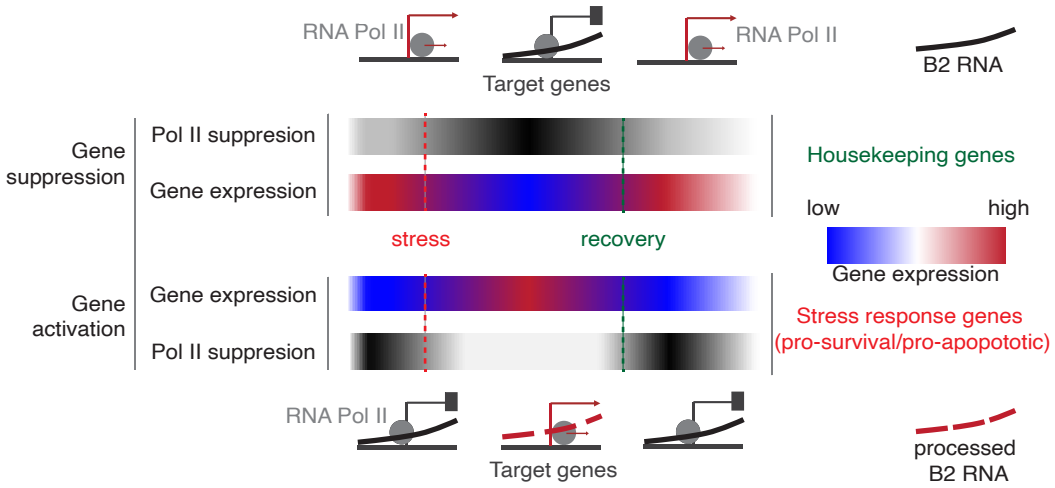
Increased human SINE Alu RNA processing in Alzheimer's disease brains.

Yubo Cheng, Luke Saville, Babita Gollen, Ana Alvarez Veronesi, Majid Mohajerani, Jeffrey T. Joseph, Athanasios Zovoilis

APPENDIX

Contents

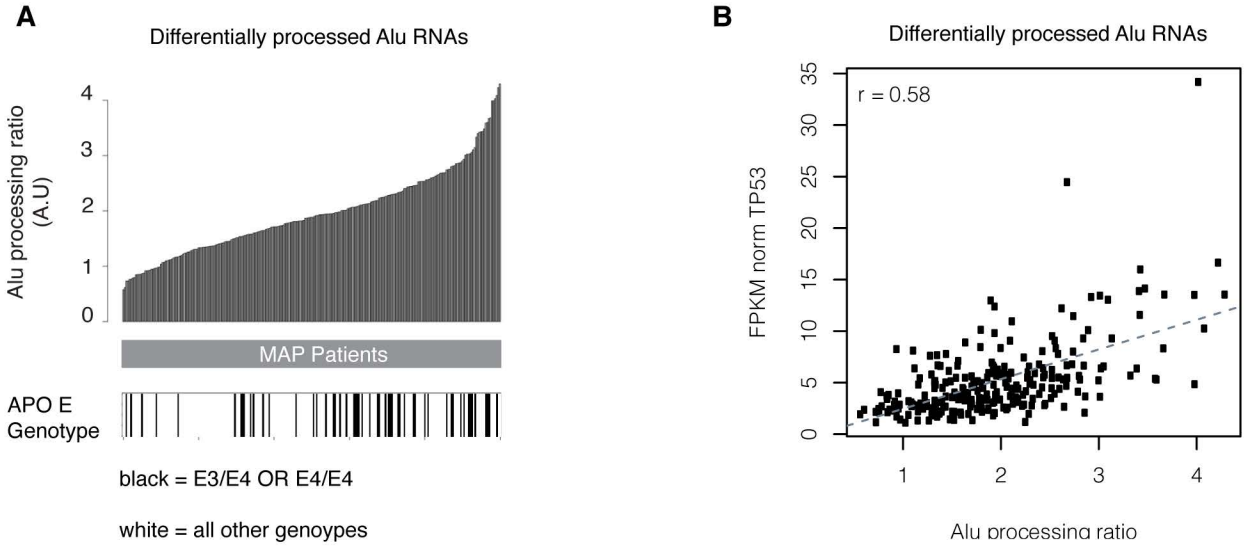
Note: Legend of each figure is included at the next page following the page of that figure.

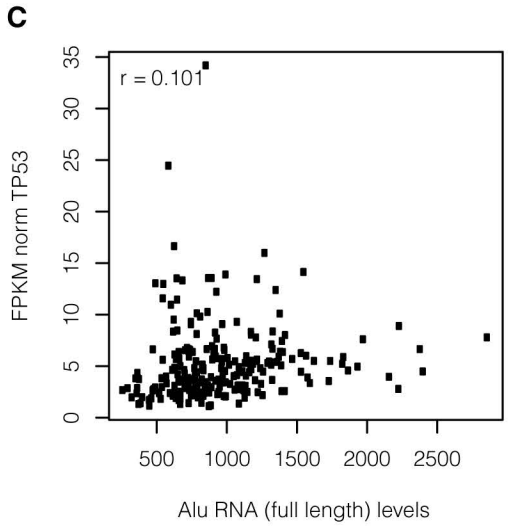


APPENDIX FIGURE S1

Appendix Figure S1.

Graphical representation of the mode of action of SINE RNAs and SINE RNA processing in mouse (B2 RNAs) with regard to transcriptome changes based on previous studies from the Kugel/Goodrich labs (Espinoza *et al*, 2007; Ponicsan *et al*, 2010, 2015; Yakovchuk *et al*, 2009) and our studies (Cheng *et al*, 2020; Hernandez *et al*, 2020; Zovoilis *et al*, 2016).





