



HAL
open science

Differential transcript usage unravels gene expression alterations in Alzheimer's disease human brains

Diego Marques-Coelho, Lukas da Cruz Carvalho Iohan, Ana Raquel Melo de Farias, Amandine Flaig, Jean-Charles Lambert, Marcos Romualdo Costa

► **To cite this version:**

Diego Marques-Coelho, Lukas da Cruz Carvalho Iohan, Ana Raquel Melo de Farias, Amandine Flaig, Jean-Charles Lambert, et al.. Differential transcript usage unravels gene expression alterations in Alzheimer's disease human brains. *npj Aging and Mechanisms of Disease*, 2021, 7 (1), pp.2. 10.1038/s41514-020-00052-5 . hal-03670336

HAL Id: hal-03670336

<https://hal.science/hal-03670336>

Submitted on 17 May 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

ARTICLE OPEN



Differential transcript usage unravels gene expression alterations in Alzheimer's disease human brains

Diego Marques-Coelho^{1,2}, Lukas da Cruz Carvalho Iohan^{1,2}, Ana Raquel Melo de Farias^{1,3}, Amandine Flaig³ and The Brainbank Neuro-CEB Neuropathology Network*, Jean-Charles Lambert^{1,3} and Marcos Romualdo Costa^{1,3}✉

Alzheimer's disease (AD) is the leading cause of dementia in aging individuals. Yet, the pathophysiological processes involved in AD onset and progression are still poorly understood. Among numerous strategies, a comprehensive overview of gene expression alterations in the diseased brain could contribute for a better understanding of the AD pathology. In this work, we probed the differential expression of genes in different brain regions of healthy and AD adult subjects using data from three large transcriptomic studies: Mayo Clinic, Mount Sinai Brain Bank (MSBB), and ROSMAP. Using a combination of differential expression of gene and isoform switch analyses, we provide a detailed landscape of gene expression alterations in the temporal and frontal lobes, harboring brain areas affected at early and late stages of the AD pathology, respectively. Next, we took advantage of an indirect approach to assign the complex gene expression changes revealed in bulk RNAseq to individual cell types/subtypes of the adult brain. This strategy allowed us to identify previously overlooked gene expression changes in the brain of AD patients. Among these alterations, we show isoform switches in the AD causal gene amyloid-beta precursor protein (APP) and the risk gene bridging integrator 1 (BIN1), which could have important functional consequences in neuronal cells. Altogether, our work proposes a novel integrative strategy to analyze RNAseq data in AD and other neurodegenerative diseases based on both gene/transcript expression and regional/cell-type specificities.

npj Aging and Mechanisms of Disease (2021) 7:2; <https://doi.org/10.1038/s41514-020-00052-5>

INTRODUCTION

Changes in gene expression characterize a multitude of human diseases and have been successfully used to predict molecular and cellular mechanisms associated with pathological processes¹. Alzheimer's disease (AD) is the most prevalent type of dementia and causes a progressive cognitive decline, for which there is no effective treatment or cure. Although expression analyses in brain diseases are generally limited by tissue availability, RNA sequencing (RNAseq) data have been generated from *postmortem* brain samples of healthy and AD individuals^{2–4}. However, a comprehensive description of the gene expression alterations in the AD brain remains elusive.

Recent work has begun to address this important gap in the study of AD pathology using bulk brain tissue RNA sequencing (RNAseq)⁵ or single-cell RNA sequencing (scRNAseq)^{6,7}. However, these studies have focused on samples obtained from different brain regions, namely the dorsolateral prefrontal^{5,7} and entorhinal cortices⁶, which could lead to important discrepancies in the results. In fact, AD pathology shows a progressive impact on different brain regions, characterized at early stages by the presence of TAU protein inclusions in the locus coeruleus, the transentorhinal and entorhinal regions (stages I and II). This is followed by the presence of TAU inclusions in the hippocampal formation and some parts of the neocortex (stages III and IV), followed by large parts of the neocortex (stages V and VI)⁸. This temporal progression of AD pathology could differently impact gene expression in those brain areas. Accordingly, a recent study has shown that changes in protein expression are much more prominent in areas affected at early and intermediate stages, such as the hippocampus, entorhinal cortex, and cingulate cortex in the

temporal lobe, compared to other brain regions affected at later stages of AD pathology, such as sensory cortex, motor cortex, and cerebellum⁹.

Another important aspect to consider is the descriptive relevance of gene expression analysis based solely on the identification of differentially expressed genes (DEG), which fails to detect dynamics in the expression of multiple related transcripts¹⁰. Recently, new approaches using transcripts-level analysis, so-called differential transcript usage (DTU), enables identification of alternative splicing and isoform switches with the prediction of functional consequences^{11,12}. Therefore, important gene expression modifications in the AD brain could occur at the transcript level and be overlooked in classical DEG analyses.

Here, we took advantage of three available RNAseq datasets, generated using samples from different brain regions, to systematically probe gene expression changes (DEG and DTU) in AD. In Mayo's clinic study, both the temporal cortex and cerebellum were used to obtain bulk RNAseq². In the Religious Orders Study (ROS) and Memory and Aging Project (MAP), henceforth called ROSMAP, the dorsolateral prefrontal cortex was used³. Finally, in the Mount Sinai/JJ Peters VA Medical Center Brain Bank (MSBB), four different Brodmann areas of the brain were studied: areas 22 and 36 from the temporal lobe, areas 10 and 44 in the frontal lobe⁴. We also added another level of complexity using an indirect approach to assign DEGs and gDTUs to unique cell types in order to identify AD gene expression signatures for neural cells, microglia, and endothelial cells. Finally, we linked these alterations with AD causal and risk genes, identifying novel isoform switches in BIN1 and APP genes of potential functional consequences for pathology progression.

¹Brain Institute, Federal University of Rio Grande do Norte, Av. Nascimento de Castro, 2155 Natal, Brazil. ²Bioinformatics Multidisciplinary Environment (BioME), Federal University of Rio Grande do Norte, Natal, Brazil. ³Unité INSERM 1167, RID-AGE-Risk Factors and Molecular Determinants of Aging-Related Diseases, Institut Pasteur de Lille, University of Lille, Lille Cedex, France. *A list of authors and their affiliations appears at the end of the paper. ✉email: marcos.costa@pasteur-lille.fr

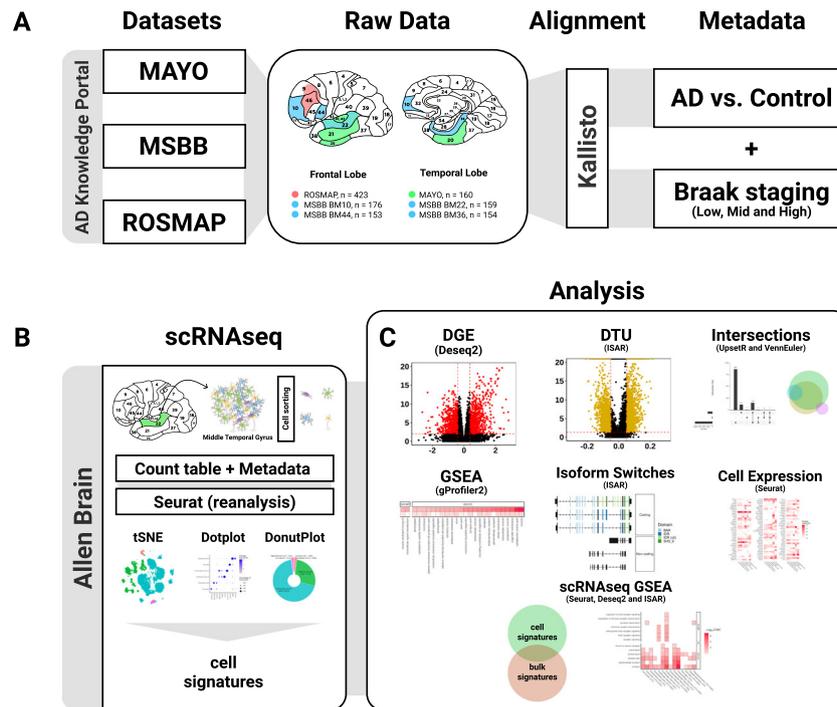


Fig. 1 Schematic summary of the methodology. **A** Datasets obtained from three consortia (Mayo, MSBB, and ROSMAP) were grouped according to the brain region sampled in the frontal lobe (FL) or temporal lobe (TL). Next, RNAseq data were pseudo-aligned using Kallisto. Clinico-pathological classifications were included as metadata. **B** scRNAseq data from the middle temporal gyrus (MTG, Allen Brain Atlas) were analyzed using the R package SEURAT. **C** Gene expression analyses were performed using the R packages DESeq2, IsoformSwitchAnalyzeR (ISAR), and gene set enrichment analysis (GSEA). Assignment of differentially expressed genes or isoform switches to specific cell types/subtypes was performed indirectly using scRNAseq signatures obtained from the MTG (B).

RESULTS

Regional gene expression alterations in the AD brain correlates with pathological progression

Several consortia have generated RNAseq data from brains of individuals with a clinical and/or pathological diagnostic of AD^{2–4}. Considering the regional progression of AD pathology⁸, we set out to identify and compare differentially expressed genes (DEG) in the temporal lobe (TL), encompassing brain regions affected at early stages of the AD such as the hippocampus and entorhinal cortex, and in the frontal lobe (FL), affected at more advanced stages of the pathology (Fig. 1). Comparisons between control and AD individuals were performed independently for each dataset and only genes with fold change >1.3 and FDR > 0.01 were considered as DEGs. We found 3348 (1244 down- and 2104 upregulated genes) and 2172 (1170 down and 999 upregulated genes in BM22 and BM36; three genes regulated in opposite directions in these two areas) DEGs in the TL of AD individuals compared to their respective controls in the MSBB_TL and Mayo datasets, respectively (Fig. 2A, B and Supplementary Table 1). Of those DEGs, 734 genes (145 down and 520 up) were commonly regulated in both Mayo and MSBB_TL (88.4% of genes altered in the same direction; 15.33% of overlap; $P = 8.56 \times 10^{-59}$, hypergeometric test). In contrast, only 327 (113 down and 214 up) and 209 (97 down and 112 up) DEGs were detected in the MSBB_FL and ROSMAP, respectively. Of those, 31 genes (18 down and 13 up) were found in both datasets (7.34% of overlap; $P = 1.67 \times 10^{-14}$, hypergeometric test) (Fig. 2A, B and Supplementary Table 1). This small number of DEGs in the FL is in agreement with previous data obtained from the DLPFC (106 down- and 158 upregulated genes with FC > 1.3)¹³. Among DEGs detected in the FL, 62.5% were also detected in the TL (Fig. 2B), suggesting that similar molecular changes occur in these brain areas, but at different stages of the disease progression. The differences in the number of DEGs

detected in the FL and TL can neither be attributed to lack of statistical power nor potential biases due to tissue processing, since the number of samples in the FL is larger than in the TL groups (Fig. 1) and differences are observed even in samples obtained from the same donors (compare MSBB_TL and MSBB_FL in Fig. 2). Thus, changes in gene expression are much more prominent in brain areas affected at the early stages of AD pathology.

To select genes consistently altered in AD brains, considering the several sources of measurement variations in RNAseq experiments¹⁴, we decided to focus only on DEGs replicated in at least two independent datasets obtained from related brain areas. This resulted in a set of 734 DEGs detected in both Mayo and MSBB TL (temporal lobe intersection—TLI), and 31 DEGs shared between ROSMAP and MSBB FL (frontal lobe intersection—FLI) (Supplementary Table 2). Among TLI DEGs, we observed ABCA1 and 2 (ATP-binding cassette subfamily A member 1 and 2), primarily involved in the maintenance of normal brain homeostasis and associated with AD and other neurological diseases¹⁵; Complement C1R and C1S, involved in the immune/inflammatory response and previously shown to be upregulated in the brain of a 3 × Tg mouse model of AD when Aβ plaques start to accumulate¹⁶; RE1 silencing transcription factor (REST), which regulates neural circuit activity during aging¹⁷; glutamate decarboxylase 1 and 2 (GAD1 and 2), solute carrier family 32 GABA vesicular transporter, member 1 (SLC32A1), calbindin 1 (CALB1), parvalbumin (PVALB), somatostatin (SST), and vasoactive intestinal peptide (VIP), all expressed in GABAergic neurons and involved in cognitive decline in AD and other neurological diseases¹⁸. Among the few DEGs common to TLI and FLI, we observed a significant downregulation of the neurosecretory protein VGF (VGF nerve growth factor inducible), recently suggested as a key regulator of Alzheimer's disease¹⁹.

