The *APOE* gene cluster responds to air pollution factors in mice with coordinated expression of genes that differs by age in humans

Alzheimer Dementia in press Oct 17 2020

Authors: Amin Haghani[1,2], Max Thorwald[1], Todd E Morgan[1], Caleb E Finch[1,3*]

1, Leonard Davis School of Gerontology, University of Southern California, Los Angeles, CA.
2, Department of Human Genetics, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, California, USA
3, Dornsife College, University of Southern California, Los Angeles, CA.

*Corresponding author. Email: cefinch@usc.edu

Highlights

- *APOE* and its gene neighbors (*APOE* cluster) have coordinated response to air pollution components
- The human brain also shows coordinate expression of the *APOE* cluster
- Aging and Alzheimer's alter *APOE* cluster expression with brain-region-specificity
- The *APOE* cluster has a conserved pattern of gene regulation in mammals
- The identified regulatory network of the *APOE* cluster gives novel links between the environment and risk of Alzheimer's disease

Abstract

Little is known of gene-environment interactions for AD risk factors. Apolipoprotein E (*APOE*) and neighbors on chromosome 19q13.3 have variants associated with risks of AD, but with unknown mechanism. This study describes a novel link between *APOE* network, air pollution, and age-related diseases. Mice exposed to air pollution nano-sized particulate matter (nPM) had coordinate responses of *Apoe-Apoc1-Tomm40* in cerebral cortex. In human, the AD vulnerable hippocampus and amygdala had stronger age decline in *APOE* cluster expression than the AD-resistant cerebellum and hypothalamus. Using consensus WGCNA, we showed that *APOE* has a conserved co-expressed network in rodent and primate brains. *SOX1,* which has AD-associated SNPs, was among the co-expressed genes in human hippocampus. Human and mouse shared 87% of potential binding sites for transcription factors in *APOE* cluster promoter, suggesting similar inducibility and a novel link between environment, *APOE* cluster and risk of AD.

Keywords: *APOE*, chromosome 19q13, air pollution, Alzheimer, aging

Narrative

Alleles of Apolipoprotein E (*APOE*) dominate the genetics of Alzheimer disease (AD) and brain aging with 7500 citations. Yet, barely twenty studies have considered GxE interactions of *APOE* alleles with air pollution or cigarette smoke, which are global AD environmental risk factors. Two population-based studies show increased risk of *APOE4* carriers to air pollution for accelerated cognitive decline and dementia [1, 2]. Besides *APOE*, we must consider its neighboring genes, which are associated with risk of AD and other diseases with increased risk of accelerated cognitive loss (Fig. 1). We do not know how the expression of these genes may change during aging and AD, or in response to air pollution and cigarette smoke.

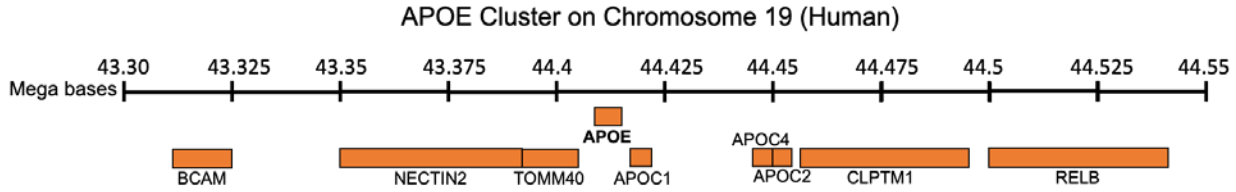


Figure 1. Human Chromosome 19q13.32, showing *APOE* and the gene neighbors we assess for expression. This locus of more than 50 genes is extensively conserved in mammals (Suppl Fig. S5). Mice have these same genes in reversed order (inverted synteny).

TOMM40 was the first *APOE* gene neighbor with AD-risk variants [3], joined by *APOC1*, *APOC2*, *APOC4*, and *NECTIN2* in complex AD haplotypes. More than 30 single nucleotide variants (SNPs) in coding and non-coding adjacent sequences are AD-associated; subsets may act in *cis* combination with *APOE* or independently [4-6]. *APOE* cluster genes encode diverse functions: lipoproteins (*APOE*, *APOC1*, *-C2, -C4*); inflammation (*RELB*, *TGFB1*); metabolism (*IGFL1*, *TOMM40*); brain development *(NTF4*); reproduction via gonadotrophins (*CGB*, *LHB*); and viral resistance (*APOE*, *NECTIN2*). Other than AD, the *APOE* cluster is associated with atherosclerosis and hyperlipidemia [7]; hypertension [8, 9]; obesity [10]; and longevity [11-13]. Because only *APOE* has been considered for GxE interaction with air pollution for risk of accelerated cognitive decline and AD dementia [1, 2], we examined other *APOE* cluster genes for response to air pollution in rodent brains. In human data, we examined APOE cluster expression in several tissues. *APOE* cluster genes are co-expressed in brain [14], as well as in liver [15]; in human hepatoma cells *APOE*, *APOC1*, and *TOMM40* mRNA were regulated by the transcription factor PPARγ [16].

Because conserved gene clusters typically have coordinated expression [17, 18], we hypothesized that the *APOE* cluster coordinate expression would extend to primates (chimpanzee, monkey) and rat, as well as mouse. Unlike humans, these species do not develop brain AD-like neurodegeneration during aging [19, 20]. Inflammatory processes associated with air pollution in AD might involve endotoxins as well as combustion products, which were experienced sequentially in human evolution [29]: pre-*Homo* was exposed to high levels of endotoxin from savannah herds, followed by domestic fire, which introduced novel toxins in smoke and charred foods that also arise during fossil fuel combustion.

Responses to inhaled air pollution components are body-wide, including arterial endothelia, plasma cytokines, myocardium, lung, liver, and brain [21-25]. While the lung receives the majority of inhaled particulate material (PM), some may enter the brain through olfactory neurons [25, 26]. The 'lung-to-brain' route is shown for systemic responses and nanoscale particles [27].

We examined brain transcriptional responses to several components of air pollution: ultrafine PM (PM0.2, <0.2 µm diameter) and bacterial endotoxins which induce inflammation and oxidative stress [21]. Urban PM0.2 are derived from fossil fuels, burning biomass, and road dust, while LPS-like endotoxins are derived from gram-negative bacteria. PM0.2 are collected for 2 months continuously on filters from an urban freeway air corridor and eluted by sonication into water. Because this subfraction excludes polyaromatic hydrocarbons [21], we designated it as *nPM* in distinction from total PM0.2 [28]. Mice were exposed to re-aerosolized nPM at 300 µg/m^3 for 15 hours per week during 8 to 15 weeks; the hourly average of 27 µg/m^3 is within the upper range of current US roadway exposures, and far below the global upper range. Mouse genotypes included both sexes of C57BL/6 wildtype ('B6') and transgenics for human *APOE3* and-E4 alleles (*APOE*-TR).