APPENDIX FIGURE S2

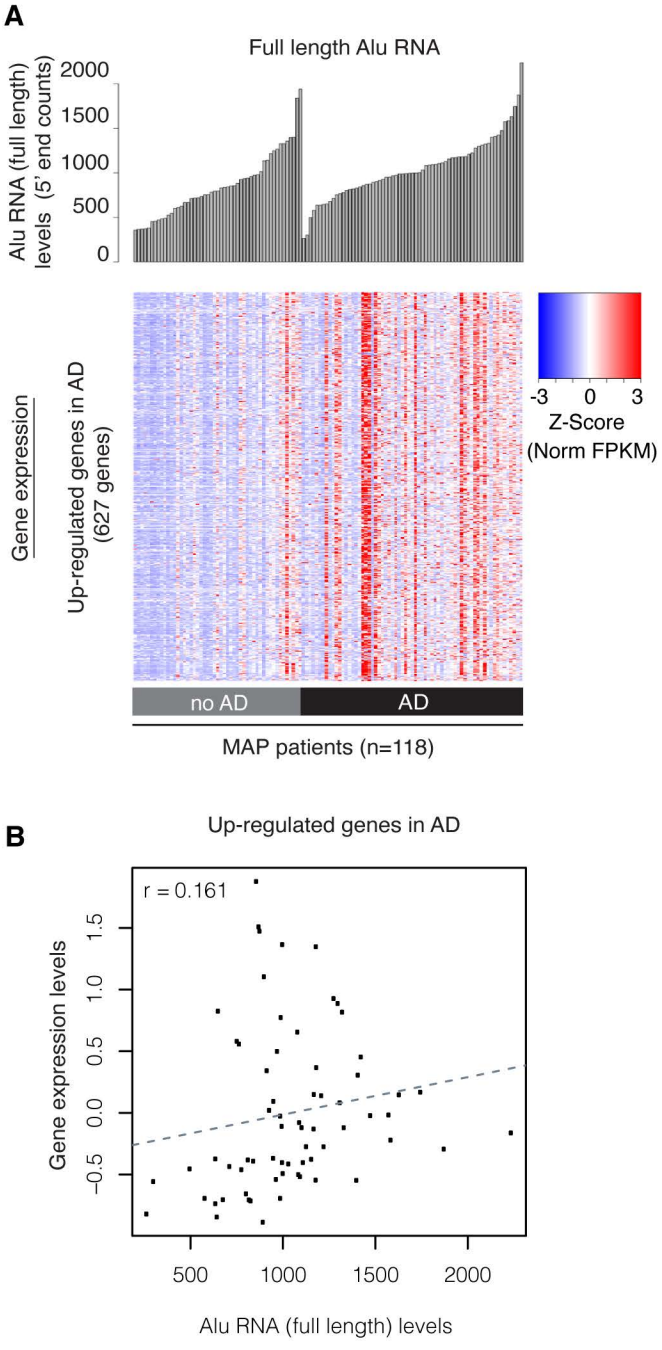
Appendix Figure S2.

(A) Association between SINE Alu RNA ratio (upper panel) and APO E genotype (lower panel). Every column in both panels corresponds to the same MAP patient (n=241). Patients are sorted from left to right in an ascending order with regard to Alu RNA processing ratio. Patients with at least one E4 allele that confers higher genetic risk to developing late onset AD are marked with black in the lower panel. The figure shows that as processing ratio climbs higher, it is more likely to identify an individual with this phenotype.

Here it should be noted though that the connection between APO E and AD, despite APO E being one of the few reliable genetic markers shown to confer some degree of higher susceptibility to AD, is not a one to one absolute match (i.e.the vast majority of people with these genotypes will not develop AD). For this reason, we also do not expect the correlation between Alu RNA processing and this genotype to be a one to one absolute match, especially since Alu RNA processing is a continuous variable while Apo E genotype is a discrete one. The reason why within the higher processing ratio, distribution of genotypes seems to be biphasic (with one density area in the middle of the range, and one towards the right end of the range) remains unknown and likely should be attributed to the large heterogeneity among AD patients and the multi-factorial nature of this disease.

(B) Scatterplot depicting the positive correlation between TP53 expression values and Alu RNA processing ratio in all MAP patients ($r=0.58, p<0.001$).

(C) Scatterplot depicting lack of correlation between TP53 mRNA expression values and full length Alu RNA levels in all MAP patients (n=241).

