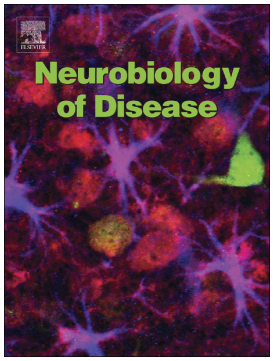


Alzheimer's brains show inter-related changes in RNA and lipid metabolism

Shahar Barbash, Benjamin P. Garfinkel, Rotem Maoz, Alon Simchovitz, Bettina Nadorp, Alessandro Guffanti, Estelle R. Bennett, Courtney Nadeau, Andreas Türk, Lukas Paul, Torsten Reda, Yan Li, Aron S. Buchman, David S. Greenberg, Alexander Seitz, David A. Bennett, Patrick Giavalisco, Hermona Soreq

PII: S0969-9961(17)30135-3
DOI: doi: 10.1016/j.nbd.2017.06.008
Reference: YNBDI 3978

To appear in: *Neurobiology of Disease*

Received date: 9 March 2017
Revised date: 17 May 2017
Accepted date: 12 June 2017

Please cite this article as: Shahar Barbash, Benjamin P. Garfinkel, Rotem Maoz, Alon Simchovitz, Bettina Nadorp, Alessandro Guffanti, Estelle R. Bennett, Courtney Nadeau, Andreas Türk, Lukas Paul, Torsten Reda, Yan Li, Aron S. Buchman, David S. Greenberg, Alexander Seitz, David A. Bennett, Patrick Giavalisco, Hermona Soreq , Alzheimer's brains show inter-related changes in RNA and lipid metabolism, *Neurobiology of Disease* (2017), doi: 10.1016/j.nbd.2017.06.008

Alzheimer's brains show inter-related changes in RNA and lipid metabolism

Shahar Barbash[1,2], Benjamin P Garfinkel[1,2], Rotem Maoz[1,2], Alon Simchovitz[2], Bettina Nadorp[2], Alessandro Guffanti[2], Estelle R. Bennett[2], Courtney Nadeau[3], Andreas Türk[3], Lukas Paul[3], Torsten Reda[3], Yan Li[4], Aron S. Buchman[5], David S. Greenberg[2], Alexander Seitz[3], David A. Bennett[5], Patrick Giavalisco[4], and Hermona Soreq[1,2]*.

[1] The Edmond & Lily Safra Center for Brain Sciences and [2]The Department of Biological Chemistry, The Hebrew University of Jerusalem, Jerusalem 9190401, Israel.

[3] Lexogen GmbH, Campus Vienna Biocenter 5, 1030 Vienna, Austria.

[4] Max Planck Institute of Molecular Plant Physiology, Am Mühlenberg 1, 14476 Potsdam-Golm, Germany.

[5] Rush Alzheimer's Disease Center, Rush University Medical Center, 600 South Paulina, Suite 1028, Chicago, IL 60612, USA.

*Correspondence to hermona.soreq@mail.huji.ac.il

Keywords: Alzheimer's Disease; Alternative Polyadenylation; Lipidomics; RNA sequencing; Cognitive decline; Neuropathology.

Abstract

Alzheimer`s disease (AD) involves changes in both lipid and RNA metabolism, but it remained unknown if these differences associate with AD's cognition and/or post-mortem neuropathology indices. Here, we report RNA-sequencing evidence of inter-related associations between lipid processing, cognition level, and AD neuropathology. In two unrelated cohorts, we identified pathway-enriched facilitation of lipid processing and alternative splicing genes, including the neuronal-enriched NOVA1 and hnRNPA1. Specifically, this association emerged in temporal lobe tissue samples from donors where postmortem evidence demonstrated AD neuropathology, but who presented normal cognition proximate to death. The observed changes further associated with modified ATP synthesis and mitochondrial transcripts, indicating metabolic relevance; accordingly, mass-spectrometry-derived lipidomic profiles distinguished between individuals with and without cognitive impairment prior to death. In spite of the limited group sizes, tissues from persons with both cognitive impairment and AD pathology showed elevation in several drug-targeted genes of other brain, vascular and autoimmune disorders, accompanied by pathology-related increases in distinct lipid processing transcripts, and in the RNA metabolism genes hnRNPH2, TARDBP, CLP1 and EWSR1. To further detect 3'-polyadenylation variants, we employed multiple cDNA primer pairs. This identified variants that showed limited differences in scope and length between the tested cohorts, yet enabled superior clustering of demented and non-demented AD brains versus controls compared to total mRNA expression values. Our findings indicate inter-related cognition-associated differences in AD's lipid processing, alternative splicing and 3'-polyadenylation, calling for pursuing the underlying psychological and therapeutics implications.

Highlights

- RNA transcripts differ between brains from Alzheimer's patients with or without dementia.
- In non-demented patients with pathology, brain neurons show facilitated RNA and lipid processing.
- Cognition-related differences included brain, vascular and autoimmune disease-related genes.
- Polyadenylation changes classify patients better than transcript expression levels.

Multiple recent studies seek causative links between Alzheimer's disease (AD)'s pathology, dementia and metabolic functions. Principally, neuronal numbers are lower in demented compared to non-demented individuals with AD Braak neuropathology (Andrade-Moraes et al., 2013; Braak and Braak, 1995), and individuals with inheritably impaired phospholipid metabolism are particularly susceptible for cognitive decline (Lacour et al., 2017). This agrees with the recently reported global sharing of physical and mental health parameters (Hagenaars et al., 2016). It also supports the co-ignition of the AD metabolic decline by β-amyloid fibril formation and tau hyper-phosphorylation (Pascoal et al., 2016), which includes altered phosphatidylcholine metabolism (Whiley et al., 2014). That AD neuropathology may be attenuated by the epigenetic SIRT1 pathway (Shah et al., 2016), and that phosphatidylcholine protects neurons from β-amyloid toxicity (Ko et al., 2016) further suggests complex inter-relationships between lipid metabolism, neuronal death and cognitive deterioration in AD.

The role and composition of brain lipids has recently emerged as a major criterion in aging and AD studies, especially for β-amyloid toxicity. Brain lipids are brain region- and species-specific, and their levels and composition are modified with aging in a human-characteristic pattern (Fu et al., 2011). Neuronal lipids regulate the location and function of synaptic membrane proteins and relay signals from the membrane to intracellular sites or to other cells, as well as impairments in depression and anxiety disorders (Muller et al., 2015). New technologies facilitate these discoveries. Thus, time of flight secondary ion mass spectrometry (TOF-SIMS) revealed cortical cholesterol overload in the AD cortex(Lazar et al., 2013), and Nonlinear microscopy enables visualization of plaque-associated lipids

in AD brain tissues (Kiskis et al., 2015), where they may invert inert β-amyloid fibrils into neurotoxic ones (Martins et al., 2008). However, it remained unknown whether lipid differences are associated with AD's cognitive deterioration.

RNA processing impairments are causally involved in various neurodegenerative events (Kim et al., 2013). Therefore, we hypothesized that they may be associated with AD's progression and lipid alterations. Pre-mRNA-processing activities, including capping, alternative splicing and 3'-alternative polyadenylation (APA) are continuously engaged in extensive crosstalk with one another (Elkon et al., 2013; Proudfoot, 2011). Specifically, alternative splicing is a major regulator of gene expression in health and disease (Di Giammartino et al., 2011). Also, impaired alternative splicing associates with decreased dendritic spine density, mal-functioning of entorhinal–hippocampal neural circuits and suppressed learning capability in mice (Berson et al., 2012). About 54% of the human genes also possess multiple APA sites (Tian et al., 2005), which can either selectively affect the 3'-untranslated region (3'-UTR) and/or change the coding region. The brain preferentially uses distal poly(A) sites (Di Giammartino et al., 2011), and functional microRNA (miRNA) recognition sites often reside within, and near both ends of the 3'-UTR (Grimson et al., 2007). This suggests co-evolution of poly(A) sites (Lee et al., 2008), miRNAs that target them (Barbash et al., 2014) and mRNA isoforms with modified repertoire of cis-regulatory elements. Supporting this notion, evolutionarily younger genes are less likely to include APA signals. For example, the longer APA variants of brain-derived neurotrophic factor (BDNF) and calmodulin-activated protein kinase 2 (Camk2a) show elevated dendritic location (An et al., 2008; Bulleit et al., 1988; Timmusk et al., 1993). Thus, both alternative splicing and APA may affect neuronal properties in more than one manner.

We, and others found both alternative splicing (Berson et al., 2012; Kolisnyk et al., 2013) and miRNA alterations (Lau et al., 2013) in AD brains with dementia. These changes may modify mRNA translation, stability and localization. However, in-depth RNA processing studies in AD are still limited, and none of the genomics, transcriptomics or proteomics resources of the AD brain cover global APA or lipidomic parameters. Therefore, it is still largely unclear whether or how RNA processing impairments are linked to the characteristic histopathology and/or cognitive and phospholipid metabolism decline, which are the major hallmarks of AD. To identify transcript differences that are separately or jointly associated with AD's dementia, neuropathology or phospholipid metabolism, we initiated an in-depth RNA-sequencing based study of temporal gyrus tissues from two independent groups of persons with or without prior evidence of cognitive impairments. We subjected polyadenylated RNA transcripts to RNA sequencing using either QuantSeq 3'-sequencing (for the first cohort), or multiple 3' primer variants-derived cDNA libraries (for the second cohort), to identify RNA metabolism and APA modifications; profiled lipids in these tissues by mass spectrometry; and examined associations between these diverse changes and the donors' cognitive and pathology states.

Results

RNA-Seq discriminates between brains of demented and non-demented Alzheimer patients

To address transcript changes that associate with AD pathology and/or cognitive status prior to death, we studied brain samples from 3 groups of 12 clinically diagnosed male participants each from the Rush Memory and Aging Project (MAP; N=36). Participants who presented no cognitive impairment (NCI), mild cognitive impairment (MCI) or AD dementia were included in each of these groups, which included 4 cases each with low (Braak 1-2), moderate (Braak 3-4) and high (Braak 5-6) levels of AD (Fig. 1a). Likewise, each of the groups with a given Braak staging level was composed of 4 samples with NCI, MCI or AD dementia. Individual brain RNA samples from these MAP groups were subjected to Illumina RNA sequencing and bioinformatics analysis (see Supplementary Tables 1 and 2

for full sample data, Fig. 1a, b for descriptions of the studied groups and workflow, and Methods for definitions).

The second group, received from the Netherland Brain Bank (NBB) was composed of 24 temporal lobe samples from demented male patients diagnosed with AD, and non-demented individuals without or with pathologic AD (Controls; 'non-demented with pathology', NDWP; AD). RNA preparations from the NBB samples were reverse transcribed into full-length 3'-tagged cDNA and segregated through a SQUARE selective exponential amplification using one universal 5'-primer and 12 selective 3'-primers, each of which hybridizes at the junction of the mRNA body and poly(A) tail. An additional group of AD diagnosed patients with early Braak 3-4 neuropathology and cognitive decline was added to the NBB groups. Sequencing files from both groups were processed and analyzed for differential expression and functional enrichment, and for the NBB groups also for APA detection (Fig. 1d). Observations were validated by high-throughput RT-PCR, enzyme activity tests and comparison to previous databases (Fig. 1e).

RNA transcript differences in MAP groups (details in Supplementary Table 2) were analyzed for association with the level of AD pathology and/or clinical diagnosis. In spite of the modest group sizes and the individual heterogeneity between samples, global transcript differences showed significant associations with either cognition or AD pathology (Supplementary Fig. 1), and the association of transcript modifications with cognitive impairment was stronger than the association to AD pathology. Gene Ontology analysis by selecting transcript groups with ANOVA P value <0.05 revealed enriched pathways that associated with cognitive impairment or with pathology (not corrected for multiple comparisons; DAVID (https://david.ncifcrf.gov/) (Supplementary Tables 3 and 4, Fig. 2a). In MAP AD tissues, the cognition-related brain pathways included RNA processing and lipid synthesis. In the 24 sample NBB groups (Fig. 2b), the deeper sequencing protocol highlighted distinct significant differences in additional, non-overlapping AD-related pathways, including alternative splicing and nuclear activities, which differed in NDWP tissues. We conclude that in both cohorts, modified transcript groups vary more distinctively throughout cognitive decline than with progressing pathology and show association with RNA processing.