Initial findings on cerebral cortex led us to examine other data sets for *APOE* cluster gene co-regulation in cultured glia and lung. Human transcriptome data included brain region specific expression during normal aging by sex, and by clinical AD stage. Potential transcription factors were then identified with multi-species co-expression networks. Lastly, we examined aging and AD for their impact on *APOE* cluster expression and identified genetic variants in *APOE*-cluster transcription factors that may modify AD risk.

Apoe gene cluster response to air pollution in mouse cerebral cortex

Five *Apoe* cluster genes with strong AD associations [4] were analyzed for transcriptional responses to air pollution- nPM: *Apoc1*, *Apoe*, *Bcam*, *Clptm1, Nectin2*, *Tomm40* (Fig. 2A,B). Cerebral cortex of wildtype (B6) and *APOE*-TR showed co-expression of *Apoe*, *Apoc1*, and *Tomm40* over a 2-fold range for controls and nPM exposed (Fig. 2A,B). No sex differences were indicated. The *APOE*-TR differed from B6 by the inverse correlation of *Bcam-Apoc1* mRNA (Fig 2A); this may be the first example of transcriptional inversion for *APOE* transgenes. The mRNA co-expression extended to protein levels of *Apoc1*, *Apoe*, and *Nectin2* (Fig. 2C).

Because air pollution activates glia [25, 29], we examined in vitro responses to nPM with mixed glia cultures from neonatal rat that contained astrocytes and microglia; LPS was included as a model for endotoxins in urban air pollution. For nPM, *Apoc1-Apoe*, and *Apoe-Tomm40* had positive co-response (Fig. S1), matching *in vivo* responses. In contrast to mixed glia, LPS responses in adult mouse brain included the positive correlation of *Apoe-Apoc1-Clptm1* expression, which paralleled the reported nPM responses (Fig. S2B). These differences may represent cell type specificity, shown for the *Apoe* promoter [30].

The diversity of responses to air pollution components was further explored with archived data from humans and rodents. Mouse lung responded with different Apoe cluster subsets to urban total air pollution [31] and coal tar [32] (Fig. S3) than cerebral cortex. Coal tar increased *Apoc1*, *Apoe*, and *Nectin2*, while ambient urban air only

induced *Apoc1*; *Tomm40* did not respond. Antioxidant and inflammatory responses of other chromosomal genes include *Nqo1* and *Il1b*, which also responded to nPM in our prior study of Nrf2 regulated phase II gene expression in lung and brain [33]. These findings give additional insights for the heterogeneity of AD risk from *APOE4* [34]. The different patterns of co-expression of *APOE* cluster genes to the above air pollution components will be further varied by the local chemistry of air pollution which can differ widely in oxidative activity and cytotoxicity for the same PM0.2 [21, 35].

Human and chimpanzee APOE gene cluster expression

Because the *APOE* cluster shows evolutionary stability between human and mouse, we reasoned that its coordinate gene expression would extend to chimpanzee, and monkey. Human RNA sequences from two databases were analyzed for brain regions and other tissues of both sexes for normal aging and AD. Gene pairs of *APOC1-APOE*, *APOE-TOMM40*, and *BCAM*-NECTIN2 were co-expressed in multiple human tissues (Fig. 3A); human and chimpanzee brain (Fig. 3B, C); and mouse brain (Fig. 2 above). Unlike humans, these species do not develop AD-like neurodegeneration with major pathway specific degeneration in the entorhinal cortex and hippocampus [19, 20].

Age and AD

Gene pair co-expression was stable up to age 79 y for *APOC1-APOE, BCAM-NECTIN2, CLPTM1-TOMM40*, and *NECTIN2-TOMM40* in the brain (Fig 3E). However, using principal component analysis (PCA), we showed that tissues and brain regions vary widely in APOE cluster expression (Fig. 4A, B). Brain was intermediate between white blood cells (highest variance) and liver (lowest). We further developed two PCs for brain regions that represent the changes in *APOE* cluster as one unit. In brain, the PC1 (60% variance, Fig. 4B) represented a positive correlation with *APOE* (r = 0.83), *APOC1* (0.64), and a negative correlation with *TOMM40* (-0.27), *BCAM* (-0.24), and *NECTIN2* (-0.15) (Figure S5B). In contrast, the PC2 (25% variance, Fig. 4B) represented a positive correlation with *APOC1* (0.97), *APOE* (0.54), and a negative correlation with *TOMM40* (-0.58), *CLPTM1* (-0.33), and *BCAM* (-0.24) (Figure S5C). *APOC1* and *APOE* had the highest factor loading for PC1 and PC2 in this brain data. In the following sections, we described the changes in *APOE* cluster PCs rather than individual genes. The individual gene expression data is reported in Figure S4.

Five brain regions differed by age for PC1 of the *APOE* cluster (Fig. 4E). By age 60, *APOE* PC1 in cerebral cortex was below other brain regions (Table S1). The *APOE* cluster PC1 was higher than age-matched controls only for ages below 80 (Fig. 4F). Amygdala and hippocampus had strong progressive decreases of PC1, while cerebral cortex declines were modest. Although these brain samples excluded gross neuropathology, nonetheless after age 70, cognitively normal elderly frequently have modest cerebrovascular pathology and pre-clinical AD pathology (Braak stages I-II and CAA) [36-38]. APOE PC2 only showed age-mediated increase in cerebellum and no other regions (Figure S5, Table S3).

We examined genes on other chromosomes that are markers for gliosis of normal brain aging in human and rodent [29, 39, 40]. Consistent with earlier findings, *GFAP*, *IBA1*, and *TNFα* increased progressively during normal aging with brain region-specificity

in hippocampus and amygdala. This parallel with the age increase of *APOE* PC1 (Fig. S4) suggests that the *APOE* cluster PC1 is representative for age changes in multiple inflammatory genes. The *APOE* cluster includes multiple transcription factors noted above that could mediate genome-wide aging changes.