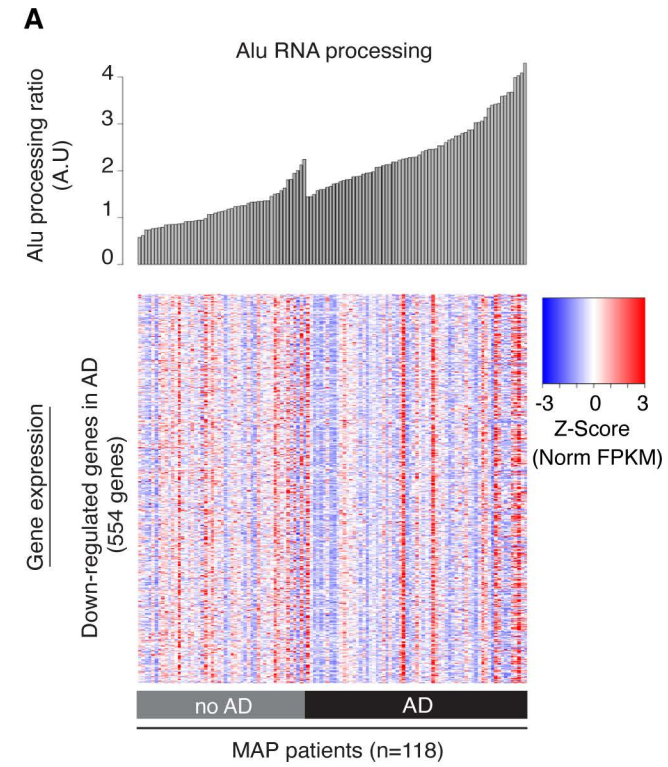


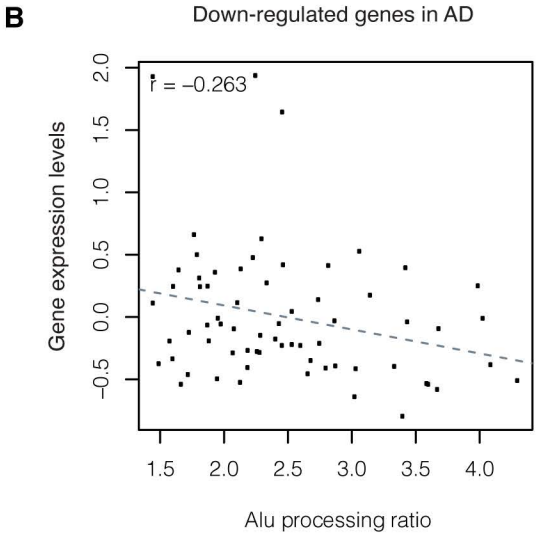
APPENDIX FIGURE S3

Appendix Figure S3.

(A) No association between full length Alu RNA levels (upper panel) and gene expression levels (lower panel) of genes up-regulated in AD (Dataset EV3, |log2FoldChange|> 0.5) is found in contrast to our observations with regard to Alu RNA processing ratio. Columns and rows as described in Figure 6 with every column in both panels corresponding to the same MAP patient sorted from left to right in an ascending order with regard to Alu RNA levels for either the no AD or AD group.

(B) Scatterplot depicting lack of correlation between average gene expression values of upregulated genes in panel A and full length Alu RNA levels in AD patients (*r=0.16*).



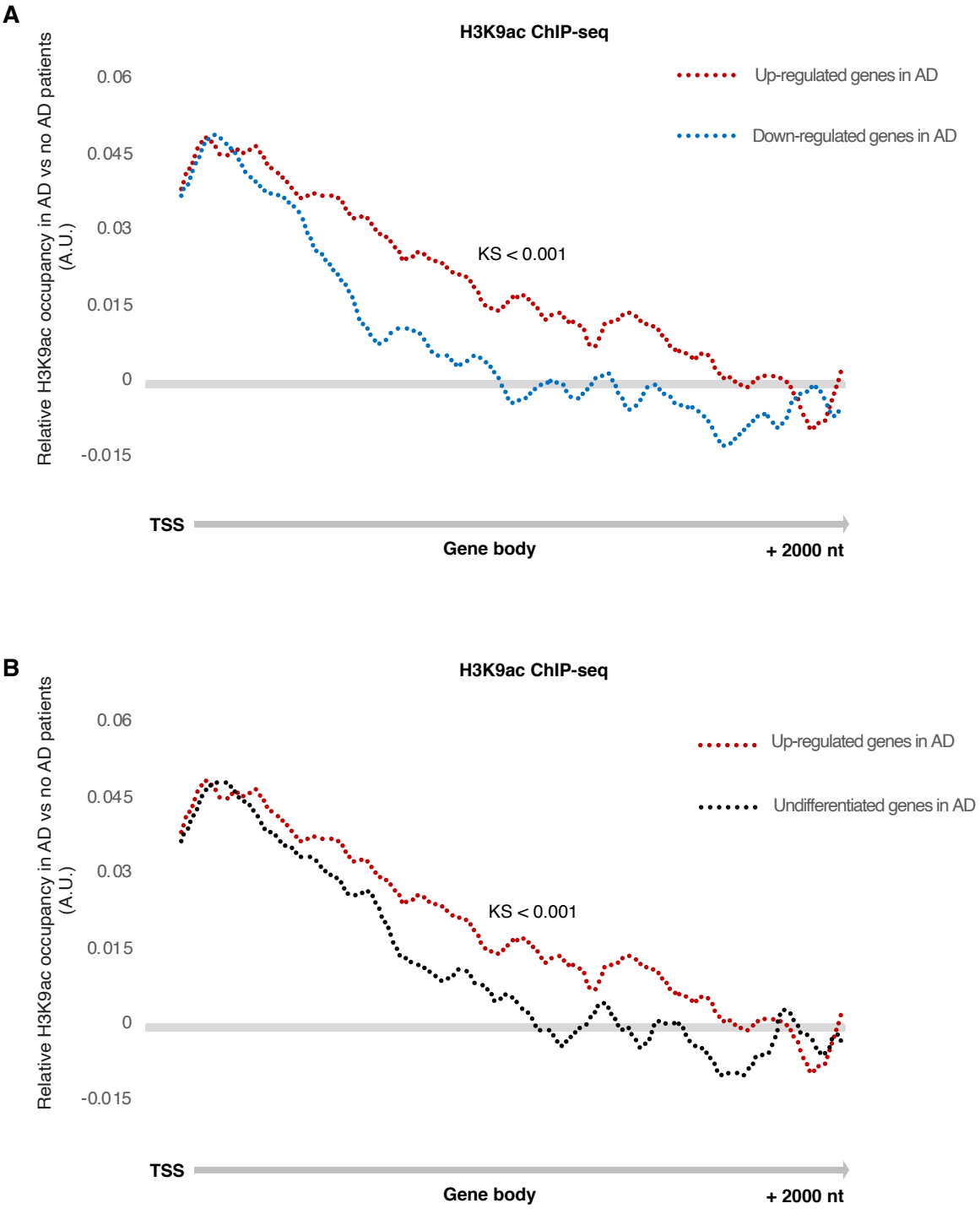


APPENDIX FIGURE S4

Appendix Figure S4.

(A) Only a weak association between Alu RNA processing ratio (upper panel) and gene expression levels (lower panel) of genes down-regulated in AD (Dataset EV4, |log2FoldChange|> 0.5) is found in contrast to our observations with regard to full length Alu RNA levels. Columns and rows as described in Figure 6 with every column in both panels corresponding to the same MAP patient sorted from left to right in an ascending order with regard to Alu RNA processing ratio for either the no AD or AD group.

