Received: 12 September 2022 | Revised: 3 January 2023 | Accepted: 6 January 2023 | Published online: 7 February 2023

DOI: 10.1002/trc2.12373

Alzheimer's & Dementia
Translational Research & Clinical Interventions

RESEARCH ARTICLE

Altered gene expression in excitatory neurons is associated with Alzheimer's disease and its higher incidence in women

A. Xavier Garcia[1,5] | Jielin Xu[2] | Feixiong Cheng[2,3,4] | Eytan Ruppin[1] | Alejandro A. Schäffer[1]

[1]Cancer Data Science Laboratory, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, Maryland, USA

[2]Genomic Medicine Institute, Lerner Research Institute, Cleveland Clinic, Cleveland, Ohio, USA

[3]Department of Molecular Medicine, Cleveland Clinic Lerner College of Medicine, Case Western Reserve University, Cleveland, Ohio, USA

[4]Case Comprehensive Cancer Center, Case Western Reserve University School of Medicine, Cleveland, Ohio, USA

[5]Weill Cornell/Rockefeller/Sloan Kettering Tri-Institutional MD-PhD Program, New York, NY, USA

Correspondence
Alejandro A. Schäffer or Eytan Ruppin, Building 15-C1, CDSL/NCI/NIH, Bethesda, MD 20892, USA.
Email: alejandro.schaffer@nih.gov or eytan.ruppin@nih.gov

Present address
A. Xavier Garcia, Weill Cornell/Rockefeller/Sloan Kettering Tri-Institutional MD-PhD Program, New York, NY, USA

Abstract

Introduction: Alzheimer's disease (AD) is a neurodegenerative disorder involving interactions between different cell types in the brain. Previous single-cell and bulk expression Alzheimer's studies have reported conflicting findings about the key cell types and cellular pathways whose expression is primarily altered in this disease. We re-analyzed these data in a uniform, coherent manner aiming to resolve and extend past findings. Our analysis sheds light on the observation that females have higher AD incidence than males.

Methods: We re-analyzed three single-cell transcriptomics datasets. We used the software Model-based Analysis of Single-cell Transcriptomics (MAST) to seek differentially expressed genes comparing AD cases to matched controls for both sexes together and each sex separately. We used the GOrilla software to search for enriched pathways among the differentially expressed genes. Motivated by the male/female difference in incidence, we studied genes on the X-chromosome, focusing on genes in the pseudoautosomal region (PAR) and on genes that are heterogeneous across individuals or tissues for X-inactivation. We validated findings by analyzing bulk AD datasets from the cortex in the Gene Expression Omnibus.

Results: Our results resolve a contradiction in the literature, showing that by comparing AD patients to unaffected controls, excitatory neurons have more differentially expressed genes than do other cell types. Synaptic transmission and related pathways are altered in a sex-specific analysis of excitatory neurons. PAR genes and X-chromosome heterogeneous genes, including, for example, *BEX1 and ELK1*, may contribute to the difference in sex incidence of Alzheimer's disease. *GRIN1*, stood out as an overexpressed autosomal gene in cases versus controls in all three single-cell datasets and as a functional candidate gene contributing to pathways upregulated in cases.

Discussion: Taken together, these results point to a potential linkage between two longstanding questions concerning AD pathogenesis, involving which cell type is the most important and why females have a higher incidence than males.

KEYWORDS
Alzheimer's disease, differential gene expression, excitatory neurons, sex difference, single-cell transcriptomics

Highlights

- By reanalyzing three, published, single-cell RNAseq datasets, we resolved a contradiction in the literature and showed that when comparing AD patients to unaffected controls, excitatory neurons have more differentially expressed genes than do other cell types.
- Further analysis of the published single-cell datasets showed that synaptic transmission and related pathways are altered in a sex-specific analysis of excitatory neurons.
- Combining analysis of single-cell datasets and publicly available bulk transcriptomics datasets revealed that X-chromosome genes, such as *BEX1, ELK1, and USP11*, whose X-inactivation status is heterogeneous may contribute to the higher incidence in females of Alzheimer's disease.

1 | INTRODUCTION

Alzheimer's disease (AD) is a progressive disease where neurons in brain regions involved in thinking, learning, and memory become damaged. In 2021, an estimated 6.2 million Americans 65 years of age and older were living with AD.[1] Alzheimer's is a complex physiological disease due to the involvement of many cell types, such as neurons, astrocytes, and microglia. We analyze published datasets of single-cell and bulk gene expression in AD to investigate two questions regarding its pathogenesis: (1) what are the cell-type–specific transcriptional alterations that are associated with AD and which cell type is the most transcriptionally altered; and (2), given the considerably larger AD incidence in females compared to males,[2] what are the underlying most notable sex-specific, cell-type–specific transcriptional alterations?

Mathys et al. published the first single-cell transcriptomics dataset from Alzheimer's patients and controls. They found that all cell types in the prefrontal cortex had transcriptional changes associated with AD.[3] They found neurons to have more differentially expressed genes (DEGs) downregulated, whereas other cell types such as astrocytes and microglia had more genes upregulated. They found excitatory neurons to have the most DEGs [3]. However, in a second, single-cell Alzheimer's dataset (GSE157827), Lau et al. found that astrocytes have more DEGs compared to other cell types.[4] Aiming to resolve this apparent quandary, we reanalyze the Mathys and Lau datasets using consistent methods. For validation, we analyze a third single-cell transcriptomics Alzheimer's dataset,[5] with the caveat that we cannot compare cell types, such as endothelial cells, that are not adequately sampled in all three datasets.

In addition to the relative importance of different brain cell types, another longstanding puzzle about AD is the observation of its substantially higher incidence in females.[2] The biology of the sex differences in AD incidence is poorly understood, even after a mouse study targeting this question.[6] Mathys et al. did sex-specific analysis and concluded that there are subtle differences in the transcriptomics between sexes in each cell type, with neurons and oligodendrocytes having the most extreme differences.[3] However, they did not look at specific genes or enriched pathways in a sex-stratified way.

To study the observed sex differences at a single-cell resolution, Belonwu and colleagues reused the Mathys dataset to perform a sex-stratified analysis to identify sex-stratified cell-type–specific perturbations in Alzheimer's patients. They found that neurons were more similar between males and females compared to glial cells, having more shared genes and pathways.[7] Their analysis was limited inadvertently to fewer than 200 genes. We instead perform a sex-stratified analysis for most genes using the same differential expression method that we use to resolve the cell-type contradiction.

To validate the sex-specific single-cell findings, we additionally analyze three bulk expression datasets from the Gene Expression Omnibus (GEO; Methods). Notably, a recent article reported a study of the role of sex differences at the transcriptome level and how it influences complex traits analyzed in the most recent (v8) version of the Genotype-Tissue Expression (GTEx) v8 project.[8] Oliva and colleagues identified sets of sex-biased genes (genes whose expression levels differ significantly between males and females) for dozens of different tissues, which we further consider in our analysis.

For our analysis of sex differences, we introduce standard terminology about X-chromosome genes. Near the Xp telomere is the pseudoautosomal region (PAR) containing a handful of genes for which females express two copies on X, whereas males express one copy on X and one or more copies on Y. The gene copies on X and Y do not recombine and hence can diverge in evolution.[9] Among the non-PAR X-chromosome genes, most have their expression between males

RESEARCH IN CONTEXT

1. **Systematic Review**: Literature search in PubMed and Google Scholar showed that females have a substantially higher incidence of Alzheimer's disease. Transcriptomics studies to investigate the sex difference have been limited. The authors searched in Synapse and GEO for single-cell and bulk transcriptomics datasets. Six suitable datasets were identified. The analyses in references 3 and 4 disagreed on whether excitatory neurons or astrocytes are the most important cell type.
2. **Interpretation**: By using appropriate methods for single-cell data, we resolved the disagreement. Excitatory neurons have the most differentially expressed genes in cases versus controls; this was validated on a third data set. Sex-specific analysis suggested an important role in female cases versus female controls for X-chromosome genes that are heterogeneous in their X-chromosome inactivation status.
3. **Future Directions**: Further analysis of X-chromosome genes should be done in transcriptomics and proteomics datasets to see how these genes contribute to the higher female incidence.

and females balanced by X-chromosome inactivation in females,[10] which is regulated primarily by the RNA gene *XIST*.[11] We partition the non-PAR genes into three categories according to their X-inactivation status in females: always inactivated, always escaping X-inactivation, and heterogeneous with respect to X-inactivation using a published classification.[12] Heterogeneity of Xinactivation may be across individuals and/or across tissues.