To validate our NBB results by an independent technology we selected 80 transcripts with high and medium expression levels in the three NBB groups and AD samples from patients with cognitive impairment and early yet discernible neuropathology (Braak stages 3 and 4), similar to that of the pathologic AD without dementia groups, for quantitation by microfluidic RT-PCR (Fluidigm, US) (Supplementary Figure 2; Primer sequences in Supplementary Table 5). Out of the 80 tested transcripts, 65 showed linear RT-PCR calibration curves in the examined expression range with a lower than 1% ratio of technical noise to biological signal (Supplementary Fig. 3 and 4). Transcript levels showed significant correlation coefficients and P values with the Fluidigm RT-PCR results for 85.5% of tested transcripts that yielded linear curves out of the 65 (Supplementary Fig. 5). Selected comparisons included global expression of CAMK2A (P< 2*10-8) and the distal and proximal SOD1 APA variants (P< 0.001) (Supplementary Fig. 6). Importantly, the Fluidigm analysis of the distal SOD1 variant showed no significant correlation with the proximal SOD1 variant, and vice versa, pointing at the high fidelity and resolution power of the results. Deep 3'-segregated RNA sequencing thus demonstrated valid levels of both global differences and identified APA variants of brain transcripts.

Impaired cholinergic transmission has long been reported in AD (Giacobini and Gold, 2013), and cholinergic neurons are the primary producers of acetylcholinesterase (AChE) (Soreq and Seidman, 2001). Hence, we performed acetylthiocholine hydrolysis measurements for the MAP group samples (Fig. 2c), and found lower values in demented AD brains compared to NDWP (pathologic AD without dementia) and healthy control samples, compatible with reported early AD-related death of cholinergic neurons (Mesulam et al., 2004). Next, we assessed AChE expression and enzymatic

activity levels in the NBB brain tissues. AChE mRNA transcript levels showed high correlation with the hydrolytic AChE activities in the tested tissues, with both transcripts and enzyme activity lower in AD dementia compared to NDWP and healthy control samples (Fig. 2d, e). Unlike AChE, measuring the enzymatic activity of the closely related butyrylcholinesterase enzyme, most of which reaches the brain from peripheral sources (Darvesh et al., 2003), showed no association with brain read counts (Supplementary Fig. 7). This is predictable by the non-brain origin of BChE, validates our protein level measurements, and indicates that these two separate cohorts are indeed comparable.

Brains of AD and NDWP donors show inversely directed lipid processing differences

Others reported lipid-related differences in early AD plasma (Di Paolo and Kim, 2011) and identified phospholipase D3 changes as an AD risk factor (Cruchaga et al., 2014). To profile the lipidomic state in the analyzed NBB tissues, we employed targeted and untargeted ultra-performance liquid chromatography-high resolution mass spectrometry (UPLC-MS) analysis (Fu et al., 2011), (Khrameeva et al., 2014). Using this methodology, we were able to separate and relatively quantify 9317 lipophilic peaks (3504 in the negative ion mode and 5813 in the positive ion mode), of which 144 lipid species were reliably annotated as free fatty acids (35), phospatidylcholines (31) or lyso-phosphatidylcholines (5), phosphatidylethanolamines (9) or lyso-phosphatidylethanolamines (7), phosphatidylglycerols (2), phosphatidylinositol (2), diacylglycerols (3) or triacylglycerols (50) (Supplementary Table 6). This methodology identified over 100 quantifiable lipid metabolites and enabled comparison to pure standards as a basis for analysis. Comparing the complete lipidomic profiles (9312 peaks) by supervised multivariate statistics (O2PLS-DS, Fig. 3a) provided an unambiguous classification of 3 distinct groups, namely early AD (eAD), late AD (lAD) and the healthy controls and NDWP groups, which were only partially separated, indicating that these two groups are possibly still too similar at the lipidomic level to be segregated. Taken together, these findings demonstrate modified patterns of cortical lipids which may associate with the capacity of some patients with AD pathology to avoid dementia longer than others having the same neuropathology state. Also, we were not able to obtain significant classification using the same supervised O2PLS-DA method with the annotated lipid data-set, indicating that the identified phospholipids, the free fatty acids and the storage lipids (TAGs) were not necessarily those lipid peaks in the complete matrix that contributed to the distinction of the AD and control samples.

Comparing the annotated lipids composition in the AD and the NDWP groups showed several significant, opposing differences, indicating relevance to the disease-related differences between these groups. Thus, cortical fatty acids were decreased in AD and unaltered in NDWP, whereas Phosphatidylcholine variants were substantially lower in AD but higher in NDWP; and the storage Triacylglycerol compounds were down regulated in NDWP and unaltered in AD (Fig 3 b). More information seems to lie in the un-annotated lipid peaks, compatible with most of the lipids so far associated to AD being sterols, sphingolipids and PIPs (phospho-inositol phosphates) (Balla, 2013; Bandaru et al., 2009).

Cell type-specific transcript differences associate with cognition, pathology and RNA processing

Modified lipids and transcripts could reflect AD-related loss of neurons (Heinemann et al., 2012; Verkhratsky et al., 2013). Therefore, we interrogated in the NBB cohort those identified transcripts whose protein products were shown to be exclusive for neurons, astroglia, microglia or oligodendroglia, which metabolically support axons and contribute to neurodegeneration (Zhang et al., 2013) (Lee et al., 2012) (Full gene lists in Supplementary Tables 7 and 8). This analysis showed higher levels of neuron-specific transcripts, in the NDWP group compared to AD and control persons (Fig. 4a). Also, oligodendroglia-specific genes were lower in NDWP compared to controls and AD (Fig. 4b); astroglia-specific transcripts were higher in the AD and NDWP groups compared to controls (Fig. 4c), and microglia-related genes were lower in both AD and NDWP compared to controls (Fig. 4d). Thus, each cell type showed distinct expression patterns in brain tissues from the tested NBB groups.

To test if these differences also reflect modified cell type composition of the analyzed tissues, we normalized the control transcript values for each cell type to -1- and re-examined human- and cell-type specific transcripts (Darmanis et al., 2015). This analysis as well showed higher neuron-specific genes in the NDWP group compared to AD and controls (Fig. 4e). Oligodendroglia-specific genes were lower in NDWP compared to controls, but not to AD (Fig. 4f). Astroglia-specific transcripts were higher in the AD and NDWP groups compared to controls (Fig. 4g), as was seen in the non-normalized analysis; whereas microglia-related genes were higher in AD, but not in NDWP compared to control levels in the normalized analysis (Fig. 4h). Thus, different analysis methods reflect an apparent association between dementia and changes in cell-type specific APA patterns and possibly cell type contents.

Identification of novel disease-related brain APA differences

In addition to the cell-type specific differences, we sought AD-related differences in alternative splicing-related genes (Barbash and Soreq, 2012; Berson et al., 2012; Kim et al., 2013; Lau et al., 2013). In both AD and NDWP compared to control brains, we found higher levels of EWSR1, implicated in amyotrophic lateral sclerosis (Couthouis et al., 2012), and HNRNPH2, whose expression differences associate with age (Mazin et al., 2013) (Fig. 4i). CLP1, which links tRNA metabolism to motor neuron loss (Hanada et al., 2013) and TARDBP, which when mutated, may induce amyotrophic lateral sclerosis (Kabashi et al., 2008) were higher in AD compared to both controls and NDWP (Fig. 4j). In comparison, NOVA1, which regulates a transcript set critical to proper neuronal functioning (Zhang et al., 2008), and the neuronal-enriched HNRNPA1, which plays a role in determining the structure of β-amyloid precursor protein isoforms (Donev et al., 2007) and is largely depleted in AD neurons (Berson et al., 2012) were both higher in NDWP compared to controls (Fig. 4k). Unlike the transcript differences, hnRNPA1 protein levels were lower in NDWP and AD as compared to Controls (Fig. 4l), together suggesting additional levels of regulation over the AD histopathology-related changes in brain RNA processing transcripts.

Next, we calculated the tendency of each studied gene to carry distinct polyadenylation signals by dividing read counts in the maximally expressed dinucleotide by the total read counts. The majority of brain-expressed genes showed several APA variants, with highly variable gene-specific patterns (Fig. 5a). The distances between the two major APA variants across all genes were generally very short, but ranged between 1 and 700 nucleotides (average: 4.9, standard deviation: 2.2; see Fig. 5b). To test if the different APA variants reflect technical noise due to primer mismatches, we calculated the correlation between the different dinucleotide primer fields across all genes for each sample separately. High mis-priming rates would be reflected in a high and positive correlation, as more copies of both the ‘correctly primed’ and the 'mis-primed' products would be similarly raised in their respective fields. However, all field comparisons showed very low correlation (Supplementary Fig. 8).

We further considered that the observed differential expression between the relatively small groups sample size (8 brains per group) could reflect a non-recorded phenotype stratification that may be entirely unrelated to the examined condition (e.g. eye color). Nevertheless, comparing calculated P values of 500 permutations across genes and fields to the simple t-test P value for each gene and field excluded this possibility. For example, the permutation P value for comparing AD to control samples in the ‘AC’ field was 0.03, reflecting the significance of the disease effect on global differences between the 186 differentially expressed genes in this field. The permutation comparison P value of controls to NDWP was 0.03 as well, supporting a dominant phenotypic effect of β-amyloid pathology on differential APA expression patterns, whereas the NDWP to AD permutation comparison P value was 0.18, representing a generally lower number of differentially expressed genes (78) within the ‘AC’ field in these two groups (Supplementary Fig. 9).

Importantly, permutation analyses for all dinucleotide fields showed significant values between the 12 AD to control samples (0.041±0.030), and controls to NDWP (0.042±0.030), but not NDWP to AD

(0.151±0.120) comparisons (Supplementary Table 9). The determined permutation P values are measures of likelihood for the group specificity as a function of the groups' sizes and the number of discovered differences. Therefore, this analysis directly quantifies the likelihood of these differentially expressed genes to be group-specific according to the described phenotype, with the risk of following up wrong leads in subsequent studies being <4.1%, 4.2% or 15.1% respectively for each of the above comparisons. We conclude that the observed APA variations were genuine and discriminate between different cognition, but not pathology states.

3'-UTR APA variations involve poly(A) consensus sequences (e.g. 'AAUAAA' and 10 less prominent consensus sequences) that are located 10-30 nucleotides upstream of the poly(A) site (Beaudoing et al., 2000). Among our 64,107 transcripts, 46,078 (72%) end reads with >1 read counts had at least one poly(A) consensus sequence. To examine if we captured genuine poly(A) sites we searched for such consensus sequences 1-50 nucleotides upstream of the identified transcript end, on those reads that included the 3' universal linker for which we could empirically determine the poly(A) site at single nucleotide resolution. We identified the most prominent consensus sequence, 'AAUAAA', at 16 nucleotides upstream of the transcript end site, on average, (Fig. 5c), which is compatible with the reported consensus sequence position.