(B) Scatterplot depicting the weak negative correlation between average gene expression values of down-regulated genes in panel A and Alu RNA processing ratio in AD patients (r=-0.26).



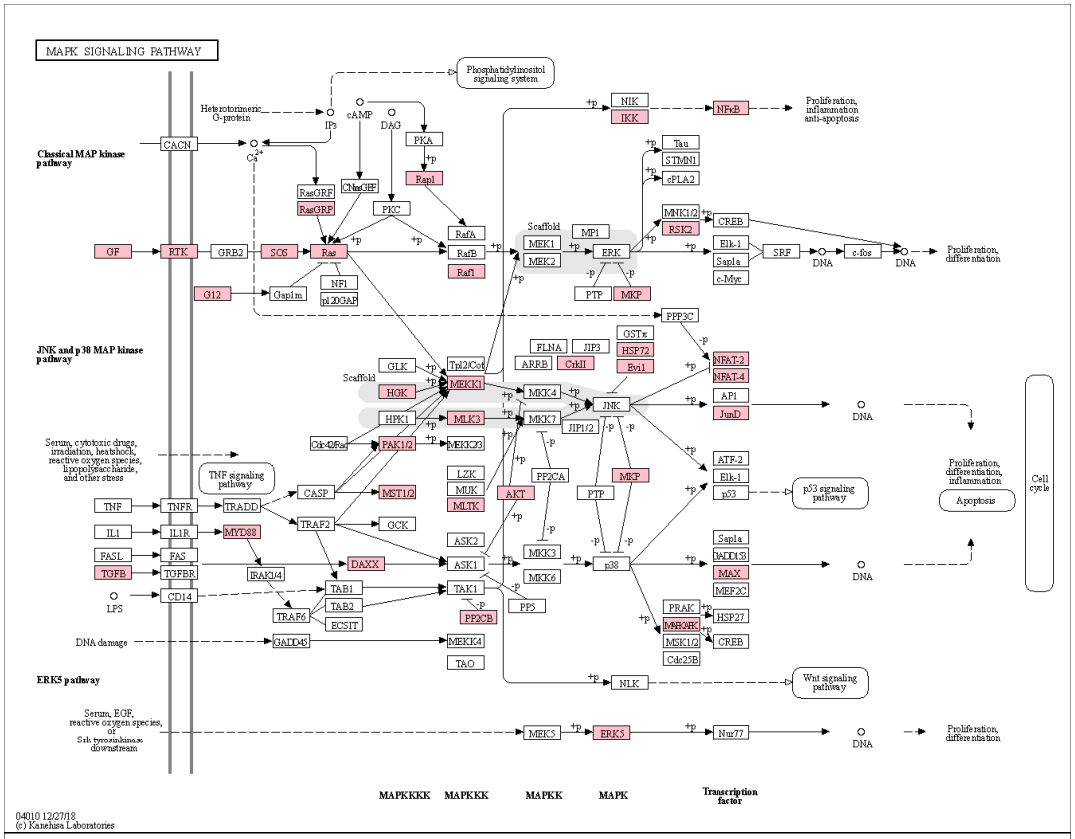
APPENDIX FIGURE S5

Appendix Figure S5. H3K9ac ChIP-seq for MAP patients.

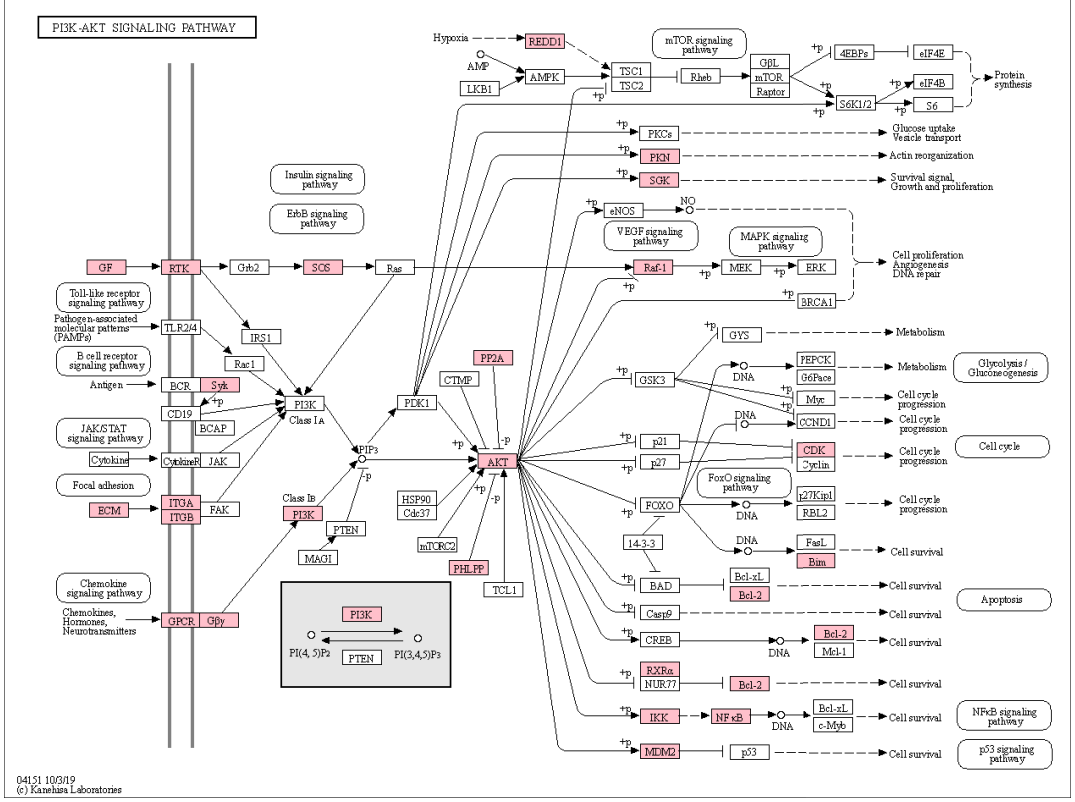
(A) Comparison of H3K9ac occupancy downstream of Transcription Start Site (TSS) between AD and no AD patients for genes up-regulated and down-regulated in AD (Dataset EV3 and Dataset EV4).