In summary, we consistently analyze three, single-cell datasets to investigate the association of brain cell-type gene expression with the progression of AD, and to further learn which cell-type–specific pathways are enriched in AD. In addition, we performed a sex-stratified enrichment analysis of DEGs in these three, single-cell datasets to identify differences in pathways that may contribute to the observed sex bias in AD. We complement this sex-stratified analysis of single-cell data with additional analysis of three existing bulk RNA datasets from the cortex to validate findings from single-cell and bulk data. Finally, we investigate the roles that X-chromosome genes may play in the sex bias of AD.

2 | METHODS

2.1 | Single-cell expression analysis

We analyzed the single-cell datasets Syn18485175 (Mathys dataset), GSE157827 (Lau dataset), and Syn21670836 (TREM2 dataset), as described in the Supplementary Text and Figure S1.

2.1.1 | Choice of MAST (Model-based Analysis of Single-cell Transcriptomics), as implemented in Seurat, as the analysis method for single-cell analysis

The data processing of the Mathys dataset done by Belonwu et al. was through a Seurat object; the Seurat package[13] is used widely for single-cell transcriptomics analysis. Therefore, we used the Belonwu et al. code up to their overaggressive filtering and continued with other Seurat options for the revised differential expression analysis. Seurat's *FindMarkers* function provides another way to identify DEGs between clusters while allowing for pre-filtering. In addition, Seurat also supports a variety of differential expression tests such as bimod, poisson, wilcox, MAST, and DESeq2.[13–16] A large-scale comparison study of DE analysis methods found Model-based Analysis of Single-cell Transcriptomics (MAST) to be the best-performing single-cell DE test.[17] MAST is a widely used tool, with Google Scholar reporting 1089 citations as of December 9, 2021. In addition, MAST has the virtue of being integrated into Seurat's *FindMarkers* function. For these reasons, we decided to use MAST for our single-cell differential expression analysis, as described in the Supplementary Text.

2.1.2 | Gene set enrichment analysis

We performed all gene set enrichment analysis using the web-based application GOrilla[18] with running mode set to "Single ranked list of genes." Genes were ranked by logFC value and pasted into the text box. *p*-value threshold was set to "10^-3" and "Run GOrilla in fast mode" was unchecked.

2.1.3 | TREM2 Ex cluster analysis

In the TREM2 dataset, Seurat's *FindMarkers* function was used to perform the differential expression analysis between Ex0 and Ex1 clusters as described in the Supplementary Text.

2.2 | Bulk expression analysis

2.2.1 | Datasets

Three datasets were obtained from Gene Expression Omnibus using GEO2R for the analysis on Alzheimer's samples.[19] These datasets are: GSE15222,[20] GSE33000,[21] and GSE44770.[22] Our processing of these datasets and the GTEx v8 data[23] are described in the Supplementary Text.

2.3 | Hypergeometric enrichment test for both single-cell and bulk analysis

For various analyses, we used hypergeometric tests to decide if one set of genes (e.g., X-chromosome-heterogeneous genes) was

TABLE 1 Summary of the six datasets re-analyzed in this study; see also Figure S1 for a more visual representation.

	Mathys		Lau		TREM2		GSE15222		GSE33000		GSE44770	
	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female
Ast	1543	1849	13922	6361	2408	3188	N/A	N/A	N/A	N/A	N/A	N/A
Ex	16151	18825	40526	18698	4147	7290	N/A	N/A	N/A	N/A	N/A	N/A
In	4539	4657	23979	10501	1470	1884	N/A	N/A	N/A	N/A	N/A	N/A
Mic	897	1023	5164	2974	900	1566	N/A	N/A	N/A	N/A	N/A	N/A
Oli	8795	9440	30677	10956	8187	15292	N/A	N/A	N/A	N/A	N/A	N/A
CTRL	12	12	6	3	4	7	98	77	86	24	16	5
AD	12	12	8	4	4	7	82	75	55	40	6	3

Note: For each single-cell dataset we show cell count per cell type and sex for the cell types present in all datasets; we show the total count of patients for each sex per dataset. Patients counted in the table were those that passed the filtering step (see Supplementary Methods).
Abbreviations: Ast, astrocytes; Ex, excitatory neurons; In, inhibitory neurons; Mic, microglia; Oli, oligodendrocytes; CTRL, unaffected controls; AD, number of individuals affected with Alzheimer's disease; N/A, not applicable.

over-represented in another set of genes (e.g., DEGs). To perform the hypergeometric enrichment test, we used the function *phyper*.

2.4 | Data availability

The Mathys dataset and the TREM2 dataset are available with permission via Synapse. The Lau single-cell dataset and the three bulk datasets are available from the Gene Expression Omnibus (GEO) without any account or extra permission. Instructions on how to reproduce our analysis are provided in ftp.ncbi.nlm.nih.gov/pub/catSMA/Alzheimer, which contains the single file README.txt and the archive Alz_final_manuscript.tar.gz.

3 | RESULTS

3.1 | Overview of the analysis

Three single-cell/ single-nucleus prefrontal cortex Alzheimer's datasets were analyzed, containing expression data for five cell types: astrocytes, excitatory neurons, inhibitory neurons, microglia, and oligodendrocytes (Figure 1, upper left, Table 1, Figure S1). Later in the analysis, we also analyzed three bulk expression datasets from the cortex (Figure 1, lower left). To investigate which cell type may contribute more to the progression of AD, we performed a cell-type differential expression analysis on each dataset to identify the cell type that have the most differentially expressed genes (DEGs, Figure 1, second panel from left, upper). Once we identified the cell type that had the most DEGs across the three datasets, we used the software program GOrilla to identify enriched pathways (Figure 1, upper right).

We also conducted sex-stratified cell-type differential expression analysis to better understand what may be contributing to the sex bias in AD (Figure 1, second panel from left, lower). For this analysis, we were interested particularly in two gene sets that can contribute to the sex difference observed: (a) GTEx cortex genes that are considered sex biased and were published by Oliva et al., and (b) X-chromosome genes (Figure 1, third panel from left). We performed a sex-stratified cell-type differential expression analysis on each dataset, and then performed a sex-stratified enrichment analysis using the software GOrilla. In addition, we performed hypergeometric enrichment analyses (Figure 1, lower right) to determine if certain sets of differentially expressed genes are enriched within the gene sets that can contribute to the observed sex bias.

To validate the role that X-chromosome genes have in the observed sex bias in AD, we analyzed three bulk cortex datasets (Figure 1, lower left) and performed an X-chromosome differential expression analysis (Figure 1, second panel from left, lowest part). After performing the differential expression analysis, we performed a meta-analysis to identify X-chromosome genes that are statistically significant and are upregulated/downregulated in the same direction across the three datasets.

3.2 | Excitatory neurons have more differentially expressed genes than other cell types in AD cases versus controls

To determine if excitatory neurons play a larger role in AD than other cell types do, we started by analyzing a single-cell dataset from Synapse (syn18485175)[3] for DEGs. We refer to these data as the "Mathys dataset."

Previously, Belonwu et al. performed a sex-stratified differential expression analysis on the Mathys dataset but reported surprisingly low two-digit numbers of DEGs for a dataset with a large sample size.[7] Using voom-limma,[24] which was developed specifically for bulk RNA-seq analysis, they filtered out more than 99% of the genes, which is too aggressive. Instead, we did the DEG analysis using (in the Seurat package) the MAST method,[14] which is used widely for single-cell transcriptomics data analysis (Methods). We identified cell-type specific DEGs for six cell types with sufficient data. Excitatory neurons have far more significant DEGs than the other cell types (Methods, Figure 2, Table S1).