Next, we measured the distance between the proximal and distal APA sites in 43 genes with at least one known alternative variant in the UCSC database (http://genome.ucsc.edu/) and for which we could detect a transcript end read that included the 3' universal linker in at least 80% of the tested groups and determine the poly(A) site with high confidence. The majority of the identified fragments showed less than 20 nucleotide length differences between the proximal and distal APA variants. Specifically, 42 of those APA sites were located within the annotated 3'-UTR regions (UTR-APA) whereas only one novel APA site protruded into the coding region (TFIP11; CR-APA; Fig. 5d). We then tested the option that expression changes between distal and proximal APA variants could change the regulation by miRNA (Lee et al., 2008). Using the microT prediction algorithm (Maragkakis et al., 2009), we measured the number of miRNA recognition elements (MREs) across the human genome that are located 5 nucleotides upstream of the putative distal poly(A) site, and could hence be affected by the majority of the identified APA variations. We found that only 2.9% of human genes have one or more MREs at this genomic location (Histogram for MRE counts across human genes is shown in Supplementary Fig. 10). Altogether, these tests demonstrated APA variations of limited length, scope and impact over miRNA regulation that were seemingly unrelated to AD's alternative splicing changes.

Group analyses support disease relevance of RNA processing and APA differences

To compare APA variants for each gene as reflected in read counts of specific dinucleotide primers, we estimated redundancy between dinucleotide fields. There was less than 1% redundancy, indicating effective separation. Figure 6a and 6b show such analysis for the gamma subunit of the growth arrest and DNA-damage-inducible protein GADD45G. Total summated GADD45G counts presented no global differences, but its APA variants differed significantly between AD and control samples. The GADD45G 'AC' and 'GA' products were higher in controls compared to AD samples and in AD compared to controls, respectively; and the global variability in GADD45G levels was considerably larger in control brains. Supplementary Figure 11 shows a correlation matrix for the GADD45G gene APA patterns between all samples, demonstrating that the AD brain alters the preferred 3'-APA variants of this growth arrest gene while also tightening the scope of variability in its levels. Parallel matrices enabled averaging these correlation matrices across all genes, with a similar outcome. Cognition-related differences in APA variants were hence small, but reproducible.

Notably, the β-amyloid precursor protein (Stein et al., 2010, Stein et al., 2012) is produced from short and long APA variants (APP-S, APP-L) (Mbella et al., 2000). APP-L, but not APP-S (which is extended by 258 3'-UTR nucleotides) interacts with brain protein extracts that promote its

translation. This exacerbates APP-L translation, suggesting a potential influence on AD pathogenesis through APA modulation of APP expression. We detected the distal and proximal transcript end sites at a single nucleotide resolution level in the dinucleotide fields 'GC' and 'AT', respectively (Fig. 6c). Also, the identified APP coding region matched the known APP gene structure in each of these two fields, and the AD but not the NDWP brain was enriched with the APP-S variant (Fig. 6d), (One-way-ANOVA $P<0.05$). Two previously identified and validated APA variants are also known for the calmodulin kinase CAMK2A gene, which displays elevated expression in AD (Wang et al., 2005). Both variants showed higher expression levels in the AD and NDWP samples as compared to healthy controls. Validated APA sites of the SPP1, TCEB2 and SOD1 genes corresponded to human brain-originated mRNA structures that were constructed from cDNA clones (http://genome.ucsc.edu/). Extracted transcript end site reads showed for each of these transcripts the known APA site variants at the reported positions (Supplementary Fig. 12).

Brain, vascular and autoimmune disease genes show cognition-related APA differences

For associating the differential polyadenylation patterns with particular AD-related processes, we matched the NBB expression differences of each APA variant to one of 6 hypothetical expression profiles: Pathology-associated increases or decreases in AD and NDWP compared to controls (1, 2); Cognition-associated increases or decreases in controls and NDWP compared to AD (3,4), and potential regulatory increases or decreases in NDWP compared to AD and controls (5,6; Fig. 6e). Transcripts were defined as matching a profile if they showed a P value<0.05 and a fold change>2, and as un-altered if they had a P value>0.5 and a fold change of less than 10%. Figure 6f presents box plots for the noted APA variants of the apoptosis-related gene LIG4, the mitochondrial gene NPY, the immune response gene SLC6A9, the synaptic transmission gene AVPI1 and the cell growth and differentiation-regulating gene TBC1D7. The NDWP-elevated fifth profile, but not the second, also showed enrichment in syntaxin-binding protein 1 (n-Sec1), important for synapse functioning (Arancillo et al., 2013). See lists of genes belonging to each of the profiles in Supplementary Table 10.

A total of 416 genes corresponded to these particular profiles: 98 genes showed increased expression in AD and NDWP compared to controls, 140 genes showed increased expression in NDWP compared to AD and controls and 45, 43, 48 and 42 genes corresponded to the 2^{nd}, 3^{rd}, 4^{th} and 6^{th} profiles, respectively (Fig. 6g). To explore differential efficiency between such APA transcripts, we selected three genes that showed significantly altered APA and in which the distal segment of the 3'-UTR was at least 60 nt long; APP, ACAT2 and GIMAP5. Luciferase assay enabled measuring translation efficiency separately for the short and long APA transcripts. Normalized luciferase activity for transcripts harboring the long (L - blue bars) or the short (S – red bars) 3'UTRs of APP, ACAT2 and GIMAP5 demonstrated higher expression of the short APP variant, whereas ACAT2 and GIMAP5 both showed considerably lower activity for their short variants (Fig. 6h; see Methods for more details). Experimental manipulations thus support disease relevance for the observed differences, at least for some transcripts.

We performed Gene Ontology analysis for each profile separately, using the DAVID bioinformatics tool (https://david.ncifcrf.gov/), with the entire expressed transcriptome as a background (Table 1). This revealed a number of profile 3 transcripts that were over-expressed in AD compared to controls and NDWP and are known to be causally associated with diverse human diseases. Examples include the LCAT, MT1F, MY1G, CLINT1 and SPP1 genes, all implicated in AD, Parkinson's disease and autoimmune diseases. Moreover, other genes in this group, such as HLA-DRA, P2RY12, S100A4, TBXAS1 and SLC6A9 emerged as already being targets of FDA-approved drugs for treating multiple sclerosis, cerebrovascular disease and schizophrenia, respectively (See Supplementary Table 11 for full details). To the best of our knowledge, this demonstrates for the first time the existence of genes whose expression is up-regulated in the brain of cognitively deteriorating AD patients but not in

histopathology-affected brains of AD patients without prior dementia; and that part of those are seemingly involved in other brain or body pathologies.

APA *patterns enable superior patient clustering compared to total transcripts*

Dendrogram plots based on Euclidean distances between APA variants of all genes (Fig. 7a-c) showed that APA ratios correctly classified AD from controls and AD from NDWP (with one misclassified patient in each comparison), but failed to distinguish between the control and NDWP groups, cautiously suggesting relevance of the cumulative APA variations to the cognitive state of tested patients. In comparison, total summated expression levels failed to cluster the different patient groups in all of the three possible comparisons (AD-control, NDWP-control and AD-NDWP, Fig. 7d-f), either in the NBB or in the Rush groups. Partial least square analysis (PLS) based on the APA variations, but not on summated sequencing similarly yielded a clear classification outcome, such that the distribution of patients in the plane of the first and second PCA components showed clear clustering (Fig. 7g). Thus, the AD group was successfully classified by both dendrogram and PLS analyses, indicating involvement of APA differences in the cognitive aspects of the disease process.

Discussion

Combining RNA-Seq with lipidomic LC-MS tests, we found inter-related RNA metabolism and lipidomic differences in AD temporal lobe tissues, which also accumulated several known disease-related transcripts of brain, vascular and autoimmune nature compared to control brains. Inversely, tissues of non-demented patients with pathology showed facilitated neuronal RNA and lipid processing, and both brain lipids and 3'-polyadenylation variants enabled superior patient classification over global transcripts, indicating cognition-related roles for the brain's RNA and lipid processing. Previous studies also identified expression changes associated with lipid metabolism as well as inflammation processes (Blalock et al., 2004; Haroutunian et al., 2009) and the involvement of ApoE genotype in these processes. Also, a very recent report argues a possibly relevant innate immune role for alternative polyadenylation (Jia et al., 2017). However, due to the known variation between AD patients and the small size of the analyzed groups, which were composed of males alone, the reproducibility of these results across studies still awaits a more thorough assessment.

Our findings indicate that the point in time in which most cellular changes accelerate is closer to the initiation of cognitive deterioration than to the start of the neuropathology phenotype; which is compatible with reports of association between cognitive decline and neuronal demise, but not AD pathology (Andrade-Moraes et al., 2013). In two different groups and sequencing strategies, cognitive impairment associated with more significant transcript differences than pathology. This may indicate distinct demise mechanisms of transcripts and/or brain cells in patients with delayed cognitive decline. Specifically, the identified AD and NDWP-associated differences in RNA processing and lipid composition may indicate a previously un-foreseen interim period. This period involves accelerated neuronal gene expression and elevated production levels of cortical lipids. It occurs prior to the initiation of neuronal death in the diseased brain, and is followed by neuronal hypo-activation and neurodegeneration. Future work should examine therapeutic possibilities to prolong this neuronal hyper activation state and by that delay cognitive deterioration (see also Supplementary Discussion).

Gene enrichment analysis showed increased apoptosis and reduced mitochondrial functioning for both AD and NDWP brains, suggesting particular relevance of these pathways in neuropathology development. AD and NDWP brains also differed from controls in several neuronal alternative splicing-related genes shown by others and by us to be causally involved in both non-familial and familial neurodegenerative syndromes (Berson et al., 2012; Hanada et al., 2013), whereas NDWP brains showed ATP synthesis-related and mitochondrial APA differences, compatible with the hypothesis of metabolically-facilitated delay of AD-characteristic dementia.

We identified APA differences by multiple 3' primer-based RNA-sequencing, with high signal to noise ratio and superior patient classification over total sequencing results. This provided an unprecedented survey of brain-expressed full-length transcripts which distinguished both advanced and early stage AD patients from non-demented donors with or without histopathology, attested to the innate-immune-related disease relevance of the APA expression status, and enabled the classification of cell type-, cognition-, lipidomic and neuropathology-associated transcript differences. Our data faithfully recapitulated and significantly exceeded previous APA information on brain and neurodegenerative disease transcripts, while finding the APA process to be tightly maintained at both the histopathology and cognitive decline stages of AD. We conclude that the AD brain presents distinct APA profile and alternative splicing differences. One example of unique interest is that AD brains over-express the more efficiently translated APP-S variant, whereas the less efficiently translated APP-L variant is the major one in NDWP brains, extending previous in vitro predictions of modulated AD risks by these APA variants.

At the translational research level, our AD transcript profiles included distinct increases in immune response genes and decreases in synaptic transmission genes compared to the other groups, indicating association of these differences with the neuro-inflammation and cognitive deterioration characteristic of AD (Heneka et al., 2015). Moreover, we discovered multiple differences in known therapeutic target genes of brain, vascular and autoimmune disorders, predicting involvement of the corresponding transcripts in the AD cognitive decline. For example, the G protein-coupled receptor gene P2RY12 is elevated in AD and interestingly is implicated in peripheral and cerebrovascular diseases (Culebras et al., 2014), reminiscent of apolipoprotein E that controls cerebrovascular integrity in healthy individuals and fails to do so in AD patients (Bell et al., 2012). The P2RY12-targeting drug Clopidogrel (Pfizer) is chronically administered to acute coronary syndrome patients. Likewise, reduced expression of S100A4, known to inhibit microtubule assembly and protein kinase C-mediated phosphorylation has been implicated in cardiomyopathies (Cmoch et al., 2012), which can be treated by the psychoactive phenothiazines. Our findings call for investigating the potential beneficial effect of these and related drugs on delaying or limiting the risk of AD-related cognitive decline.