For AD, co-expression differed from age-matched normal with stronger co-expression of *BCAM-CLPTM1* and *APOC1-TOMM40* (Fig. 3D). *APOE* and *APOC1* mRNA paralleled *APOE* cluster PC1 in AD brains (Fig. S7). In general, the differences in APOE cluster between AD and non-AD brain was age-specific (Fig 5F, S7). Before age 80, AD and non-AD brains had larger difference in *APOE* PC1 (Figure 5F). At age <80, AD brains had higher mRNA levels for *APOE*, *APOC1*, and lower expression in *TOMM40* and *CLPTM1* (Fig. S7). After age 80, AD brains only had lower expression in *NECTIN2* and *CLPTM1*. These findings confirm associations of elevated *APOE* mRNA in AD brains [41, 42], but indicate a need for a more detailed analysis of brain regions by age.

Regulatory networks of the APOE cluster

The shared co-expression of *APOE* cluster genes cluster in several species suggested the possibility of shared regulatory networks. First, we analyzed gene modules in the consensus weighted gene co-expression network (WGCNA) for the brain transcriptomes of human, chimpanzee, and mouse. Next, *APOE* cluster genes were screened for transcription factor (TF) binding sites in promoters of human and mouse. Four WGCNA modules (ME1-4) were shared in human, chimpanzee, and mouse (Fig 5A,C). The top upstream regulators of modules ME1-4 included TGF-β3 (transforming growth factor-β3), CLPP (caseinolytic mitochondrial matrix peptidase proteolytic subunit), and NFKBIA (NF-κB inhibitor alpha) (Fig 5B).

APOE cluster gene expression differed by module. *CLPTM1*, *BCAM*, and *NECTIN2* were positively correlated in all modules, while correlations of *APOC1, APOE*, and *TOMM40* expression were restricted to subsets (Fig. 5C). The modules mediate diverse activities: protein homeostasis (protein ubiquitination), development (HIPPO signaling, WNT/β catenin signaling, stem cell pluripotency), DNA repair (non-homologous end-joining repair), immune system (osteoarthritis, STAT3, necroptosis), and metabolism (sirtuin signaling, TCA cycle II, purine biosynthesis).These four gene modules further document shared co-expression in human, chimpanzee, and mouse, which is consistent with the highly conserved gene synteny.

Next we searched for shared TF binding motifs in the promoters of *APOE* cluster genes in human and mouse using the TRANFAC database [43]. Promoters of *APOE* cluster genes in human and mouse shared 105 potential TF binding sites, comprising most (87%) of the identified TF motifs (Fig. 6B). The highest ranked binding sites included *KLF6, ING4*, *and Sox-related factors*, which are proximal to the transcription start sites (Fig. 6C). Notably, the *APOE* promoter lacked the CREB group and NR-DR of its gene neighbors (Fig. 5E).

TF gene candidates with identified binding motifs were screened in GTEx brain data (Fig. 6B). These genes were ranked by their correlation with *APOE* cluster brain PC1. The initial screen included all brain regions; secondary analysis focused on hippocampus and amygdala for relevance to AD. Some of the top TFs with strong correlation with *APOE* PC1 included POU3F4, SOX2, and MLX (Table 1). At a relaxed criteria of lower correlation, several TF of the *MAF* gene group (Table 1) showed inverse

correlation with *APOE* PC1: *BACH1, MAFB, MAFG, NFE2,* and *NFE2L3* (Suppl, excel file). These genes also responded to air pollution in mouse brain (*Nhlh2,* -30%; *Mafg,* +30%), which mediate oxidative stress responses through NRF2.

Contrary to expectations from prior studies of human hepatocytes [16] noted earlier, we did not find a role for PPARγ or RXRA binding sites in the *APOE* cluster of cerebral cortex. Some of the *APOE* cluster upstream regulators (*TGFB1*, *NFKBIA,* Figure 5B) have *PPARγ* and *RXRA* binding sites (identified by TRANSFAC), which suggest an indirect effect of either TF. Moreover, there may be binding sites further down-stream or upstream of defined promoter regions. A direct comparison of mouse and human *APOE* promoters showed limited shared sequence beyond -180 nucleotide upstream [30]. Several TF binding sites in human were absent in mouse for PPARγ and a cluster of AP1/SP1/AP2.

Candidate regulatory genes were screened for AD-associated variants, using the GWAS of the International Genomics of Alzheimer's project (IGAP). At a stringent criteria for genome wide significance ($p < 5 \times 10^{-8}$), there were no SNPs for our genes of interest. Relaxed significance ($p<0.003$) showed 10 SNPs for AD risk in *ELK4*, and 2 SNPs in *SOX1*. *SOX1* correlated positively with *APOE* cluster PC1 in hippocampus. For additional gene candidates, see Supplement excel file.

Promoter evolution

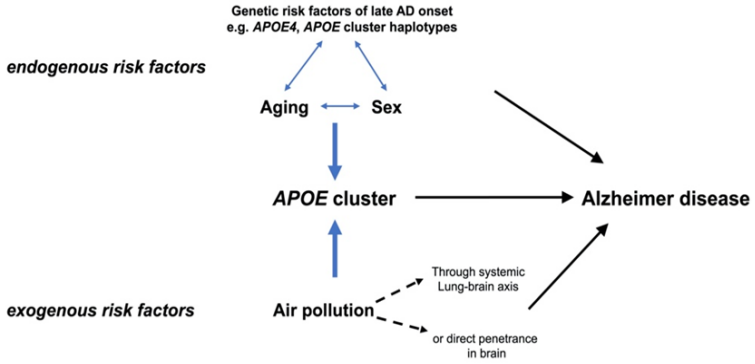
Because human and mouse shared 89% of TF binding motifs (Fig 5A), we compared the promoters of *APOE*, *APOC1*, *BCAM*, *CLPTM1*, and *NECTIN2* of human with neanderthal, chimpanzee, macaque; mouse and rat (rodents); dog and pig (omnivore) (Fig.6). The human and primate promoters were clearly separated from promoters of mouse-rat, and dog-pig, as expected. The human and mouse *APOE* promoters were more conserved than the other four *APOE* cluster genes (Figure 6). In contrast, the *TOMM40* promoter was the least conserved between humans and rodents. The human and mouse *APOE* promoters were more conserved than the other four *APOE* cluster genes examined: human-mouse phylogenetic differences were ranked in ascending order of distance *APOE*, 1.1 (26.6% sequence similarity); *APOC1*, 1.4 (22.5%); *NECTIN2*, 1.4 (17.7%); *CLPTM1*, 1.6 (19.7%); *BCAM*, 1.8 (15.6%); *TOMM40*, 2.1 (19%).

The shared co-regulation of these six genes may have been a factor in the synteny that persisted in multiple lineages that diverged at least 100 million years ago, as observed for other co-regulated gene clusters [18]. Because topologically associated domains (TAD) of chromatin are conserved particularly for syntenic regions [44], the *APOE* gene cluster is likely to be within a TAD.