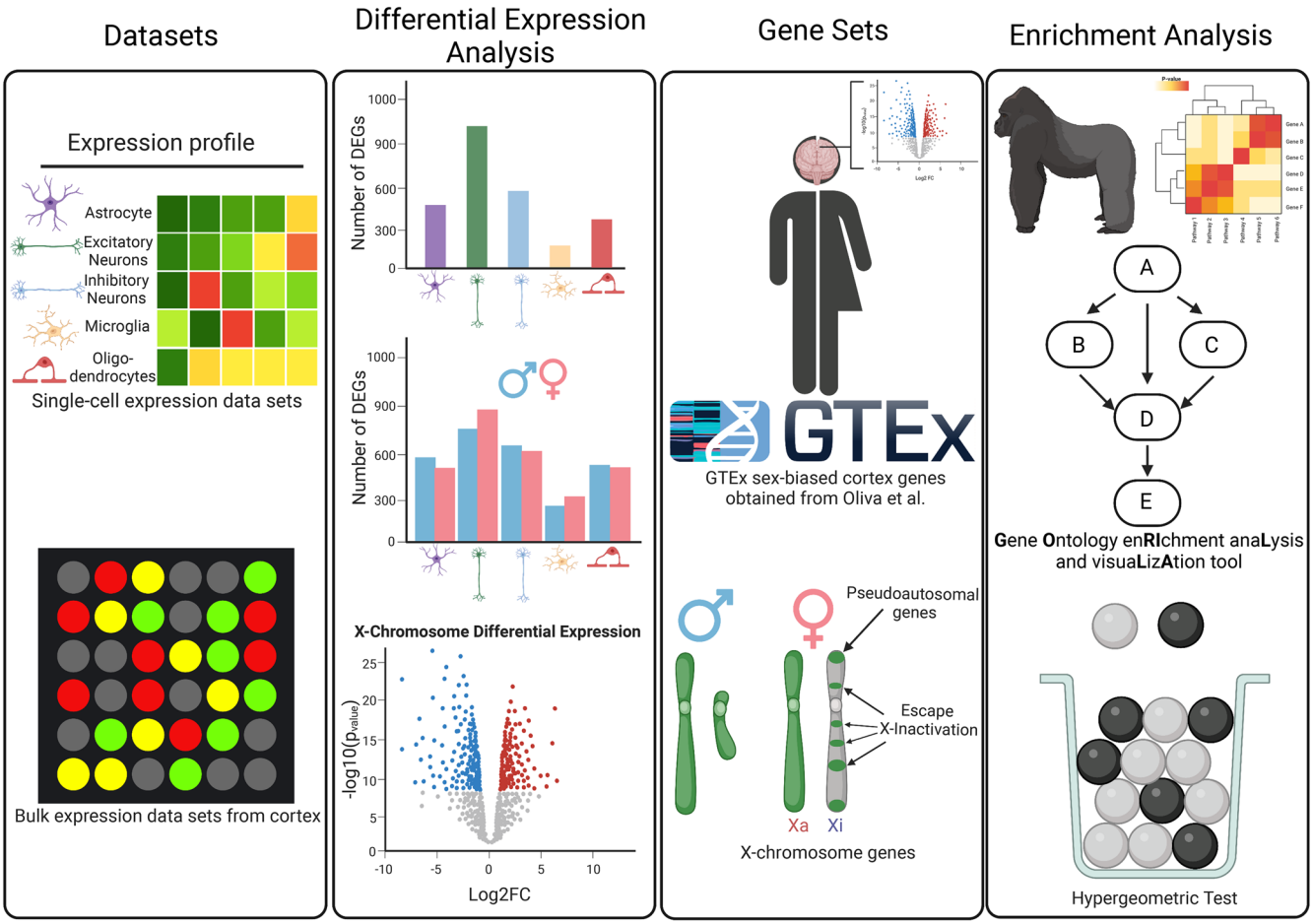

FIGURE 1 Overview of the analysis. Single-cell/-nucleus transcriptomics and bulk gene expression Alzheimer's datasets were obtained for this project (left panel). We performed three different types of differential expression analyses: cell-type specific, sex-stratified cell-type specific, and X-chromosome genes (second panel from the left). Some of our analyses on sex bias incorporated published classifications of sex-biased genes from the GTEx project and a published classification of X-chromosome genes according to X-inactivation status (third panel from the left). After performing the differential expression analysis, we used the software GOrilla to perform the enrichment analysis (upper right). In addition, we performed hypergeometric tests (lower right) to identify enrichment of two gene sets in Alzheimer's patients: GTEx sex-biased cortex genes obtained from reference 8, and the classification of X-chromosome genes.

To investigate whether the larger number of significant DEGs found in excitatory neurons happens because they have a larger cell count, we down-sampled the cells for each cell type and repeated the analysis, taking the mean across 100 replicates (Supplementary Text). Excitatory neurons continued to have more significant DEGs than all other cell types (Figure 1, Table S2). Three hundred genes are differentially expressed in more than 50 of 100 down-sampling replicates, with *RASGEF1B*, *LINGO1*, and *SLC26A3* appearing in all 100 (Table 2), overall testifying that the notable transcriptional alterations observed in the excitatory neurons are likely to reflect the biology of the disease. Of interest, *LINGO1* has been implicated in numerous neurodegenerative disorders and has been proposed as a potential for AD due to its critical role in the pathophysiology of AD by favoring the β cleavage of APP and the generation of amyloid beta (Aβ) fragments.[25,26] Belonwu et al. also detected *LINGO1* as significantly differentially expressed in the Mathys dataset.[7]

TABLE 2 Top 10 genes ranked by the number of times each gene appeared to be differentially expressed across the 100 subset replicates for the down-sampling analysis of the Mathys dataset.

Gene	Replicate count
RASGEF1B	100
LINGO1	100
SLC26A3	100
NGFRAP1	99
DHFR	97
GRIN1	93
PDE10A	89
BEX1	88
SPARCL1	88
IDS	87

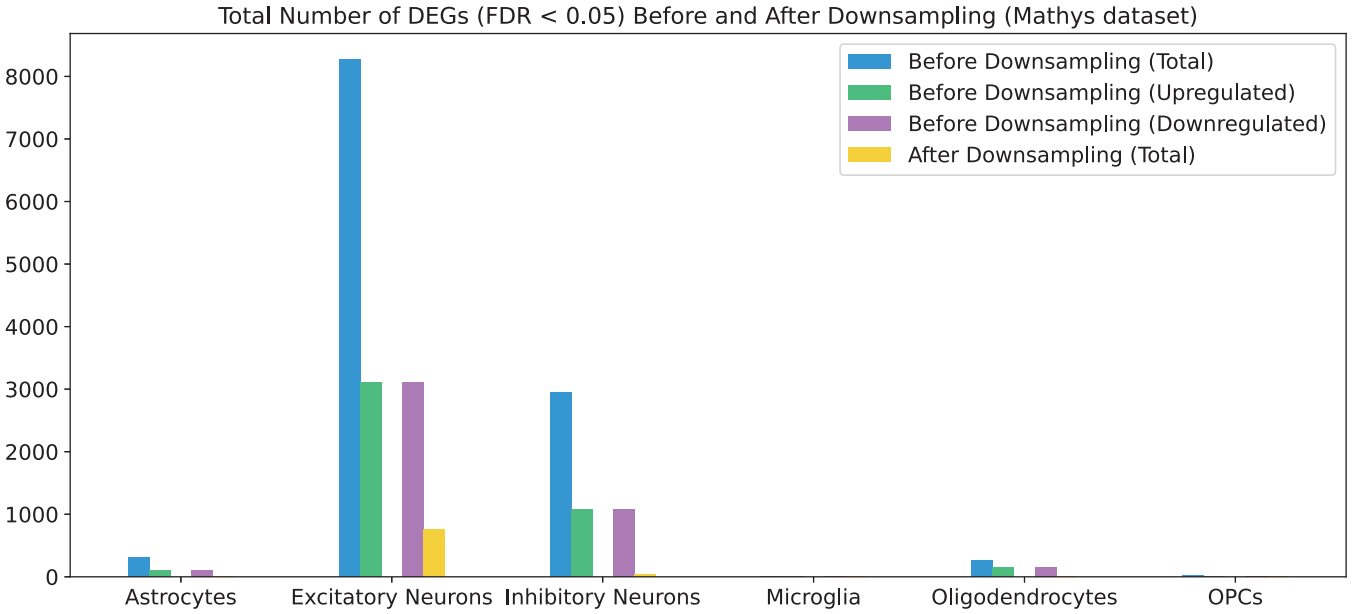

FIGURE 2 Total number of DEGs with an FDR-adjusted *p*-value less than 0.05 before and after down-sampling in the Mathys dataset. The exact numbers are shown in Table S1.