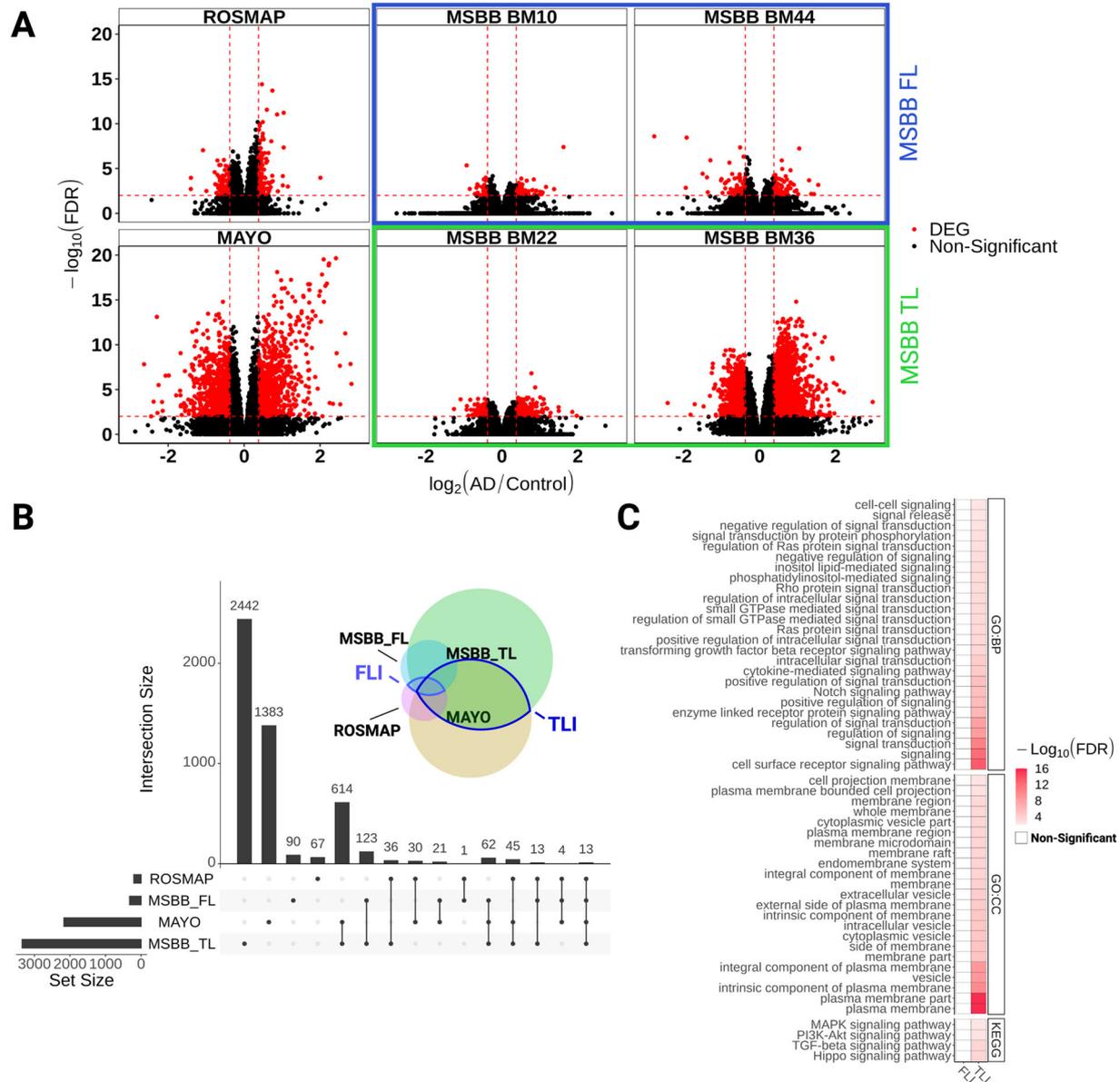


Fig. 2 Gene expression alterations are more prominent in the temporal than the frontal lobe of AD patients. **A** Volcano plots showing differentially expressed genes (DEG, red dots; FC > 1.3 and FDR < 0.01) in the frontal lobe (ROSMAP and MSBB FL - BM10 and BM44) and temporal lobe (MAYO and MSBB TL - BM22 and BM36). **B** Upset plot showing the total number of DEGs identified in each dataset (horizontal bars) and the number of DEGs exclusive of one dataset (first four vertical bars) or shared by different datasets (other vertical bars). Black dots below vertical bars indicate datasets quantified. Venn diagram illustrates the same results in colors and circle sizes. **C** Gene ontology terms enriched for DEGs identified in the TL or FL intersections (TLI and FLI, respectively).

Next, we used gene set enrichment analyses (GSEA) to assess the functional profile of the DEGs identified in our analysis. Again, we used only genes commonly altered in two datasets (TLI or FLI) to avoid inaccurate results associated with the use of large gene sets in functional analysis²⁰. We found that TLI DEGs were significantly enriched for terms (GO:BP, GO:CC and KEGG) associated with generic biological processes, such as cell-signaling pathways and cell-cell signaling, whereas the small number of DEGs in the FLI were not significantly enriched for any term (Fig. 2C and Supplementary Table 3). The limited number of significant gene set enrichment observed in our analysis after inputting DEGs is in disagreement with results reported by Canchi et al.¹³. This discrepancy can likely be explained by the use of stringent criteria to detect TLI DEGs in our study (only genes detected in at least two independent datasets with FC > 1.3 and

FDR < 0.01), which significantly reduce the number of genes used in the GSEA.

Differential transcript usage analysis reveals novel genes associated with AD pathology

Gene-level expression analysis lacks the sensitivity to detect possible changes at the transcript-level caused, for example, by alterations in alternative splicing^{10,21}. To overcome this limitation, we used differential transcript usage (DTU) analysis to identify additional alterations of gene expression in the AD brains compared to controls. We observed 2509 and 1843 genes with differential transcript usage (gDTU) in the temporal lobe of AD brains studied in the MAYO and MSBB datasets, respectively (Fig. 3A, B and Supplementary Table 1). Similar to what we observed for DEGs, a much smaller number of gDTUs were

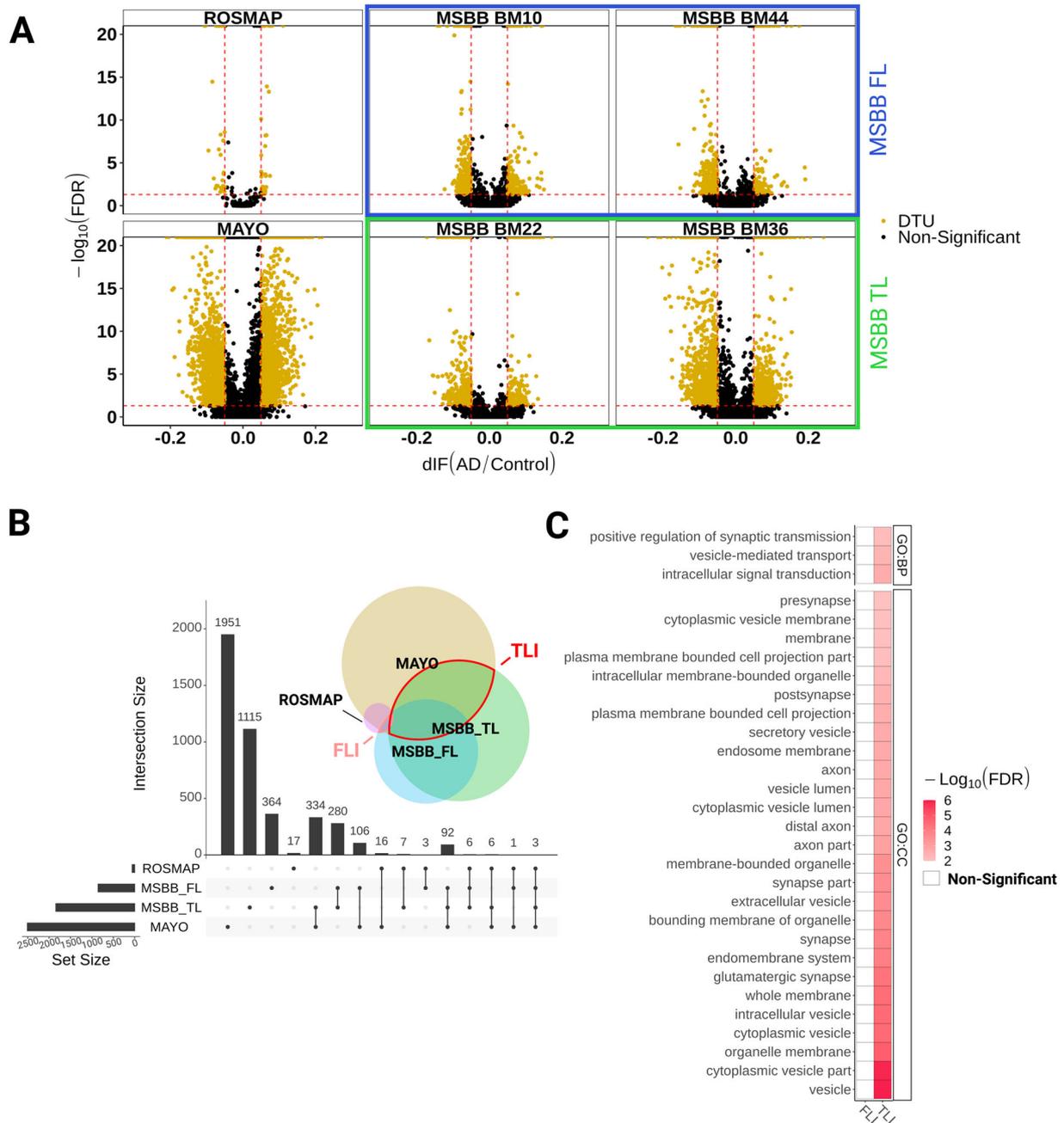


Fig. 3 Differential transcript usage analysis identifies gene expression alterations in AD associated with synapse transmission. **A** Volcano plots showing genes with differential transcript usage (gDTU, yellow dots; Differential isoform fraction (dIF) >0.05 and FDR <0.05) in the frontal lobe (ROSMAP and MSBB_FL - BM10 and BM44) and temporal lobe (Mayo and MSBB_TL - BM22 and BM36). **B** Upset plot showing the total number of gDTUs identified in each dataset (horizontal bars) and the number of gDTUs exclusive of one dataset (first four vertical bars) or shared by different datasets (other vertical bars). Black dots below vertical bars indicate datasets quantified. Venn diagram illustrates the same results in colors and circle sizes. **C** Synapse-related terms enriched for gDTUs in the TLI are not observed in the FLI.

detected in the frontal lobe, both in ROSMAP and MSBB studies (59 and 855 genes with transcripts altered, respectively). We found 435 gDTUs in TLI (11.1% of overlap; $P = 6.16 \times 10^{-25}$, hypergeometric test) and 13 gDTUs in FLI (1.47% of overlap; $P = 2.56 \times 10^{-3}$, hypergeometric test) (Supplementary Table 2). In TLI, most gDTUs did not overlap with DEGs (TL—34 gDTUs that are DEGs out of 435 gDTUs, Fig. 4A), whereas in FLI, we found no overlap at all. Consistent with this small overlap, GSEA using only DEGs, only gDTUs or both showed complementary results (Fig. 4B). GSEA using gDTUs (alone or in combination with DEGs) showed significant enrichment for vesicle-mediated transport and other

synapse-related terms, which were not observed while inputting only DEGs (Figs. 3C and 4B; Supplementary Table 3). The functional enrichment annotation using both DEGs and gDTUs is in agreement with previous studies using scRNAseq to identify gene expression alterations in unique cell types^{6,7} and clearly improves the annotation observed using only DEGs, suggesting that the use of DTU analysis could contribute to unraveling gene expression alterations overlooked in the classical DEG analysis.

