trajectories (**Extended Data Fig. 1f**). Subclasses with increasing trends were enriched for pathways involved in neurogenesis, neuronal fate specification, and cell differentiation, while decreasing subclasses showed enrichment for glial apoptotic processes, inflammation-related signaling, and negative regulation of proliferation.

To validate the subclass-specific compositional trends inferred from our modeling, we employed an imaging-based approach using RNAscope in situ hybridization across a series of postmortem human DLPFC samples spanning the lifespan. We targeted marker genes representative of subclasses with divergent trajectories: SST for inhibitory neurons with a log-decreasing trend and GPR37 for oligodendrocytes with a log-increasing trend, alongside DAPI to identify total nuclei. Quantification of RNAscope signals revealed a decline in SST+ inhibitory neurons and a rise in GPR37+ oligodendrocytes across age, in agreement with the modeled logarithmic trends (**Fig. 1g** and **Extended Data Fig. 1d,e**). These results provide orthogonal experimental validation of the predicted age-associated shifts in cellular abundance and further support the biological relevance of the identified log-scale trajectories in the DLPFC. Together, these results provide a robust statistical and biological justification for using a logarithmic model to capture the nonlinear dynamics of DLPFC cellular composition across the human lifespan.

The relevant snippet from the discussion section is provided below:

"Our atlas reveals that DLPFC nuclei composition follows a global non-linear trajectory across the lifespan, with the most pronounced changes occurring during early development, a period marked by gliogenesis, synaptic remodeling, and apoptosis, followed by stabilization in adulthood. Age 24 emerged as an inflection point, after which the composition of most neuronal and glial subclasses remained stable, with the exception of IN_SST and OPCs, which continued to exhibit age-associated changes. This pattern suggests that the major wave of compositional changes is during early development, while aging induces selective and limited shifts restricted to specific subclasses. Importantly, these non-linear trends were validated by RNAscope, which confirmed a logarithmic decline in IN_SST and an increase in oligodendrocyte abundance across the lifespan."

2.9. The persistent use of the disease ontology gene sets is somewhat confounding to the "healthy aging" nature of this dataset. Since the genes either go up or go down with age, the coordinated enrichment of the GWAS loci with these patterns suggests Schizophrenia (among others) is manifesting in your cohort. However, this is highly subject to the massive changes that occur among the first 20 years of life, in this dataset, so the meaning of these enrichments is puzzling. The authors should reconsider where they use these disease ontology gene set enrichments and give a clearer discussion as to their interpretation and meaning.

Response: We appreciate the reviewer's comment. To clarify, disease ontology enrichments from MAGMA and scDRS in this study are interpreted as molecular overlap with disease-associated loci, not evidence of disease manifestation in our neurotypical cohort. Our dataset is intended as a baseline molecular atlas of the human DLPFC across the lifespan, enabling assessment of how normative age-related programs intersect with genetic risk architecture.

To address the concern that early-life changes could dominate these enrichments, we analyzed disease associations in multiple complementary ways:

1) Stratified analyses separating developmental and adulthood samples (**Fig. 1j** and **Supplementary Fig. 7g**). This revealed that developmental enrichments primarily reflect normative neurodevelopmental programs overlapping with psychiatric risk loci, while late-life enrichments—particularly in glial lineages—reflect aging-associated programs overlapping with neurodegenerative pathways.

2) MAGMA on pseudotime-inferred gene modules — Modules were stage-specific (development, maturation, aging) within each lineage (**Figs. 3f**, **4c**,**g**, and **Extended Data Fig. 5e**,**i**). Developmental modules showed enrichment for psychiatric trait loci (e.g., SCZ), while aging modules, particularly in glia, showed enrichment for neurodegenerative trait loci (e.g., AD). This design ensures that late-life enrichments are not driven by early-life transcriptional changes.

3) scDRS along continuous pseudotime trajectories — Single-cell disease relevance scores were computed for each cell along lineage-specific pseudotime (**Figs. 3d**, **4g**, **Extended Data 4b**, **5e**,**i**). This revealed enrichment peaks at both early developmental and late-life segments, confirming that late-life associations persist independently of early-life dynamics.

These complementary approaches consistently demonstrate that observed enrichments reflect stage-specific overlaps between normative age-related programs and disease-associated loci, rather than disease manifestation in our neurotypical cohort.

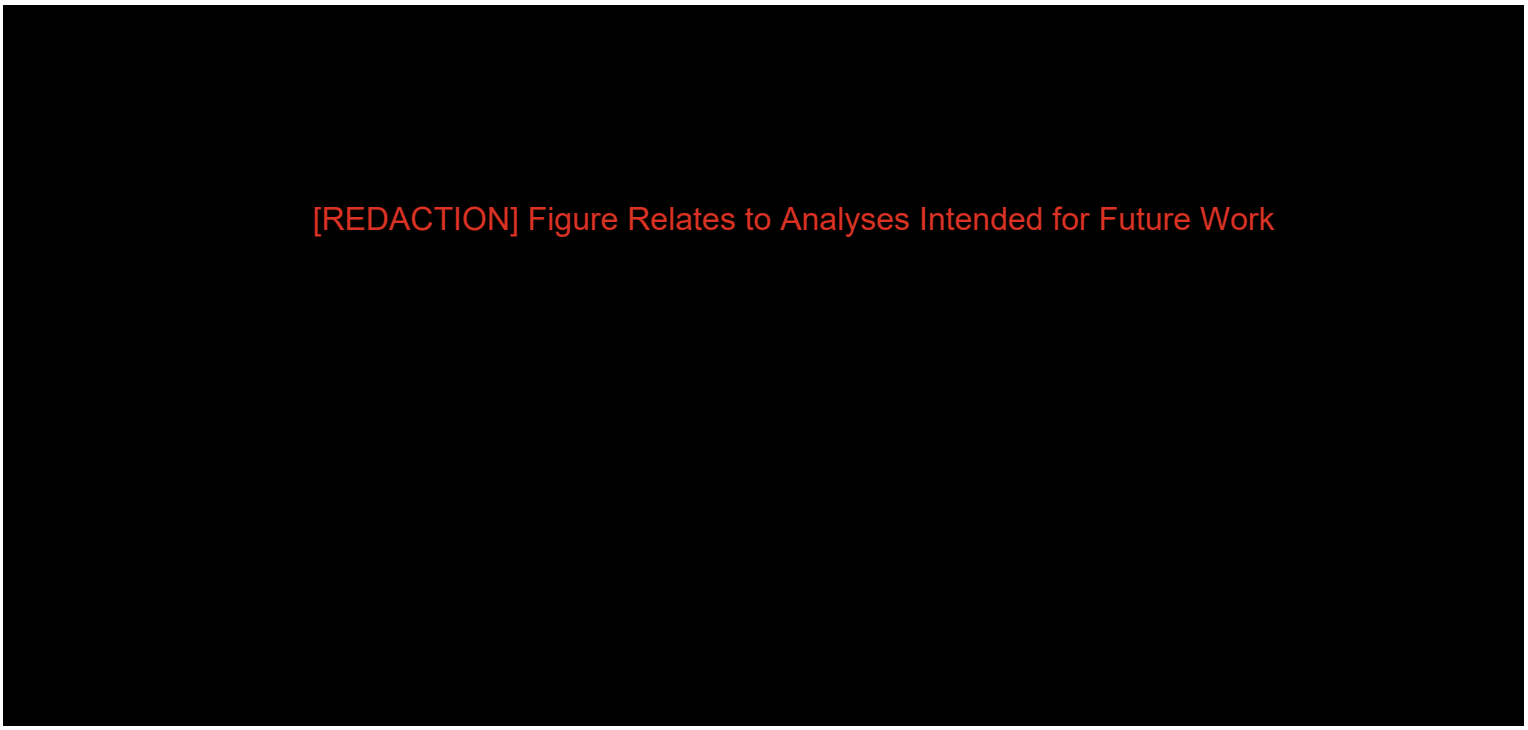
2.10. It is known that nuclei do not always represent the whole cell transcriptome. There is evidence to suggest that microglia, specifically, give different results depending on the library preparation (cell vs. nuclei). https://www.cell.com/cell-reports/fulltext/S2211-1247(20)31178-5. Given the size of this dataset, it may be a good idea to have a few samples of single-cell RNA-Seq to corroborate the patterns found in single nuclear RNA-Seq. Understandably, this will not work for neurons, but you will still capture most of the non-neuronal populations.