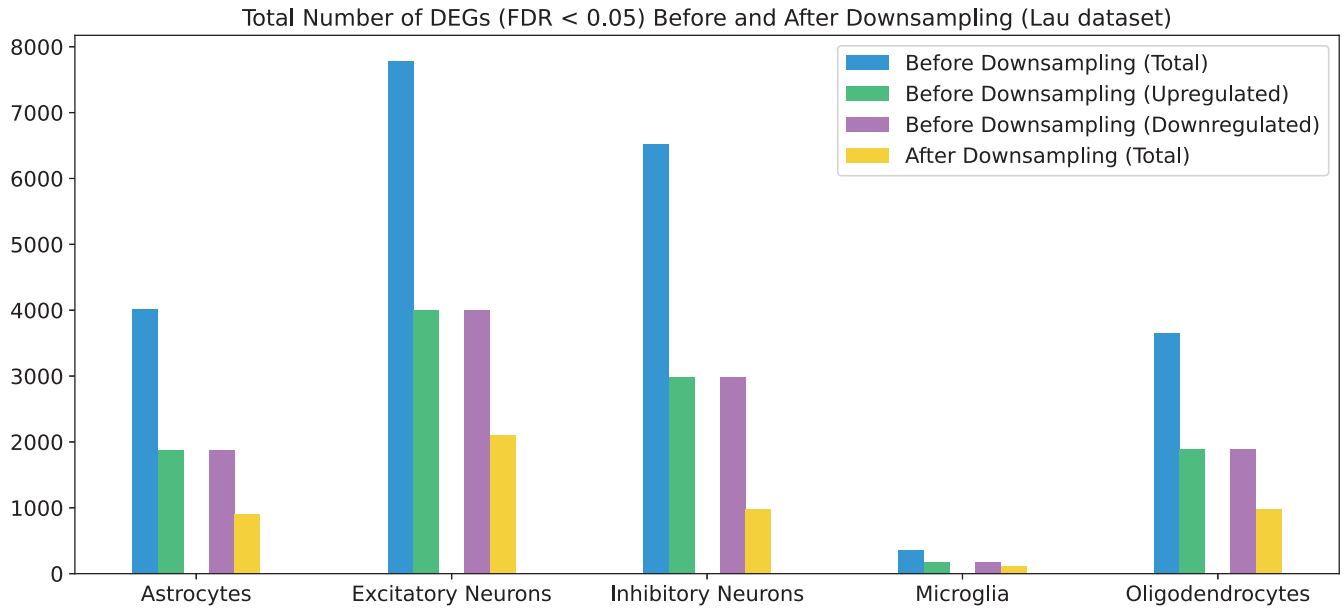

FIGURE 3 Total number of DEGs with an FDR-adjusted *p*-value less than 0.05 before and after down-sampling in the Lau dataset. The exact numbers are shown in Table S2.

Again, we used MAST within Seurat for the most important analysis steps (Methods). Reassuringly, the percentages of cells of each type that we found correspond well overall with those reported originally by Lau et al. (Table S2). Using our cell classifications, Figure 3 and Table S3 show the number of DEGs found per cell type for each of the five pertaining cell types; we excluded endothelial cells because they were removed in Mathys analysis (Supplementary Text). These results reinforce our previous findings in the Mathys dataset that excitatory neurons had more significant DEGs compared to other cell types. We performed the same down-sampling test that we performed on Mathys dataset, finding that excitatory neurons contained more significant DEGs than all other cell types and validating Table 2.

Next, we studied the cell type question in a third dataset, syn21670836,[5] which we refer to as the "TREM2 dataset" (Supplementary Text). The metadata provided by Zhou et al. classified cells by cell type and provided the sex and age for each patient. One unusual aspect

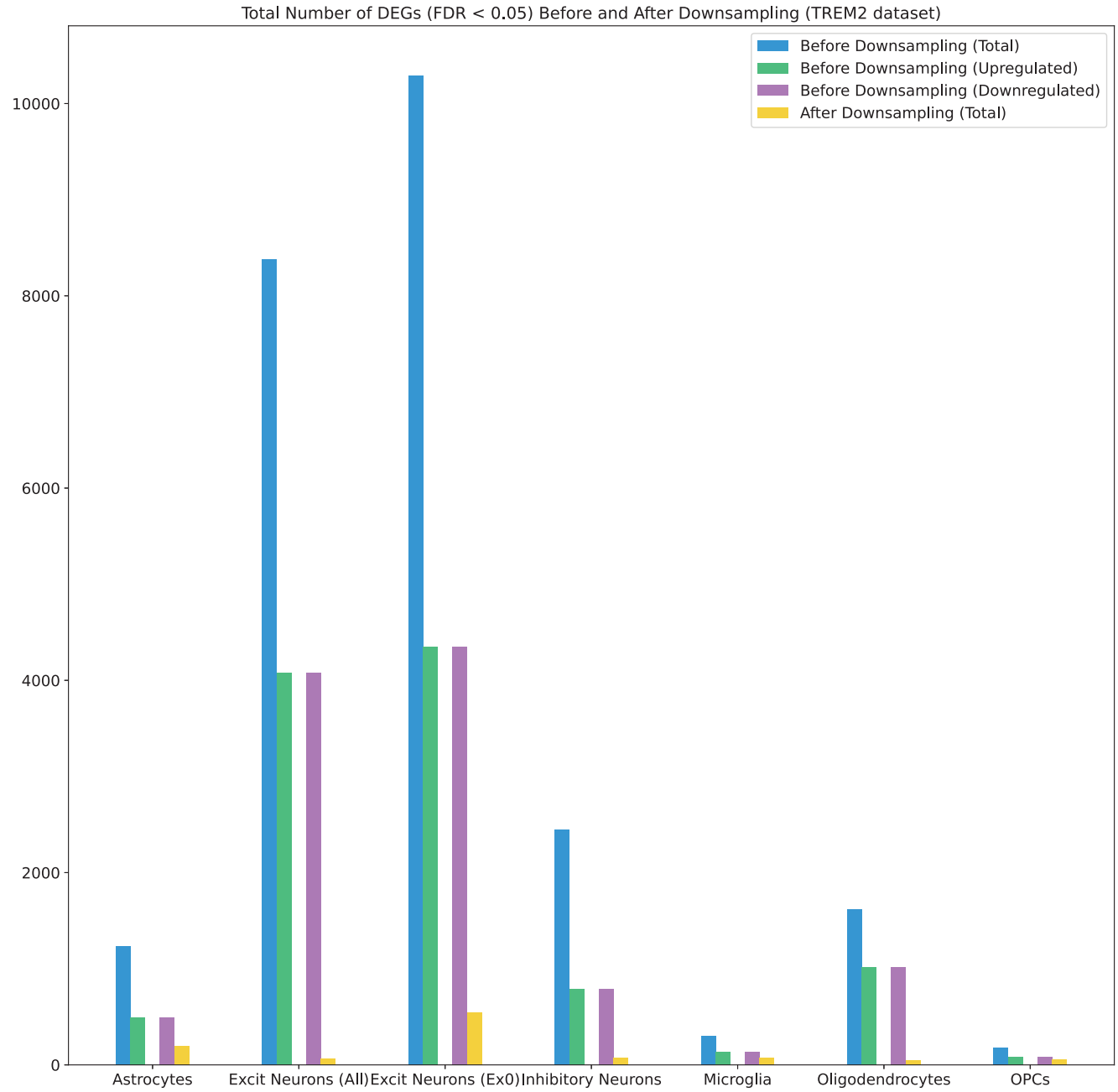

FIGURE 4 Total number of DEGs with an FDR-adjusted *p*-value less than 0.05 before and after down-sampling in the TREM2 dataset. The exact numbers are shown in Table S4.

of this data set is that there were two clusters of excitatory neurons, which the authors denoted by Ex0 and Ex1.