Among genes with isoform switches enriched in synaptic-related terms, we observed the AD causal gene APP, previously associated with regulation of synapse transmission and long-term

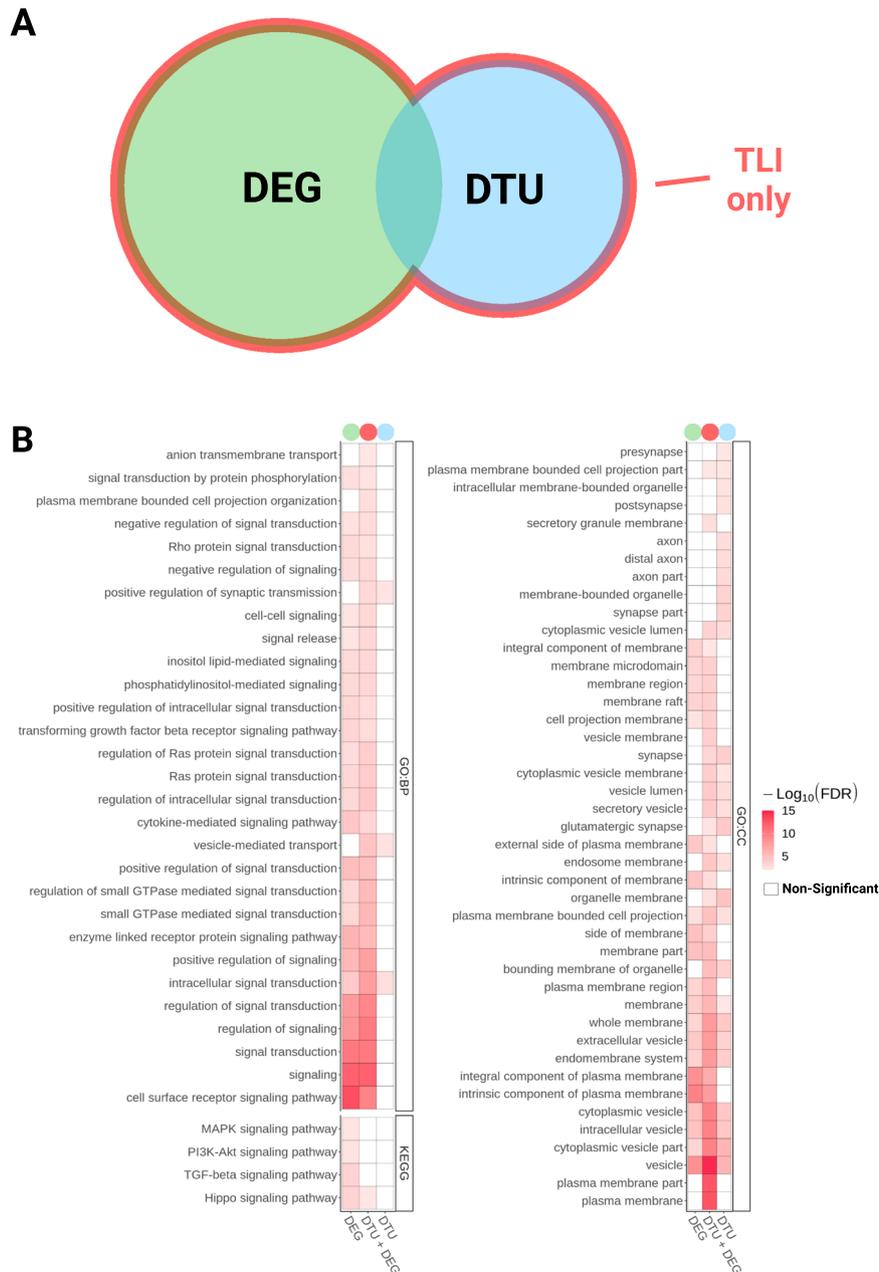


Fig. 4 Differential transcript usage analysis in AD brains reveals gene expression alterations overlooked in DEG analysis. **A** Venn diagram showing DEGs and gDTUs identified in the TLI. **B** Comparison of GO and KEGG terms enriched for DEG, gDTU, or DEG + gDTU identified in the TLI.

plasticity in AD²²; neuronal vesicle trafficking associated 1 (NSG1), which has been implicated in the regulation of AMPA receptors (AMPA) and APP trafficking, thus affecting synaptic transmission, plasticity, and A β production^{23,24}; RELN (Reelin) that plays important role in synaptic transmission and has been associated with AD²⁵; gamma-aminobutyric acid type A receptor subunit alpha 1 (GABRA1), which encodes for a subunit of the main ionotropic GABA receptor in the brain and has previously been shown to be downregulated in the AD brain²⁶.

Alternative splice events in AD brains and functional consequences

To identify the causes subjacent to gene isoform switches in the AD brain, we quantified the frequency of splicing events

associated with the isoform switches detected in AD compared to control brains (Fig. 5 and Supplementary Table 4). We found that alternative transcription start site (ATSS), alternative transcription termination site (ATTS), and exon skipping (ES) were the most frequent splicing events in AD brains (Fig. 5B). Other common splicing events observed were alternative 3' or 5' splice sites (A3 and A5, respectively), multiple exons skipping (MES) and intron retention (IR) (Fig. 5B). These observations suggest that changes in alternative splicing could be implicated in AD pathogenesis, corroborating previous analyses in the ROSMAP cohort using intronic usage ratios to identify abnormal splicing events in the AD brain⁵.

Alternative splicing events may have diverse functional consequences for protein expression, such as shifting the frequency of transcripts containing introns (noncoding) or mRNA

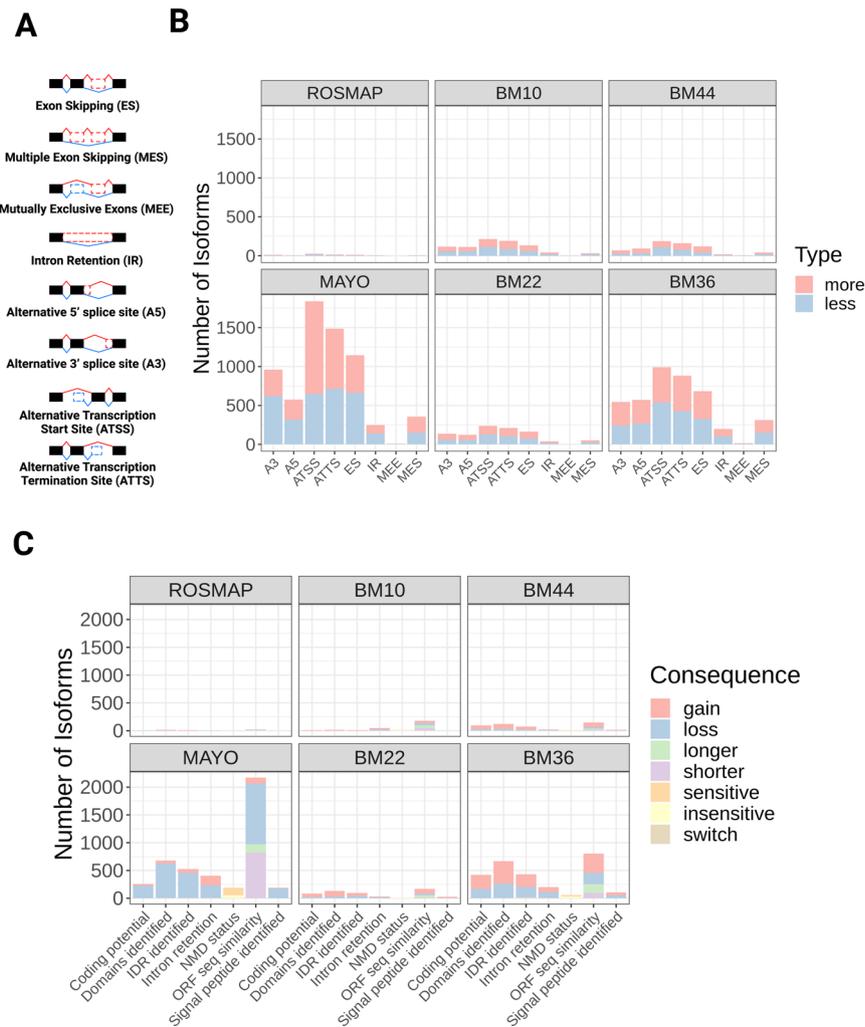


Fig. 5 Alternative splicing mechanisms associated with isoform switches and consequences for protein expression. **A** Schematic showing different splicing events that can lead to gene isoform switches. **B** Quantification of the number of isoforms showing more or less splicing events in AD compared to controls for each dataset. **C** Quantification of the number of isoforms showing (i) gain or loss of coding potential, domains/signal peptides identified, intrinsically disordered regions (IDR), intron retention, open-reading frame (ORF) sequencing similarity; (ii) switch (simultaneous gain and loss) of domains identified or IDR; (iii) sensitive or insensitive to nonsense-mediated decay (NMD); and (iv) longer or shorter ORF sequencing similarity.

stability (nonsense-mediated decay) or leading to gain/loss of protein domains, intrinsically disordered regions, or signaling peptides¹². Quantification of these consequences revealed some interesting differences between Mayo and MSBB BM36 (Fig. 5C), the two datasets with the largest numbers of gDTUs. Whereas in the Mayo dataset, a high number of isoforms showed loss of coding potential and protein domains, in the MSBB BM36 isoforms showed an even distribution of loss and gain of coding potential or protein domains (Fig. 5B). These differences could be at least partly explained by the larger number of gDTUs detected in the Mayo compared to MSBB TL (Fig. 3) and are likely related to the different median read depth of these datasets (Mayo—12.58 billion bases; MSBB BM22—3.23 billion bases; MSBB BM36—3.56 billion bases)²⁷.

Differential expression of genes involved in alternative splicing correlates with isoform switches during disease progression

To evaluate whether the emergence of gDTUs could be correlated with AD pathology hallmarks, we quantified the total of gDTUs observed at different disease stages in the MSBB dataset using the Braak classification (Fig. 6 and Supplementary Table 5). For this

purpose, we subdivided samples into three groups: low Braak (0, 1, and 2)—196 samples (clinical diagnosis: 15 AD and 181 controls); mid-Braak (3 and 4)—133 samples (clinical diagnosis: 58 AD and 75 controls); and high Braak (5 and 6)—308 samples (clinical diagnosis: 305 AD and 3 controls). Next, we evaluated the number of gDTUs when comparing individuals at these different stages (Fig. 6). We observed that most gDTUs were detected only while comparing high with either low or mid-Braak stages (Fig. 6A–D). This pattern was observed both in the FL (BM10 and BM44) and TL (BM22 and BM36), suggesting that gene isoform switches positively correlate with AD pathology progression.