Another potential advantage of our study involves our findings of novel cognition-disrupting targets for therapeutic intervention. We found larger distances and distinct transcript differences between AD patients and controls compared to NDWP and controls, cautiously suggesting more pronounced involvement of RNA metabolism in AD cognitive deterioration than in its neuropathology. Specific examples include SLC6A9, which contributes to NMDA potentiation (Pich et al., 2012), and GRIN2B, encoding the NMDA receptor NR2B subunit that harbors a single AD-associated nucleotide polymorphism (SNP rs10845840) (Stein et al., 2010; Stein et al., 2012). Also, aberrant cholesterol processing in AD (Pooler et al., 2006) implicates the cholesterol-processing gene LCAT in the disease profile (Demeester et al., 2000); and a recent lipid profiling study identified AD-antecedent memory impairments, supporting relevance of the lipid processing-associated gene TBXAS1 (Mapstone et al., 2014). Likewise, MT1F and MT1G function in metal storage that is impaired both in Parkinson`s disease and AD (Barnham and Bush, 2008). That these genes are up-regulated in AD but not in healthy controls or NDWP suggests their involvement in AD's cognitive deterioration and makes them potentially useful therapeutic targets for delay of deterioration, either alone or in multi-drug combinations, and distinct from the current approach of palliative treatment. In summary, in spite of the limitation involved in the small size cohorts that were analyzed, our findings offer both predictable and unexpected findings of great basic and translational significance.

Methods

Tissue samples and diagnosis

Two sets of human male brain tissues, of varying gyri of the temporal lobe, were received from Rush University (N=36) and from the Netherland Brain Bank (NBB; N=24). The Rush groups comprised temporal lobe brain tissue from participants in the Memory and Aging Project. All participants enroll without known dementia and agree to annual details clinical evaluation and brain donation. Details of the clinical and pathologic evaluation have been described (Bennett et al., 2012). Participants signed an informed consent, an anatomical gift act, and a consent to place the data and bio-specimens in a repository for future use. The procedures were approved by the institutional review board of Rush University Medical Center. Frozen material was stored in 1cm slabs in a -80^0 freezer until pulled and dissected for this study. Full patient data is in Supplementary Tables 1 and 2. Potential confounding measurements differential expression between the groups are shown in Supplementary Figure 13.

Before homogenization samples were transferred onto ice, weighed quickly, and 10-fold volumes of ice cold lysis buffer (10mM Tris–Cl pH 7.4, 1M NaCl, 1mM EGTA, 1% Triton-X100) were added containing phosphatase inhibitors (20mM NaF, 1mM β-glycerophosphate, 1mM orthovanadate) and protease inhibitors (1:200 of Calbiochem cocktail set III) for Western Blot protein samples. Samples were lysed with a Kontes pellet pestle for 10-15 seconds, incubated for 30 minutes on ice and centrifuged for 30 minutes, 4°C, at 17,900 rcf. Supernatants were transferred to clean test tubes and stored at −20°C until use. For Western Blotting, samples were separated by SDS-PAGE (10µg total protein per lane) and transferred to nitrocellulose. hnRNPA1 and neuron-specific beta III tubulin were detected with primary antibodies (Santa Cruz sc-56700 at 1:200 and Abcam ab-78078 at 1:500, respectively), HRP-conjugated secondary antibodies (Jackson 115-035-062, 1: 20:000) and ECL (SuperSignal West Femto, Thermo 34094), using the myECL Imager and myImage software (Thermo Scientific) for detection, analysis and quantification. Hydrolytic activity of cholinesterases was measured using Ellman assay as previously described (Berson et al., 2012).

QuantSeq

Library preparation and sequencing

Per sample, 20 ng of the purified full-length total RNA were used for performing the 3’ QuantSeq Reverse kit (Lexogen GmbH, Vienna, Austria) resulting in NGS libraries which originate from the 3’ end of polyadenylated RNA. Library generation was started by oligo(dT) priming with primers already containing the Illumina-compatible linker sequence for Read 1. After first strand synthesis, the RNA was removed before the second strand synthesis was initiated by random primers which contained the corresponding Illumina-compatible linker sequence. 3’ QuantSeq generated just one fragment per transcript at the very 3’ end. The libraries were PCR amplified and barcoded in 18 cycles. Equal amounts of the 72 libraries were combined to one lane mixture.

Illumina HiSeq 2500 (Illumina, Inc.) was used in rapid run mode to sequence the 3’ QuantSeq library mixture in two lanes in 150 bp single read mode. Because Read 1 started directly from the poly(A) tail the Multiplex Read 1 Sequencing Primer needed to be replaced by a Custom Sequencing Primer to cover the oligo(T) stretch at the beginning for achieving cluster calling and prevent dephasing.

Read alignment

TopHat 2.0.11 and Cufflinks 2.1.1 with the GRCh37 ENSEMBL hg19 genome sequence and associated ENSEMBL transcript annotation in gtf format were used for read mapping and aligning. Tables of read counts per gene were generated from the alignments using the HTSEQ package. The minimal sequence base quality value selected for further processing was 20 (Phred score). Bases with a quality value below this parameter were replaced with 'N'. Progressive alignment method was selected. Only primary alignments were considered for gene counts and quantification. The minimal identity seed for alignment extension was 25 nucleotides.

SQUARE

RNA extraction and 5'P RNA removal

RNA was extracted with the SPLIT RNA extraction kit (Lexogen GmbH, Vienna, Austria) yielding the large RNA fraction with a lower cut-off size of 150 nt. Evaluation on a Bioanalyzer RNA Nano chip (Agilent Technologies) showed medium to high RNA quality (RIN of 6.2 – 8.3), and samples with different RIN values showed similar RNA-seq qualities. 4 µg RNA was incubated with Terminator 5´-phosphate-dependent exonuclease (Epicentre, Madison, WI, USA) to remove degraded RNA while leaving the capped, full-length mRNA intact.

Reverse transcription and 3'-selective amplification

The complete procedure is offered as SQUARE^{TM} service (Selective Quantitative Amplification of RNA) by Lexogen (Vienna, Austria; work-flow available on www.lexogen.com.) Briefly, 4 µg Terminator-treated RNA was full-length reverse-transcribed, and 1/40th of the purified cDNA was then amplified in one of 12 parallel PCR reactions. Each PCR reaction was set up with one universal 5' primer and a 3' primer selective for one of the 12 possible combinations of the two terminal nucleotides of the mRNA body, immediately upstream of the poly(A) tail. This generated a so-called SQUARE matrix with 12 fields. Each of these matrix fields contains the amplification products of an mRNA sub-population with a given 3'-terminal dinucleotide, and all matrix fields combined correspond to the entire (mRNA) transcriptome.

Library preparation & lane mixing

200 ng of all 288 PCR products (24 RNA samples amplified separately in 12 matrix fields) were heat-fragmented and then ligated to SOLiD-compatible adapters. The libraries were PCR-amplified in 17 cycles, with PCR primers indexing each sample with one out of 96 bar codes. Three lane mixes were created, each from 96 libraries corresponding to 24 samples amplified in 4 matrix fields, dedicating equal molarities to each of the libraries. The three lane mixes were 1) for selective 3' primer nucleotides AC, AG, CA, GT, 2) AA, AT, CC, CG and 3) CT, GA, GC, GG; and were then further size-selected on a Pippin Prep automated gel elution system (Sage Biosciences, Beverly, MA, USA) in the range of 170 to 400 bp.

RNA Sequencing
The automated SOLiD EZ Bead System and SOLiD EZ Bead E80 System Consumables (Life Technologies Corp.) were applied for the template preparation. The SOLiD 5500xl System and single-end chemistry for RNA sequencing was applied (Life Technologies Corp.). Sequencing data is available from NCBI's Gene expression omnibus (GEO), Series record GSE57152.

Next generation sequencing (NGS) libraries made from the NBB temporal gyrus samples yielded an average of 6.0*106 (STD=2.0*106) uniquely aligned 75 base pair (bp) single end reads, or approximately 7.0*107 (STD=1.8*107) total read counts when combining all 12 SQUARE fields. These reads were mapped against the GRGCh37/hg19 version of the Homo sapiens genome (http://genome.ucsc.edu/). Transcripts with more than 1 read per kilobase per million (RPKM) per SQUARE field were defined as being detected, whereas any RPKM below 1 was considered as reflecting leakiness of the transcription (Hebenstreit et al., 2011). An average of 6610±1367 genes per field were detected across the 12 fields (details in Supplementary Table 10). Expression criteria were set to RPKM>1 in at least one of the SQUARE fields, in at least 80% of the tested donor groups (Supplementary Table 12). Under these criteria, 10,885 genes were expressed, 2365 with a single poly(A) site and the remaining 8520 with two or more poly(A) sites. Comparison to an expressed sequence tags (EST) database; PolyA_DB (Zhang et al., 2005), http://exon.umdnj.edu/polya_db/), yielded 5373 expressed genes in the human temporal lobe which had two poly(A) sites or more

according to the EST-based database. SQUARE detected more than one poly(A) site in 4353 out of these 5373 genes (81% overlap between these two resources, which may also differ due to the tissue selected) while identifying 4167 novel APA transcripts, showing a high overlap with previous APA data. Sequencing data is available from NCBI's GEO, Series record GSE70424

Read alignment

Sequence and quality files, in Lifetech proprietary Color Space format, were mapped against the GRGCh37/hg19 version of the *Homo sapiens* genome using the Lifetech Lifescope 2.5.1 whole Transcriptome analysis pipeline separately for single and merged SQUARE matrix fields. The files produced by this analytical pipeline were coverage; alignment (.bam files); exon junction; exon expression in RPKM; gene expression in RPKM with reference to the RefSeq gene structure; read counts with reference to each gene. Mapping and sequence / quality control metrics were generated both with the Lifetech suite and the Integromics SeqSolve analysis suite on all the samples. Read lengths were 75 nucleotide fragments, with a percentage of genome alignments over the whole sequence length over 80%. The minimal sequence base quality value selected for further processing was 20 (Phred score). Bases with a quality value below this parameter were replaced with 'N'. Progressive alignment method was selected. The minimum genome alignment quality value for an alignment to be processed was again 10 (Phred value). Only primary alignments were considered for gene count and quantification. The minimal identity seed for alignment extension was 25 nucleotides. For a comparison between QuantSeq and SQUARE see Supplementary Discussion.

Fluidigm RT-PCR

Samples were analyzed by quantitative RT-PCR (qPCR) using the Fluidigm 96.96 Dynamic Array Integrated Fluidic Circuits (IFC) and the Biomark System (www.fluidigm.com/biomark-system.html). All cDNA samples were pre-amplified by combining each sample with the primer pool and TaqMan Pre-Amp Mastermix Kit (PN 4384266, Applied Biosystems), total of 6.25 ng in 5ul reactions, following the Fluidigm Specific Target Amplification (STA) protocol (14 amplification cycles) enriching samples for loci of interest. Samples were subsequently treated with exonuclease 1 to remove any single-stranded material. Finally, samples were diluted 1:5 in Tris-EDTA. A Fluidigm 96.96 Dynamic Array IFC was prepared according to the manufacturer's instructions. Quantitative PCR was performed using EvaGreen binding dye (SsoFast EvaGreen Supermix with Low ROX, Bio-Rad) according to the Biomark System protocol.

Standard RT-PCR

Duplicate real-time reverse transcriptase (RT)-PCR tests involved CyberGreen master mix (Quante). PCR was performed using Taq DNA polymerase (Sigma). Primer sequences are listed in Supplementary Table 4. Annealing temperature was 60°C for all primers. Serial dilution of samples served to evaluate primer-efficiency and the cDNA concentration that yields linear differences. Absence of genomic DNA was verified using NRT controls.

Lluciferase assays

The complete 3'-UTRs of human APP, ACAT2 and GIMAP5 were amplified from genomic DNA, and cloned downstream of the Renilla luciferase gene in psiCHECK2 (Clonetech) using XhoI and NotI. To create vectors expressing the short 3'-UTR alone, this region of the 3'- UTR from each of the cloned transcripts was similarly sub-cloned into psiCHECK2. To obtain expression of the long 3' UTR isoform alone, the proximal polyadenylation site was mutated from AAUAAA to ACUCAA (APP, ACAT2) and AAUAGA to AAUCGA (GImAP5) using QuickChange Site Directed Mutagenesis (Agilent). HEK-293 cells were cultured in DMEM supplemented with 10% fetal bovine serum and transfected with TransIT-X2 (Mirus) according to the manufacturer's instructions. Cells were then incubated overnight before performing luciferase assays. Luciferase activity was assessed using the Dual-Glo system (Promega)

performed according to the manufacturer's instructions. Renilla fluorescence was normalized to firefly signal, and results are presented as this ratio.