Inflammatory processes associated with air pollution might involve endotoxins as well as combustion products that were experienced sequentially in human evolution: pre-*Homo* was exposed to high levels of endotoxin from savannah animal herds, followed by domestic fire, which introduced novel toxins in smoke and charred foods that also arise during fossil fuel combustion [45]. The major histocompatibility complex (MHC) of mammals is another ancient ensemble of genes that mediate metabolism and reproduction, as well as in adaptive immunity [46]. The *APOE* and *MHC* clusters exemplify 'life history gene complexes' that mediate reproductive success through pleiotropic networks of metabolism, host defense, and reproduction [46, 47]. The *APOE* cluster includes two gonadotrophins (noted above, but not shown on Fig.1).

The body-wide impact of the *APOE* gene cluster is enabled by extensive pleiotropies with systemic interactions of lipids (*APOE*, *APOC1*), energy (*CYP2A*, *BCL3*, *OPA3*, *TOMM40*); inflammation (*C5a*, *IGFL1*, *IRF2B1*, *TGFβ1*); immunity (*CLPTM1, BCAM, NECTIN2*); viral binding (*APOE*, *NECTIN2*); blood brain barrier (*NECTIN2*); gonadotrophins (*CGB*, *LHB);* transcription factors (*BCL3*, *RELB*, *ZNF*). The *APOE* cluster may also modulate AD risk via viral host defense (Table S2). *APOE4* carriers show higher risk of infections for hepatitis B [48], herpes simplex virus 1 [49], and COVID19 [50]. These complex exogenous and endogenous interactions are outlined in Figure 7.

A.



B.

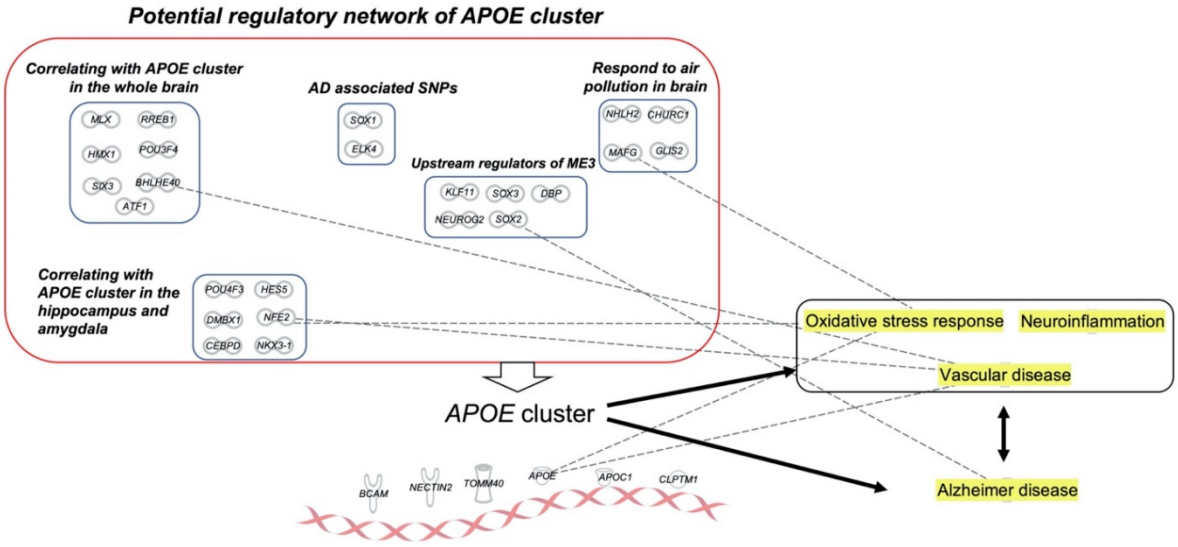


Figure 7. The *APOE* cluster is a potential link between endogenous and exogenous AD risk factors. A) Schema of the complex interplay of AD risk factors through APOE cluster. B) Transcriptional network of *APOE* cluster (Table 1), and the relationship with vascular disease, AD, neuroinflammation, and oxidative stress response. Dashed lines are links based on the IPA database.

Besides haplotypes of variants associated with dementia, the body mass index (BMI) differs by haplotypes of *APOE* and *TOMM40* [51]. The *APOE* cluster presents a regulatory nexus for infections, as well as non-infectious diseases. Both air pollution and *APOE4* may increase vulnerability to COVID-9 infection [50, 52]. We anticipate expanding roles of the *APOE* gene cluster in global environmental hazards.

Vascular disease factors in the *APOE* gene cluster may contribute to the associations of air pollution and AD. *APOE4* is a risk factor for ischemic heart disease, stroke, weak blood-brain barrier, microinfarcts (microbleeds), and hypertension [11, 53-56]. Air pollution is also a leading preventable risk factor of ischemic heart disease and stroke [57, 58]. Other *APOE* cluster genes associated with AD have vascular associations (Table S2). *APOC1* modulates the AD risk factor of low HDL [59] by inhibiting the cholesterol ester transfer protein (CETP). Both NFKB and NRF2 pathways are involved in atherosclerosis [60, 61]. BHLHE40 (transcription factor, alternate name DEC1) also influences blood pressure [62]. Correlation of BHLE40 gene with the *APOE* cluster PC1 (Table 1, Fig. 7) suggests links between circadian rhythm, blood pressure, vascular disease, and AD progression.

A major unknown is how age changes in *APOE* cluster expression may influence its AD-associations [4, 5]. We need a new set of transgenic mice carrying AD-associated variants of the human *APOE* cluster for expanded GxE analysis and for testing of drug interventions. Transgene responses to exogenous and endogenous factors in the AD exposome [63] should be a priority in further development of *APOE* mice. We need a more comprehensive systems approach that combines multiple targets with the individual genotype and lifestyle of the patient [63, 64]. Cigarette smoking was recognized as an AD risk factor before air pollution [65], but is rarely considered in subject selection for drug trials. The identified transcriptional factor candidates could be combined with diagnostic or therapeutic ongoing studies (Fig. 7B).

In conclusion, we showed that the rodent *Apoe* gene cluster responds to external stimuli with coordinated gene expression changes that also differ for the human *APOE* cluster by age, sex, and stages of AD. The rodent gene responses to air pollution nPM in brain and lung may be useful biomarkers for human responses to air pollution. The transcription factor regulators of this network suggest links between changes in *APOE* gene cluster expression for cognitive aging and AD, as well as other age-related conditions. Future studies should consider screening for *APOE*-cluster coordinated changes in relationship to AD stages. We cannot remain content with mouse models of single AD genes without considering the GxE.