Next, we set out to evaluate alterations in the expression of genes encoding for proteins of the splicing machinery between the same Braak stages. We found that among 441 genes related to “splicing” or “spliceosome” terms (Supplementary Table 6), 79 were DEGs at high compared to low or mid-Braak stages (Fig. 6E). In contrast, we could not detect any DEG in the comparison of mid vs low Braak stages. Among DEGs detected in the comparison between high and low/mid-Braak stages, we observed that several genes specifically associated with the neuronal splicing regulatory network²⁸, such as RBF1 and 2 (RNA binding Fox-1 homolog 1 and 2), ELAVL2 (ELAV like RNA binding protein 2), MBNL3

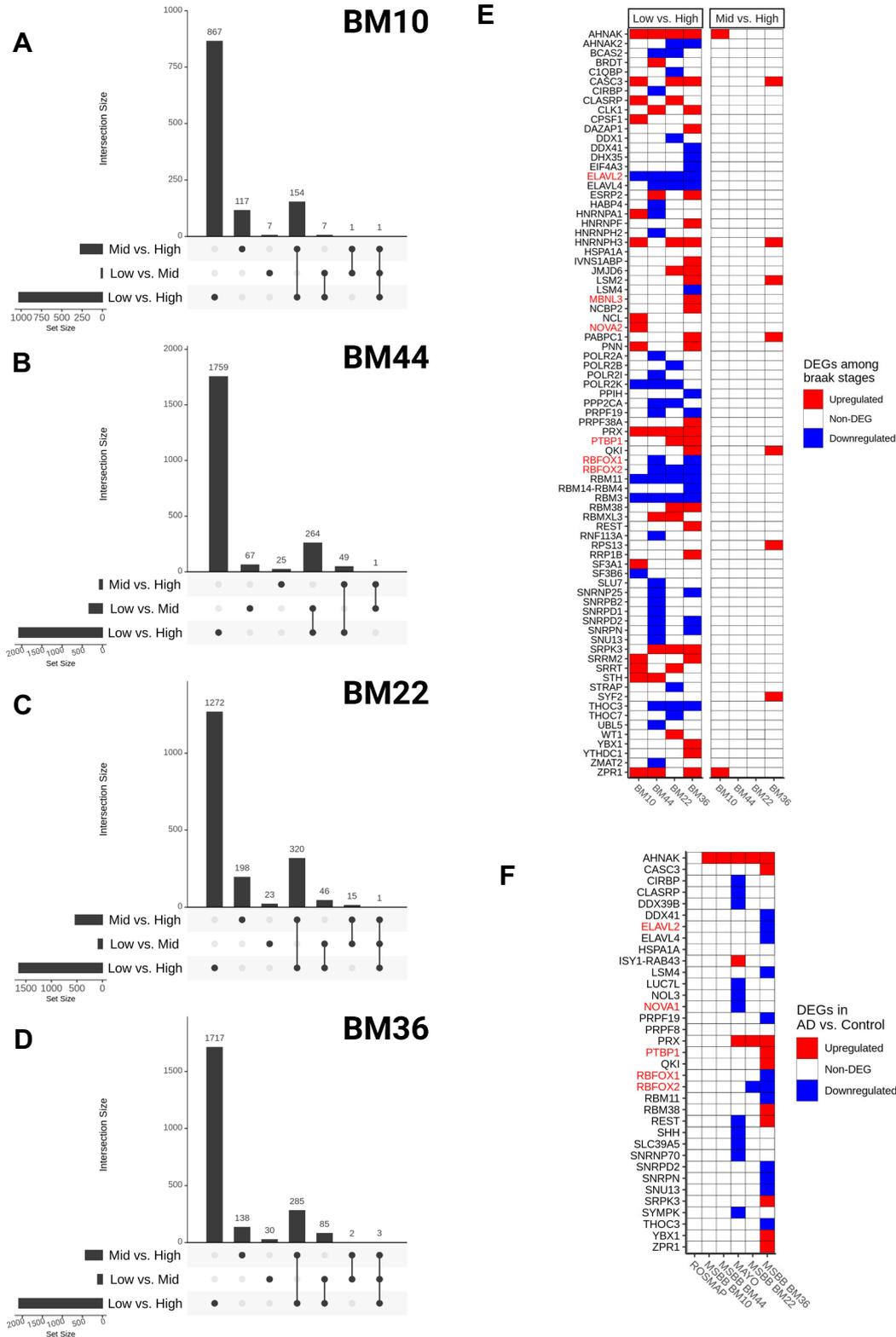


Fig. 6 Coincidence between altered expression of splicing-related genes and gDTUs in advanced pathologic stages of AD. **A–D** Upset plots showing the total number of gDTU identified in the comparison between different Braak stages (low vs. high, low vs. mid, and mid vs. high) in BM10 (**A**), BM44 (**B**), BM22 (**C**), or BM36 (**D**). Horizontal bars show the total number of gDTUs identified in each comparison (low vs. high, low vs. mid, and mid vs. high), whereas vertical bars indicate the gDTUs exclusive or common to different comparisons. Black dots below vertical bars indicate stages analyzed. **E, F** Differential expression of genes associated with splicing/spliceosome after comparison of different Braak stages (**E**) or AD vs controls in different datasets (**F**). Red and blue squares indicate, respectively, up- and downregulated genes. Gene symbols highlighted in red indicate genes belonging to the neuronal splicing machinery.

(muscleblind like splicing regulator 3), PTBP1 (polypyrimidine tract binding protein 1), and NOVA2 (NOVA alternative splicing regulator 2) (Fig. 6E, highlighted in red). A similar correlation between pathological burden and differential expression of the same 441 splice-related genes was observed in the comparison between all AD versus control subjects of the different datasets (Fig. 6F). Changes in the expression of those genes were hardly observed in FL (low number of gDTUs—Fig. 3), but were frequent in TL samples (high number of gDTUs—Fig. 3), albeit to a lesser extent than that observed in the comparison between different Braak stages (likely due to the effects of combining low, mid and high Braak stages in the AD group). Remarkably, the majority of the splicing-related genes with altered expression in the Mayo dataset was not reproduced in the MSBB BM36 dataset, and vice versa (Fig. 6F). This could help to explain the dissimilar consequences of alternative splicing events observed in those datasets (Fig. 5C) and suggest that a myriad of proteins could be involved in altered splicing in the AD brains.

Differential gene expression in separate cell types of the human brain

Considering the cellular diversity in the brain, we took an indirect approach to sort DEGs and gDTUs according to individual cell types. To that, we used scRNAseq data obtained from the adult human brain to identify cell types expressing the genes altered in our DEG/gDTU analysis (Fig. 7 and Supplementary Fig. 1). We found that, out of the 1135 genes with altered expression, i.e., gDTU + DEG, in the TLI (Figs. 2 and 3), 839 were found in at least one cell-type using as cutoff the expression in more than 10% of cells assigned for a specific cell-type (Supplementary Table 7). From these, 239 were identified in unique cell-types/subtypes, 396 in multiple (2–4 cell-types), and 211 in all cell-types analyzed (Fig. 7A, Supplementary Figs. 1 and 2, Supplementary Table 7). Confirming the efficacy of our strategy, GO analyses using cell-type-specific genes revealed that DEGs/gDTUs in the TLI of AD patients were significantly enriched for biological processes associated with inflammation in microglial cells, whereas those associated with cell adhesion were enriched in endothelial cells (Fig. 7B and Supplementary Table 8). Similarly, DEGs/gDTUs identified in neuronal cells were enriched for GO terms such as synaptic signaling, synaptic plasticity, and synapse vesicle cycle (Fig. 7C). Notably, these enrichments were more significant in GABAergic neurons, which could suggest a more pronounced pathological burden on these cells compared to glutamatergic neurons (Fig. 7C). Comparison of the cell-type gene expression signatures identified in our work with previous studies using scRNAseq in AD^{6,7} showed a similar degree of overlap (Supplementary Figs. 3 and 4; Supplementary Table 9), further supporting the effectiveness of our strategy to assign gene expression alterations to unique cell types in the AD brain.

DEG/gDTU analyses identify cell-type-specific alterations in AD risk/causal genes

Genomic association studies have revealed about 45 loci containing variants related to an increased or decreased probability of developing AD^{29,30}. However, the functional variants and their target genes remain mostly elusive³¹. To contribute to the identification of target genes, we first evaluate the expression of 176 genes located within the 45 loci associated with the AD risk (Supplementary Table 10)³¹ and 3 causal AD genes—PSEN1, PSEN2, and APP—in individual cell types of the adult human brain. We found that 116 out of the 179 AD risk/causal genes were expressed by at least one of the major cell types identified in the brain (Fig. 8A and Supplementary Table 11). Subsets of these genes were exclusively expressed either in microglial cells (14 out of 116), neurons (12), astrocytes (2), oligodendrocytes (6), or

endothelial cells (6), suggesting cell-type specific roles for these AD risk/causal genes.

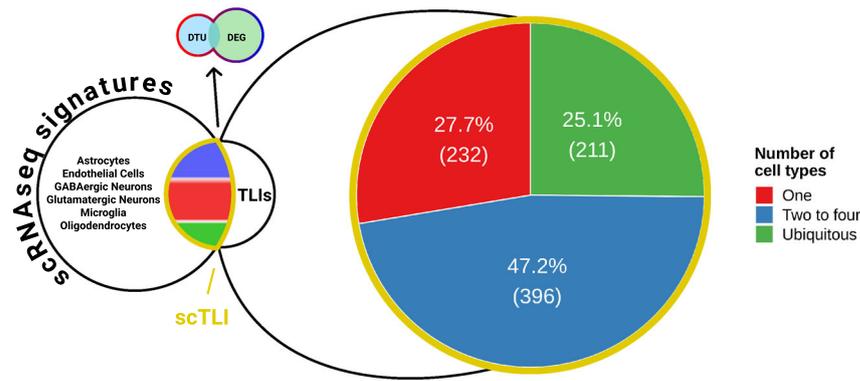
Next, we set out to evaluate the differential expression or transcript usage for these genes. Out of the 116 AD risk/causal genes expressed by brain cell types (Fig. 8A), we observed that 54 were also DEGs/gDTUs in at least one of the bulk RNAseq datasets analyzed. Among those genes, two were exclusively identified in the FL (Fig. 8B). We, therefore, decided to focus on the 52 AD risk/causal genes identified in the temporal lobe for further analyses. In this region, we identified 27 and 17 DEGs/gDTUs in the MSBB_TL and Mayo datasets, respectively, including some well-characterized AD risk genes, such as ADAM10 (ADAM metallopeptidase domain 10), BIN1, CLU (Clusterin), and TREM2 (triggering receptor expressed on myeloid cells 2), and the causal AD genes APP, PSEN1 and 2 (presenilin 1 and 2) (Fig. 8A, B). Eight genes were altered in both datasets (Fig. 8B, yellow box; 15,38% of overlap) and were selected for further analysis of isoform switch. Using ISAR to identify the isoforms altered in the AD brains compared to controls, we observed some patterns of isoform switch that could have important functional relevance (Fig. 8C, D). For instance, while BIN1 transcripts ENST00000316724.9 (NP_647593.1—isoform 1) and ENST00000409400.1 (NP_647600.1—isoform 9) were downregulated, transcripts ENST00000393040.7 (NP_647598.1—isoform 6) and ENST00000462958.5, ENST0000046611.5 and ENST00000484253.1 (intron retention) were upregulated (Fig. 8C). This pattern could lead to a decrease of the neuronal-specific BIN1 isoform 1 expression³², given that retained introns are noncoding sequences. Using western blotting analysis, we confirmed this decrease of BIN1 isoform 1 protein in the frontal cortex and hippocampus of AD brain samples compared to controls (Supplementary Fig. 5).

We also observed isoform switches in the AD causal gene APP with possible functional consequences in neuronal cells. While two APP isoforms were downregulated (ENST00000348990 and ENST00000354192), the isoforms ENST00000346798 and ENST00000357903 were upregulated in Mayo and MSBB datasets (Fig. 8D). Noteworthy, significantly downregulated APP isoforms lack exon 7, which contains the Kunitz protease inhibitor (KPI) domain. KPI is one of the main serine protease inhibitors and increased KPI(+) APP mRNA and protein expression levels have been described in AD brains and are associated with increased amyloid-beta deposition^{33–35}. At the exception of ENST00000354192, the other transcripts are mostly expressed in neurons (Marques-Coelho and Costa, unpublished data), indicating that these cells may have a selective increase in the expression of KPI(+)-APP and, consequently, enhanced production of A β 1–42.