Statistical analysis

To check significance of total expression and expression in specific fields, across the three patient groups (AD, Con, NDWP), One-Way-ANOVA was performed followed by post-hoc Tukey`s test. To determine significant change between distributions, the Kolmogorov-Smirnov test was performed. In the Gene Ontology enrichment analysis gene groups were downloaded from the Gene Ontology Consortium (http://www.geneontology.org/) and statistical tests (as described in each section) to assess the effect on these gene groups were performed (see Supplementary Discussion).

GC-MS experiment

Metabolites were extracted from the frozen brain tissue powder by methanol: methyl-tert-butyl-ether (1:3 (v/v)) extraction (Khrameeva et al., 2014). In brief, approximately 25 mg of frozen powdered tissue was re-suspended in 1 ml extraction solution containing two internal standards (1 mg of 1,2-diheptadecanoyl-sn-glycero-3-phosphocholine (PC 34:0)). The samples were incubated for 10 min at 4 °C on an orbital shaker, before subjecting them to ultra-sonication for 10 min in an ice-cooled bath-type sonicator. The insoluble tissue material (including proteins) was pelleted by a centrifugation step (5 min; 14,000g) and the supernatant was transferred to a fresh 2ml Eppendorf tube. To separate the organic from the aqueous phase, 650 ml of an H_2O: methanol mixture (3:1(v/v)) was added to the supernatant, mixed by vortexing and centrifuged (5 min; 14,000g). Five hundred microliters of the upper lipid- (methyl-tert-butyl-ether) phase was transferred to a fresh 1.5-ml Eppendorf tube, concentrated in a speed vacuum and the pellet was re-suspended in 100 µl of an acetonitrile: isopropanol mixture (7:3 (v/v)) before liquid chromatography–mass spectrometry analysis. For the LC-MS analysis two times 5 µl of lipid the re-suspended lipid extract was injected onto the ultra- performance liquid chromatography C_8-reversed phase column (BEH C8, Waters), connected to an Orbitrap Exactive mass spectrometer and analyzed once in positive and once in negative ion mode (Giavalisco et al., 2011; Hummel et al., 2011; Khrameeva et al., 2014). Data was annotated, processed and normalized as previously described in (Giavalisco et al., 2011; Khrameeva et al., 2014).

Author Contributions

S.B. designed and performed experiments, analyzed data and wrote the paper; H.S. initiated and guided this study, designed experiments, analyzed data and wrote the paper. E.R.B., B.N., C.N., and L.P. performed experiments. A.S., A.T., T.R., Y.B.P. and A.G. analyzed data. D.G. wrote the paper.

Acknowledgments

This work was supported by the European Research Council Advanced Award [grant number 321501], the Legacy Heritage Science Initiative (LHSI) of the Israel Science Foundation [grant number 378/11], the Nofar program of the Israel Innovation Authority of the Ministry of Economics and Industry [grant number 56802], the Ministry of Science, Technology and Space [grant no. 53140] and the Austrian Research Promotion Agency (FFG Bridge1 project) [Grant numbers 853294 and 6104542]. S.B. was an incumbent of the TEVA National Network of Excellence in Neuroscience—NNE fellowship. B.G. and R.M. won post-doctoral and pre-doctoral fellowships by the Edmond and Lily Safra Center For Brain Sciences.

Legends to Figures

Figure 1: Research outline. (a) Rush groups (b) NBB groups, with three sets of brain tissues from donors at three levels of dementia and pathology: apparently healthy controls (Braak stage =0), documented early pathology (Braak stages 3-4) but no discernible dementia (NDWP) and advanced AD pathology and cognitive decline (Braak stages 5,6). Numbers of analyzed samples are noted above. (c) Sequencing data from the two cohorts were generated with two methods of library construction for RNA sequencing: 3'-QuantSeq, generating one most 3'-directed read per transcript, and SQUARE, identifying transcript variants with alternative 3'-UTR-APA by the 3'-segregation approach (scheme), based on 12 reverse-transcription sets using distinct poly(A)-adjacent 3'-primers (Methods), followed by barcoding and sequencing. (d) Schematic analytic design. Sequencing xsq files derived in both methods were processed with Liftech Lifescope software. Aligned bam files served to calculate RPKM values and detect APA sites, differential expression and enrichment results. (e) Validation tests involved: high-throughput Microfluidic RT-PCR, specific protein blots and enzyme activity tests and comparison to other databases.

Figure 2: Identifying strong cognition-associated expression differences. (a) Shown are histograms of the specific P values of 2-Way-ANOVA in which pathology and cognition served as the two parameters for the Rush groups. The Y axis shows cumulative counts of all transcripts. Circles show the numbers of genes changed with cognition (yellow) or pathology (red) in AD patients, and the overlap between them. Columns show the GO categories ranked by the P value of their enrichment score. (b) Venn diagram of differentially expressed genes in the AD and NDWP NBB groups and their associated GO enrichment terms, with the scale showing enrichment scores. (c) Cholinesterase activity changes in the NBB and Rush groups. (d) Expression fold change of AChE transcript levels in the three NBB patient groups (Mean ±SEM). One-way-ANOVA $P<0.05$. (e) Scatter plot and correlation across the NBB patient groups of AChE RNA level with AChE enzymatic activity.

Figure 3: Inverse patterns of lipid profile differences in NDWP and AD. (a) Cumulative distribution functions of specific lipid subgroup levels for the AD and NDWP groups, each versus controls. KS= Kolmogorov Smirnov test for comparing distributions. (c) Partial least Square and discriminant analysis for lipid levels shown as scatter plot on the first and second components of the analysis. Controls (green) are inseparable from NDWP (yellow) while the early (blue) and late AD(red) groups are separable from one another.

Figure 4: In-depth sequencing identifies cell type-specific, cognition- and pathology-associated transcript differences. (a-h) NDWP NBB brains show up-regulated neuronal transcripts and down-regulated transcripts in other cell types. (a-d) Shown are cumulative distribution functions (CDFs) for the global change between AD or NDWP and control (dashed lines, red and blue, correspondingly) for the subgroups of cell type-specific genes, following Cahoy et al. (Cahoy et al., 2008) (solid lines, same colors). (a) Neuronal genes are sharply up-regulated in NDWP (a; Kolmogorov Smirnov (KS) P value < 0.0001). Columns: APA variant of the calcium sensor SYT1 with a 3'-CA terminus. (b) Oligodendroglia genes were down regulated in NDWP and up-regulated in AD (KS P value < 0.001). Columns: a 3'-AC-terminated variant of the central nervous system myelin-associated CLDN11 gene. (c) Astroglial genes were moderately up-regulated in both disease groups (KS P value <0.01). Columns: a CA variant of the metabolic regulator ATP1A2. (d) Microglia-related genes were up regulated in both NDWP and AD. Columns: a 3'- GA variant of the differentiation marker CD53 and following Zhang (Zhang et al., 2013). (e-h) Parallel box plot presentation of normalized values for a different list of human cell-type characteristic genes, following Darmanis (Darmanis et al., 2015) with healthy control levels referred to as 1. Note parallel differences in this analysis as well. (i) AD-related differences in RNA metabolism genes: normalized RPKM values across the three patient groups

(mean±SEM) for EWSR1 and HNRNPH2 (One-way-ANOVA P values=0.02 and 0.04, correspondingly), CLP1 and TARDBP (One-way-ANOVA P values=0.005 and 0.02, correspondingly), and NOVA1 and HNRNPA1 (One-way-ANOVA P values=0.005 and 0.02, correspondingly).

Figure 5: Characteristics of APA variations in the NBB groups. (a) Most transcripts show diverse APA variations. The histogram shows the % contribution of the maximal field to the observed reads for each transcript. Insets show distribution of expression across the fields for three examples. (b) Distance in nucleotides from RefSeq TES for the first (field A) and second (field B) major products, for a single patient.. Right graph is zoom-in of left graph as shown on the X and Y axes. (c) Number of the poly(A) site consensus sequence 'AAUAAA' as distance from poly(A) site in sequencing reads. (d) 3'-UTR regions for representative genes with at least 2 poly(A) sites identified, aligned according to their transcription end site (Gaugler et al., 2014). Proximal and distal APA variants are marked as red triangles. Right side part: zoom-in for the TES-adjacent 20 nucleotide region for each of the genes shown. Note that the APA distance is shorter than 20 nucleotides for over 80% of these transcripts.

Figure 6: APA variants classify patient groups better than cumulative transcripts from all fields together. (a) Expression levels of GADD45G in AD patients and controls. (b) APA pattern of GADD45G in AD (red) and controls (blue). Shown are averages ± SEM of expression in each SQUARE field in the patient groups, note switch between the proximal and distal products from controls to AD. Fields 'CA' and 'GA', but none of the others neither the global transcript counts show significant change. (c) APP gene structure with two APA variants; one with a long (APP-L) and one with a short 3'-UTR (APP-S). Numbers of sequencing reads across APP exons in SQUARE fields 'GC' and 'AT'. The unique 3'-UTR region of APP-L is marked in a dashed rectangle. (d) Number of reads in specific SQUARE fields divided by the number of total transcript reads in each sample for APP-L and APP-S. In both cases One-way-Anova-P value has been <0.05. (e) Predicted profiles of expression differences between the groups may reflect association with pathology, cognitive decline or NDWP-specific features. (f) Mean ±95% confidence level for the exemplary genes LIG4, NPY, SLC6A9, AVPI1, TBC1D7, and NDUFA3, each presenting a distinct profile (colors as in a). (g) Bar graph for transcript APA counts in each profile (colors as in e). **(h):** Normalized luciferase activity for transcripts harboring the long (L - blue bars) or the short (S – red bars) 3'UTRs of APP, ACAT2 and GIMAP5. Briefly, the complete 3'-UTRs of human APP, ACAT2 and GIMAP5 were amplified from genomic DNA, and cloned downstream of the Renilla luciferase gene in psiCHECK2 (Clonetech) using XhoI and NotI. To create vectors expressing the short 3'-UTR alone, this region of the 3'- UTR from each of the cloned transcripts was similarly sub-cloned into psiCHECK2. To obtain expression of the long 3' UTR isoform alone, the proximal polyadenylation site was mutated from AAUAAA to ACUCAA (APP, ACAT2) and AAUAGA to AAUCGA (GImAP5) using QuickChange Site Directed Mutagenesis (Agilent). HEK-293 cells were cultured in DMEM supplemented with 10% fetal bovine serum and transfected with TransIT-X2 (Mirus) according to the manufacturer's instructions. Cells were then incubated overnight before performing luciferase assays. Luciferase activity was assessed using the Dual-Glo system (Promega) performed according to the manufacturer's instructions. Renilla fluorescence was normalized to firefly signal, and results are presented as this ratio.

Figure 7: Alternative polyadenylation segregates AD from NDWP and controls better than the summated total expression values. (a-c) Dendrograms based on APA ratio correlation for AD vs. controls (a), NDWP vs. controls (b) and AD vs. NDWP (c) based on all SQUARE fields and therein segregated transcripts. (d-f) Dendrograms based on total expression values obtained by pooling of all SQUARE fields for AD vs. controls (d), NDWP vs. controls (e) and AD vs. NDWP (f). (g) Patients' PLS differentiates AD from NDWP and controls based on alternative polyadenylation values.