(B) Comparison of H3K9ac occupancy downstream of Transcription Start Site (TSS) between AD and no AD patients for genes up-regulated and a random set of non-differentiating genes (Dataset EV3 and Dataset EV5).

A



B



APPENDIX FIGURE S6

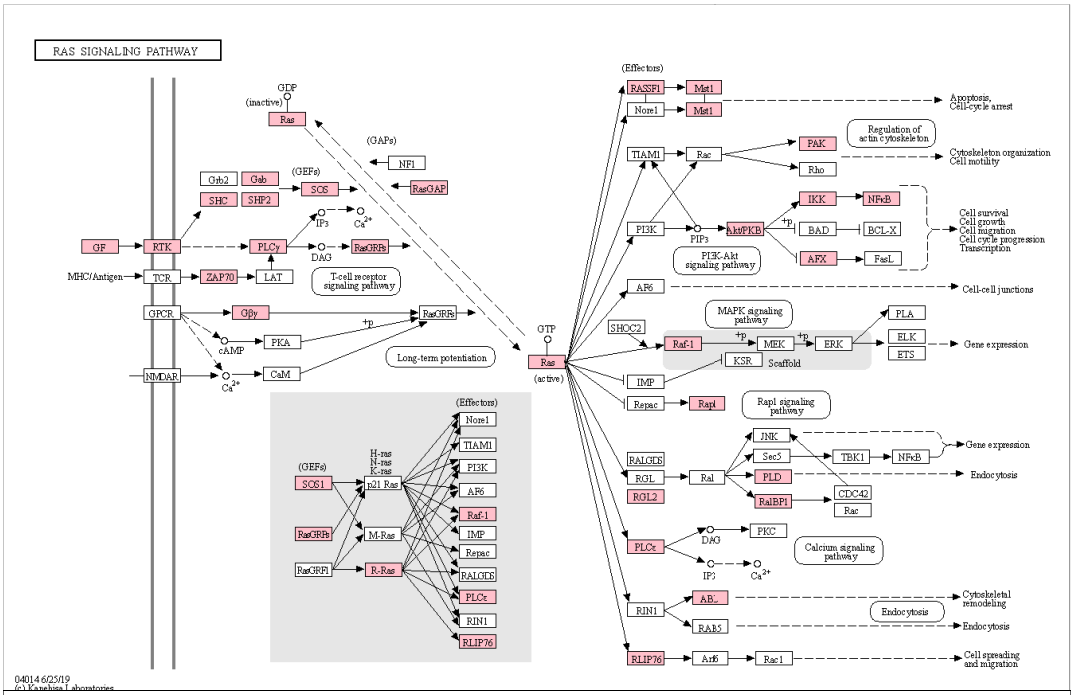
Appendix Figure S6. Pathways impaired in AD up-regulated genes that are strongly correlated with Alu RNA processing ratio.

Pathways identified by KEGG pathway term enrichment analysis of AD upregulated genes the expression of which is strongly correlated with Alu RNA processing ratio ($r > 0.5$, p value <0.05 Pearson correlation, Dataset EV8). Selected pathways are among those with a padj< 0.05 in the KEGG enrichment term analysis by the DAVID platform (see methods). Pathways have been selected based on their association with functional terms such as cellular stress, cell survival, cell death and proliferation.

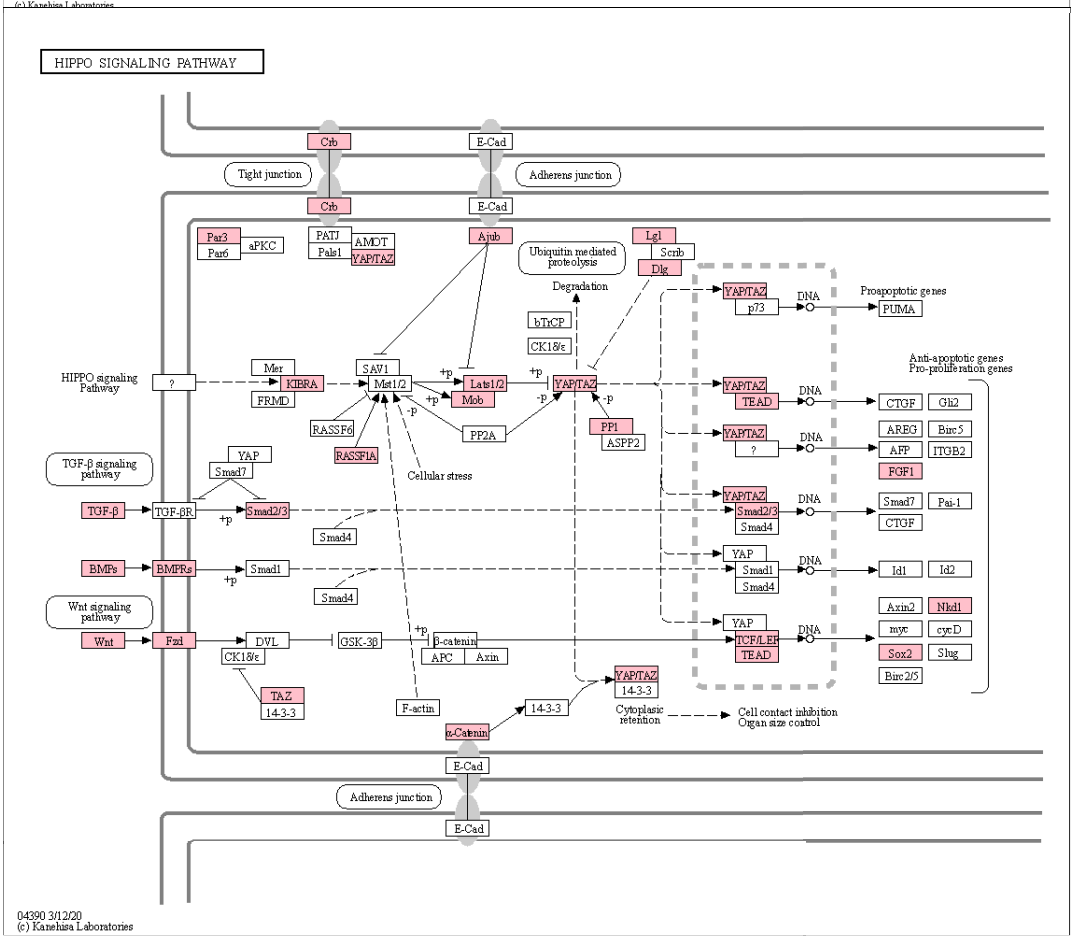
(A) MAPK signalling pathway. Multiple members (colored with pink) belong to the set of AD upregulated genes correlated with Alu RNA processing ratio. This pathways is upstream of the TP53 signalling pathway and cell death and proliferation.