Response: We fully agree that microglial transcriptional profiles can be influenced by library preparation methods, and we appreciate the reviewer highlighting this important consideration. As noted in the cited study [20] there are known differences between scRNA-seq and snRNA-seq, particularly in the representation of cytoplasmic transcripts. However, several factors limited our ability to directly incorporate single-cell RNA-seq (scRNA-seq) data into the current study:

1. Availability of matched scRNA-seq data: Unfortunately, large-scale, high-quality scRNA-seq datasets from aging human PFC are currently not available. Most published human scRNA-seq studies either rely on fetal tissue, dissociation-sensitive regions, or lack sufficient age diversity. In contrast, our dataset includes deeply phenotyped postmortem donors spanning a broad adult age range, which is only feasible through the use of frozen tissue and snRNA-seq. As such, no matched or appropriately scaled scRNA-seq data currently exist for direct comparison.

2. Evidence for concordance between scRNA-seq and snRNA-seq in microglia: Recent work by Lee et al., 2024 compared aging-associated microglial profiles derived from fresh-tissue scRNA-seq and frozen-tissue snRNA-seq. Despite differences in platform and tissue preservation, the authors observed broadly consistent subtype structures and aging-related transcriptional patterns across both methods. In particular, they found positive correlations in subtype-specific log fold changes between platforms, including for adaptive and proliferative microglial states[21] (Lee et al. 2024, Fig. 5B-C).

3. Support from our own analysis: In our dataset, we observe similar structured and reproducible associations between microglial subtypes and age (**Response to reviewer Fig. 1a**). Moreover, when comparing subtype-level aging signatures between our snRNA-seq data and the fresh-tissue scRNA-seq results from Lee et al., we observe consistent directional trends and cross-platform correlations (**Response to reviewer Fig. 1b**). These observations further support the conclusion that key aging-related microglial states are robustly captured by snRNA-seq.



[redacted]

While we fully recognize the importance of evaluating platform-specific biases, the current lack of large-scale, age-matched scRNA-seq datasets from human PFC precludes a comprehensive replication of our findings at the single-cell level. As a partial validation, we performed exploratory analyses focused specifically on the myeloid compartment, comparing subtype-level aging signatures between our snRNA-seq data and available scRNA-seq datasets. These analyses revealed encouraging concordance; however, as microglial subtypes are not the focus of the main manuscript, these comparisons were not included in the primary results and are presented here in response to the reviewer's comment.

2.11. The authors conducted a gene set pathway analysis, but they did not identify key transcription factors or regulatory networks driving the gene expression changes across life stages, especially in development and aging. Although this is somewhat captured in Figures 2 and 3, the results are mainly coming from pseudobulked analyses. This represents a missed opportunity to identify regulatory niches that dictate the normal and possibly aberrant aging programs in the PFC. TFs involved in glial and neuronal trajectories should be uncovered to supplement the work shown in Figures 4 and 5.

Response: We thank the reviewer for emphasizing the importance of identifying transcription regulators driving lifespan-associated programs. To address this, we performed a systematic analysis to uncover key transcription factors (TFs) underlying each pseudotime-driven gene module (traDEG) across glial and neuronal trajectories.

We identified candidate TFs from two sources: (1) TFs directly present among module genes, and (2) additional regulators inferred using decoupleR[22] applied to meta-cell expression profiles along lineage-specific pseudotime. To contextualise these regulators, we constructed protein-protein interaction (PPI) networks using STRING-db[23], combining module genes and decoupleR-inferred TFs. We then extracted TFs present in connected components (5 nodes) with node degree ≥ 3, identifying 44 high-confidence TFs as potential hubs driving lifespan programs across cell lineages.

These TFs are summarized in **Extended Data Fig. 3d**, and the top three key TFs per module are annotated above the functional enrichment plots (**Fig. 3g**, **Fig. 4d**,**h** for glia; **Fig. 5e**, **i** for neurons). For example, the developmental excitatory neuron module (EN dev) includes a densely connected subnetwork containing ESR1, TP53, MYC, and AR (**Fig. 5e** and **Supplementary Fig. 23**), consistent with their roles in neural stem cell proliferation, excitatory neuron differentiation, and synaptic formation[24–27]. In astrocyte aging modules (**Fig. 3g** and **Extended Data Fig. 3c**), we identified STAT3 and AP-1 family members (JUN, FOS), known mediators of glial activity and neuroinflammation[28,29].

To evaluate the biological relevance of these regulators, we compared our 44 key TFs to 582 enhancer-driven regulons (eRegulons) from SCENIC+ analysis of the developing human neocortex[30]. Notably, 29 of the 44 TFs (65.91%) overlapped with eRegulon TFs, and 11 of 12 TFs assigned to developmental modules showed enriched expression in matching neocortical lineages (**Extended Data Fig. 3d**,**e**).

Together, these analyses identify key transcriptional regulators shaping normal lifespan programs in major glial and neuronal lineages of the DLPFC, with supporting data provided in updated **Fig. 3g**, **4d**,**h**, **5e**,**i, Extended Data Fig. 3c-e, 4d**, and **Supplementary Fig. 22-24**.

2.12. In Figure 5, the analysis primarily centers on psychiatric diseases and pathway analysis. However, a more in-depth investigation is warranted. Specifically, it would be beneficial to explore which genes are involved in neuron maturation, and which genes are significantly associated with or contribute to psychiatric diseases. Additionally, identifying the genes that regulate the branching into upper-IT, deep-IT, and deep-non-IT neurons would provide valuable insights and enhance the overall analysis.