Table 1: Pathway enrichment analysis for Figure 6:

		Clusters	enrichment score
Profiles	1	Programed cell death, apoptosis.	1.74
		RNA splicing, RNA binding, RNA processing, spliceosome, ribonucleoprotein complex.	1.7
		Hydrolase, protease, exopeptidase and endopeptidase activity.	1.31
	2	Mitochondrion, mitochondrion inner membrane, mitochondrion envelope, respiratory chain, Huntington`s disease, Parkinson's disease, Alzheimer`s disease, oxidation reduction.	2.91
		Ribosomal protein, ribosome, translation, ribonucleoprotein.	1.72
	3	MHC class II, immune response, cell adhesion molecules (CAMs), antigen processing and presenting, Asthma, autoimmune thyroid disease.	2.31
		Defense response, response to wounding, inflammatory response.	2.14
		Peptide binding, positive regulation of immune system.	1.65
		Glycoprotein, disulfide bind.	1.43
	4	Secretion, exocytosis vesicle-mediated transport.	2.36
		Synaptic transmission, neurotransmitter transport, neurological system process.	1.86
		Golgi apparatus	1.3
	5	Generation of precursor metabolites and energy, ATP biosynthetic process, ATPase activity, ion transport.	1.65
		Nucleotide binding, ATP binding.	1.63
		Unfolded protein binding, chaperone, protein folding.	1.62
		Melanosome, pigment granule, cytoplasmic vesicle	1.48
		Mitochondrion, mitochondrial membrane.	1.36
	6		

References

An, J. J., et al., 2008. Distinct role of long 3' UTR BDNF mRNA in spine morphology and synaptic plasticity in hippocampal neurons. Cell. 134**,** 175-87.

Andrade-Moraes, C. H., et al., 2013. Cell number changes in Alzheimer's disease relate to dementia, not to plaques and tangles. Brain. 136**,** 3738-52.

Arancillo, M., et al., 2013. Titration of Syntaxin1 in mammalian synapses reveals multiple roles in vesicle docking, priming, and release probability. J Neurosci. 33**,** 16698-714.

Balla, T., 2013. Phosphoinositides: tiny lipids with giant impact on cell regulation. Physiol Rev. 93**,** 1019-137.

Bandaru, V. V., et al., 2009. ApoE4 disrupts sterol and sphingolipid metabolism in Alzheimer's but not normal brain. Neurobiol Aging. 30**,** 591-9.

Barbash, S., et al., 2014. Global coevolution of human microRNAs and their target genes. Mol Biol Evol. 31**,** 1237-47.

Barbash, S., Soreq, H., 2012. Threshold-independent meta-analysis of Alzheimer's disease transcriptomes shows progressive changes in hippocampal functions, epigenetics and microRNA regulation. Curr Alzheimer Res. 9**,** 425-35.

Barnham, K. J., Bush, A. I., 2008. Metals in Alzheimer's and Parkinson's diseases. Curr Opin Chem Biol. 12**,** 222-8.

Beaudoing, E., et al., 2000. Patterns of variant polyadenylation signal usage in human genes. Genome Res. 10**,** 1001-10.

Bell, R. D., et al., 2012. Apolipoprotein E controls cerebrovascular integrity via cyclophilin A. Nature. 485**,** 512-6.

Bennett, D. A., et al., 2012. Overview and findings from the religious orders study. Curr Alzheimer Res. 9**,** 628-45.

Berson, A., et al., 2012. Cholinergic-associated loss of hnRNP-A/B in Alzheimer's disease impairs cortical splicing and cognitive function in mice. EMBO Mol Med. 4**,** 730-42.

Blalock, E. M., et al., 2004. Incipient Alzheimer's disease: microarray correlation analyses reveal major transcriptional and tumor suppressor responses. Proc Natl Acad Sci U S A. 101**,** 2173-8.

Braak, H., Braak, E., 1995. Staging of Alzheimer's disease-related neurofibrillary changes. Neurobiol Aging. 16**,** 271-8; discussion 278-84.

Bulleit, R. F., et al., 1988. Conserved and variable regions in the subunits of brain type II Ca2+/calmodulin-dependent protein kinase. Neuron. 1**,** 63-72.

Cahoy, J. D., et al., 2008. A transcriptome database for astrocytes, neurons, and oligodendrocytes: a new resource for understanding brain development and function. J Neurosci. 28**,** 264-78.

Cmoch, A., et al., 2012. S100A proteins in propagation of a calcium signal in norm and pathology. Postepy Biochem. 58**,** 429-36.

Couthouis, J., et al., 2012. Evaluating the role of the FUS/TLS-related gene EWSR1 in amyotrophic lateral sclerosis. Hum Mol Genet. 21**,** 2899-911.

Cruchaga, C., et al., 2014. Rare coding variants in the phospholipase D3 gene confer risk for Alzheimer's disease. Nature. 505**,** 550-4.

Culebras, A., et al., 2014. Summary of evidence-based guideline update: prevention of stroke in nonvalvular atrial fibrillation: report of the Guideline Development Subcommittee of the American Academy of Neurology. Neurology. 82**,** 716-24.

Darmanis, S., et al., 2015. A survey of human brain transcriptome diversity at the single cell level. Proc Natl Acad Sci U S A. 112**,** 7285-90.

Darvesh, S., et al., 2003. Neurobiology of butyrylcholinesterase. Nat Rev Neurosci. 4**,** 131-8.

Demeester, N., et al., 2000. Characterization and functional studies of lipoproteins, lipid transfer proteins, and lecithin:cholesterol acyltransferase in CSF of normal individuals and patients with Alzheimer's disease. J Lipid Res. 41**,** 963-74.

Di Giammartino, D. C., et al., 2011. Mechanisms and consequences of alternative polyadenylation. Mol Cell. 43**,** 853-66.

Di Paolo, G., Kim, T. W., 2011. Linking lipids to Alzheimer's disease: cholesterol and beyond. Nat Rev Neurosci. 12**,** 284-96.

Donev, R., et al., 2007. A role for SC35 and hnRNPA1 in the determination of amyloid precursor protein isoforms. Mol Psychiatry. 12**,** 681-90.

Elkon, R., et al., 2013. Alternative cleavage and polyadenylation: extent, regulation and function. Nat Rev Genet. 14**,** 496-506.

Fu, X., et al., 2011. Rapid metabolic evolution in human prefrontal cortex. Proc Natl Acad Sci U S A. 108**,** 6181-6.

Gaugler, T., et al., 2014. Most genetic risk for autism resides with common variation. Nat Genet. 46**,** 881-5.

Giacobini, E., Gold, G., 2013. Alzheimer disease therapy--moving from amyloid-beta to tau. Nat Rev Neurol. 9**,** 677-86.

Giavalisco, P., et al., 2011. Elemental formula annotation of polar and lipophilic metabolites using (13) C, (15) N and (34) S isotope labelling, in combination with high-resolution mass spectrometry. Plant J. 68**,** 364-76.

Grimson, A., et al., 2007. MicroRNA targeting specificity in mammals: determinants beyond seed pairing. Mol Cell. 27**,** 91-105.

Hagenaars, S. P., et al., 2016. Shared genetic aetiology between cognitive functions and physical and mental health in UK Biobank (N=112 151) and 24 GWAS consortia. Mol Psychiatry.

Hanada, T., et al., 2013. CLP1 links tRNA metabolism to progressive motor-neuron loss. Nature. 495**,** 474-80.

Haroutunian, V., et al., 2009. Transcriptional vulnerability of brain regions in Alzheimer's disease and dementia. Neurobiol Aging. 30**,** 561-73.

Hebenstreit, D., et al., 2011. RNA sequencing reveals two major classes of gene expression levels in metazoan cells. Mol Syst Biol. 7**,** 497.

Heinemann, U., et al., 2012. Blood-brain barrier dysfunction, TGFbeta signaling, and astrocyte dysfunction in epilepsy. Glia. 60**,** 1251-7.

Heneka, M. T., et al., 2015. Neuroinflammation in Alzheimer's disease. Lancet Neurol. 14**,** 388-405.

Hummel, J., et al., 2011. Ultra performance liquid chromatography and high resolution mass spectrometry for the analysis of plant lipids. Front Plant Sci. 2**,** 54.

Jia, X., et al., 2017. The role of alternative polyadenylation in the antiviral innate immune response. Nat Commun. 8**,** 14605.

Kabashi, E., et al., 2008. TARDBP mutations in individuals with sporadic and familial amyotrophic lateral sclerosis. Nat Genet. 40**,** 572-4.

Khrameeva, E. E., et al., 2014. Neanderthal ancestry drives evolution of lipid catabolism in contemporary Europeans. Nat Commun. 5**,** 3584.

Kim, H. J., et al., 2013. Mutations in prion-like domains in hnRNPA2B1 and hnRNPA1 cause multisystem proteinopathy and ALS. Nature. 495**,** 467-73.

Kiskis, J., et al., 2015. Plaque-associated lipids in Alzheimer's diseased brain tissue visualized by nonlinear microscopy. Sci Rep. 5**,** 13489.

Ko, M., et al., 2016. Phosphatidylcholine protects neurons from toxic effects of amyloid beta-protein in culture. Brain Res. 1642**,** 376-83.

Kolisnyk, B., et al., 2013. Forebrain deletion of the vesicular acetylcholine transporter results in deficits in executive function, metabolic, and RNA splicing abnormalities in the prefrontal cortex. J Neurosci. 33**,** 14908-20.

Lacour, A., et al., 2017. Genome-wide significant risk factors for Alzheimer's disease: role in progression to dementia due to Alzheimer's disease among subjects with mild cognitive impairment. Mol Psychiatry. 22**,** 153-160.

Lau, P., et al., 2013. Alteration of the microRNA network during the progression of Alzheimer's disease. EMBO Mol Med. 5**,** 1613-34.

Lazar, A. N., et al., 2013. Time-of-flight secondary ion mass spectrometry (TOF-SIMS) imaging reveals cholesterol overload in the cerebral cortex of Alzheimer disease patients. Acta Neuropathol. 125**,** 133-44.

Lee, J. Y., et al., 2008. Phylogenetic analysis of mRNA polyadenylation sites reveals a role of transposable elements in evolution of the 3'-end of genes. Nucleic Acids Res. 36**,** 5581-90.

Lee, Y., et al., 2012. Oligodendroglia metabolically support axons and contribute to neurodegeneration. Nature. 487**,** 443-8.

Mapstone, M., et al., 2014. Plasma phospholipids identify antecedent memory impairment in older adults. Nat Med. 20**,** 415-8.

Maragkakis, M., et al., 2009. Accurate microRNA target prediction correlates with protein repression levels. BMC Bioinformatics. 10**,** 295.

Martins, I. C., et al., 2008. Lipids revert inert Abeta amyloid fibrils to neurotoxic protofibrils that affect learning in mice. EMBO J. 27**,** 224-33.

Mazin, P., et al., 2013. Widespread splicing changes in human brain development and aging. Mol Syst Biol. 9**,** 633.

Mbella, E. G., et al., 2000. A GG nucleotide sequence of the 3' untranslated region of amyloid precursor protein mRNA plays a key role in the regulation of translation and the binding of proteins. Mol Cell Biol. 20**,** 4572-9.

Mesulam, M., et al., 2004. Cholinergic nucleus basalis tauopathy emerges early in the aging-MCI-AD continuum. Ann Neurol. 55**,** 815-28.

Muller, C. P., et al., 2015. Brain membrane lipids in major depression and anxiety disorders. Biochim Biophys Acta. 1851**,** 1052-65.

Pascoal, T. A., et al., 2016. Amyloid-beta and hyperphosphorylated tau synergy drives metabolic decline in preclinical Alzheimer's disease. Mol Psychiatry.

Pich, E. M., et al., 2012. Biomarkers for antipsychotic therapies. Handb Exp Pharmacol. 339-60.

Pooler, A. M., et al., 2006. The 3-hydroxy-3-methylglutaryl co-enzyme A reductase inhibitor pravastatin enhances neurite outgrowth in hippocampal neurons. J Neurochem. 97**,** 716-23.

Proudfoot, N. J., 2011. Ending the message: poly(A) signals then and now. Genes Dev. 25**,** 1770-82.