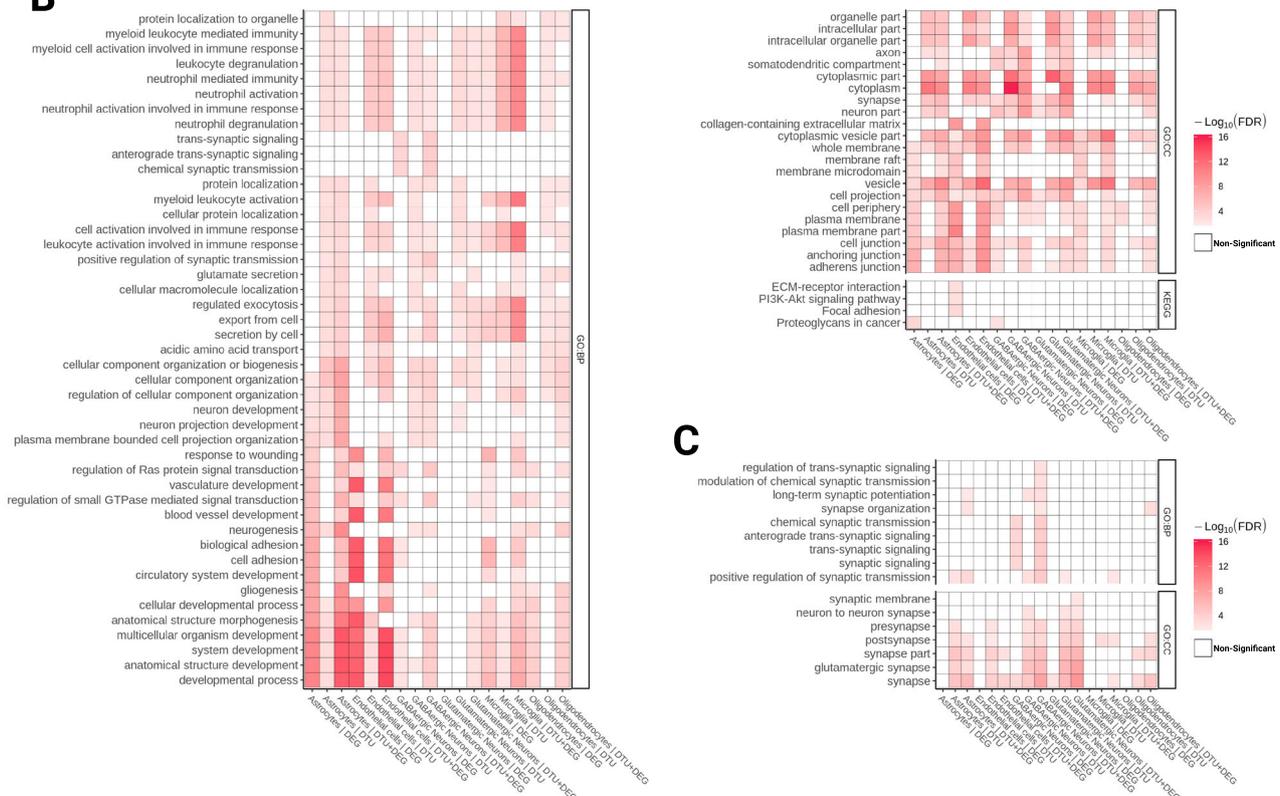
DISCUSSION

Comprehensive knowledge of gene expression alterations associated with the onset and progression of human diseases is a key step toward the understanding of their cellular and molecular mechanisms³⁶. In this work, we provide a novel framework to identify cell-type-specific gene expression alterations in AD using patient-derived bulk RNAseq. Comparing RNA sequencing data obtained from distinct brain regions of control and AD patients, we show that changes in gene expression are more significant in the temporal than frontal lobe. We also show that a large number of genes present isoform switches without changes in the global expression levels. As a consequence, these genes are overlooked in classical differential expression analysis but can be detected through differential transcript usage analysis. Gene isoform switches are mostly evident at late stages of the pathology and correlate with altered expression of genes encoding for splicing-related proteins. Using an indirect approach to assign genes to unique cell types, we are also able to map DEGs/gDTUs to unique cell populations of the adult brain, and our results are comparable to previously published scRNAseq data^{6,7}. Finally, we show that a

A



B



C

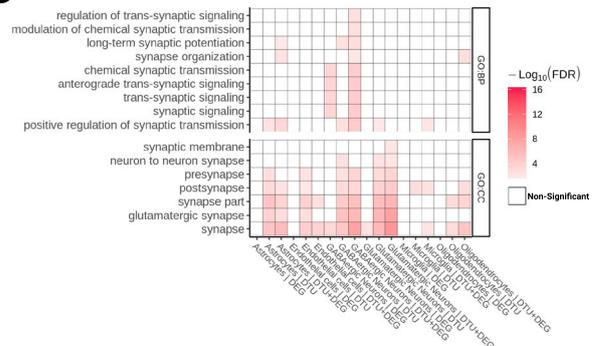


Fig. 7 Cell-type expression pattern for genes altered in AD brains. **A** Schematic representation showing our strategy to assign DEGs and gDTUs identified in the TLI to specific cell types of the adult human brain (see also Supplementary Fig. 1). Out of 839 single-cell TLI genes (scTLI), 281 were expressed in a unique cell-type, 249 in 2–4 cell-types, and 77 in all cell-types/subtypes analyzed. **B** Gene ontology terms enriched for scTLI DEGs, gDTUs, or both per cell type. **C** Selected GO terms associated with synaptic transmission.

subset of AD causal/risk factors such as APP or BIN1 is differentially expressed in the AD brain. Altogether, our work provides a comprehensive description of regional and cell-type-specific gene expression changes in the AD brain and suggests that alternative splicing could be an important mechanism for pathological progression.

Despite the availability of RNAseq datasets generated from healthy subjects and AD patients^{2–4}, a systematic evaluation of the gene expression changes in the AD brain, as well as comparisons of these changes in distinct brain regions, was missing. To the best of our knowledge, only one study aimed at comparing gene expression levels in different AD brain regions³⁷, but this work was based on microarray data which has limited gene coverage.

We show, using bulk tissue RNAseq data, that alterations in gene expression are highly prominent in biological samples obtained from the temporal lobe, which harbors the first brain regions affected in the AD pathogenesis⁸. Conversely, few changes are present in biological samples derived from the frontal lobe, where cells are affected only at advanced stages of AD. These observations are in line with recent data showing that changes in protein expression levels in AD brains are much more prominent in the temporal lobe (hippocampus, entorhinal cortex, and cingulate gyrus) than in the frontal lobe (motor cortex)⁹. They can also help to explain the low number of DEGs identified in scRNAseq data obtained from the frontal lobe⁷ compared to a similar study in the entorhinal cortex⁶.

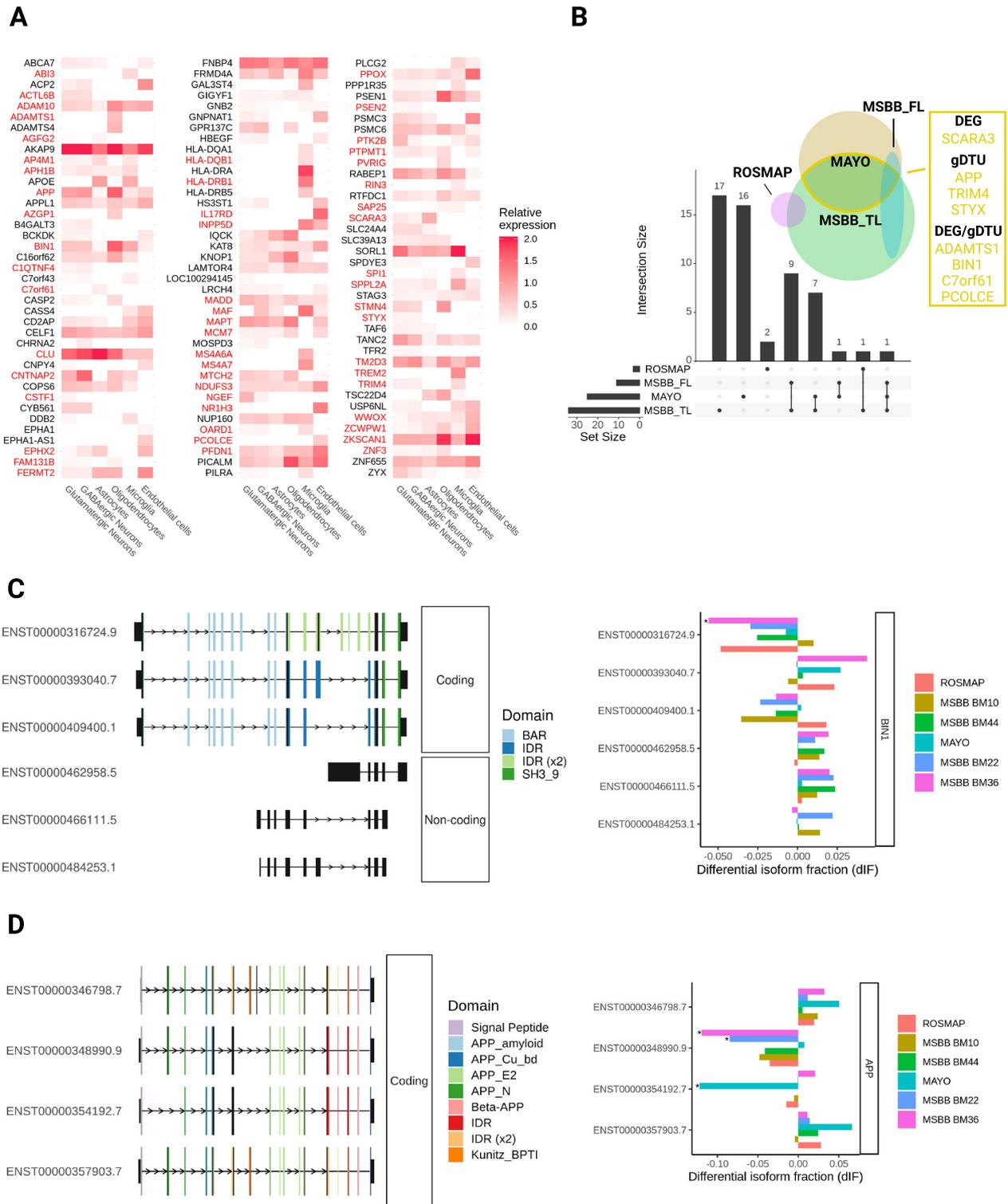


Fig. 8 Expression of AD risk/causal genes is mostly altered in the TL of patients. A Heatmap showing the expression of predicted AD risk/causal genes in different cell types of the adult human brain. DEGs and gDTUs in at least one dataset are highlighted in red. **B** Venn diagram showing the number of AD risk/causal DEGs or gDTUs identified in the different datasets analyzed. The intersection between Mayo and MSBB TL is highlighted in yellow, and genes identified are shown in the yellow box. **C** Representation of the 6 most significant BIN1 isoforms altered (left) and quantification of the differential isoform fraction (dIF) in AD brains compared to controls (right). Main protein domains are indicated with different colors. **D** Similar representation for APP. * dIF >0.05 and FDR >0.01.

In order to minimize the variability in RNAseq experiments³⁶, we here focused on DEGs (genes with FC > 1.3 and FDR < 0.01 in AD versus control) detected independently in at least two datasets containing samples of similar brain regions (TLI or FLI). These stringent criteria limited the number of DEGs used in subsequent analyses, but still allowed the uncovering of several genes previously associated with AD pathologies, such as ABCA1, ABCA2, CALB1, C1R, C1S, GAD1/2, PVALB, REST, SLC32A1, SST, VGF, and VIP^{15–19}. The reduced number of DEGs in FLI and TLI likely explains our failure to detect functional annotations associated with synaptic transmission and immune response in GSEA, as previously reported¹³. However, this study analyzed only the ROSMAP dataset and considered genes with FDR < 0.05 as significant, regardless of the fold change, identifying 1722 DEGs in AD versus control brains. Besides the questionable meaning of DEGs with very small fold changes, the use of such a large set of genes for GSEA can artificially increase the number of significantly enriched functional annotations and is not advised²⁰.