*

Methods

Data

This study examined RNAseq datasets from mouse, human, chimpanzee, and primary mixed glial culture to screen for the changes in 5 APOE gene-neighbors. Mouse datasets included cerebral cortex of young adults of both sexes from three mice genotypes: C57BL/6J (B6) and transgenic for human APOE3 or APOE4, by targeted replacement (APOE-TR). Mice were exposed to 300 μg/m3 nanosized particulate matter (nPM, 5 h/day, 3 d/wk) or filtered air for 8 (B6) or 15 wk (APOE-TR). Study design, collection of air pollution, and chemical composition are described in [21] and Figure S10. The nPM subfraction of ultrafine PM0.2 is depleted in polycyclic aromatic hydrocarbons. RNA datasets were produced from mRNA libraries using TRUseq Stranded mRNA Kit (Illumina), and single end-sequencing (> 50 nt) using Illumina NextSeq500. Preprocessing used Partek flow software platform [66]. Rhe reads were aligned and quantified using mouse reference genome (mm10) with Tophat2 (v2.0.8b). Counts per million (CPM) were normalized using trimmed mean of M values (TMM) [67]. Data are

accessible in NCBI GEO (GSE142066). Other used datasets include: Mouse lungs exposed to coal tar (GSE87690); Mouse lungs exposed to ambient air pollution (GSE41698); Rat fetal brain with maternal LPS challenge (GSE34058); and Adult mouse brain after LPS challenge (GSE3253).

Human data was accessed from Genotype-Tissue Expression (GTEx) database representing multiple tissues and brain regions of 651 men and women, aged 20-80 yr [68, 69]. The data was available as transcripts per kilobase million (TPM). Expression changes were further validated in a human brain microarray (GSE48350) [70].

Brain RNA data for chimpanzee (Pan troglodytes) included 11 brain regions from 2 female and 1 male, age 15-34 yr (GSE7540) [71]. Data were generated by human oligonucleotide arrays (GENECHIP Human Genome U95Av2 arrays (Affymetrix, Santa Clara, CA), representing 10,000 genes and 12,625 probes. Data from a genome wide association meta-analysis by the International Genomics of Alzheimer's project (IGAP; 17,008 AD cases, 37,154 age-matched controls) was used to screen for potential AD single nucleotide mutations in the identified genes [72]. Mixed glial cDNA data was generated by Affymetrix Rat Whole Genome 230.2 array in our prior study [29].

Protein analysis APOE, APOC1, and NECTIN2 protein levels were analyzed by mouse specific ELISA (LSBio: LS-F33290-1, LS-F5921-1, LS-F4714-1).

Consensus Weighted Gene Co-expression Analysis (WGCNA) The consensus co-expression network was formed from the four brain expression datasets: GTEx (limited to human brain), chimpanzee and mouse (C57BL/6J, APOE-TR). We examined 7061 genes shared by all four datasets. Consensus co-expression networks were identified following methods previously described [73]. Briefly, the adjacency matrices (correlation) were constructed using log2 expression in each data set. The matrices were converted to scale free networks using the soft threshold power of six. Results were converted to topological overlap matrices (TOM); these were merged to form a consensus tree network using a hierarchical clustering of dissimilarity matrix (1-TOM). Modules of $\geq$ 30 genes were formed using a dynamic tree-cut algorithm. Singular value decomposition method was used to calculate the maximum amount of variance per module. Modules containing APOE gene clusters were identified by Ingenuity Pathway Analysis (IPA, Qiagen). Hub genes of modules were selected for eigengene connectivity (kME) of genes in each module.

Transcriptional factor binding sites Transcription factor (TF) binding motifs in APOE gene-neighbors were predicted from TRANSFAC database [74, 75]. We screened for TF binding motifs with http://www.genexplain.com in APOE cluster genes of human and mouse, using default parameters of genexplain defined promoter boundaries: -10,000, +1000 nucleotides of transcription start sites (TSS). Genes from TF groups were manually extracted from genexplain for expression screening.

Statistical analysis The Pearson correlation analysis, principle component analysis, and data management used Rstudio. Multiple sequence alignment and phylogenic analysis of the promoter sequences used MAFFT online tool [76]. Sequences (-5,000, +1000 nucleotides of the TSS) were from ENSEMBL reference genomes [77].

Results

Apoe gene cluster response to air pollution

Five genes of the mouse *Apoe* cluster shared transcriptional responses of cerebral cortex to air pollution nanosized particulate matter (nPM): *Bcam*, *Nectin2*, *Tomm40*, *Apoe*, *Apoc1*, and *Clptm1*. The strongest correlations were found for *Apoe*, *Apoc1*, and *Tomm40* (Fig. 2A). Positive correlations of *Apoe-Apoc1* (r = 0.57, p<0.01), and *Apoe-Tomm40* (r = 0.72, p<0.0001) were generally consistent among mouse strains (Fig 2A-B). However, wildtype B6 differed from *APOE*-TR with inversely correlated expression of *Bcam-Apoc1*. While *APOE* mRNA showed a higher baseline in female APOE-TR mice, there was no sex difference in nPM response. Proteins of B6 male mice had corresponding responses to nPM for *Apoe*, *Apoc1*, and *Nectin2* in. APOE protein was lowered (25%) by exposure to three nPM levels (Fig. 2C). Levels of APOE, APOC1, and NECTIN2 were positively correlated (r = 0.36-0.46).