Shah, S. A., et al., 2016. Novel osmotin inhibits SREBP2 via the AdipoR1/AMPK/SIRT1 pathway to improve Alzheimer's disease neuropathological deficits. Mol Psychiatry.

Soreq, H., Seidman, S., 2001. Acetylcholinesterase--new roles for an old actor. Nat Rev Neurosci. 2**,** 294-302.

Stein, J. L., et al., 2010. Genome-wide analysis reveals novel genes influencing temporal lobe structure with relevance to neurodegeneration in Alzheimer's disease. Neuroimage. 51**,** 542-54.

Stein, J. L., et al., 2012. Identification of common variants associated with human hippocampal and intracranial volumes. Nat Genet. 44**,** 552-61.

Tian, B., et al., 2005. A large-scale analysis of mRNA polyadenylation of human and mouse genes. Nucleic Acids Res. 33**,** 201-12.

Timmusk, T., et al., 1993. Multiple promoters direct tissue-specific expression of the rat BDNF gene. Neuron. 10**,** 475-89.

Verkhratsky, A., et al., 2013. Astroglia in neurological diseases. Future Neurol. 8**,** 149-158.

Wang, Y. J., et al., 2005. The expression of calcium/calmodulin-dependent protein kinase II-alpha in the hippocampus of patients with Alzheimer's disease and its links with AD-related pathology. Brain Res. 1031**,** 101-8.

Whiley, L., et al., 2014. Evidence of altered phosphatidylcholine metabolism in Alzheimer's disease. Neurobiol Aging. 35**,** 271-8.

Zhang, B., et al., 2013. Integrated systems approach identifies genetic nodes and networks in late-onset Alzheimer's disease. Cell. 153**,** 707-20.

Zhang, H., et al., 2005. PolyA_DB: a database for mammalian mRNA polyadenylation. Nucleic Acids Res. 33**,** D116-20.

Zhang, Z., et al., 2008. SMN deficiency causes tissue-specific perturbations in the repertoire of snRNAs and widespread defects in splicing. Cell. 133**,** 585-600.