Response: We thank the reviewer for this insightful comment regarding genes involved in neuron maturation, psychiatric disease risk, and regulatory programs of excitatory neuron branching. In response, we expanded our analyses to address these key points:

1. Neuron maturation: Our pathway enrichment analysis of neuronal traDEG modules highlighted biological processes relevant to neuronal development and maturation (**Supplementary Data 15)**. Notably, the Deep-layer-spec mat module is enriched for "neuron maturation" and "synapse maturation", with genes such as CAMK2B, NRXN1, NEURL1, APP, LGI4 implicated in dendritic arborization, synaptic organization, and cell-cell signaling critical for postmitotic neuronal maturation[31,32].

2. Psychiatric disease associations: To provide more granular insights, we extracted the significant genes (FDR < 0.05) from MAGMA gene-level outputs for each trait and overlapped them with pseudotime-inferred gene modules. This allows detailed exploration of gene-level contributions to trait enrichment, now provided in **Supplementary Data 16**. These gene lists complement the heatmaps shown in **Extended Data Fig. 5e**, enabling readers to directly investigate which genes within each module drive enrichment for psychiatric and other brain-related traits.

3. Regulatory control of excitatory neuron branching: To investigate transcriptional regulators driving the divergence of excitatory neuron into upper-IT, deep-IT, and deep-non-IT trajectories, we examined TF-gene regulatory networks inferred for each EN traDEG module (**Fig. 5e** and **Supplementary Fig. 23**). For example, in the Upper-layer-spec mat-aging module (aligned with the upper-IT branch), a TF-centered regulatory subnetwork involving CTNNB1, ESR1, SMAD3 and TCF7L2 was identified, consistent with roles in epithelial-to-mesenchymal transition and synaptic modulation in upper cortical layers[30,33]. In contrast, deep-layer neuron modules (Deep-layer-spec dev-mat and Deep-layer-spec mat) feature subnetworks dominated by CTNNB1, MYC, SREBF2 and NOS1, with target genes enriched for protein folding and circadian entrainment–programs consistent with deep-layer neuron specification and integration into subcortical projection pathways[34,35].

Together, these expanded analyses provide a more comprehensive view of neuron maturation, psychiatric disease associations, and the regulatory programs driving excitatory neuron diversity during lifespan.

2.13. A lot of the results, especially in Figure 3 are not new. As the author even cites, a lot of work to "corroborate" their results. However, some focus should be paid to the results which have not been reported previously. This makes it so that the discussion almost reads as a different study than the results.

Response: We thank the reviewer for this comment. While prior studies have reported transcriptomic convergence with age in bulk tissue[36] and in model systems[37], our study substantially extends this work by offering, to our knowledge, the first systematic analysis of cell subclass–specific transcriptional convergence across the human lifespan in the DLPFC using snRNA-seq. Our *mashr*[15] based framework, applied across 9 EN, 7 IN, and 4 glial subclasses and stratified by four age groups, enables us to identify patterns of gene sharing not resolvable in prior bulk or animal studies. For instance:

1) We demonstrate that transcriptional convergence is strongest in ENs, with a clear age-dependent increase in composite sharing probabilities ($P_E > 0.9$), a result not previously shown at subclass resolution in human brain tissue (**Extended Data Fig. 2b, Supplementary Fig. 18a**).

2) We link convergence to functional categories, revealing that ubiquitous, low cell-specificity genes are more likely to be shared (**Supplementary Fig. 18c**), and that these genes encode proteins involved in core aging mechanisms, such as DNA repair and immune surveillance as shown in (**Extended Data Fig. 2g-h)**.

Together, these analyses advance our understanding of age-related transcriptomic dynamics beyond previous studies by integrating lifespan stratification, cell subclass resolution, empirical Bayes modeling, and pathway-informed mechanistic insights. We have revised the results text to more clearly highlight these novel contributions and distinguish them from corroborative findings.

MINOR COMMENTS

2.14. Why are down DEGs with age in the extended data figures but up DEGs with age are in the supplemental?

Response: We thank the reviewer for pointing this out. In the revised manuscript, we have moved the analysis of genes upregulated with age into the **Fig. 1i,j**, alongside the downregulated gene results, to present both directions of change together for clarity and consistency.

2.15. Much of this article is relegated to extended data figures. So much so, that the primary (or at least some of the most interesting) results are almost entirely contained in the extended data figures. Perhaps the authors might consider re-organizing as there are some results in the extended data figures that are more interesting than some of those in the main figures.

Response: We appreciate the reviewer's feedback. In response, we carefully reviewed all Extended Data figures and identified results of highest conceptual and mechanistic interest. These have been moved into the main figures to ensure that the most impactful findings are highlighted in the primary narrative. This restructuring improves the logica flow of the Results section.

2.16. The authors state they optimize the model of lifespan trajectories from 334,689 genes. Is this coming from ~12-13k genes per class? This is a strange way to word it if so. It is suggested to edit the text to emphasize the per-class comparison. There are apparently not 334,689 genes in the genome.

Response: We appreciate the reviewer's careful reading and agree that the original wording could be misleading. The reported number of 334,689 does not refer to unique genes but to the total number of gene–subclass combinations evaluated across 26 subclasses. To improve clarity, we have revised the manuscript to read:

"To model non-linear age-dependent expression patterns, we used subclass-level pseudobulk matrices encompassing 334,689 gene-subclass combinations (median=13,227 genes per subclass). and evaluated twelve alternative aging models. "

2.17. I am not sure why the model incorporates a logarithmic age, since age (as it is cataloged in this dataset) is count data and linear. Also, gene expression is not a polynomial function. It has been shown many times that non-logged normalized gene expression data is best modeled on a linear function.

Response: We appreciate the reviewer's feedback and the opportunity to clarify our modeling rationale.

For the pseudobulk trajectories presented in Fig. 2, our objective was to determine whether non-linear models of age more accurately capture age-related gene expression dynamics compared to a purely linear model. To avoid assumptions and ensure a data-driven approach, we systematically tested twelve models for each gene–subclass combination. These included: 1) linear age, 2) log-transformed age and 3) polynomial terms of degree 1-5, applied to both linear and log-transformed age. As shown in ***Supplementary Fig. 11a***, the model incorporating poly($\log_2$(Age + 1), 2) consistently produced the lowest mean BIC values across gene–subclass pairs, particularly in neuronal subclasses, which comprise approximately 60% of the dataset. Based on this empirical evidence, we selected this model as the optimal framework to represent age-associated expression trends.

In response to the concern that gene expression is not a polynomial function, we fully agree. We do not assume gene expression itself follows a polynomial form. Rather, the polynomial term is applied to the age variable to flexibly model non-linear relationships between age and gene expression.