(B) PI3K-AKT pathway. This pathway intersects with various signalling pathways regulating TP53 signalling and cell cycle. Pink color as in A.

A



B

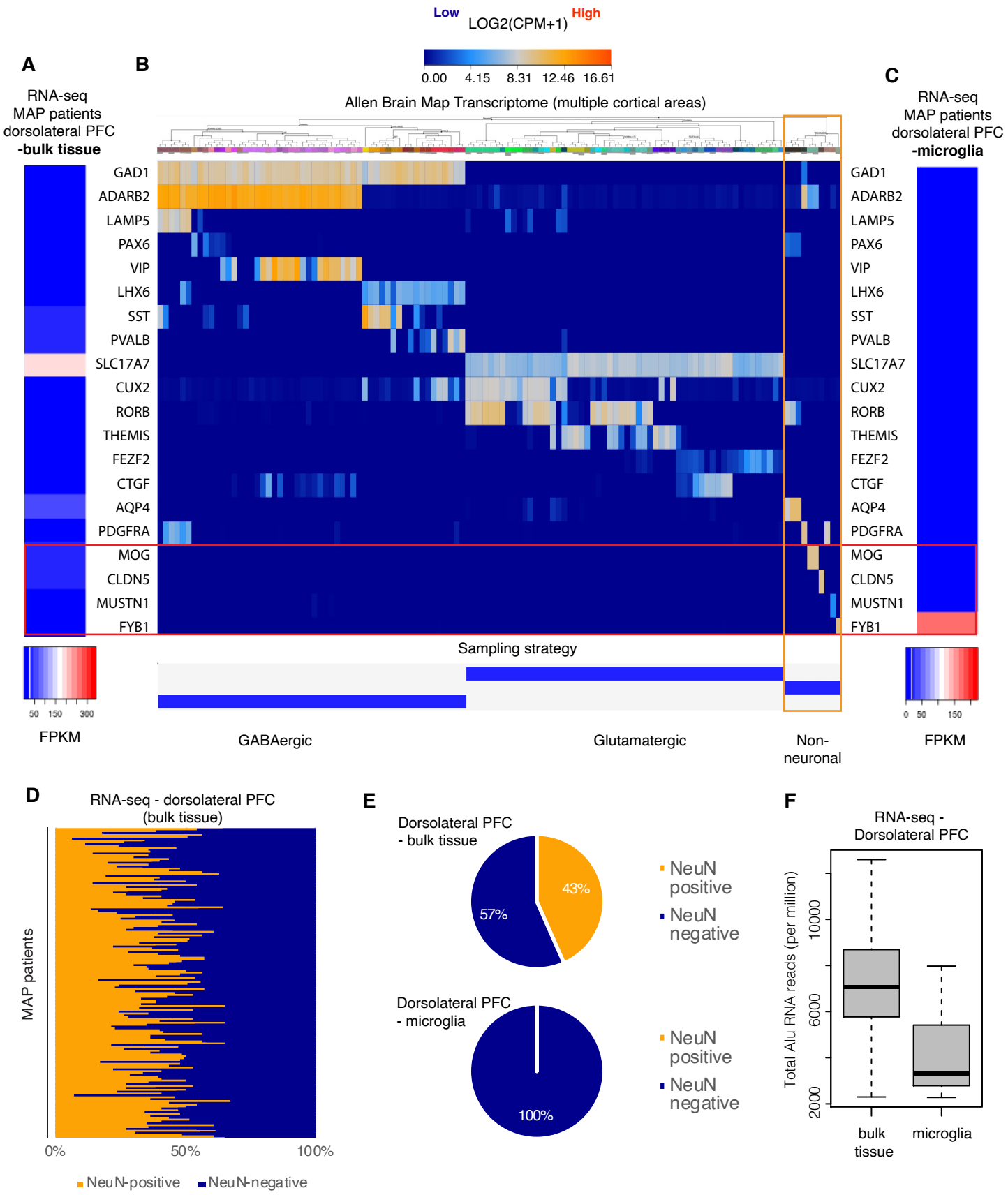


APPENDIX FIGURE S7

Appendix Figure S7. Additional pathways impaired in AD up-regulated genes that are strongly correlated with Alu RNA processing ratio.

(A) RAS signalling pathway. Multiple members (colored with pink) belong to the set of AD upregulated genes correlated with Alu RNA processing ratio. This pathway is important for cell survival and also upstream TP53.