We focused our analysis on the Alzheimer's patients with TREM2-CV, a common *TREM2* variant, because they were sex-matched with the controls. As shown in Figure 4 and Table S4, we again found that excitatory neurons have far more significant DEGs compared to the other cell types, if we combine Ex0 and Ex1. However, this time the down-sampling test (**Supplementary Text**) revealed that astrocytes contained the most significant DEGs (Figure 4 and Table S4, second row).

We reasoned that combining Ex0 and Ex1 for DEG analysis may be ill-advised and, therefore, we re-performed the down-sampling analysis on cluster Ex0, which has a much larger cell count (Supplementary Text). This analysis showed that Ex0 cells have more DEGs than each other cell types, reaffirming the findings on the first two datasets (Table S4).

Zhou et al. did not provide any biological characterization of the difference between Ex0 and Ex1; they just accepting these as distinct clusters produced automatically by Seurat.[5] To try to find a biological difference, we performed a differential expression analysis that we will present later in Results.

Our primary threshold for determining that a gene is differentially expressed is that the false discovery rate (FDR)-adjusted *p*-value is < 0.05. In addition, we did sensitivity analysis for the logFC threshold (Supplementary Text and Tables S5–S7). Tables S8–S15 list for

TABLE 3 Total number of DEGs with an FDR-adjusted *p*-value less than 0.05 before and after down-sampling in the Lau dataset.

	Astrocytes	Excitatoryneurons	Inhibitoryneurons	Microglia	Oligodendrocytes
Pre-down-samplingDEGs	4021	7,784	6515	350	3655
Post-down-samplingDEGs	908	2106	984	118	975

each of astrocytes, excitatory neurons, microglia, and oligodendrocytes, the genes that are consistently upregulated and consistently downregulated in the three datasets.

3.3 | Pathway enrichment analysis of DEGs

After discovering that excitatory neurons have the most DEGs, we performed an enrichment analysis to determine which of their pathways are upregulated or downregulated in Alzheimer's patients (Methods, Figures S2 and S3). Overall, we find different key enriched pathways in the different datasets we have studied (Discussion). One factor that could contribute to the dissimilar pathway enrichment in the datasets is that the Lau dataset contains substantially more cells from males than females, whereas the TREM2 dataset contains more cells from females than males (Tables S16–S18). In the Mathys dataset, pathways involved in synaptic signaling are significantly upregulated in Alzheimer's patients, which bears relevance to the core homeostatic machinery theory proposed by Frere and Slutsky.[27] The top pathways that are significantly downregulated are involved in the electron transport chain, which also contributes to the stability of the homeostatic machinery.[27] Analysis for the Lau and TREM2 datasets is summarized in the Supplementary Text and Figures S4–S7. Pathway analysis for astrocytes and microglia is in the Supplementary Text and Tables S19–S63. The difference across datasets in synaptic signaling regulation is surprising, so we analyzed individual, widely studied, synaptic genes and confirmed inconsistencies across datasets (Supplementary Information and Tables S64–S65).

To search for common potential AD gene targets, we looked for genes that were among the top 500 DEGs for excitatory neurons in each dataset ranked by adjusted *p*-value and that are members of the dysregulated pathways identified by GOrilla. We focused on overexpressed genes, since they may be easier to target. One gene that stood out is *GRIN1*, which is one of the most consistent DEGs in our down-sampling of the excitatory neurons in the Mathys dataset (Table 3). This gene, also known as *GluRN1*, encodes a glutamate receptor and is a functional candidate because of the key roles of glutamatergic synapses in the pathogenesis of AD.[28–30] Overexpression of the GRIN1 protein in AD has been found in at least two studies,[31,32] but this has been challenging to study in bulk samples because the glutamatergic neurons tend to die early in the disease.[28,33] *GRIN1* contributes to the following upregulated pathways (GOrilla rankings in parentheses) in the Mathys dataset excitatory neurons GO:0050804 modulation of chemical synaptic transmission (1st), GO:0099177 regulation of trans-synaptic signaling (2nd), GO:0043269 regulation of ion transport (11th), GO:0032879 regulation of localization (16th), and GO:0099537 trans-synaptic signaling (32nd). Similarly, in the Lau dataset, *GRIN1* contributes to the top two ranked upregulated pathways: GO:0050804 modulation of chemical synaptic transmission and GO:0099177 regulation of trans-synaptic signaling.

TABLE 4 Top 10 DEGs between Ex0 and Ex1 ranked by descending logFC.

Gene	logFC	adj.*p*.value
MEG3	4.387	0
MALAT1	3.905	0
MIAT	3.532	0
MIR124-1HG	2.765	0
XIST	2.740	0
PNISR	2.521	0
LINC00632	2.520	0
KCNIP4-IT1	2.485	0
RNPC3	2.471	0
AH1	2.463	0

Note: The second row is the average of 100 replicates rounded to the nearest integer. A positive logFC value indicate higher expression in Ex0 cluster. Zero adj. *p*-values indicate a value less than 1.00e-310.

Finally, to try to find a biological difference between the two excitatory neuron clusters in the TREM2 dataset, we performed a MAST differential expression analysis between Ex0 and Ex1 clusters, controlling for sex, age, and condition. **Table 4** shows the top 10 DEGs between Ex0 and Ex1. Of interest, the fifth gene that was expressed higher in Ex0 is *XIST*, which is known to control X-inactivation.[34] Based on this finding, we were interested in seeing if this means that Ex1 would have more upregulated X-chromosome inactivation heterogeneous genes upregulated than Ex0. To test this hypothesis, we analyzed only female samples. Indeed, we found that Ex1 had more X-chromosome inactivation heterogeneous genes upregulated compared to Ex0 (100 and 63, respectively). Because heterogeneity in X-chromosome inactivation occurs only in females, this finding led us to wonder whether these X-chromosome genes are relevant to the difference in AD incidence between males and females. To pursue the potential contributions of excitatory neurons to sex differences, we performed sex-stratified differential expression analysis.

3.4 | Sex-stratified enrichment analysis of differentially expressed pathways in excitatory neurons

We performed a sex-stratified enrichment analysis to determine if there are sex differences in enriched pathways (Supplementary Text). In each dataset we compared male cases to male controls and

female cases to female controls (we did not compare male cases to female cases because this would not be a properly controlled comparison). In the Mathys dataset, the top pathways upregulated in Alzheimer's males and females were involved in cell adhesion and synaptic transmission, with trans-synaptic signaling being more prominent in females. Females also had upregulation of cellular component organization, which was not upregulated in males. Top pathways downregulated in both Alzheimer's males and females were involved with cellular metabolism and the electron transport chain. In addition, males had several downregulated pathways including ferric iron transport, transferrin transport, immune effector process, and immune system process, which were not enriched in females. Further analyses of the other two datasets and specific genes can be found in the Supplementary Text.

Key pathways that appear most often in the sex-specific comparisons are GO:0099177 regulation of trans-synaptic signaling and GO:0050804 modulation of chemical synaptic transmission. Genes in these pathways may be both upregulated and downregulated, consistent with the hypothesis that loss of homeostasis is the key to Alzheimer's pathogenesis.[27] When comparing upregulated and downregulated pathways, the most striking observation was that the TREM2 Ex0 cluster has only upregulated pathways in males and only downregulated pathways in females. Furthermore, among the 178 significant downregulated pathways in females and the 90 significantly upregulated pathways in males, 56 of 178 and of 90 are the same pathways, but changing in opposite directions in the two sex-specific analyses. This reinforces our previous conclusion that the Seurat split between the Ex0 and Ex1 clusters in the TREM2 dataset unmasks some clues about sex differences in AD. The Mathys dataset has mostly upregulated pathways in both males and females, which presents an opportunity to compare datasets in the pathways upregulated in Mathys dataset males and TREM2 dataset males, but not in Mathys dataset females. There are three such pathways: GO:0009653 anatomical structure morphogenesis, GO:0045597 positive regulation of cell differentiation, and GO:0007156 homophilic cell adhesion via plasma membrane adhesion molecules.

3.5 | Revisiting the single-cell sex-specific findings in a bulk expression analysis

After discovering sex differences in enriched pathways in excitatory neurons, we decided to perform a differential expression analysis on bulk data to see if bulk data analysis would provide additional information on the observed sex differences. For this analysis, we were especially interested in X-chromosome PAR genes and X-chromosome heterogeneous for escaping X-inactivation and their potential role contributing to AD. We analyzed three datasets for AD that had gene expression from the cortex: GSE15222, GSE33000, and GSE44770.[20–22] We performed linear modeling of the association of their DE with AD, controlling for sex and age (Supplementary Text).