2.18. In Fig. 1e, the classification switches from the finer age-group distinctions (e.g., neonatal, childhood, etc.) to broader classifications (development, young, middle, and late adulthood). This lack of consistency might confuse readers. Ensure consistent classification schemes throughout the manuscript or provide a clear explanation for the switch.

Response: We appreciate the Reviewer's observation regarding the classification scheme in **Fig. 1c** (previously Fig.1e) and agree this requires further explanation. We have now clarified the rationale for this switch in the manuscript text and figure legends. Specifically, the infancy, childhood, and adolescence age groups were merged into a single "development" group in **Fig 1a** due to the limited number of nuclei and samples in these individual categories. This merging increases statistical power and allows for a more robust comparison with the young, middle, and late adulthood groups.

Here is the snippet from "A single nuclei RNA-seq atlas of the DLPFC across the human lifespan" results section

*"For downstream analyses (**Fig. 1b**), infancy, childhood, and adolescence were consolidated into a single development group to increase the nucleus counts for early stage, thus yielding four broad age groups."*

2.19. Some methods, like variancePartition and crumblr, are not explained in detail, in the main text. These methods are essential for interpreting age-related cellular composition changes but may not be familiar to all readers. Provide brief descriptions of these tools in the main text or supplement, explaining how they were applied and what insights they provide.

Response: We thank the reviewer for the useful feedback and agree that a more detailed description of the methods used for key analyses in the main text or supplement is warranted. As suggested, we modified the Results and Methods (for the case of VariancePartition[16] and other methods (See response 2.4)). Here is the snippet of text from results section:

A single nuclei RNA-seq atlas of the DLPFC across the human lifespan

*To quantify age-related transcriptional variation across subclasses and age groups, we applied a linear mixed model using the variancePartition[15] tool (**Methods**) on pseudobulk gene expression aggregated at the donor-by-subclass level. This approach estimates the relative contribution of biological covariates (e.g., age, sex) and technical covariates to expression variance for each gene across all subclasses.*

Age-related changes in nuclei composition across the lifespan

*We next applied the crumblr[27] framework (**Methods**) to model nuclei abundance changes across each age group while adjusting for sex, postmortem interval, and sample source (**Supplementary Fig. 5a**).*

2.20. The manuscript does not specify the developmental stage from which the spatial data were prepared. Providing this information is crucial for understanding the context of the data and its relevance to the study. Please clarify the developmental stage used for the spatial data.

Response: We thank the reviewer for pointing out the need to clarify the developmental stage of the spatial transcriptomics data. In the revised manuscript, we now clarify that the Visium data analyzed are from 10 donors aged 33–62 years from a previously published dataset[4], covering young, middle, and late adulthood stages. This information has been added to the Methods section.

Here is the updated text from the Methods:

"Validation of traDEG modules with spatial transcriptomics

To validate spatial expression patterns of traDEG modules, we analyzed Visium spatial transcriptomics data from 10 adult donors (ages 33-62 years; three cortical sections per donor, 30 sections total) from a publicly available dataset[53]. Data was accessed via the spatialLIBD R package (v1.19.11)[96] using fetch_data(type="spatialDLPFC_Visium")."

Referee #3

Referee #3 (Remarks to the Author):

In this manuscript, Yang et al., studied the gene expression patterns associated with aging at single cell resolution by generating snRNA-seq libraries from the dorsolateral prefrontal cortex (DLPFC) of 284 neurotypical male and female controls spanning from infancy to late adulthood. By doing this, they were able to observe that the largest aging-dependent effect on gene expression occurred in childhood followed by late adulthood. Interestingly, neuronal differences were most prominent in the differences observed early in development, while non-neuronal cells contributed the majority of gene expression changes observed in late adulthood. Beyond performing differential expression gene (DEG) analysis across the age spectrum, the researchers also performed several secondary analyses on this generated dataset. First, they conducted trajectory analysis which revealed several non-linear gene expression trends that, to varying degrees were specific or shared to certain cell-types and life periods. They then sought to understand the degree of sharing of age-related changes and found that there was a significant degree of overlap in late adulthood of gene expression changes compared to development across

cell types. Third, the researchers performed pseudotime analysis to characterize the dynamics of glial and neuronal cell types. In the glial domain, they highlight two distinct astrocyte pseudotime trajectories with protoplasmic astrocytes maturing later than fibrous astrocytes. For neuronal cells, they describe maturation of deep layer excitatory neurons being observed earlier compared to upper excitatory neurons. Lastly, leveraging the time of death the authors claim to have uncovered a disruption in the rhythmicity of circadian gene expression in late adulthood.

3.1. Overall, this work is an impressive resource to have generated, but the paper is lacking in two major ways. First, the authors do not perform validation of any of their age-related findings. While they perform spatial transcriptomic analysis of astrocytes to further work-up their astrocyte pseudotime trajectory, this finding only confirmed a fact that has long been known: that Gfap expression is high in white matter astrocytes and absent in gray matter astrocytes, and they did not leverage this novel technology to shed insights or validate their pseudotime differences.

Response: We thank the reviewer for this important comment. While GFAP localization reflects the well-known enrichment of fibrous astrocytes in white matter, our spatial transcriptomics analyses extended far beyond this single marker. Specifically, we leveraged public Visium data[4] to validate the branch-specific astrocyte pseudotime modules identified in our trajectory analysis. These modules revealed distinct transcriptional programs with preferential localization to gray or white matter (**Fig. 3h** and **Extended Data Fig. 3f**), supporting functional divergence of astrocyte subtypes across the lifespan.

Importantly, we applied this same spatial validation framework to other major lineages. In microglia and oligodendrocytes, spatial data recapitulated the predicted aging-associated transcriptional shifts from pseudotime, including the transition from gray matter enrichment for development modules to white matter localization for aging modules (**Fig. 4i**,**j** and **Extended Data Fig. 4e**). For excitatory neurons, spatial localization confirmed the layer-specific distribution of pseudotime modules: the Upper-layer-specific mat-aging module localized predominantly to superficial cortical layers, while Deep-layer-spec modules aligned with deeper cortical layers (**Fig. 5d** and **Extended Data Fig. 5d**). This spatial correspondence reinforces the pseudotime-derived branching trajectories of excitatory neuron subtypes.

These lineage- and stage-specific spatial correspondences provide orthogonal validation of the pseudotime-derived lifespan programs, strengthening both the biological interpretation and robustness of our findings. We have clarified this role of the spatial analyses in the revised Results and added further explanation in the Discussion.

3.2. Second, there is very little if any novel biology that has been elucidated by their findings. All of the above points have already been established in the field, and in fact, cited by the authors as evidence to the external validation of their discoveries. Together these two limitations result in a lack of significant enthusiasm for this work, outside of a resource that others in the field may use to probe age-related gene expression changes in the human DLPFC.

Response: We appreciate the reviewer's concern regarding the perceived lack of novel biology and the emphasis on the resource aspect of our study. While our dataset indeed provides a valuable resource, it also reveals multiple new biological insights into human DLPFC that, to our knowledge, have not been previously described at this scale or resolution.