(B) HIPPO signalling pathway. This pathway intersects with various signalling pathways regulating cellular response to stress. Pink color as in A.



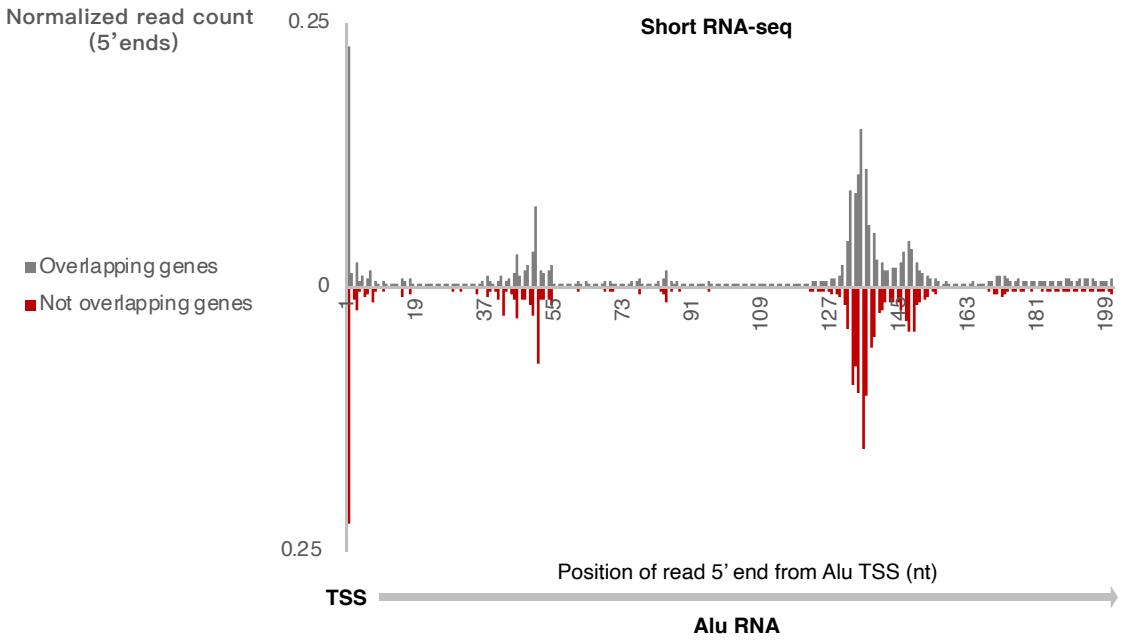
APPENDIX FIGURE S8

Appendix Figure S8. Expression of known hippocampal markers and cellular deconvolution in the MAP RNA-seq data

(A-B-C) Expression levels comparison between RNA-seq data from MAP RNA-seq data used in our study (panel A), RNA-seq data from the Allen Brain Atlas (Panel B) and RNA-seq data generated from microglia cells within the ROSMAP study (Panel C). Comparison is done for the standard panel genes of reference genes used by Allen Brain Atlas. Every column in panel B represents a different cell type in brain with cell types belonging to neuronal and non neuronal cells depicted at the lower part of panel B labelled as GABAergic, Glutamatergic and Non-neuronal. Markers that are enriched in neuronal vs non-neuronal cells were extracted from: [https://celltypes.brain-map.org/rnaseq/human_ctx_smart-seq] based on (Hodge *et al*, 2019; Tasic *et al*, 2018).

(D-E) Cellular deconvolution of RNA-seq data from MAP patients (bulk tissue-dorsolateral prefrontal cortex) and microglia isolated cells within the ROSMAP study. D represents the percentage of neuronal (NeuN positive) vs non-neuronal (NeuN negative cells) per sequenced MAP sample used in our study. E represents the overall percentage of each of these categories in each RNA-seq data set.

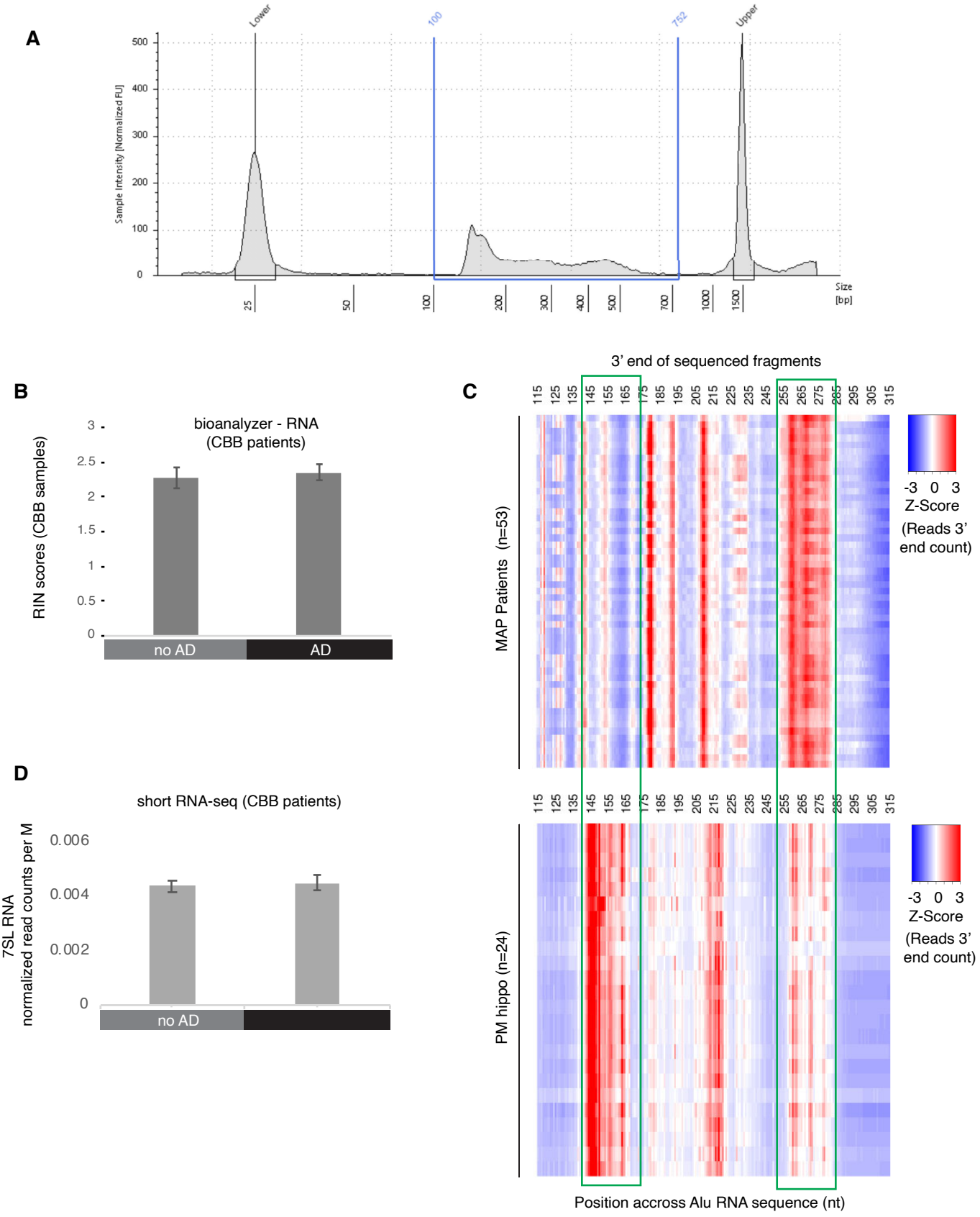
(F) Alu RNA levels comparison between the MAP patients (bulk tissue-dorsolateral prefrontal cortex) and microglia isolated cells within the ROSMAP study.