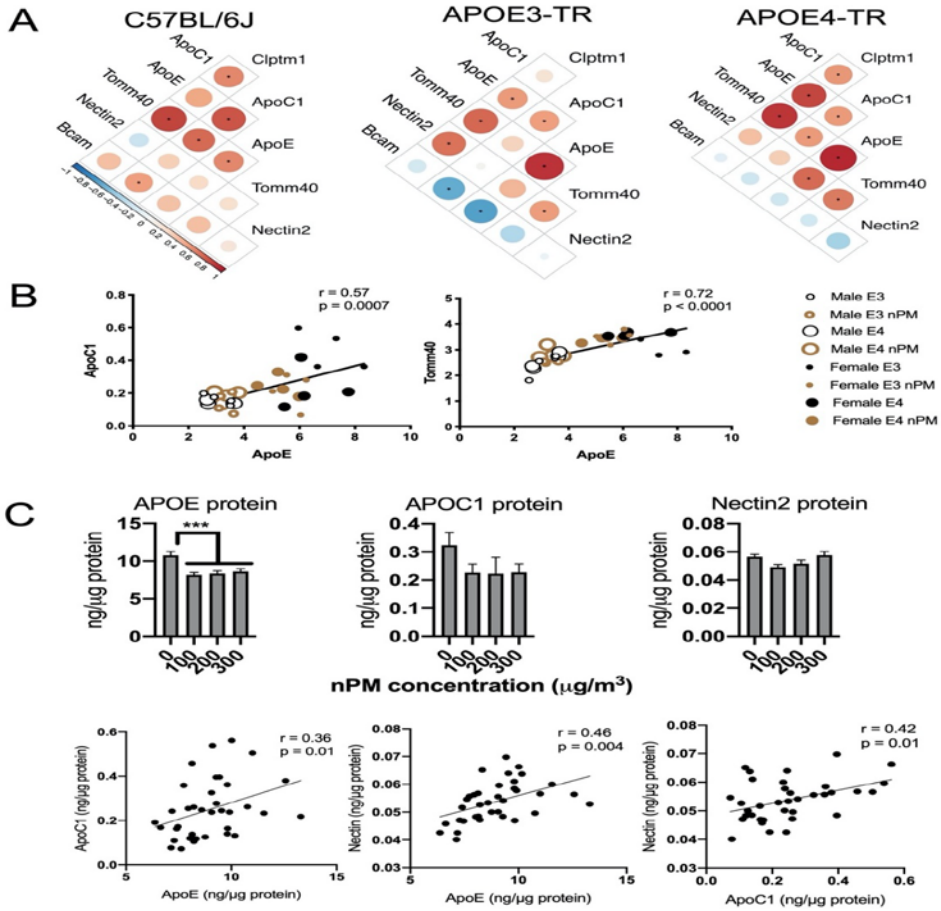


Figure 2. Transcriptional response of the *Apoe* gene cluster to air pollution nPM in cerebral cortex of mice. A) Pearson correlation heatmap of nPM mediated gene expression changes in *Apoe* cluster of C57BL/6J (B6), *APOE3*-TR, and *APOE4*-TR mice (both sexes). Mice were exposed to 300 µg/m^3 nPM at 8 wk (B6) or 15 wk (*APOE*-TR). N=16 per genotype (4/sex /treatment). * p < 0.05, Bonferroni multiple test correction. B) Scatter plots of *Apoe-Apoc1* and *Apoe-Tomm40* in *APOE3* and *APOE4*-TR mice (both sexes) exposed to nPM or filtered air for 15 weeks; linear regression modeling of. Data source: GSE142066. C) Protein responses of male B6 mice in nPM (100-300 µg/m^3 nPM for 3 weeks; 5 h/d, 3 d/wk). *** p < 0.001 in ANOVA test after FDR multiple test correction. N = 9/group.

Because air pollution activates astrocytes and microglia [25, 29], we examined response of mixed glial cultures (astrocyte: microglia, 3:1) to air pollution nPM, or to LPS as a model for endotoxins in urban air pollution (Fig. S1A). Glial mRNA for *Apoe-Apoc1* and *Apoe-Tomm40* were positively correlated (r = 0.78) (Fig. S1A), as observed for in vivo exposure. There was a strong correlation for glial *Tomm40-Nectin2* response to LPS (r = 0.76) similar to in vivo. However, there was no in vitro response to nPM, in contrast to vivo. For *Apoe-Tomm40*, response to LPS, the expression was inversely correlated (r=-0.7), again in contrast to nPM responses in mouse brain and in vitro. In contrast to mixed glia, LPS responses in adult mouse brain included the positive correlation of *Apoe-Apoc1-Clptm1* expression, which paralleled the reported nPM responses (Fig. S2B). These differences may represent cell type specificity, shown for the *Apoe* promoter [30].

Lung was also examined for *APOE* cluster responses to air pollution components because responses mediated by NFKB and NRF2 are systemic [33]. Since RNAseq was not available for mouse lung we accessed cDNA datasets for two other air pollutant exposures: ambient urban air by inhalation [31] or coal tar by gavage (Labib et al 2017). Each treatment induced different subsets of the *Apoe* cluster (Fig. S1). Coal tar increased *Apoc1*, *Apoe*, and *Nectin2*, while ambient urban air only induced *Apoc1*; *Tomm40* did not respond. Antioxidant and inflammatory responses of genes located elsewhere include *Nqo1* and *Il1b*, which also responded to nPM in our study of Nrf2 regulated phase II gene expression in lung and brain [33].

Human and chimpanzee APOE gene cluster expression

Findings on rodent brain and glia were extended to human and chimpanzee RNA from three public databases: the human Genotype-Tissue Expression (GTEx) and GSE48350 databases for brain regions and other tissues (both sexes, ages 20-80 y, normal aging, and AD); and adult chimpanzee brain (GSE7540).

Humans had strong co-expression relationships in brain and other tissues for *APOE-APOC1*, *APOE-TOMM40*, and *BCAM-NECTIN2* (Fig. 3 A, B). Other correlations included *TOMM40-CLPTM1* (r = 0.8), *BCAM-NECTIN2* (r = 0.7), and *APOE-APOC1* (r = 0.65) (Fig. 3E). These relationships were consistent across adult ages 20 to 80. Chimpanzee brain also showed positive correlations among *BCAM-CLPTM1*, *BCAM-NECTIN2, CLPTM1-NECTIN2* (Fig. 3C). The co-expression of *APOE-APOC1* and *TOMM40-CLPTM1* is shared, for human, chimpanzee, and rodent; (Fig. 2, 3). The apparent lack of species differences in the brain expression of *APOE* cluster is not conclusive because of the limited samples for chimpanzee.

Human brain age changes were corroborated with an independent cDNA dataset of brains with carefully defined AD neuropathology (GSE48350)[70]. For normal brain aging before age 80 with minimal AD changes, both datasets showed consistent co-expression of *APOE-APOC1, TOMM40-CLPTM1, TOMM40-BCAM-NECTIN2* (Fig. 3D). AD may alter expression of *APOE* cluster, suggested in the lack of correlation in *TOMM40-BCAM* and *BCAM-NECTIN*, caveat the small sample.