In the revised Introduction, Results, and Discussion, we now highlight key biological findings enabled by our dataset and analysis:

1) Identification of distinct lifespan transcriptomic trajectories across all major DLPFC cell types, revealing maturation-linked neuronal programs tied to psychiatric vulnerability and late-life glial programs enriched for immune and unfolded protein response pathways relevant to neurodegeneration.

2) Lineage-resolved disease risk mapping, linking the maturation and aging of specific neuronal and glial lineages to neuropsychiatric and neurodegenerative genetic architectures for the first time at single-cell resolution.

3) Spatial validation of lifespan programs across multiple lineages, showing layer-specific localization of excitatory neuron modules, gray-vs-white matter enrichment of protoplasmic and fibrous astrocyte modules, and distinct life stage-

specific spatial patterns in microglia and oligodendrocyte, with developmental modules enriched in gray matter and aging modules enriched in white matter.

4) Discovery of circadian reprogramming in late adulthood, with loss of rhythmicity in neuronal core clock genes and gain of rhythmicity in microglia and oligodendrocytes, enriched for stress-adaptive pathways—representing the first such observation in the adult human DLPFC.

These findings go beyond corroborating prior knowledge, offering new mechanistic hypotheses and conceptual frameworks for understanding normative human DLPFC. For a detailed outline of the broader innovations in this work, please see our detailed response to Comment 1.1.

3.3. Midway through the paper, the researchers choose to expand their dataset to include a published dataset from the DLPFC spanning from mid-gestation to adulthood, without a clear rationale of why this was done, and if so, why it wasn't done for all of their earlier work. Additionally, if this additional dataset spans adulthood, a powerful way to have used it would be to validate their findings in this established dataset. Do their gene expression changes observed in the dynamic developmental period recapitulate what is found in this already published work?.

Response: We thank the reviewer for highlighting the need to clarify the rationale for integrating the fetal-to-adult DLPFC dataset.

Our primary dataset focuses on the postnatal lifespan (0 to 97 years), enabling detailed analysis of age-associated changes in cellular composition and gene expression from infancy through aging. For these analyses (e.g., nuclei composition, aDEGs, and lifespan trajectories), it was critical to maintain a uniform processing and sampling pipeline specific to postnatal donors to minimize technical variation and ensure robust age-related comparisons.

In contrast, pseudotime trajectory analyses aimed to reconstruct the full continuum of glial and neuronal lineages maturation, from progenitor stages in gestation to fully differentiated adult states. Pseudotime inference benefits from continuous cellular transitions, and the inclusion of fetal nuclei allowed more complete reconstruction of early developmental processes that are underrepresented in postnatal data alone.

To address the reviewer's suggestion about validation, we note that the fetal dataset was used specifically for trajectory completeness and not for validation of age-related findings. Instead, we performed validation using an independent external snRNA-seq dataset spanning the postnatal lifespan, harmonized from three recent studies[1–3]. This external dataset independently reproduced key findings, including developmental and late-adulthood aDEGs (**Supplementary Fig. 10**) and nonlinear trajectory-based transcriptional changes (**Fig. 2i,j** and **Supplementary Fig. 15**). Together with spatial transcriptomics validation for pseudotime-inferred modules, these analyses reinforce the robustness and biological relevance of our conclusions.

We have clarified this rationale in the revised Results:

*"Recognizing the variance observed during development and late adulthood (**Fig. 1c** and **Extended Data Fig. 2b**), we performed pseudotime trajectory analysis to delineate cellular dynamics of DLPFC lineages across the lifespan. To capture early developmental transitions, we integrated our dataset with published snRNA-seq spanning gestation to adulthood[11], resulting in a combined dataset of 1,454,617 nuclei from 311 individuals (**Supplementary Fig. 21a-d,** see **Methods**)."*

3.4. There is a very skewed representation of adulthood in this dataset with nearly 10-fold greater cells analyzed in late/middle adulthood vs. childhood/neonatal periods. How does this disparity in representation affect the claims and findings of the researchers? This was not discussed at all.

Response: We thank the reviewer for raising this important point. We agree that the imbalance in cellular representation across age groups, particularly the larger number of cells in middle and late adulthood, warrants careful evaluation. To address this, we have now conducted both theoretical and empirical statistical power analyses, as detailed in the revised Methods ("Statistical Power Analysis" and **Supplementary Fig. 8**). Please see response to **Reviewer #2 comment 2.2** for details.

Referee #4

Referee #4 (Remarks to the Author):

In this manuscript, Yang and colleagues constructed a comprehensive single-cell transcriptomic atlas of the human dorsolateral prefrontal cortex, encompassing over 1.3 million nuclei from 284 postmortem samples across the human lifespan (0-97 years). They identified distinct phases of transcriptomic activity, revealed ten different trajectories of the transcriptome across all cell types, unveiled crucial gene clusters with dynamic expression changes linked to development, maturation, aging, and brain diseases using pseudotime analysis. One interesting finding of this study is to reveal significant reprogramming of circadian rhythms in late adulthood, marked by disrupted rhythmicity of core clock genes and the emergence of new patterns, especially in microglia and oligodendrocytes. Overall, this extensive atlas provides a crucial foundation for understanding the molecular changes occurring from development to successful aging in the human dorsolateral prefrontal cortex. However, there is a lack of experimental validation for key findings, which is essential to substantiate the conclusions. Below are the questions that the authors need to address.

Major points:

4.1. This comprehensive study involves a substantial sample size and an extensive number of cells, allowing for a robust analysis of the data. The intensive and thorough data analysis underscores the meticulous approach taken in this research. However, we believe that validation experiments, such as RNA scope and immunostaining, are necessary to enhance the credibility and ensure a more accurate interpretation of the data.

Response: We thank the reviewer for recognizing the scale and depth of our analysis and for the helpful suggestion to further strengthen our findings through orthogonal validation. In response, we incorporated multiple complementary approaches, including independent external datasets, spatial transcriptomics, and RNAscope in situ hybridization, to validate both compositional shifts and transcriptional patterns across the human lifespan. For details please see response to **Reviewer #1 comment 1.2**.

4.2. The observation of circadian reprogramming in late adulthood is very interesting. The authors should perform a randomization analysis across time-of-day (TOD) and age groups to validate the robustness of the findings. This will help confirm that the observed patterns are not due to chance.

Response: We thank the Reviewer for their suggestion and have performed rhythmicity analyses on 100 permutations of our data with scrambled TOD values for both the YA+MA and LA groups. Using these permutations, we found that all genes that previously survived multiple comparison testing (**Fig. 6b**) also had an empirical p-value of $p = 0.0099$ (0 permutation R2 values greater than observed R2 value; empirical $p = (0 + 1)/(100 + 1)$) (**Supplementary Data 18**), confirming the robustness of these findings. Q-Q plots of the observed and representative permutations are supplied in the supplement (**Supplementary Fig. 25**) and confirm that the most robust rhythmicity findings were in the YA+MA ENs, LAMP5_LHX6 and SST INs, PCs, VLMCs, and Endos.