Figure 1

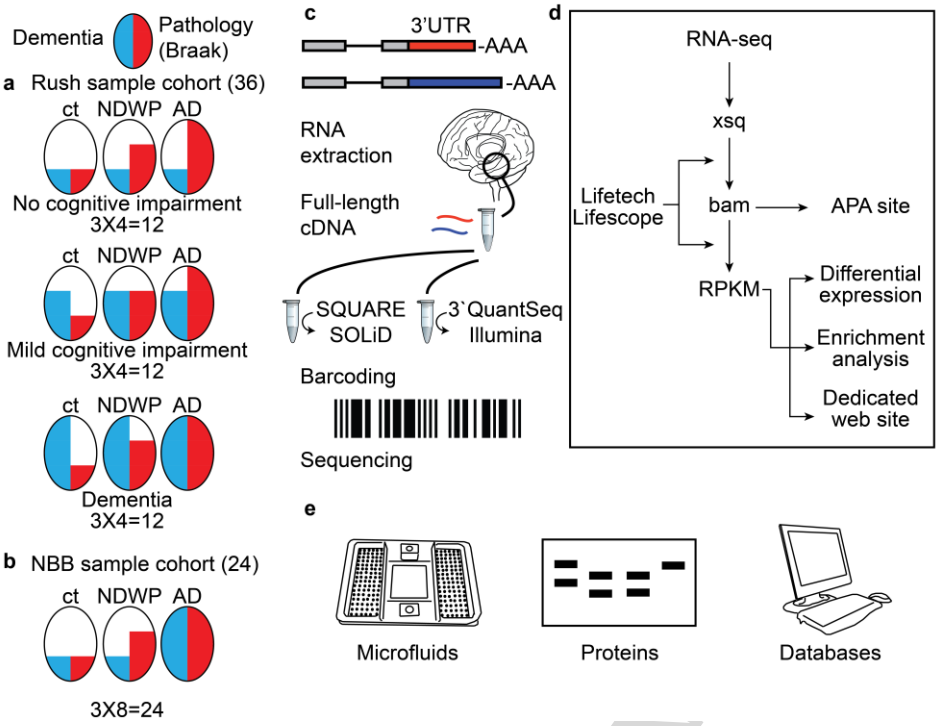


Figure 2

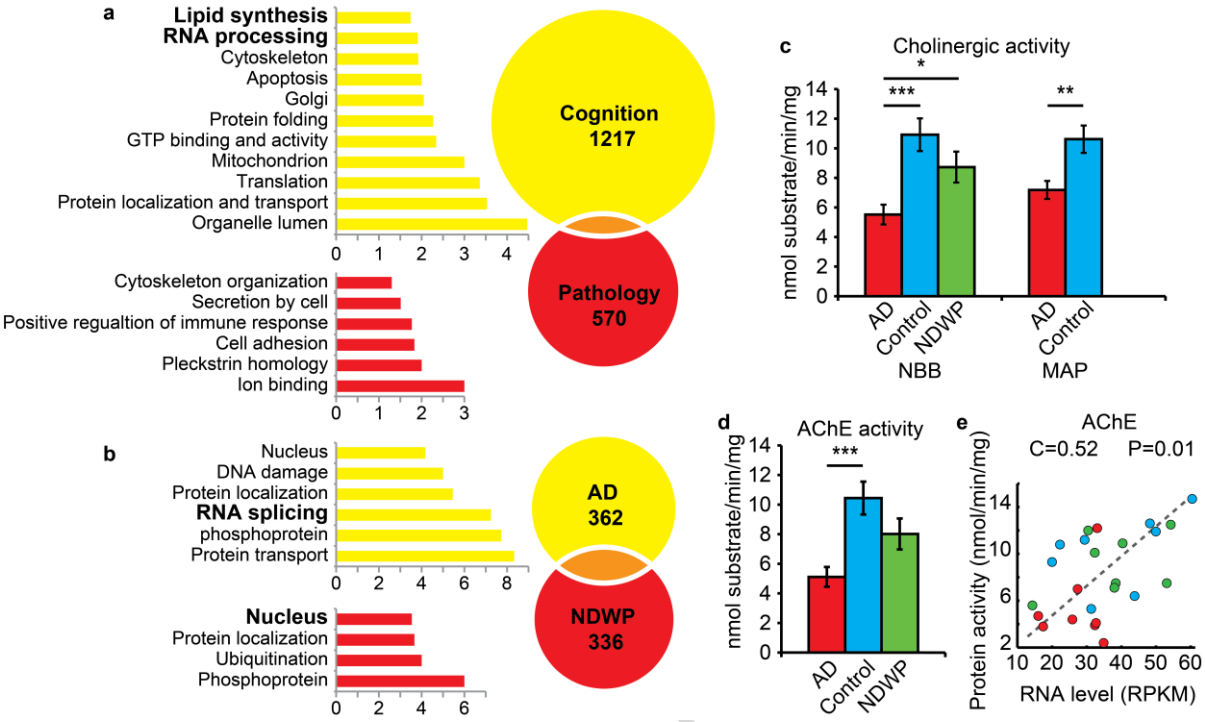


Figure 3

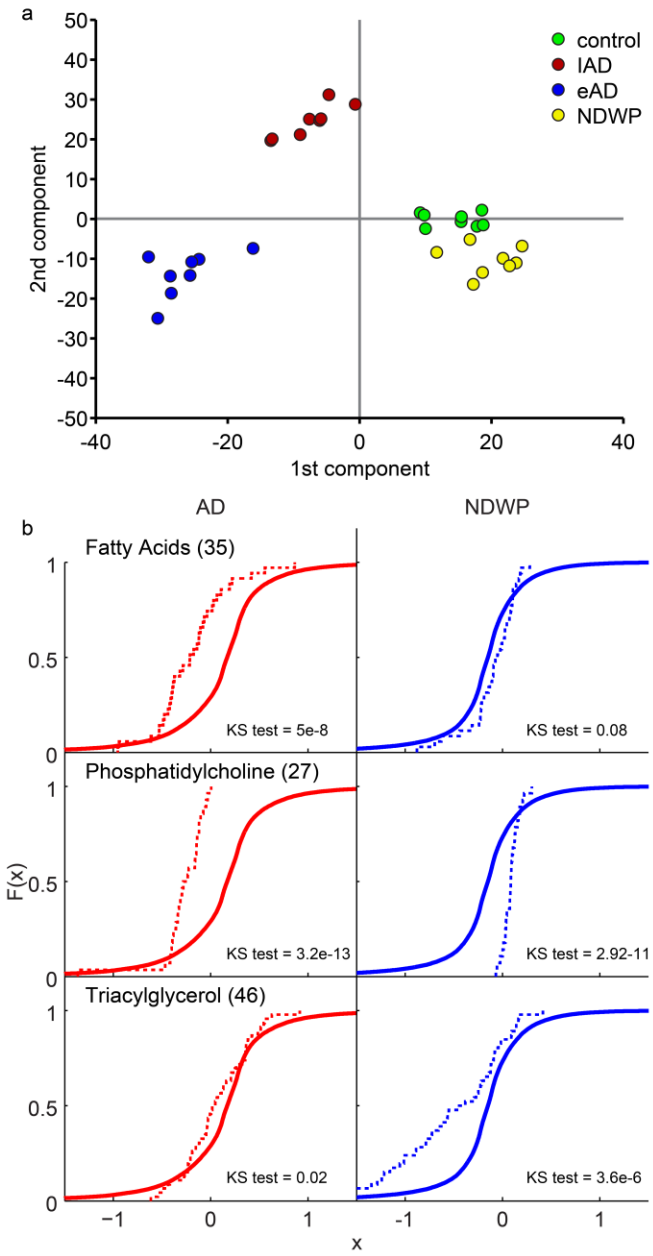


Figure 4

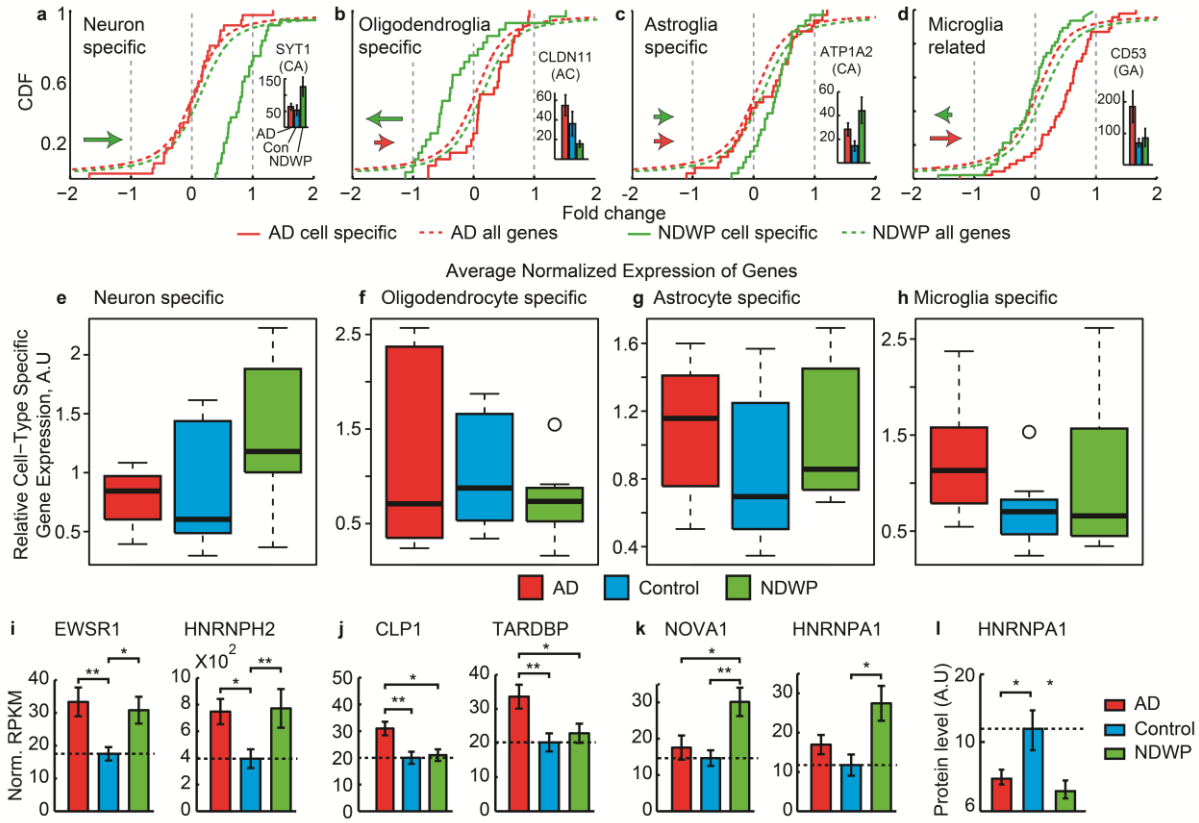


Figure 5

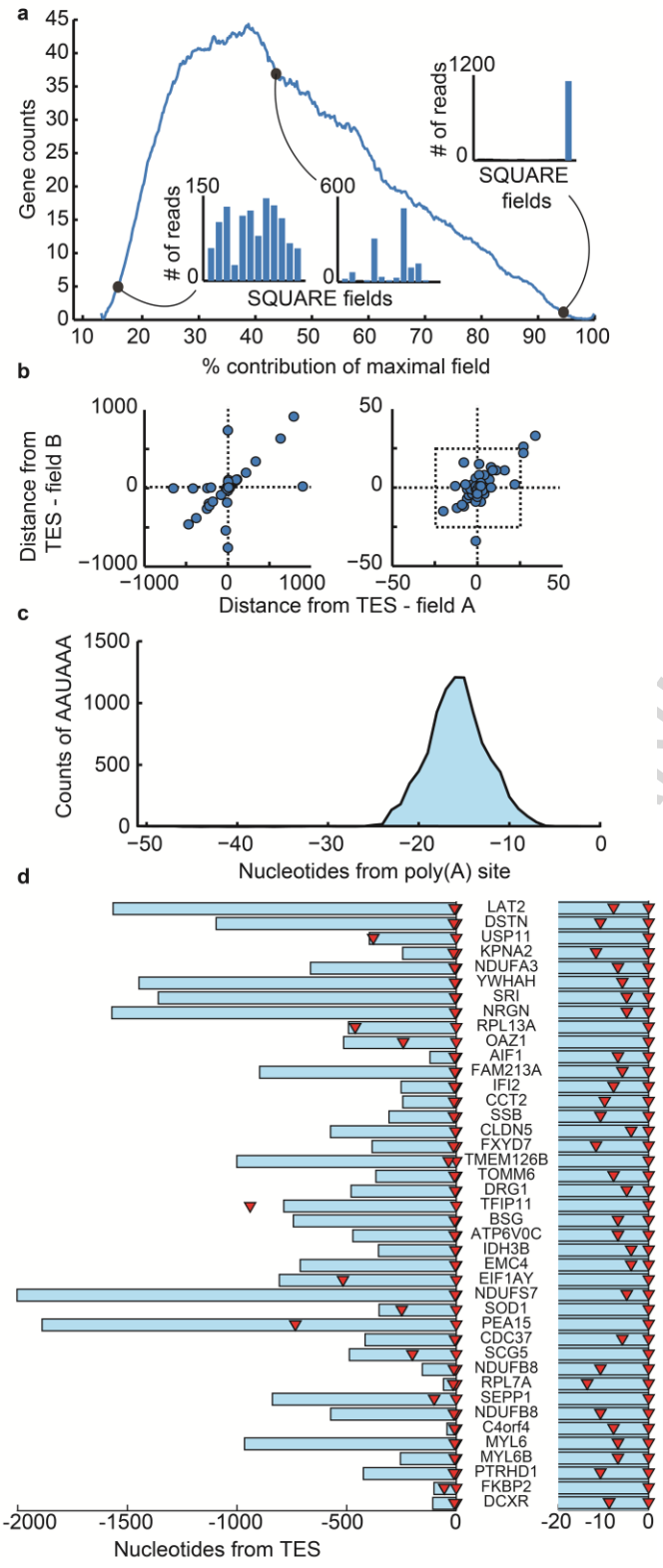


Figure 6

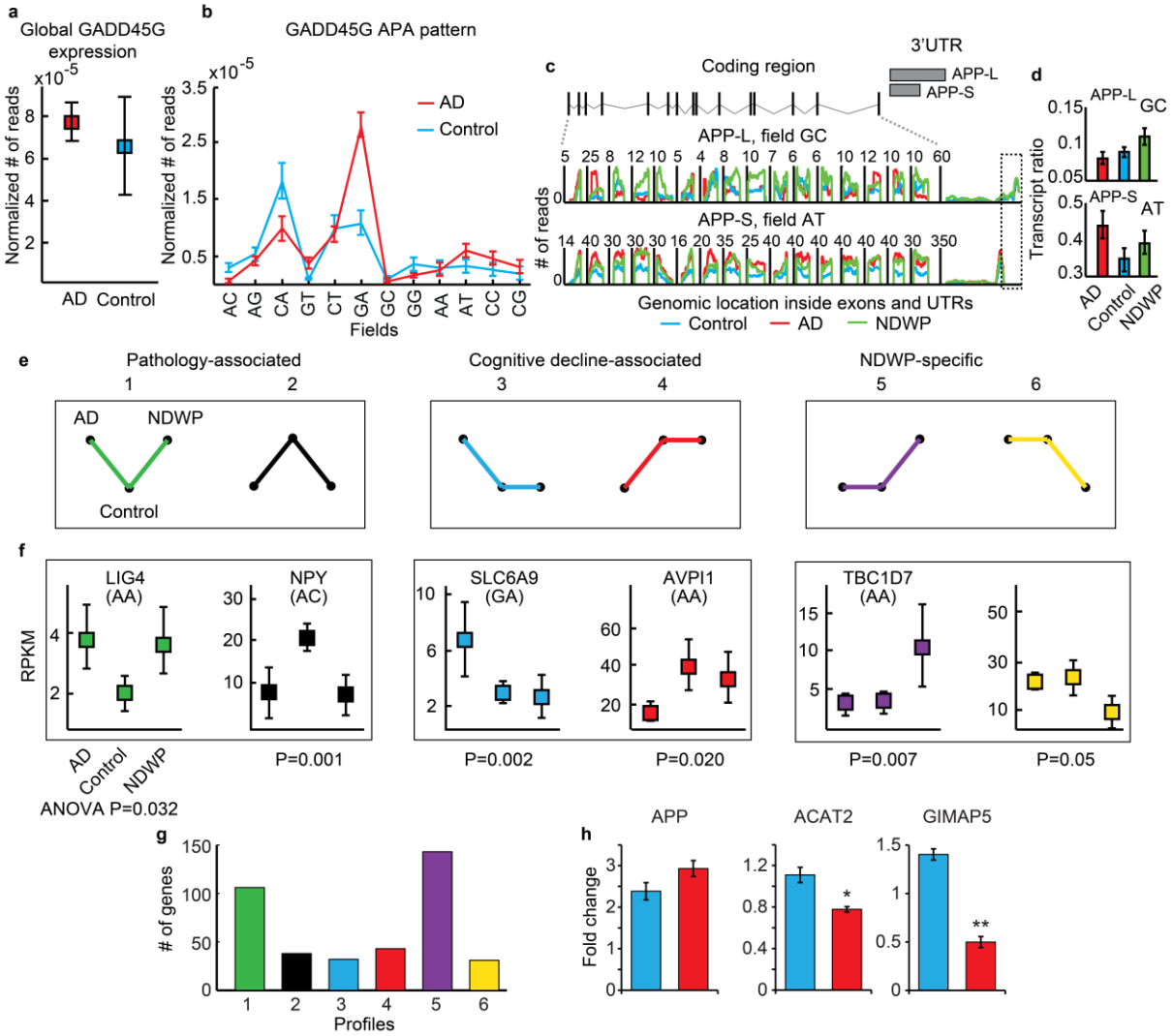
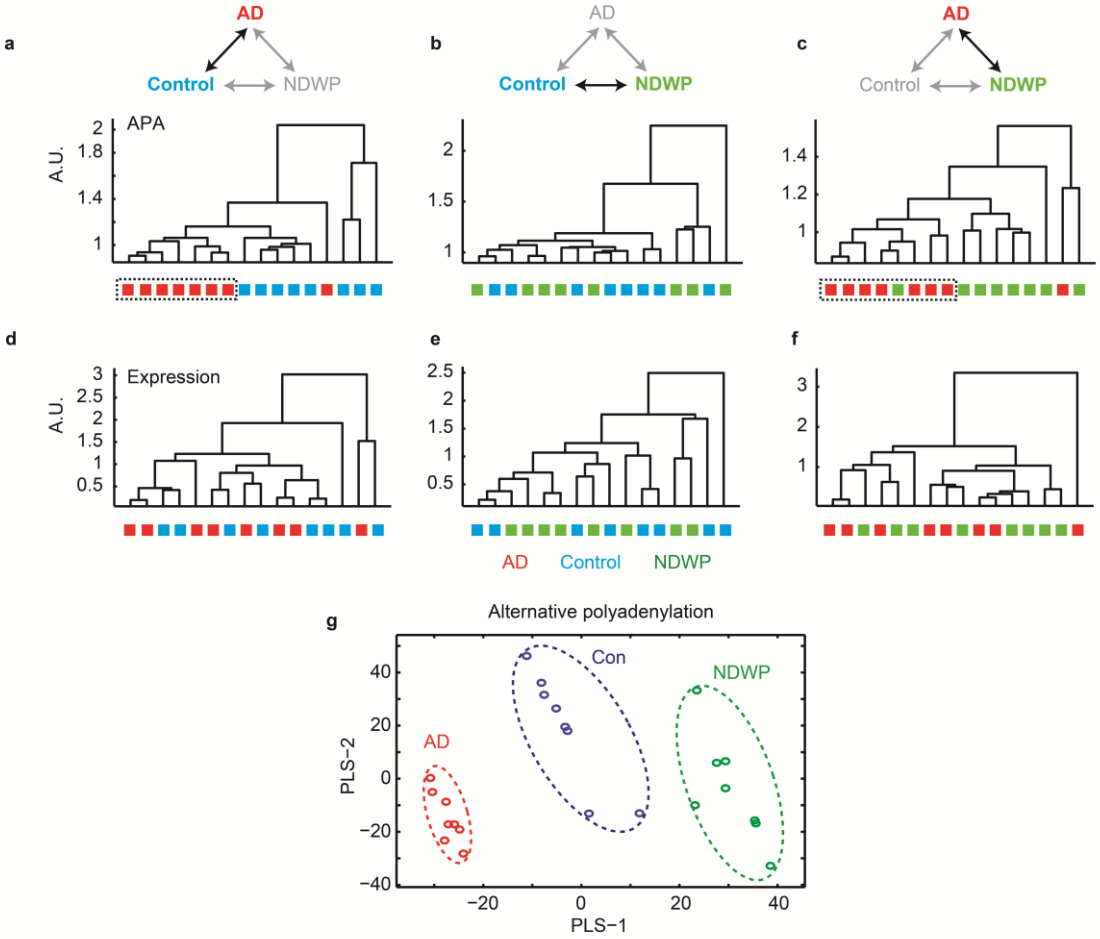


Figure 7



Highlights for: Alzheimer's brains show inter-related changes in RNA and lipid metabolism

- RNA transcripts differ between brains from Alzheimer's patients with or without dementia.
- In non-demented patients with pathology, brain neurons show facilitated RNA and lipid processing.
- Cognition-related differences included brain, vascular and autoimmune disease-related genes.
- Polyadenylation changes classify patients better than transcript expression levels.