Nevertheless, our failure to detect key functional annotations associated with AD pathology while inputting TLI DEGs is puzzling and could suggest that DEG analysis fails to detect relevant alterations in gene expression in the AD brain. Indeed, classical DEG analysis using DESeq or edgeR, which rank all gene transcripts, including noncoding sequences³⁸, are insensitive to the dynamics of gene expression that could, for example, lead to isoform switches with important functional consequences²¹. Therefore, important gene expression alterations could occur at the level of transcripts, without significant changes in the global expression of genes. According to this possibility, we provide convincing evidence that a high number of genes in the AD brain show isoform switches (DTU) but are not detected by DEG analysis, including several genes associated with the regulation of synapse transmission, such as APP, NSG1, RELN, GABRA1^{22–26}. Moreover, gDTUs identified in two independent datasets (TLI), alone or in combination with TLI DEGs, were enriched for key biological processes involved in AD pathogenesis, such as synaptic communication, immune response, inflammation, endocytosis, and cell-signaling³⁹. Similar gene set enrichment has been described using the analysis of co-expression modules in bulk RNAseq^{27,40} or DEG analysis DEGs in unique cell types in scRNAseq^{6,7}. This could suggest that the combination of DEG and DTU to analyze bulk RNAseq is comparable to scRNAseq regarding the sensitivity to detect gene expression alteration in AD brains. In agreement with this possibility, we were able to assign DEGs and gDTUs to unique cell types and confirm the similarities among cell-type-specific functional annotations observed in our work compared to previous scRNAseq studies^{6,7}.

Notably, we show that several DEGs/gDTUs associated with AD pathogenesis, such as NSG1, CALB1, RELN^{23–25,41} are exclusively assigned to GABAergic neurons. These genes may be particularly relevant for AD pathogenesis, given the central role of GABAergic neurons for the generation of oscillatory rhythms, network synchrony, and memory in different animal models of AD⁴². Isoform switches in the APP gene could particularly affect GABAergic neurons, which express high levels of that gene, contributing to AD pathogenesis. According to this possibility, conditional knockout of APP/APLP2 only in GABAergic forebrain neurons using DlxCre mice leads to cognitive deficits in hippocampus-dependent spatial learning and memory tasks, associated with altered neuronal morphology and synaptic plasticity⁴³. It is tempting to speculate that GABAergic neurons could be particularly vulnerable in AD, contributing to the increased neuronal activity and synapse downscaling observed in AD brains^{39,44}.

The high number of gDTUs observed in AD brains compared to controls can likely be explained by altered expression of genes encoding for proteins of the splicing machinery, affecting

alternative splicing. According to this interpretation, we show that a high number of isoform switches is associated with alternative transcription start site, alternative transcription termination site, exon skipping, alternative 3' or 5' splice sites, multiple exon skipping and intron retention. Moreover, we show that several genes encoding for proteins of the splicing machinery have their expression altered in AD brains, especially those showing a high degree of pathology (Braak > 5). Also in agreement with the regional differences in gene expression described above, alterations in the splicing machinery are more prominent in the TL than in the FL, which could help to explain the low number of gDTUs in the latter brain region identified in our work and in the previous study using a different strategy to detect isoform switch⁵.

Particularly interesting, several genes encoding for proteins involved in the control of alternative splicing in neurons are differently expressed in the TL of AD brains. For instance, RBFOX1 and 2 are downregulated in the MSBB BM36 and could contribute to the altered rate of exon skipping observed in this region^{28,45}. Noteworthy, reduced expression of RBFOX1 has been associated with increased inclusion of exon 7 in the APP gene, leading to an enhanced expression of APP isoforms 770 and 751 containing the KPI domain⁴⁵. A similar switch in the APP isoforms has also been associated with somatic gene recombination in AD⁴⁶, indicating that increased ratios of APP isoforms containing the KPI domain could be detrimental to neurons. Considering these findings and the well-established associations between KPI(+)-APP expression levels, amyloid plaque deposition, and AD pathology progression^{33–35}, it is tempting to speculate that controlling APP isoform switches by manipulating RBFOX family proteins could be a potential therapeutic strategy to hamper disease progression.

Altered exon skipping could also help to explain the isoform switch observed for BIN1, which is a major risk factor for AD^{29,30}. BIN1 comprises a N-BAR domain involved in membrane curvature sensing, an SH3 domain that binds to proline-rich motifs, and a clathrin-binding domain (CLAP) specific of the neuronal isoform 1³². We show that the transcript encoding for this latter isoform is significantly reduced in the temporal lobe, suggesting that expression of BIN1 isoform 1 in neurons could be reduced. This observation is in line with decreased BIN1 isoform 1 protein expression in the AD brain compared with controls (our own results)⁴⁷. This would be also in agreement with the observation that overexpression of the BIN1 isoform 1 may be protective in a model of Tauopathy⁴⁸.

Although we cannot formally rule out that a stage-dependent increase in the number of DEGs and gDTUs could be due to the loss of neuronal cells in brain regions affected by the pathology, several lines of evidence indicate that this is not the most parsimonious explanation for the data described here. First, we observe that the percentage of up- and downregulated genes in GABAergic and glutamatergic neurons are close to 50%, ruling out the possibility that changes in cell numbers could explain these changes. Secondly, previous scRNAseq studies in AD observed a consistent fraction of cell types isolated across control and AD individuals^{6,7}, ruling out significant changes in the cellular composition of AD brains. Lastly, a large number of genes with total expression levels unchanged but presenting isoform switches in the AD brains may likely presuppose a steady cellular composition of the tissue.

Altogether, our work proposes a novel strategy to analyze bulk RNAseq data and identify meaningful gene expression alterations in the diseased brain. It also corroborates previous work implicating alternative splicing in AD susceptibility⁵ and suggests that isoform switches in the gene BIN1 are involved in the reduced expression of the main neuronal BIN1 isoform 1 in AD brains.

Table 1. Summary of clinical, demographic, and technical variables of samples analyzed from different datasets.

Diagnosis	Sex	<i>n</i>	Braak	AOD	CDR	RIN	PMI
<i>Mayo-TL</i>							
AD	F	49	5.55 ± 0.53	83.33 ± 7.23	–	8.58 ± 0.58	6.34 ± 5.8
AD	M	33	5.42 ± 0.55	81.67 ± 8.16	–	8.59 ± 0.5	8.6 ± 5.83
Control	F	37	2.03 ± 0.78	84.81 ± 8.04	–	7.59 ± 1	7 ± 7.69
Control	M	41	2.36 ± 0.87	80.27 ± 9.13	–	7.64 ± 1.05	5.51 ± 6.42
<i>MSBB-BM10</i>							
AD	F	86	5.18 ± 1.21	85.94 ± 6	3.17 ± 1.35	6.05 ± 1.64	6.01 ± 3.97
AD	M	53	5.48 ± 0.99	79.26 ± 7.67	3.23 ± 1.28	6.19 ± 1.42	7.38 ± 4.73
Control	F	52	2.02 ± 1.09	83.04 ± 7.43	0.81 ± 1.1	6.37 ± 1.28	8.69 ± 5.73
Control	M	42	1.62 ± 0.94	78.93 ± 8.56	1.23 ± 1.45	6.73 ± 1.14	11.31 ± 7.47
<i>MSBB-BM22</i>							
AD	F	88	5.25 ± 1.2	85.81 ± 5.95	3.17 ± 1.39	6 ± 1.5	5.34 ± 3.71
AD	M	64	5.41 ± 0.97	79.44 ± 7.67	3.48 ± 1.35	5.53 ± 1.28	7.2 ± 4.8
Control	F	45	2 ± 1.26	82.09 ± 8.1	0.94 ± 1.29	6 ± 1.23	9.18 ± 6.18
Control	M	41	1.72 ± 0.97	79.29 ± 8.66	1.22 ± 1.42	5.72 ± 1.06	11.01 ± 7.29
<i>MSBB-BM36</i>							
AD	F	95	5.35 ± 1.1	86.79 ± 5.2	3.22 ± 1.28	5.59 ± 1.71	5.96 ± 4.14
AD	M	53	5.47 ± 0.91	79.26 ± 7.98	3.42 ± 1.41	5.69 ± 1.61	6.65 ± 4.9
Control	F	44	2.14 ± 1.15	83.45 ± 7.23	0.86 ± 1.28	6.15 ± 1.32	8.3 ± 5.66
Control	M	40	1.65 ± 0.92	78.28 ± 8.79	0.85 ± 1.18	6.2 ± 1.2	12.47 ± 7.92
<i>MSBB-BM44</i>							
AD	F	84	5.2 ± 1.29	85.62 ± 6.02	3.11 ± 1.41	6.9 ± 2.51	5.74 ± 3.82
AD	M	51	5.49 ± 0.98	79.69 ± 7.59	3.2 ± 1.25	6.57 ± 2.41	7.19 ± 4.58
Control	F	45	2.11 ± 1.19	81.82 ± 9.06	0.93 ± 1.26	7.43 ± 2.15	8.19 ± 5.74
Control	M	47	1.82 ± 0.95	78.13 ± 8.16	0.96 ± 1.22	7.35 ± 2.14	11.48 ± 7.3
<i>ROSMAP-DLPFC</i>							
AD	F	147	4.18 ± 0.93	88.55 ± 3.05	–	6.97 ± 0.88	7.25 ± 4.51
AD	M	64	3.92 ± 1.28	87.6 ± 3.18	–	7 ± 1.09	6.58 ± 4.24
Control	F	118	3.04 ± 1.07	85.66 ± 4.83	–	7.31 ± 0.99	7.25 ± 5.18
Control	M	74	2.58 ± 1.4	83.69 ± 5.69	–	7.19 ± 1.08	7.34 ± 4.05

AOD age of death, CDR clinical dementia rate, *n* number of samples, RIN RNA integrity number, PMI postmortem interval (in hours), AD Alzheimer disease, F female, M, male.

Values are mean ± SD. See supplementary Table 13 for individual values.

METHODS

Bulk RNAseq data from human control and disease banks

RNAseq datasets obtained from different brain regions were used (Mayo²; MSBB³; ROSMAP⁴). Datasets were downloaded from AMP-AD Knowledge Portal (<https://www.synapse.org>) following all terms and conditions to use the data. The brain area analyzed and the number of individuals per condition was the following: Mayo - Temporal cortex, which neuroanatomically subdivides into the inferior, middle, and superior temporal gyri (STG), and cytoarchitectonically can be subdivided into Brodmann areas (BM, instead of BA) 20/21/22/41/42⁴⁹, *N* = 160 subjects (82 AD and 78 controls); MSBB - BM22, which is part of the Wernicke's area in the STG, *N* = 159 subjects (98 AD and 61 controls); MSBB BM36, corresponding to the lateral perirhinal cortex, *N* = 154 subjects (88 AD and 64 controls); MSBB BM10, corresponding to the anterior prefrontal cortex, *N* = 176 subjects (105 AD and 71 controls); MSBB BM44, corresponding to the inferior frontal gyrus, *N* = 153 subjects (90 AD and 63 controls); and ROSMAP - Dorsolateral prefrontal cortex (DLPFC), containing BM46 and part of BM9, *N* = 423 subjects (222 AD and 201 controls). Unless stated otherwise, data obtained from different analyses were grouped in "temporal lobe" (TL) - Mayo, MSBB BM22 and MSBB BM26; or "frontal lobe" (FL) - ROSMAP, MSBB BM10 and MSBB BM44.

Metadata obtained from each study was used to classify patients into Control and Alzheimer's disease groups (Supplementary Table 12). Briefly, for the MSBB dataset, we used patients with Neuropathology Category (NP.1) labeled as "Control" and "definitive Alzheimer". For the Mayo

dataset, we used the "Diagnosis" column of the metadata, selecting only "AD" and "Control" patients. For the ROSMAP dataset, we also used the column "Diagnosis" of the metadata, selecting only "Control" (value = 1) and "Alzheimer with no other conditions" (value = 4). In all those datasets, subjects marked as "AD" showed Braak stage values higher than 4. In the MSBB dataset, CDR scores of AD patients were consistently higher than 2. In the Mayo and ROSMAP datasets, all AD patients had also a definitive diagnosis according to NINCDS criteria. Covariates such as "Postmortem interval (PMI)", "RNA integrity number (RIN)", "Age of death", and "Sex" were balanced among the different groups (Table 1; Chi-square test, *p* > 0.05). We used RIN and PMI as covariates in our model to control for possible "batch effects" (linear regression). For detailed information of all individual samples used in this study, please refer to Supplementary Table 13.