APPENDIX FIGURE S9

Appendix Figure S9

Plotting of the position of the first base (5' end) of Alu RNA fragments in a CBB patient against Alus that fall within known Ensemble genes (primary within introns) (upper panel) and Alus outside genic areas (intergenic space) (lower panel). The x axis represents an Alu RNA metagene aligned at the TSS of the Alu elements and the y axis shows the 5' end count for Alu RNA fragments aligning to any position downstream of position +1.



APPENDIX FIGURE S10

Appendix Figure S10

(A) Agilent Bioanalyzer electropherogram of one of the short-RNA-seq libraries we produced from RNA from the CBB patients' hippocampi. It shows a fragment range of 150-650nt, that corresponds to initial RNA inserts of 30nt to approx. 530 nt (adapter is approx. 120nt).

(B) RIN scores of RNAs sequenced from CBB patients.

(C) Plotting of the position of the 3'end of Alu RNA fragments across the end of the consensus Alu metagene to depict potential 3' end positions of Alu RNAs in all post-mortem hippocampal tissues (PM hippo) from patients from the CBB (lower panel) and in MAP patients (upper panel). Each row in the heatmap depicts the distribution of counts of the 3'ends of reads mapped across the Alu metagene for each patient. The x-axis represents a metagene combining all unique Alu RNA sequences (ALUome) as in Figure 1. Heatmap density corresponds to normalized counts of the 3'end of the reads with red corresponding to higher density of these 3'ends at a specific position. Orange rectangles mark expected end positions for full length fragments (right) and fragments generated by the XR1 processing point (left).

(D) 7SL RNA levels based on short-RNA-seq in CBB patients used for the normalization of total Alu RNA base coverage in Figure 2.

Cheng Y, Saville L, Gollen B, Isaac C, Belay A, Mehla J, Patel K, Thakor N, Mohajerani MH, Zovoilis A (2020) Increased processing of SINE B2 ncRNAs unveils a novel type of transcriptome deregulation in amyloid beta neuropathology. *Elife* 9

Espinoza CA, Goodrich JA, Kugel JF (2007) Characterization of the structure, function, and mechanism of B2 RNA, an ncRNA repressor of RNA polymerase II transcription. *RNA* 13: 583-596

Hernandez AJ, Zovoilis A, Cifuentes-Rojas C, Han L, Bujisic B, Lee JT (2020) B2 and ALU retrotransposons are self-cleaving ribozymes whose activity is enhanced by EZH2. *Proc Natl Acad Sci U S A* 117: 415-425

Hodge RD, Bakken TE, Miller JA, Smith KA, Barkan ER, Graybuck LT, Close JL, Long B, Johansen N, Penn O *et al* (2019) Conserved cell types with divergent features in human versus mouse cortex. *Nature* 573: 61-68

Ponicsan SL, Kugel JF, Goodrich JA (2010) Genomic gems: SINE RNAs regulate mRNA production. *Curr Opin Genet Dev* 20: 149-155

Ponicsan SL, Kugel JF, Goodrich JA (2015) Repression of RNA Polymerase II Transcription by B2 RNA Depends on a Specific Pattern of Structural Regions in the RNA. *Noncoding RNA* 1: 4-16

Tasic B, Yao Z, Graybuck LT, Smith KA, Nguyen TN, Bertagnolli D, Goldy J, Garren E, Economo MN, Viswanathan S *et al* (2018) Shared and distinct transcriptomic cell types across neocortical areas. *Nature* 563: 72-78

Yakovchuk P, Goodrich JA, Kugel JF (2009) B2 RNA and Alu RNA repress transcription by disrupting contacts between RNA polymerase II and promoter DNA within assembled complexes. *Proc Natl Acad Sci U S A* 106: 5569-5574
Zovoilis A, Cifuentes-Rojas C, Chu HP, Hernandez AJ, Lee JT (2016) Destabilization of B2 RNA by EZH2 Activates the Stress Response. *Cell* 167: 1788-1802 e1713