After obtaining the differential expression analysis on each Alzheimer's dataset, meta-analysis was performed to identify the significant pseudoautosomal and inactivation-heterogeneous X-chromosome genes across all three datasets and the *p*-values were corrected for multiple testing (Supplementary Text). Among significant consistently differentially expressed genes genome-wide were five pseudoautosomal genes: *CD99* (genome-wide FDR adjusted $p = 2.7e\text{-}11$), *ZBED1* ($p = 5.2e\text{-}11$), *IL3RA* ($p = 2.5e\text{-}09$), *ASMTL* ($p = 6.0e\text{-}09$), and *GTPB6* ($p = 8.6e\text{-}05$). **Table 5** shows the top 10 statistically significant differentially expressed inactivation-heterogeneous X-chromosome genes adjusting for multiple testing (Methods). The top three genes were *BEX1*, *PRKX*, and *TSR2*. Of interest, *BEX1* was one of the most consistent DEGs in our sampling of the excitatory neurons in the Mathys dataset (**Table 4**), is significantly downregulated in cases versus controls in all three single-cell datasets, and functionally has been shown to be involved in the regeneration of axons[35] and is downregulated in Alzheimer's patients compared to controls. *PRKX* encodes a serine threonine protein kinase and has been shown to play a crucial role in neural development.[36] *TSR2* has been shown to inhibit the transcriptional activity of NF-κB, which is one of the key transcription factors for the homeostatic model,[27] and *TSR2* was found to be downregulated in Alzheimer's patients.[37] Of note, while this manuscript was under review, Yan et al., published a beautiful study implicating the X-chromosome heterogeneous gene *USP11* in the pathogenesis of AD in females[38]; this coincidentally adds functional data evidence to the results shown in **Table 5**.

Because we saw that X-chromosome heterogeneous genes were significantly differentially expressed between Alzheimer's and control in bulk data, we went back to the single-cell transcriptomics dataset and tested whether X-chromosome heterogeneous genes are over-represented among male or female excitatory neurons DEGs. In the Mathys dataset we did not find significant enrichment of the X-chromosome heterogeneous genes in the DEGs for males (p-value = 0.21). However, quite strikingly, a hypergeometric test for females resulted in a p-value of 0, indicating that enrichment is statistically significant with a p-value less than 1.00e-310. This difference makes sense, since some X-chromosome heterogenous genes differ in the expression between female individuals.

In the Lau dataset, we again did not find significant enrichment of the X-chromosome heterogeneous genes in the DEGs for males (p-value = 0.10). In the females, the hypergeometric test again resulted in a p-value of 0, indicating that enrichment is statistically significant with a p-value less than 1.00e-310.

In the TREM2 Ex0 cluster, we did not find significant enrichment of the X-chromosome heterogeneous genes in the DEGs for either males or females (p-value 0.51 and 0.21, respectively). This was expected based on the above analysis that puts the X-chromosome heterogeneous DEGs preferentially in the Ex1 cluster.

Another source of candidate genes to be involved in diseases, such as Alzheimer's, with a difference in prevalence by sex, comes from the analysis of Oliva et al. of the newly published Genotype-Tissue

TABLE 5 Top 10 significant differentially expressed X-chromosome heterogeneous genes between Alzheimer's and controls meta-analysis.

Gene	GSE15222 logFC	GSE33000 logFC	GSE44770 logFC	Combined adj.*p*.value
BEX1	-0.580	-0.065	-0.087	4.45E-24
PRKX	0.545	0.081	0.080	2.50E-23
TSR2	-0.254	-0.047	-0.046	2.13E-22
FOXO4	0.598	0.080	0.046	2.61E-21
ELK1	0.333	0.092	0.107	5.57E-21
USP11	-0.537	-0.040	-0.054	3.15E-20
TBL1X	0.429	0.061	0.030	8.83E-20
ATP6AP2	-0.375	-0.037	-0.039	1.82E-19
IDS	-0.524	-0.026	-0.049	8.71E-19
MCTS1	-0.474	-0.022	-0.039	9.10E-18

Note: Meta-analysis was done using Fisher's method with FDR correction for those X-chromosome inactivation heterogeneous genes that had logFC (log fold change) with the same sign in each of the three data-sets. The rightmost column has the combined adjusted p-value after applying Fisher's method.

Expression (GTEx) v8 project data. Of the 112 sex-bias genes reported in that study for the cortex, 31 are either X-chromosome inactivation heterogeneous or pseudoautosomal genes (19 and 12, respectively), with an enrichment *p*-value of 1.21e-40. Of interest, 8 of the 19 X-chromosome inactivation heterogeneous genes are also identified in our X-chromosome meta-analysis in the bulk data. Of the 31 genes mentioned, 21 were found to be differentially expressed in the Mathys dataset, 18 were found to be differentially expressed in the Lau dataset, and 24 were found to be differentially expressed in the TREM2 dataset (Supplementary Table S66). Among these genes, *NAP1L3* and *CHM* are heterogeneous genes that are consistently downregulated in cases versus controls.

Taken together, our results suggest that among the commonly studied cell types, excitatory neurons have the most differentially expressed genes, resolving a previous contradiction in the literature. Analysis of the excitatory neuron data in three single-cell datasets and cortex data in three bulk datasets suggest that the differentially expressed genes in this cell type and this tissue disproportionately include X-chromosome heterogeneous genes in females and pseudoautosomal genes when analyzing all cases versus all controls. Furthermore, different pathways are upregulated and downregulated in the two sexes in excitatory neurons. Thus the longstanding questions of which cell type is most important in Alzheimer's pathogenesis and why females have higher incidence than males can be connected logically and biologically by analyzing differential gene expression in single-cell data.

4 | DISCUSSION

In this study, we focused on different neuronal cell types and their expression in the pathogenesis of AD. In contrast, other studies have investigated the possible pathogenetic role of synaptic alterations[28,39] or genetic factors.[40] Our study has become feasible due to publicly available Alzheimer's single-cell transcriptomics datasets. Following previous work,[3,4,7] we considered the number of differentially expressed genes (or DEGs) between Alzheimer's and controls in specific cell types as an indicator of cell type importance. Analyzing these datasets in a uniform manner, we asked two fundamental questions: (1) what are the key transcriptomics differences between the brains of Alzheimer's patients and controls and (2) what are the sex-dependent differences in the cell types and genes that are differentially altered in Alzheimer's disease?

Regarding the first question, previous studies[3,4] reached contrasting conclusions, finding that excitatory neurons[3] or astrocytes,[4] have the most DEGs between Alzheimer's and controls. Our first key result resolves this discrepancy by homogeneously using the MAST[14] method on both datasets, showing that excitatory neurons have more DEGs. We validated this result on a third single-cell dataset.[5]

To attempt a formal replication, we deliberately analyzed the three sets separately and, to the extent possible, respected the definition of cell types in the original publications. In addition, to check that cell type identification was similar across datasets, we further integrated the three single-cell datasets using Seurat[13] (Supplementary Methods, Table S67). We collected marker genes of cell types from the original manuscript of each dataset. We present the dot plots of marker genes of all cell types regarding four scenarios: (1) integration of three datasets; (2) Mathys dataset (separate from the integration results); (3) Lau dataset (separate from the integration results); (4) TREM2 dataset (separate from the integration results) (Figures S8–S11). We found empirically that cell types are consistent across different datasets. Using excitatory neurons as an example, the genes *RALYL*, *KCNIP4*, *CBLN2*, *LDB2*, and *KCNQ5* were markers of excitatory neurons suggested for the Lau dataset, whereas *GRIN1*, *SYT1*, *RBFOX2*, *PDE1A*, and *ETL4*, etc., were marker genes of excitatory neurons, as suggested for the TREM2 dataset. Our integration dot plot (Figure S8) clearly identified excitatory neurons by both marker gene sets. Altogether, we can successfully re-identify clusters after integration of three snRNA-seq datasets via Seurat across different marker gene sets for excitatory neurons reported by the original studies. As more and larger

single-cell datasets are collected for AD, it would help downstream analysis to have more homogeneous sequencing protocols and marker gene sets to define cell types.