For single-cell RNA sequencing (scRNAseq) analyses, we used processed data obtained from the middle temporal gyrus (MTG), available at the Allen Brain Atlas consortium (<https://celltypes.brain-map.org/rnaseq>).

Realignment of human reads into single pseudoaligner pipeline
Using human GRCh38 cDNA release 94 (<ftp://ftp.ensembl.org/pub/release-94>) as a reference, we built an index to align all our fastq files. Next, we used pseudoaligner Kallisto⁵⁰ (version 0.43.1) with our pre-built index to align fastq files.

Differential gene expression analyses. Differentially expressed genes (DEGs) were identified using differential gene expression at transcript-level using DESeq2 R library^{51,52}. To facilitate kallisto output import, transcript-level estimated counts, length, and abundance were extracted using *tximport* function⁵³. As described by Michael Love group, transcript-level differential gene expression enhances analysis resolution⁵². Using *DESeqDataSetFromTximport*, a DESeq2 object was created and filtered using rows with sum of all counts bigger than 10. Next, *DESeq* function was used with default values. Using the *results* function, we selected all genes with a false discovery rate (FDR) < 0.01 and fold change (FC) > 1.3. We also used RIN and PMI as covariates (linear regression).

Differential transcript usage (DTU) analysis was performed using the R library *IsoformSwitchAnalyzeR*¹². Following pipeline instructions, kallisto abundance tables were imported using *importIsoformExpression* and *importRdata* functions to create a *switchAnalyzeRList* object. Same cDNA release used in kallisto alignment and correspondent annotation (ftp://ftp.ensembl.org/pub/release-94/gtf/homo_sapiens/Homo_sapiens.GRCh38.94.chr_patch_hapl_scaff.gtf.gz) were applied as input. We filtered data using a gene expression cutoff = 10, isoform expression cutoff = 3, differential isoform fraction (dIF) cutoff = 0.05 and removed single isoform genes. Although DEXSeq is recommended to test differential isoform usage, it does not work efficiently for large datasets (more than 100 samples)¹¹. For that reason, we chose *isoformSwitchAnalysisPart1* function using DRIM-Seq⁵³ to test differential transcript usage. Using part1 fasta files, all external analysis was performed and used as input to *isoformSwitchAnalysisPart2* function. We used CPC2, Pfam, SignalIP and Netsurfp2 as indicated in the pipeline. Next, we performed a confirmation stage using stageR¹⁴ to generate isoforms overall false discovery rate (OFDR). We selected all isoforms with OFDR < 0.01 and dIF > 0.05. RIN and PMI metadata were used as covariates (linear regression).

Statistical significance of the intersections among different datasets was calculated using the hypergeometric test (phyper).

Splicing events and event consequences

We used *extractSplicingSummary* and *extractConsequenceSummary* functions to quantify gain/loss of predicted splicing events (such as exon skipping and intron retention); and gain/loss (also sensitive/insensitive, shorter/longer and switch) of predicted functional consequences (such as coding potential and domain identified), respectively.

Single-cell RNAseq

Using R library *seurat*⁵⁴, we created a *seurat* object (*CreateSeuratObject*), normalized data (*NormalizeData*), found variable genes (*FindVariableFeatures*), and rescaled data using a linear model (*ScaleData*, *use.umi = F*). After that, we generated 50 PCs (*RunPCA*) but only used 35 of them based on the PC's visualization distribution (*ElbowPlot*). Since Allen data were already annotated, we only used tSNE (*RunTSNE*) to facilitate visualization. A group classified as "None" by Allen metadata were removed from our analysis. This strategy generated 7 main different cell types: Astrocytes, Endothelial cells, Glutamatergic Neurons, GABAergic Neurons, Microglia, Oligodendrocytes and oligodendrocyte precursor cells (OPCs). To assign genes to specific cell types, we used the *AverageExpression* function. Using *pct.exp* bigger than 0.1, we created a list of genes that were expressed by each cell type.

Gene set enrichment analysis (GSEA)

For gene ontology analysis, R library *gprofiler2*⁵⁵ was used. Using *gost* function, *correction_method = "fdr"* and *significant = TRUE*. To minimize the enrichment of gene ontologies based on a small set of genes, we used three conditions for significance assessment: false discovery rate (FDR) < 0.01; intersection size (intersection between gene set vs. a number of genes in a term) > 3; and precision (intersection size divided by gene set) > 0.03. We used Gene Ontology (GO or by branch GO:MF, GO:BP, GO:CC), Kyoto Encyclopedia of Genes and Genomes (KEGG), Reactome (REAC), WikiPathways (WP), TRANSFAC (TF), miRTarBase (MIRNA), Human Protein Atlas (HPA), CORUM (CORUM), Human phenotype ontology (HP) as sources. For improved visualization, we plotted results only for GO:BP, GO:CC and KEGG and show only FDR related to terms reaching all criteria of significance.

Selection of splicing-associated genes

To select splicing-related genes, we searched for terms containing the words "splicing" or "spliceosome" in gProfiler bank (<https://biit.cs.ut.ee/gprofiler/gost>). Taking only GO and WP datasets, 25 terms and 441 genes related to those terms were selected (Supplementary Table 6).

Selection of AD risk/causal genes

The complete list of AD risk/causal genes used in this study is described in Supplementary Table 10. Briefly, AD risk loci were selected from previous work using genome-wide association studies and whole exome sequencing^{29,30}. AD risk genes within these loci were determined based on regional association plots, assuming that the functional risk variants are located in the vicinity of the SNP producing the top signal and taking into account the linkage disequilibrium patterns and the recombination peaks within the loci of interest³¹. Early-onset AD causal genes used in this study are APP, PSEN1, and PSEN2.

Western blotting

Frozen brain samples obtained from the frontal cortex (FCx) and hippocampus (hip) of three non-pathology (age: 80.33 ± 3.78 years; Braak: 2.66 ± 1.15; PMI: 37.33 ± 22.50 h) and six AD patients (age: 79.57 ± 6.70 years; Braak: 6; PMI: 26.57 ± 13.40 h) were lysed with RIPA buffer and sonicated at 100% during 10 s before use for the western blotting analyses. The controls for BIN1 isoforms 1 (Iso1) and 9 (Iso9) were obtained using HEK cells transiently transfected with 1 µg/ml DNA solution containing plasmids encoding for BIN1 isoforms mixed with the transfection reagent FuGENE HD (Promega) at the ratio 1:3. Cells were lysed using RIPA buffer 48 h after transfection and frozen for further analyses.

Protein quantification was performed using the BCA protein assay (Thermo Scientific). In total, 10–20 µg of total protein from extracts were separated in SDS–polyacrylamide gels 4–12% (NuPAGE Bis-Tris, Thermo Scientific) and transferred to nitrocellulose membranes (Bio-Rad). Next, membranes were incubated in milk (5% in Tris-buffered saline with 0.1% Tween-20, TTBS) to block non-specific binding sites during 1 h at RT, followed by several washes with TTBS. Immunoblotting was carried out with primary antibodies anti-BIN1 (Abcam, ab182562), anti-β-ACTIN (Sigma-Aldrich, A1978), and anti-GAPDH (Millipore, AB2302) overnight at 4 °C on 20 RPM. The membranes were washed three times in TTBS, followed by incubation with HRP-conjugated secondary antibodies (Jackson, anti-mouse 115-035-003 and anti-rabbit 111-035-003; Thermo Scientific, anti-chicken A16054) overnight at 4 °C on 20 RPM agitation. The membranes were washed three times in TTBS, and the immunoreactivity was revealed using the ECL chemiluminescence system (SuperSignal, Thermo Scientific) and imaged using the Amersham Imager 600 (GE Life Sciences). Optical densities of bands were quantified using "Gel Analyzer" plugin in Fiji⁵⁶.

All western blot experiments were performed in compliance with relevant guidelines provided by the Neuro-CEB - Biological Resources Platform and the protocols were approved by the National Institute of Health and Medical Research (INSERM), at the Institut Pasteur de Lille, University of Lille.

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

DATA AVAILABILITY

MSBB RNAseq³: <https://doi.org/10.7303/syn3388564>; ROSMAP RNAseq⁴: <https://doi.org/10.7303/syn3157743>; MAYO RNAseq²: <https://doi.org/10.7303/syn5550404>.

CODE AVAILABILITY

Codes used for analyses are available, differentially expressed genes and differential transcript usage in individual datasets can be assessed using <https://diegomscelho.github.io/AD-IsoformSwitch/>.

Received: 5 May 2020; Accepted: 12 November 2020
Published online: 04 January 2021