4.3. Generate a quality control (QC) matrix comparing metrics before and after filtering to demonstrate sample variation and data quality. Key metrics should include total cell count, mean reads per cell, mean genes detected per cell. This will help assess both the technical consistency across samples and identify any outliers requiring attention.

Response: We thank the Reviewer for their comment. The Aging cohort analyzed in this study is a subset of the larger "PsychAD dataset," which comprises over 6 million nuclei. This dataset is discussed in more detail in a separate manuscript currently under revision (Lee et al., 2024; Fullard et al., 2024)[7,17]. We have cited this other paper, and we describe the extensive quality control (QC) process applied to the original full dataset in that manuscript. However, we understand that this reference might not be sufficient for clarity. To address this, we have added a supplementary figure (**Supplementary Fig. 2**) in the current manuscript that outlines the QC steps for the Aging subset. Additionally, we have updated the Methods section to clarify the preprocessing steps and QC measures applied to both the original PsychAD dataset and the Aging cohort. This includes detailed information on computational processing, genotype-based demultiplexing, QC thresholds, and data integration procedures. Please see **Reviewer #1 comment 1.7** for detailed reply.

4.4. The authors need to explain why the shape of UMAP in Figure 1b and Extended Data Figure 1b is different.

Response: We appreciate the reviewer's comment and would like to clarify the difference between **Fig. 1a** and **Supplementary Fig. 1b** (formerly Extended Data Fig. 1b). As indicated in the respective legends, **Fig. 1b** shows a UMAP embedding, while **Supplementary Fig. 1b** presents a UMAT (UMAP of MATuration) embedding, as described in Herring et al. (2022)[1].

For clarity, UMAT is a specialized UMAP embedding designed to compare transition probabilities and developmental trajectories by restricting neighbor selection to cells from adjacent developmental stages. This ensures that the resulting embedding reflects maturation dynamics across stages. A detailed description of the UMAT approach is provided in Herring et al., 2022, and the relevant reference is cited in our Methods section.

Here is the snippet from main text:

"We utilized UMAP of MATuration (UMAT)[11], which restricts neighbor selection to adjacent stages, thereby enhancing the precision of representing transitional processes across the lifespan."

References

1. Herring, C. A. *et al.* Human prefrontal cortex gene regulatory dynamics from gestation to adulthood at single-cell resolution. *Cell* **185**, 4428–4447.e28 (2022).
2. Ling, E. *et al.* A concerted neuron-astrocyte program declines in ageing and schizophrenia. *Nature* **627**, 604–611 (2024).
3. Emani, P. S. *et al.* Single-cell genomics and regulatory networks for 388 human brains. *Science* **384**, eadi5199 (2024).
4. Huuki-Myers, L. A. *et al.* A data-driven single-cell and spatial transcriptomic map of the human prefrontal cortex. *Science* **384**, eadh1938 (2024).
5. Hahn, O. *et al.* Atlas of the aging mouse brain reveals white matter as vulnerable foci. *Cell* **186**, 4117–4133.e22 (2023).
6. Fullard, J. F. *et al.* Population-scale cross-disorder atlas of the human prefrontal cortex at single-cell resolution. *Sci. Data* **12**, 954 (2025).
7. Lee, D. *et al.* Single-cell atlas of transcriptomic vulnerability across multiple neurodegenerative and neuropsychiatric diseases.
8. Kaminow, B., Yunusov, D. & Dobin, A. STARsolo: accurate, fast and versatile mapping/quantification of single-cell and single-nucleus RNA-seq data. *bioRxiv* 2021.05.05.442755 (2021) doi:10.1101/2021.05.05.442755.
9. Hoffman, G. E. GabrielHoffman/crumblr: to zenodo. Zenodo (2024); https://doi.org/10.5281/zenodo.12752107.
10. Hoffman, G. E. & Roussos, P. Fast, flexible analysis of differences in cellular composition with crumblr. *bioRxivorg* (2025) doi:10.1101/2025.01.29.635498.
11. Hoffman, G. *Crumblr: Count Ratio Uncertainty Modeling Base Linear Regression*. (Github).
12. Hoffman, G. E. *et al.* Efficient differential expression analysis of large-scale single cell transcriptomics data using dreamlet. *bioRxiv* 2023.03.17.533005 (2023) doi:10.1101/2023.03.17.533005.
13. Zhang, M. J. *et al.* Polygenic enrichment distinguishes disease associations of individual cells in single-cell RNA-seq data. *Nat. Genet.* **54**, 1572–1580 (2022).
14. de Leeuw, C. A., Mooij, J. M., Heskes, T. & Posthuma, D. MAGMA: generalized gene-set analysis of GWAS data. *PLoS Comput. Biol.* **11**, e1004219 (2015).
15. Urbut, S. M., Wang, G., Carbonetto, P. & Stephens, M. Flexible statistical methods for estimating and testing effects

in genomic studies with multiple conditions. *Nat. Genet.* **51**, 187–195 (2019).

16. Hoffman, G. E. & Schadt, E. E. variancePartition: interpreting drivers of variation in complex gene expression studies. *BMC Bioinformatics* **17**, 483 (2016).
17. Fullard, J. F. *et al.* Population-scale cross-disorder atlas of the human prefrontal cortex at single-cell resolution.
18. Xu, C. *et al.* Probabilistic harmonization and annotation of single-cell transcriptomics data with deep generative models. *Mol. Syst. Biol.* **17**, e9620 (2021).
19. Mathys, H. *et al.* Single-cell atlas reveals correlates of high cognitive function, dementia, and resilience to Alzheimer's disease pathology. *Cell* **186**, 4365–4385.e27 (2023).
20. Thrupp, N. *et al.* Single-nucleus RNA-seq is not suitable for detection of microglial activation genes in humans. *Cell Rep.* **32**, 108189 (2020).
21. Lee, D. *et al.* Plasticity of human microglia and brain perivascular macrophages in aging and Alzheimer's disease. *medRxiv* 2023.10.25.23297558 (2024) doi:10.1101/2023.10.25.23297558.
22. Badia-I-Mompel, P. *et al.* decoupleR: ensemble of computational methods to infer biological activities from omics data. *Bioinform. Adv.* **2**, vbac016 (2022).
23. Szklarczyk, D. *et al.* The STRING database in 2025: protein networks with directionality of regulation. *Nucleic Acids Res.* **53**, D730–D737 (2025).
24. Gegenhuber, B., Wu, M. V., Bronstein, R. & Tollkuhn, J. Gene regulation by gonadal hormone receptors underlies brain sex differences. *Nature* **606**, 153–159 (2022).
25. Marin Navarro, A. *et al.* P53 controls genomic stability and temporal differentiation of human neural stem cells and affects neural organization in human brain organoids. *Cell Death Dis.* **11**, 52 (2020).
26. Zinin, N. *et al.* MYC proteins promote neuronal differentiation by controlling the mode of progenitor cell division. *EMBO Rep.* **15**, 383–391 (2014).
27. La Rosa, P. *et al.* Androgen Receptor signaling promotes the neural progenitor cell pool in the developing cortex. *J. Neurochem.* **157**, 1153–1166 (2021).
28. Sarmiento Soto, M. *et al.* STAT3-mediated astrocyte reactivity associated with brain metastasis contributes to neurovascular dysfunction. *Cancer Res.* **80**, 5642–5655 (2020).
29. Gadea, A., Schinelli, S. & Gallo, V. Endothelin-1 regulates astrocyte proliferation and reactive gliosis via a JNK/c-Jun signaling pathway. *J. Neurosci.* **28**, 2394–2408 (2008).