Identifying the cellular pathways that are over-represented among the DEGs, we found that synaptic signaling (upregulated in the Mathys dataset but downregulated in the Lau and TREM2 datasets) and mitochondrial functions related to energy production and ion transport were downregulated (in the Mathys and TREM2 datasets). In addition, synaptic signaling was downregulated in female patients by themselves but not in male patients in the TREM2 dataset. These findings are consistent with the homeostatic model of,[27] in which "firing homeostasis" is central and depends partly on "proteostasis" and "energy homeostasis" for which the mitochondria play key roles.

We next investigated the pathogenic role of sex differences by analyzing three single-cell transcriptomics datasets and three bulk gene expression datasets from cortex, and incorporating the recent genome-wide characterization of tissue-specific sex-biased genes published as part of the release of GTEx v8.[8] Based on the analysis of the TREM2 dataset, we hypothesized and validated that two gene sets on the X-chromosome—PAR genes and heterogeneously X-inactivated genes—are significantly over-represented among the DEGs between Alzheimer's male and female patients, both in single-cell and bulk data. A recent functional study established the importance of the PAR gene *IL3A*, in Alzheimer's pathogenesis.[41]

Our hypotheses about the possible role of X-chromosome genes can be pursued further by reconsidering published GWAS datasets in females only via meta-analysis, as was done recently by Chung and colleagues, leading to the identification of variants in the autosomal gene *MGMT* as a female-specific risk factor.[42] Oliva and colleagues attempted to connect the sex-biased genes they identified in GTEx to diseases by searching the NHGRI-EBI GWAS catalog,[43] but they found no associations for AD. Early GWAS ignored the X-chromosome due to lack of statistical methods; this has been overcome with newer methods such as XWAS,[44] but the NHGRI-EBI catalog does not include X-chromosome-specific analyses.

Our study has several imitations. First, our analysis is focused on differential expression and statistical associations, not causal mechanisms. Second, our analyses of upregulated and downregulated gene pathways uncovered quite different results for each of the three single-cell datasets, possibly due to heterogeneity in their study designs and wet lab transcriptomics data collection methods. Third, some cell types, such as endothelial cells could not be compared and other cell types, such as microglia, could be formally compared across datasets but have relatively few cells sampled relative to excitatory neurons. Fourth, to have larger numbers of samples, we ignore important covariates, such as the apolipoprotein E (*APOE*) genotype. Finally, our analysis is focused on transcriptomic alterations and hence has limited ability to shed light on classical and prevailing theories of Alzheimer's pathogenesis, involving many post-transcriptional mechanisms concerning the formation of amyloid plaques and neurofibrillary tangles and beyond.[45–47]

In sum, we resolved the contradiction between the study of Mathys et al.[3] and the study of Lau et al.[4] showing that when comparing AD patients to unaffected controls, excitatory neurons have more DEGs than do astrocytes and other cell types. Analysis of enriched pathways in excitatory neurons points to differences in synaptic transmission and related pathways between males and females. Further sex-specific analysis of differentially expressed genes between cases and controls in single-cell and bulk transcriptomics datasets suggests that PAR genes and X-chromosome heterogeneous genes may contribute to the difference in sex incidence of AD. Identifying cell types and genes whose expression is different between female and male patients opens new possibilities for understanding the cellular and molecular etiology of this vexing illness.

ACKNOWLEDGMENTS

This research was supported in part by the National Institutes of Health (NIH Intramural Research Program, National Cancer Institute. The research of A.X.G. was supported in part by the National Academy Enrichment Program at the NIH. This work was primarily supported by the National Institute on Aging (NIA) under Award Number R56AG074001, U01AG073323, R01AG076448, R01AG066707, 3R01AG066707-01S1, and 3R01AG066707-02S1 to F.C. This work utilized the computational resources of the NIH HPC Biowulf cluster (hpc.nih.gov).

The results in this manuscript are in whole or in part based on data (Mathys dataset and TREM2 dataset) obtained from the AD Knowledge Portal (adknowledgeportal.org). Samples for this study (for the Mathys dataset) were provided by the Rush Alzheimer's Disease Center, Rush University Medical Center, Chicago. Data collection was supported through funding by NIA grants P30AG10161, R01AG15819, R01AG17917, R01AG30146, R01AG36836, U01AG32984, and U01AG46152; the Illinois Department of Public Health; and the Translational Genomics Research Institute.)

Figure 1 and Figure S1 were created with Biorender.com (biorender.com/)

CONFLICT OF INTEREST STATEMENT

E.R. is a co-founder of MedAware, Metabomed, and Pangea Biomed (divested), and an unpaid member of Pangea Biomed's scientific advisory board. The other authors have no competing interests. Author disclosures are available in the supporting information.

ORCID

Alejandro A. Schäffer https://orcid.org/0000-0002-2147-8033

REFERENCES

1. 2021 Alzheimer's disease facts and figures. *Alzheimers Dement*. 2021;17(3):327-406. https://doi.org/10.1002/alz.12328
2. Ullah MF, Ahmad A, Bhat SH, Abu-Duhier FM, Barreto GE, Ashraf GM. Impact of sex differences and gender specificity on behavioral characteristics and pathophysiology of neurodegenerative disorders. *Neurosci Biobehav Rev*. 2019;102:95-105. https://doi.org/10.1016/j.neubiorev.2019.04.003
3. Mathys H, Davila-Velderrain J, Peng Z, et al. Single-cell transcriptomic analysis of Alzheimer's disease. *Nature*. 2019;570(7761):332-337. https://doi.org/10.1038/s41586-019-1195-2
4. Lau SF, Cao H, Fu AKY, Ip NY. Single-nucleus transcriptome analysis reveals dysregulation of angiogenic endothelial cells and