REFERENCES

1. Masters, C. L. et al. Alzheimer's disease. *Nat. Rev. Dis. Primers* **1**, 15056 (2015).
2. Allen, M. et al. Human whole genome genotype and transcriptome data for Alzheimer's and other neurodegenerative diseases. *Sci. Data* **3**, 160089 (2016).
3. De Jager, P. L. et al. Data descriptor: a multi-omic atlas of the human frontal cortex for aging and Alzheimer's disease research. *Sci. Data* **5**, 180142–180142 (2018).
4. Wang, M. et al. The Mount Sinai cohort of large-scale genomic, transcriptomic and proteomic data in Alzheimer's disease. *Sci. Data* **5**, 1–16 (2018).
5. Raj, T. et al. Integrative transcriptome analyses of the aging brain implicate altered splicing in Alzheimer's disease susceptibility. *Nat. Genet.* **50**, 1584–1592 (2018).
6. Grubman, A. et al. A single-cell atlas of entorhinal cortex from individuals with Alzheimer's disease reveals cell-type-specific gene expression regulation. *Nat. Neurosci.* **22**, 2087–2097 (2019).
7. Mathys, H. et al. Single-cell transcriptomic analysis of Alzheimer's disease. *Nature* **570**, 332–337 (2019).
8. Braak, H. & Braak, E. Neuropathological staging of Alzheimer-related changes. *Acta Neuropathologica* **82**, 239–259 (1991).
9. Xu, J. et al. Regional protein expression in human Alzheimer's brain correlates with disease severity. *Commun. Biol.* **2**, 1–15 (2019).
10. Yi, L., Pimentel, H., Bray, N. L. & Pachter, L. Gene-level differential analysis at transcript-level resolution. *Genome Biol.* **19**, 53 (2018).
11. Anders, S., Reyes, A. & Huber, W. Detecting differential usage of exons from RNA-seq data. *Genome Res.* **22**, 2008–2017 (2012).
12. Vitting-Seerup, K. & Sandelin, A. IsoformSwitchAnalyzeR: analysis of changes in genome-wide patterns of alternative splicing and its functional consequences. *Bioinformatics* **35**, 4469–4471 (2019).
13. Canchi, S. et al. Integrating gene and protein expression reveals perturbed functional networks in Alzheimer's disease. *Cell Rep.* **28**, 1103–1116.e4 (2019).
14. Van den Berge, K., Soneson, C., Robinson, M. D. & Clement, L. stageR: a general stage-wise method for controlling the gene-level false discovery rate in differential expression and differential transcript usage. *Genome Biol.* **18**, 151 (2017).
15. Abuznait, A. H. & Kaddoumi, A. Role of ABC transporters in the pathogenesis of Alzheimers disease. *ACS Chem. Neurosci.* **3**, 820–831 (2012).
16. Benoit, M. E. et al. C1q-induced LRP1B and GPR6 proteins expressed early in Alzheimer disease mouse models are essential for the C1q-mediated protection against amyloid- β neurotoxicity. *J. Biol. Chem.* **288**, 654–665 (2013).
17. Zullo, J. M. et al. Regulation of lifespan by neural excitation and REST. *Nature* **574**, 359–364 (2019).
18. Prévot, T. & Sibille, E. Altered GABA-mediated information processing and cognitive dysfunctions in depression and other brain disorders. *Mol. Psychiatry* 1–17 (2020) <https://www.nature.com/articles/s41380-020-0727-3>.
19. Beckmann, N. D. et al. Multiscale causal networks identify VGF as a key regulator of Alzheimer's disease. *Nat. Commun.* **11**, 18 (2020).
20. Subramanian, A. et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* **102**, 15545–15550 (2005).
21. Vitting-Seerup, K. & Sandelin, A. The landscape of isoform switches in human cancers. *Mol. Cancer Res.* **15**, 1206–1220 (2017).
22. Kamenetz, F. et al. APP processing and synaptic function. *Neuron* **37**, 925–937 (2003).
23. Alberi, S. et al. The endosomal protein NEEP21 regulates AMPA receptor-mediated synaptic transmission and plasticity in the hippocampus. *Mol. Cell. Neurosci.* **29**, 313–319 (2005).
24. Norstrom, E. M., Zhang, C., Tanzi, R. & Sisodia, S. S. Identification of NEEP21 as a β -amyloid precursor protein-interacting protein in vivo that modulates amyloidogenic processing in vitro. *J. Neurosci.* **30**, 15677–15685 (2010).
25. Yu, N. N., Tan, M. S., Yu, J. T., Xie, A. M. & Tan, L. The role of reelin signaling in Alzheimer's disease. *Mol. Neurobiol.* **53**, 5692–5700 (2016).
26. Limon, A., Reyes-Ruiz, J. M. & Miledi, R. Loss of functional GABA A receptors in the Alzheimer diseased brain. *Proc. Natl Acad. Sci. USA* **109**, 10071–10076 (2012).
27. Wan, Y. W. et al. Meta-analysis of the Alzheimer's disease human brain transcriptome and functional dissection in mouse models. *Cell Rep.* **32**, 107908 (2020).
28. Raj, B. & Blencowe, B. J. Alternative splicing in the mammalian nervous system: recent insights into mechanisms and functional roles. *Neuron* **87**, 14–27 (2015).
29. Lambert, J. C. et al. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat. Genet.* **45**, 1452–1458 (2013).
30. Kunkle, B. W. et al. Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates A β , tau, immunity and lipid processing. *Nat. Genet.* **51**, 414–430 (2019).
31. Dourlen, P., Kilinc, D., Malmanche, N., Chapuis, J. & Lambert, J. C. The new genetic landscape of Alzheimer's disease: from amyloid cascade to genetically driven synaptic failure hypothesis? *Acta Neuropathologica* **138**, 221–236 (2019).
32. Zhou, Y. et al. Intracellular clusterin interacts with brain isoforms of the bridging integrator 1 and with the microtubule-associated protein Tau in Alzheimer's disease. *PLoS ONE* **9**, e103187 (2014).
33. Kitaguchi, N., Takahashi, Y., Tokushima, Y., Shiojiri, S. & Ito, H. Novel precursor of Alzheimer's disease amyloid protein shows protease inhibitory activity. *Nature* **331**, 530–532 (1988).
34. Johnson, S. A., Rogers, J. & Finch, C. E. APP-695 transcript prevalence is selectively reduced during Alzheimer's disease in cortex and hippocampus but not in cerebellum. *Neurobiol. Aging* **10**, 755–760 (1989).
35. Tanzi, R. E. et al. Protease inhibitor domain encoded by an amyloid protein precursor mRNA associated with Alzheimer's disease. *Nature* **331**, 528–530 (1988).
36. Lee, T. I. & Young, R. A. Transcriptional regulation and its misregulation in disease. *Cell* **152**, 1237–1251 (2013).
37. Haroutunian, V., Katsel, P. & Schmeidler, J. Transcriptional vulnerability of brain regions in Alzheimer's disease and dementia. *Neurobiol. Aging* **30**, 561–573 (2009).
38. Costa-Silva, J., Domingues, D. & Lopes, F. M. RNA-Seq differential expression analysis: an extended review and a software tool. *PLoS ONE* **12**, e0190152 (2017).
39. Canter, R. G., Penney, J. & Tsai, L. H. The road to restoring neural circuits for the treatment of Alzheimer's disease. *Nature* **539**, 187–196 (2016).
40. Milind, N. et al. Transcriptomic stratification of late-onset Alzheimer's cases reveals novel genetic modifiers of disease pathology. *PLoS Genet.* **16**, e1008775 (2020).
41. Odero, G. L. et al. Evidence for the involvement of calbindin D28k in the presenilin 1 model of Alzheimer's disease. *Neuroscience* **169**, 532–543 (2010).
42. Verret, L. et al. Inhibitory interneuron deficit links altered network activity and cognitive dysfunction in alzheimer model. *Cell* **149**, 708–721 (2012).
43. Mehr, A. et al. Lack of APP and APLP2 in GABAergic forebrain neurons impairs synaptic plasticity and cognition. *Cereb. Cortex* **30**, 4044–4063 (2020).
44. Dörbaum, A. R., Alvarez-Castelao, B., Nassim-Assir, B., Langer, J. D. & Schuman, E. M. Proteome dynamics during homeostatic scaling in cultured neurons. *eLife* **9**, e52939 (2020).
45. Alam, S., Suzuki, H. & Tsukahara, T. Alternative splicing regulation of APP exon 7 by RBFOX proteins. *Neurochem. Int.* **78**, 7–17 (2014).
46. Lee, M. H. et al. Somatic APP gene recombination in Alzheimer's disease and normal neurons. *Nature* **563**, 639–645 (2018).
47. Glennon, E. B. C. et al. BIN1 is decreased in sporadic but not familial Alzheimer's disease or in aging. *PLoS ONE* **8**, e78806 (2013).
48. Sartori, M. et al. BIN1 recovers tauopathy-induced long-term memory deficits in mice and interacts with Tau through Thr348 phosphorylation. *Acta Neuropathologica* **138**, 631–652 (2019).
49. Strotzer, M. One century of brain mapping using Brodmann areas. *Clin. Neuroanatomy* **19**, 179–186 (2009).
50. Bray, N. L., Pimentel, H., Melsted, P. & Pachter, L. Near-optimal probabilistic RNA-seq quantification. *Nat. Biotechnol.* **34**, 525–527 (2016).
51. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
52. Love, M. I., Soneson, C. & Patro, R. Swimming downstream: statistical analysis of differential transcript usage following Salmon quantification. *F1000Research* **7**, 952–952 (2018).
53. Soneson, C., Love, M. I. & Robinson, M. D. Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences [version 2; referees: 2 approved]. *F1000Research* **4**, 1521 (2016).
54. Butler, A., Hoffman, P., Smibert, P., Papalexi, E. & Satija, R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* **36**, 411–420 (2018).
55. Kolberg, L., Raudvere, U., Kuzmin, I., Vilo, J. & Peterson, H. gprofiler2—an R package for gene list functional enrichment analysis and namespace conversion toolset g:Profiler. *F1000Research* **9**, 709 (2020).
56. Schindelin, J. et al. Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).

ACKNOWLEDGEMENTS

We thank Charles Duyckaerts (Neuroceb, GIE Neuro-CEB BB-0033-00011) for providing the brain samples. We would also like to thank Dr. Tarciso Velho, Pierre Dourlen, and Fabien Delahaye for reading and suggestions on the paper. The results published here are in whole or in part based on the data obtained from the AMP-AD Knowledge Portal (<https://adknowledgeportal.synapse.org/>). Study data were provided by the following sources: The Mayo Clinic Alzheimer's Disease Genetic Studies, led by Dr. Nilufer Taner and Dr. Steven G. Younkin, Mayo Clinic, Jacksonville, FL using

samples from the Mayo Clinic Study of Aging, the Mayo Clinic Alzheimer's Disease Research Center, and the Mayo Clinic Brain Bank. Data collection was supported through funding by NIA grants P50 AG016574, R01 AG032990, U01 AG046139, R01 AG018023, U01 AG006576, U01 AG006786, R01 AG025711, R01 AG017216, R01 AG003949, NINDS grant R01 NS080820, CurePSP Foundation, and support from Mayo Foundation. Study data include samples collected through the Sun Health Research Institute Brain and Body Donation Program of Sun City, Arizona. The Brain and Body Donation Program is supported by the National Institute of Neurological Disorders and Stroke (U24 NS072026 National Brain and Tissue Resource for Parkinson's Disease and Related Disorders), the National Institute on Aging (P30 AG19610 Arizona Alzheimer's Disease Core Center), the Arizona Department of Health Services (contract 211002, Arizona Alzheimer's Research Center), the Arizona Biomedical Research Commission (contracts 4001, 0011, 05-901, and 1001 to the Arizona Parkinson's Disease Consortium) and the Michael J. Fox Foundation for Parkinson's Research. Study data were also provided by the Rush Alzheimer's Disease Center, Rush University Medical Center, Chicago. Data collection was supported through funding by NIA grants P30AG10161 (ROS), R01AG15819 (ROSMAP; genomics and RNAseq), R01AG17917 (MAP), the Illinois Department of Public Health (ROSMAP), and the Translational Genomics Research Institute (genomic). For ROSMAP, additional phenotypic data can be requested at www.radc.rush.edu. Part based on the data generated from postmortem brain tissue collected through the Mount Sinai VA Medical Center Brain Bank and were provided by Dr. Eric Schadt from Mount Sinai School of Medicine. This work was co-funded by the European Union under the European Regional Development Fund (ERDF) and by the Hauts de France Regional Council (contrat no.18006176), the MEL (contract_2016_ESR_05), and the French State (contract no. 2018-3-CTRL_IPL_Phase2) to MRC. This work was also funded by the Lille Métropole Communauté Urbaine and the French government's LABEX DISTALZ program (Development of innovative strategies for a transdisciplinary approach to Alzheimer's disease) to J.C.L. D.M.C. and L.I.C.C. are supported by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) scholarships.

AUTHOR CONTRIBUTIONS

M.R.C. conceived of the presented idea and supervised the experiments. D.M.-C. performed all bioinformatic analyses and prepared the figures and tables of the paper. L.I.C.C. contributed with bioinformatic analyses. A.F. prepared protein lysates and A.R.M.F. performed western blot analysis in human samples. D.M.-C., J.-C.L., and

M.R.C. wrote the paper. All authors revised and approved the final version of the paper.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information is available for this paper at <https://doi.org/10.1038/s41514-020-00052-5>.

Correspondence and requests for materials should be addressed to M.R.C.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021

THE BRAINBANK NEURO-CEB NEUROPATHOLOGY NETWORK

Franck Letournel⁴, Marie-Laure Martin-Négrier⁵, Françoise Chapon⁶, Maxime Faisant⁶, Catherine Godfraind⁷, Claude-Alain Maurage⁸, Vincent Deramecourt⁸, Mathilde Duchesne⁹, David Meyronnet¹⁰, Nathalie Streichenberger¹⁰, André Mauès de Paula¹¹, Valérie Rigau¹², Fanny Vandembos-Burel¹³, Charles Duyckaerts¹⁴, Danielle Seilhean¹⁴, Serge Milin¹⁵, Dan Christian Chiforeanu¹⁶, Annie Laquerrière¹⁷, Florent Marguet¹⁷ and Béatrice Lannes¹⁸

⁴CHU Hôpital St Jean, Angers, France. ⁵CHRU Pellegrin, Bordeaux, France. ⁶CHU Côte de Nacre, Caen, France. ⁷CHU Gabriel Montpied, Clermont-Ferrand, France. ⁸CHRU Hôpital Roger Salengro, Lille, France. ⁹CHRU Hôpital Dupuytren, Limoges, France. ¹⁰Groupe Hospitalier Est. Hospices Civils de Lyon, Lyon, France. ¹¹CHU Hôpital de La Timone, APHM, Marseille, France. ¹²CHU de Montpellier, Montpellier, France. ¹³CHU Hôpital Pasteur, Nice, France. ¹⁴GH Pitié-Salpêtrière, APHP, Paris, France. ¹⁵CHU La Milétrie, Poitiers, France. ¹⁶CHU Hôpital Pontchaillou, Rennes, France. ¹⁷CHU Hôpital Charles Nicolle, Rouen, France. ¹⁸CHRU Hôpital de Haute-pierre, Strasbourg, France.