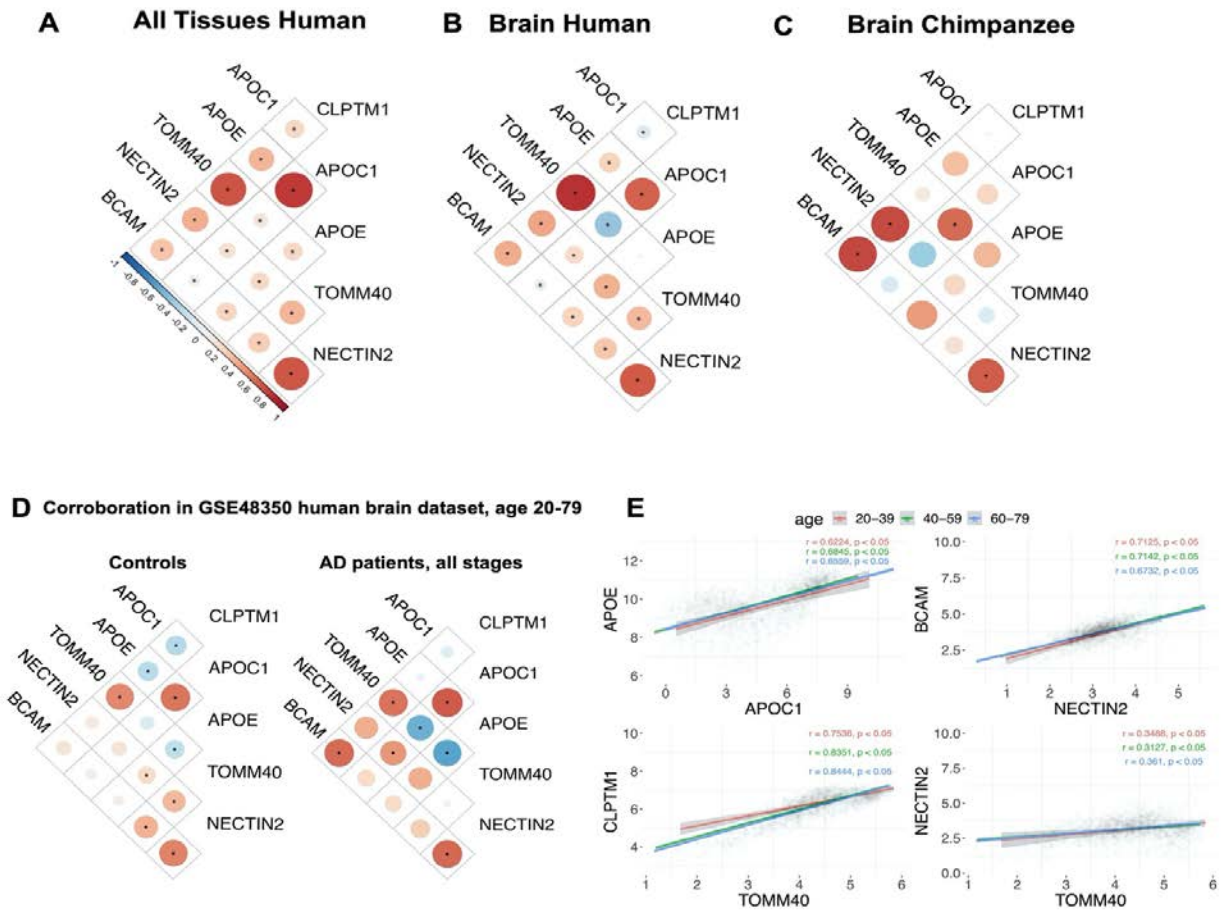


Figure 3. Coordinated expression of the human *APOE* cluster in human and chimpanzee. A) Human, all tissues, age 20-80 y: Pearson correlation heatmaps of *APOE* cluster expression (12,283 samples from 651 individuals); data from GTEx. B) Human brains, age 20-80 y: heatmaps of *APOE* cluster expression (2,112 brain regions, 321 individuals). Data from GTEx. C) Chimpanzee brain: Pearson correlation of *APOE* cluster expression (11 brain regions; 2 female, 1 male, 15-34 y). Data from GSE7540. * $p < 0.05$, after Bonferroni multiple test correction. D) Corroborating data for human brain, normal and AD, age 20-79 [70]: heatmaps of *APOE* cluster expression, control (119 brain regions, 35 individuals); AD brains (18 brain regions, 7 AD). E) Scatter plot and regression analysis of *APOE-APOC1*, *NECTIN2-BCAM*, *CLPTM1-TOMM40*, and *NECTIN2-TOMM40* in GTEx for all brain regions and ages 20-39, 40-59, 60-79 y.

APOE cluster gene expression varied widely by tissues in principal component analysis (PCA). At the extremes, white blood cells had the highest variance in *APOE* cluster expression (Fig. 4A), versus liver with the lowest variance. In brain, the PC1 (60% variance, Fig. 4B) represented a positive correlation with *APOE* (r = 0.83), *APOC1* (0.64), and a negative correlation with *TOMM40* (-0.27), *BCAM* (-0.24), and *NECTIN2* (-0.15) (Figure S5B). In contrast, the PC2 (25% variance, Fig. 4B) represented a positive correlation with *APOC1* (0.97), *APOE* (0.54), and a negative correlation with *TOMM40* (-0.58), *CLPTM1* (-0.33), and *BCAM* (-0.24) (Figure S5C). *APOC1* and *APOE* had the highest factor loading for PC1 and PC2 in this brain data. In the following sections, we described the changes in *APOE* cluster PCs rather than individual genes. The individual gene expression data is reported in Figure S4.