30. Wang, L. *et al.* Molecular and cellular dynamics of the developing human neocortex. *Nature* 1–10 (2025).

31. Kennedy, M. B. Signal-processing machines at the postsynaptic density. *Science* **290**, 750–754 (2000).

32. Müller, U. C., Deller, T. & Korte, M. Not just amyloid: physiological functions of the amyloid precursor protein family. *Nat. Rev. Neurosci.* **18**, 281–298 (2017).

33. Wang, J. *et al.* A Ctnnb1 enhancer regulates neocortical neurogenesis by controlling the abundance of intermediate progenitors. *Cell Discov.* **8**, 74 (2022).

34. Mutch, C. A., Funatsu, N., Monuki, E. S. & Chenn, A. Beta-catenin signaling levels in progenitors influence the laminar cell fates of projection neurons. *J. Neurosci.* **29**, 13710–13719 (2009).

35. Kuwahara, A. *et al.* Wnt signaling and its downstream target N-myc regulate basal progenitors in the developing neocortex. *Development* **137**, 1035–1044 (2010).

36. Izgi, H. *et al.* Inter-tissue convergence of gene expression during ageing suggests age-related loss of tissue and cellular identity. *Elife* **11**, e68048 (2022).

37. Jin, K. *et al.* Brain-wide cell-type-specific transcriptomic signatures of healthy ageing in mice. *Nature* **638**, 182–196 (2025).

Here we provide a detailed, point-by-point response to all remaining comments raised during the review process. We have carefully addressed each concern and made substantial revisions to strengthen the manuscript accordingly.

Referee #1 (Remarks to the Author):

1.1 The authors have adequately addressed all of my comments.

Response: We thank the reviewer for their thorough review and are pleased that the revisions adequately address all their comments. Their feedback was valuable in improving the manuscript.

Referee #2 (Remarks to the Author): The authors have addressed our previous comments to a satisfactory degree, and the revisions have improved the clarity and statistical robustness of the manuscript. Upon careful review, substantial edits have been made with additional robust statistical tests supporting their decisions. We appreciate that the authors made a committed effort to not just address all comments with a rebuttal but also performed new tests for most comments. We have a few minor points that still need to be considered:

2.1.Regarding logarithmic age, while I appreciate the author's rebuttal, the truth of the matter is that testing multiple models and exponential functions with age modeling is still "move it until it fits". I appreciate that variables may not linearly relate to age, but polynomial fitting is likely an overfitting problem, as increasing polynomial functions have a predisposition to fit with higher power values because they "wiggle" more. I believe the trend that is presented truly exists but likely looks "cleaner" with a polynomial function as opposed to a more logically defined kernel functions (linear, exponential, or logarithmic). This is not a major issue because multiple kernel functions were tested, and the selected model is only a second-order polynomial, despite testing a third-order function as well.

Response: We thank the reviewer for raising an important concern regarding potential overfitting in age modeling. Our model selection strategy was to optimize fit within individual subclasses, followed by a single parsimonious model that generalizes across diverse subclasses exhibiting distinct age-related trajectories. In practice, BIC values varied substantially across subclasses. While simpler functional forms (e.g., log-transformed age) provided comparable fits for some subclasses (e.g., excitatory neurons and glia), they performed poorly for inhibitory neuron subclasses, for which a second-order polynomial of log2(age+1) consistently achieved substantially lower BIC values. Importantly, rather than selecting different models for each subclass, we adopted a single second-order polynomial model that provides robust performance across all subclasses. Lastly, our cross-dataset validation (**Fig. 2i** and **Supplementary Fig. 15**) shows strong concordance of aging trajectories between discovery and replication datasets, confirming robustness. We acknowledge that simpler models (log, exponential, linear) may be valid alternatives that capture the dominant trends. However, the heterogeneity in optimal fits across subclasses justifies the additional complexity of the polynomial model.

2.2 In the figure legends, the authors report p-values, but the statistical test used to compute these values is not specified. Please clarify the test applied.

Response: We appreciate this comment. All figure legends have been revised to explicitly state the statistical tests used to compute p-values.

2.3. For Fig. 1g, showing only a representative image is not sufficient. A quantitative analysis (e.g., boxplot or violin plot) with appropriate statistical testing should be provided to support the findings.

Response: We appreciate the reviewer's request for quantitative validation of the RNAscope findings. Figure 1d presents a representative RNAscope image for illustrative purposes, whereas the quantitative

analysis with statistical testing is shown in **Extended Data Figure 1e**. Specifically, the figure displays the fraction of SST+ and GPR37+ nuclei across age groups measured by RNAscope, based on 3 regions of interest per sample from 8 biologically independent donors spanning development to late adulthood (developmental: 0.33, 0.75, and 9 years; young adult: 38 years; middle adult: 51 and 53 years; late adult: 65 and 86 years). For IN_SST, we detected a significant age-related decline ($p = 0.0030689$) with strong effect size. For oligodendrocytes, we detected a significant age-related increase ($p = 0.01569$). These quantitative findings directly support and extend the RNAscope validation shown in **Figure 1g**.

2.4 Fig. 2i, the bars representing lifespan discovery (R^2 at trajectories 5 and 10) extend to the very top of the figure panel. Please adjust the figure layout by adding some spacing above the tallest bars to improve readability.

Response: We thank the reviewer for this helpful suggestion regarding figure readability. We have adjusted **Figure 2i** to add vertical spacing above R^2 bars for Trajectories 5 and 10.

2.5 In Supplementary Fig. 2d, there is a typographical error ("deveopmenr" instead of "development").

Response: We thank the reviewer for catching this typographical error. We have corrected "deveopmenr" to "development" in Supplementary Figure 2d.

Referee #3 (Remarks to the Author): For their revised submission, Yang et al. should be commended for the significant efforts that went into responding to reviewer's critiques regarding statistical methodology and biological validation.