neuroprotective glia in Alzheimer's disease. *Proc Natl Acad Sci U S A*. 2020;117(41):25800-25809. https://doi.org/10.1073/pnas.2008762117
5. Song WM, Andhey PS, et al. Human and mouse single-nucleus transcriptomics reveal TREM2-dependent and TREM2-independent cellular responses in Alzheimer's disease. *Nat Med*. 2020;26(1):131-142. https://doi.org/10.1038/s41591-019-0695-9
6. Zhao N, Ren Y, Yamazaki Y, et al. Alzheimer's risk factors age, APOE genotype, and sex drive distinct molecular pathways. *Neuron*. 2020/06/03/ 2020;106(5):727-742.e6. doi: https://doi.org/10.1016/j.neuron.2020.02.034
7. Belonwu SA, Li Y, Bunis D, et al. Sex-stratified single-cell RNA-Seq analysis identifies sex-specific and cell type-specific transcriptional responses in Alzheimer's disease across two brain regions. *Mol Neurobiol*. 2022;59(1):276-293. https://doi.org/10.1007/s12035-021-02591-8
8. Oliva M, Munoz-Aguirre M, Kim-Hellmuth S, et al. The impact of sex on gene expression across human tissues. *Science*. 2020;369(6509):eaba3066. https://doi.org/10.1126/science.aba3066
9. Olney KC, Brotman SM, Andrews JP, Valverde-Vesling VA, Wilson MA. Reference genome and transcriptome informed by the sex chromosome complement of the sample increase ability to detect sex differences in gene expression from RNA-Seq data. *Biol Sex Differ*. 2020;11(1):42. https://doi.org/10.1186/s13293-020-00312-9
10. Lyon MF. Gene action in the X-chromosome of the mouse (Mus musculus L.). *Nature*. 1961;190:372-373. https://doi.org/10.1038/190372a0
11. Clerc P, Avner P. Role of the region 3' to Xist exon 6 in the counting process of X-chromosome inactivation. *Nat Genet*. 1998;19(3):249-253. https://doi.org/10.1038/924
12. Zhang Y, Castillo-Morales A, Jiang M, et al. Genes that escape X-inactivation in humans have high intraspecific variability in expression, are associated with mental impairment but are not slow evolving. *Mol Biol Evol*. 2013;30(12):2588-2601. https://doi.org/10.1093/molbev/mst148
13. Hao Y, Hao S, Andersen-Nissen E, et al. Integrated analysis of multimodal single-cell data. *Cell*. 2021;184(13):3573-3587 e29. https://doi.org/10.1016/j.cell.2021.04.048
14. Finak G, McDavid A, Yajima M, et al. MAST: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data. *Genome Biol*. 2015;16:278. https://doi.org/10.1186/s13059-015-0844-5
15. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15(12):550. https://doi.org/10.1186/s13059-014-0550-8
16. McDavid A, Finak G, Chattopadyay PK, et al. Data exploration, quality control and testing in single-cell qPCR-based gene expression experiments. *Bioinformatics*. 2013;29(4):461-467. https://doi.org/10.1093/bioinformatics/bts714
17. Soneson C, Robinson MD. Bias, robustness and scalability in single-cell differential expression analysis. *Nat Methods*. 2018;15(4):255-261. https://doi.org/10.1038/nmeth.4612
18. Eden E, Navon R, Steinfeld I, Lipson D, Yakhini Z. Gorilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. *BMC Bioinformatics*. 2009;10:48. https://doi.org/10.1186/1471-2105-10-48
19. Davis S, Meltzer PS. GEOquery: a bridge between the Gene Expression Omnibus (GEO) and BioConductor. *Bioinformatics*. 2007;23(14):1846-1847. https://doi.org/10.1093/bioinformatics/btm254
20. Webster JA, Gibbs JR, Clarke J, et al. Genetic control of human brain transcript expression in Alzheimer disease. *Am J Hum Genet*. 2009;84(4):445-458. https://doi.org/10.1016/j.ajhg.2009.03.011
21. Narayanan M, Huynh JL, Wang K, et al. Common dysregulation network in the human prefrontal cortex underlies two neurodegenerative diseases. *Mol Syst Biol*. 2014;10:743. Doi: 10.15252/msb.20145304
22. Zhang B, Gaiteri C, Bodea LG, et al. Integrated systems approach identifies genetic nodes and networks in late-onset Alzheimer's disease. *Cell*. 2013;153(3):707-720. https://doi.org/10.1016/j.cell.2013.03.030
23. GTEx Consortium. The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science*. 2020;369(6509):1318-1330. https://doi.org/10.1126/science.aaz1776
24. Ritchie ME, Phipson B, Wu D, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. 2015;43(7):e47. https://doi.org/10.1093/nar/gkv007
25. Fernandez-Enright F, Andrews JL. Lingo-1: a novel target in therapy for Alzheimer's disease? *Neural Regen Res*. 2016;11(1):88-89. https://doi.org/10.4103/1673-5374.175048
26. de Laat R, Meabon JS, Wiley JC, Hudson MP, Montine TJ, Bothwell M. LINGO-1 promotes lysosomal degradation of amyloid-β protein precursor. *Pathobiol Aging Age Relat Dis*. 2015;5:25796. https://doi.org/10.3402/pba.v5.25796
27. Frere S, Slutsky I. Alzheimer's Disease: From Firing Instability to Homeostasis Network Collapse. *Neuron*. 2018;97(1):32-58. https://doi.org/10.1016/j.neuron.2017.11.028
28. Griffiths J, Grant SGN. Synapse pathology in Alzheimer's disease. *Semin Cell Dev Biol*. 2022;https://doi.org/10.1016/j.semcdb.2022.05.028
29. Hardy J, Cowburn R, Barton A, et al. Region-specific loss of glutamate innervation in Alzheimer's disease. *Neurosci Lett*. 1987;73(1):77-80. https://doi.org/10.1016/0304-3940(87)90034-6
30. Sokolow S, Luu SH, Nandy K, et al. Preferential accumulation of amyloid-beta in presynaptic glutamatergic terminals (VgluT1 and VgluT2) in Alzheimer's disease cortex. *Neurobiol Dis*. 2012;45(1):381-387. https://doi.org/10.1016/j.nbd.2011.08.027
31. Leuba G, Vernay A, Kraftsik R, Tardif E, Riederer BM, Savioz A. Pathological reorganization of NMDA receptors subunits and postsynaptic protein PSD-95 distribution in Alzheimer's disease. *Curr Alzheimer Res*. 2014;11(1):86-96. https://doi.org/10.2174/1567205013106660170
32. Zeidán-Chuliá F, de Oliveira BH, Salmina AB, et al. Altered expression of Alzheimer's disease-related genes in the cerebellum of autistic patients: a model for disrupted brain connectome and therapy. *Cell Death Dis*. 2014;5(5):e1250. https://doi.org/10.1038/cddis.2014.227
33. Leng K, Li E, Eser R, et al. Molecular characterization of selectively vulnerable neurons in Alzheimer's disease. *Nat Neurosci*. 2021;24(2):276-287. https://doi.org/10.1038/s41593-020-00764-7
34. Panning B. X-chromosome inactivation: the molecular basis of silencing. *J Biol*. 2008;7(8):30. https://doi.org/10.1186/jbiol95
35. Khazaei MR, Halfter H, Karimzadeh F, Koo JH, Margolis FL, Young P. Bex1 is involved in the regeneration of axons after injury. *J Neurochem*. 2010;115(4):910-920. https://doi.org/10.1111/j.1471-4159.2010.06960.x
36. Blaschke RJ, Monaghan AP, Bock D, Rappold GA. A novel murine PKA-related protein kinase involved in neuronal differentiation. *Genomics*. 2000;64(2):187-194. https://doi.org/10.1006/geno.2000.6116
37. He H, Zhu D, Sun J, Pei R, Jia S. The novel protein TSR2 inhibits the transcriptional activity of nuclear factor-kappaB and induces apoptosis. *Mol Biol (Mosk)*. 2011;45(3):496-502.
38. Yan Y, Wang X, Chaput D, et al. X-linked ubiquitin-specific peptidase 11 increases tauopathy vulnerability in women. *Cell*. 2022;185(21):3913-3930 e19. https://doi.org/10.1016/j.cell.2022.09.002
39. Styr B, Slutsky I. Imbalance between firing homeostasis and synaptic plasticity drives early-phase Alzheimer's disease. *Nat Neurosci*. 2018;21(4):463-473. https://doi.org/10.1038/s41593-018-0080-x
40. Fenoglio C, Scarpini E, Serpente M, Galimberti D. Role of genetics and epigenetics in the pathogenesis of Alzheimer's disease and frontotemporal dementia. *J Alzheimers Dis*. 2018;62(3):913-932. https://doi.org/10.3233/jad-170702

41. McAlpine CS, Park J, Griciuc A, et al. Astrocytic interleukin-3 programs microglia and limits Alzheimer's disease. *Nature*. 2021;595(7869):701-706. https://doi.org/10.1038/s41586-021-03734-6
42. Chung J, Das A, Sun X, et al. Genome-wide association and multi-omics studies identify MGMT as a novel risk gene for Alzheimer's disease among women. *Alzheimers Dement*. 2022;https://doi.org/10.1002/alz.12719
43. Buniello A, MacArthur JAL, Cerezo M, et al. The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res*. 2019;47(D1):D1005-D1012. https://doi.org/10.1093/nar/gky1120
44. Gao F, Chang D, Biddanda A, et al. XWAS: A software toolset for genetic data analysis and association studies of the X chromosome. *J Hered*. 2015;106(5):666-671. https://doi.org/10.1093/jhered/esv059
45. Binder LI, Guillozet-Bongaarts AL, Garcia-Sierra F, Berry RW. Tau, tangles, and Alzheimer's disease. *Biochim Biophys Acta*. 2005;1739(2-3):216-223. https://doi.org/10.1016/j.bbadis.2004.08.014
46. Bloom GS. Amyloid-β and tau: the trigger and bullet in Alzheimer disease pathogenesis. *JAMA Neurol*. 2014;71(4):505-508. https://doi.org/10.1001/jamaneurol.2013.5847
47. Zhang YW, Thompson R, Zhang H, Xu H. APP processing in Alzheimer's disease. *Mol Brain*. 2011;4:3. https://doi.org/10.1186/1756-6606-4-3

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Garcia AX, Xu J, Cheng F, Ruppin E, Schäffer AA. Altered gene expression in excitatory neurons is associated with Alzheimer's disease and its higher incidence in women. *Alzheimer's Dement*. 2023;9:e12373. https://doi.org/10.1002/trc2.12373