3.1 First, the methodology and statistical analyses of their dataset is significantly improved and more thoroughly detailed in the Methods section. The inclusion of a power analysis, and careful justification of additional datasets to support the author's claims significantly strengthened the manuscript.Second, we also appreciate the validation of some of their findings via orthogonal methods. Using RNAscope, the authors now verify cell-type proportion changes identified in their data, specifically a decrease in SST interneurons and increase in oligodendrocytes across the lifespan. It should be noted however that age-dependent decreases in SST interneurons (Chen et al., PMID: 36841202) and age-dependent increases in oligodendrocytes (Sams, PMID: 34290887) have been well-described in the literature. Additionally, they now provide a more detailed analysis of their spatial transcriptomic data showing distinct astrocytic programs (beyond GFAP expression) that distinguish protoplasmic from fibrous astrocytes. They also mined their spatial transcriptomic data to validate additional pseudotime trajectories in microglia and oligodendrocytes. At a high level, the paper is undoubtedly an impressive resource for the field, but remains limited in terms of the novel biological conclusions made with this rich dataset. We appreciate the detailed response with four key biological findings uncovered by this manuscript. And while each of those discoveries is supported by the paper, and at single cell resolution, experimental observations using more traditional methods for each of the points abound in the literature. For example, age-related circadian changes were first described almost 10 years ago by Chen et al. (PMID: 26699485) and these findings have been further validated in both human and mouse models. Glial programs specific to aging that are enriched for immune and unfolded protein response are the topic of a rich body of literature; for review see Latham et al. (37649972) and Wodrich et al. (35283733).

Response: We thank the reviewer for their careful and thoughtful evaluation of the revised manuscript and for recognizing the substantial improvements in statistical methodology, analytical rigor, and orthogonal validation. We also appreciate the acknowledgment that the dataset constitutes an impressive and potentially valuable resource for the field.

We agree that several of the biological phenomena highlighted in our study, such as age-associated decreases in SST interneurons and increases in oligodendrocytes, have been reported previously using bulk transcriptomics, histological approaches, or targeted experimental designs (e.g., Chen et al., PMID: 36841202; Sams et al., PMID: 34290887). Our intent is not to claim these observations as entirely novel

per se, but rather to recontextualize them within a unified, lifespan-resolved single-cell framework that enables precise subclass-level quantification, trajectory inference, and integration with transcriptional state changes across all major cortical lineages within the same human brain region.

Similarly, prior studies have documented age-related circadian changes and glial programs enriched for immune and unfolded protein response pathways (e.g., Chen et al., PMID: 26699485; Latham et al., PMID: 37649972; Wodrich et al., PMID: 35283733). What distinguishes our work is that these processes are simultaneously modeled across the entire human lifespan at single-cell resolution, allowing us to define when such programs emerge, which cell subclasses are most affected, and how these processes relate to one another temporally within a consistent analytical framework. These relationships cannot be readily inferred from bulk tissue analyses, animal models, or studies focused on narrow developmental or aging windows. For circadian reprogramming specifically, a major question in the field is whether what we see at the bulk transcriptomic level during aging is due to individual cells undergoing circadian reprogramming or different populations of cells becoming rhythmic. This study suggests a combination of both processes is occurring in the DLPFC, depending on each individual cell type- indicating that future research will need to take these differences into account.

Importantly, the single-nucleus and lifespan scope of our data reveal features that extend beyond prior work, including non-linear age trajectories that differ markedly across subclasses, late-life transcriptional convergence among excitatory neuronal populations, and cell-type-specific coupling between circadian dysregulation and immune or stress-response programs in glia. These findings are not accessible through traditional experimental approaches and require both the scale and temporal continuity of the dataset presented here.

With respect to spatial transcriptomic analyses, while astrocytic heterogeneity and aging-associated glial states have been previously discussed, our study integrates spatial profiling with pseudotime and single-nucleus trajectories to distinguish protoplasmic and fibrous astrocytic programs beyond canonical GFAP-based definitions, and to validate lineage-specific aging trajectories in microglia and oligodendrocytes directly within tissue context. This integrative validation strengthens the biological interpretation of inferred trajectories and anchors them anatomically.

Taken together, while the individual components of the biological processes described here have antecedents in the literature, our study, to our knowledge, provides the first normative, cell-type-resolved, lifespan-scale framework for human cortical aging. Rather than presenting isolated observations, the manuscript defines how canonical cortical cell subclasses change, stabilize, converge, or reprogram across development, midlife, and late adulthood. We view this integrative and temporal synthesis as the primary conceptual advance of the work and as a foundation for future mechanistic and disease-focused investigations.

Referee #4 (Remarks to the Author):

4.1 The authors have satisfactorily addressed all of my concerns in this revised manuscript, and I do not have additional technical questions at this stage. However, I remain concerned about the overall novelty and impact of the study. While the sample size is big, the analyses are carefully conducted, there are already several comprehensive datasets available for human the prefrontal cortex, and the current work does not appear to provide substantial advances beyond what has been published. As such, the incremental contribution of this study may be limited.

Response: We thank the reviewer for their careful evaluation of the revised manuscript and for acknowledging that the technical concerns have been satisfactorily addressed. We appreciate the thoughtful perspective regarding novelty and impact. While several human prefrontal cortex datasets (PMID: 37824647, PMID: 37192616, PMID: 33723434, PMID: 36318921) are indeed available, the primary contribution of our study lies not in the generation of another static reference dataset, but in the systematic

modeling of age-associated transcriptomic dynamics across the human lifespan at single-cell resolution, using a harmonized and uniformly annotated cohort spanning prenatal development through late adulthood. Importantly, prior studies are largely development-centric or treat adulthood as a terminal state, whereas our work explicitly models aging as a distinct biological phase rather than a continuation of development. To our knowledge, no existing DLPFC resource integrates this breadth of age coverage with consistent single-nucleus processing, unified subclass annotation, and explicit modeling of non-linear age trajectories across all major subclasses (**Fig 1a**).

Importantly, our analyses reveal subclass-specific patterns that are not apparent in prior datasets, including differential timing of developmental versus aging-related transcriptional changes, subclass-specific non-linear trajectories, and a relative resilience of neuronal transcriptional programs contrasted with pronounced late-life vulnerability in glial populations (**Fig 2,3,4,5**). Notably, we observe transcriptional convergence among excitatory neuronal subclasses in late life, accompanied by coordinated activation of maintenance and stress-response pathways, a phenomenon that cannot be inferred from datasets focused primarily on developmental divergence (**Extended Data Fig. 2**).

In addition, the lifespan scope of this dataset enables the systematic characterization of circadian gene expression programs across subclasses, revealing age-associated disruptions in clock gene rhythmicity that are not readily discernible in datasets with narrower age ranges (**Fig. 6)**. By focusing exclusively on neurotypical individuals across the lifespan, our study provides a normative reference for human cortical aging, establishing a baseline that will be critical for disentangling disease-associated signals from normal age-related changes in future studies. These insights, together with the scale and functional breadth of the analyses, provide a complementary and conceptually distinct contribution to existing resources.