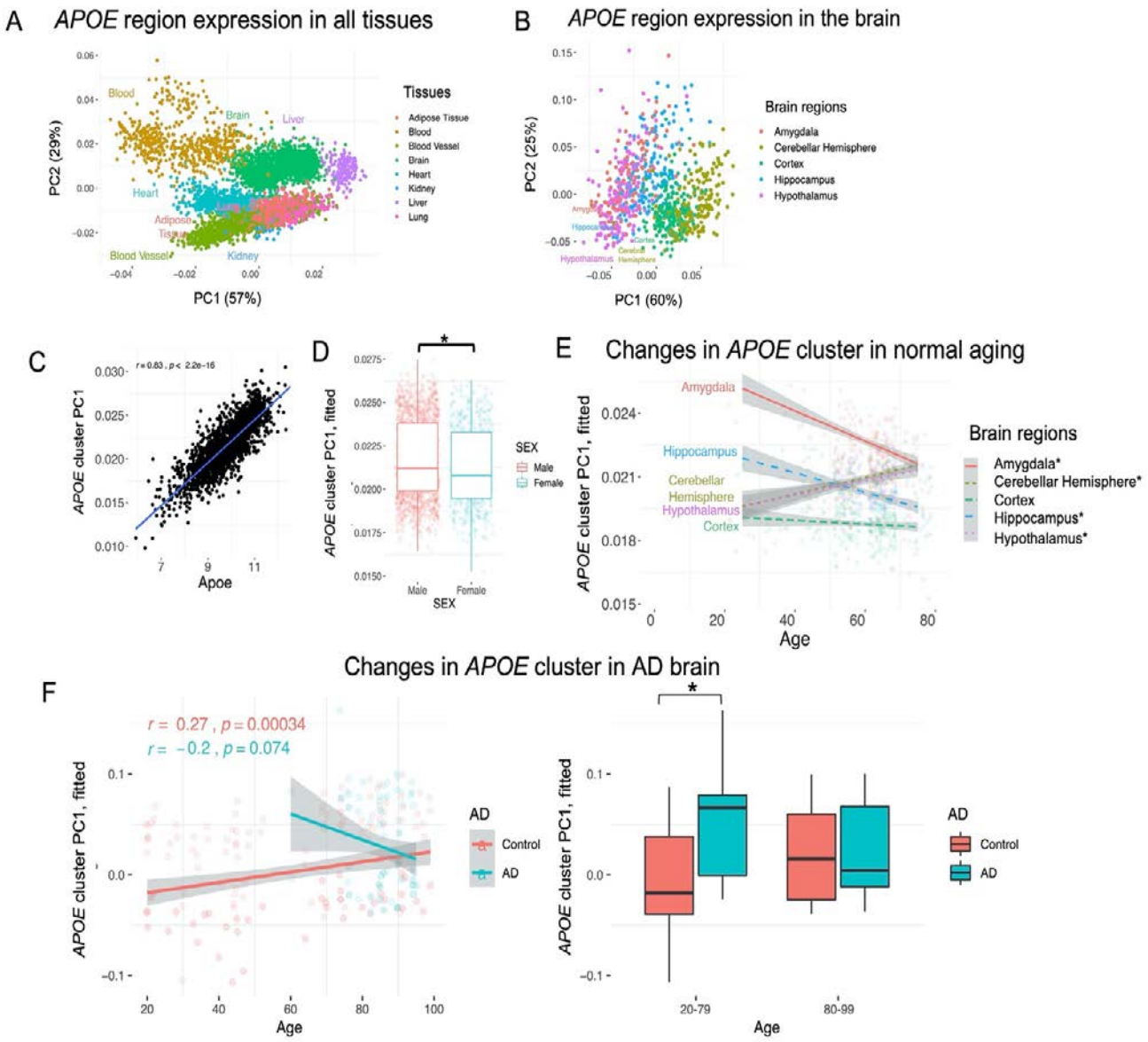


Figure 4. The human *APOE* gene cluster has different expression by tissue, brain region, and age. A) Principal components of *APOE* cluster expression in multiple tissues (12,283 samples from 651 individuals). Blood, heart, and blood vessels had higher variance in *APOE* region expression than other tissues. Blood vessel: combined data from aorta, coronary, and tibial arteries. B) Principal components of *APOE* region expression for brain regions (2,112 brain regions, 321 individuals). C) Positive correlation of *APOE* expression and *APOE* cluster PC1 in brain.D) Association of sex and *APOE* cluster expression in brain. E) Age was associated with *APOE* cluster gene expression changes in five brain regions, tested by a mixed effects model; fixed effects: sex, age, brain regions, and interaction of age by brain region; random effect: subjects; outcome: *APOE* cluster PC1. *, significant interaction of age by brain region. Dataset for panels A-E: GTEx. F) Age-dependent changes in *APOE* cluster PC1 in normal aged and AD brains [70]. This association was tested by a mixed-effects model with fixed effects for sex, age group, and brain region (entorhinal cortex, superior frontal gyrus, postcentral gyrus, hippocampus), AD and interaction of AD, and age, and random effect, subjects. Outcome: *APOE* cluster PC1. *, significant interaction of age by AD. AD dataset: GSE48350. N = 57 controls (age <80 y, 35; >80, 22) and 28 AD brains (age <80, N= 7; >80, N= 21).

Age and sex were examined for brain region differences in expression of the *APOE* cluster PC1. A mixed effect model was used to adjust for subject random effects; age was centered at age 60 to capture the baseline differences at this age. The GTEx brain samples were curated to exclude extensive pathology. At age 60, cerebral cortex had a lower baseline *APOE* PC1 than other brain regions (β: ranged 6-fold, from 0.005 in nucleus accumbens to 0.0008 in spinal cord; Table S1). Five brain regions differed by age for expression of the *APOE* cluster PC1. At older ages, cerebellar hemisphere and hypothalamus had higher *APOE* PC1. Contrarily, amygdala and hippocampal expression decreased with age. Cerebral cortex age changes were modest. Sex had a minor fixed effect on *APOE* cluster PC1 (3% sex difference, $p<0.05$) (Fig 4D; Table S1, full analysis).

We examined other genes outside of the APOE complex that are associated with the gliosis of normal brain aging in human and rodent [29, 39, 40]. Consistent with these earlier findings, *GFAP*, *IBA1*, and *TNFα* showed brain region-specific age-dependent increases, particularly in the hippocampus and amygdala, paralleling the age increase of *APOE* PC1 (Fig. S4). The *APOE* cluster PC1 is an indicator of change for all genes in the cluster. Individual

mRNA changes with age differed by brain region: *APOE*, decreased in hippocampus; *APOC1*, decreased in hippocampus and amygdala, increased in cerebellar hemisphere; *CLPTM1*, decreased in amygdala, increased in hypothalamus; *BCAM*, decreased in amygdala and hypothalamus; NECTIN2, decreased in amygdala, cerebellar hemisphere, hypothalamus; TOMM40, no strong age change (Fig. S4).

Because the GTEx dataset lacked ages >80 y, we examined older decades in the cDNA human brain dataset analyzed above [70] (GSE48350). Surprisingly, this oldest age group had no correlation of *APOE*-*CLPTM1*, *CLPTM1*-*TOMM40*, and *NECTIN2*-*TOMM40*. The increase of *APOE* PC1 in hippocampus (Fig. S6) warrants further study with larger samples. AD brains also differed by age below 80 yr. The *APOE* cluster PC1 was higher than age-matched controls only at ages <80 (Fig. 4F). Ages >80 y, had similarly higher expression of *APOE* in AD and controls than for young ages, suggesting an age ceiling for relationship of AD in the *APOE* cluster. *APOE* and *APOC1* mRNA paralleled *APOE* cluster PC1 in the AD brains (Fig. S6). In contrast, some mRNA changes differed by the age of the AD patients. For AD >80 years, NECTN mRNA was higher than normal controls, while *TOMM40* mRNA was lower for younger AD <80 years old. *CLPTM1* mRNA had a lower level in AD brains in both age groups. *BCAM* did not differ from AD in old age.

Regulatory networks of the APOE cluster in relation to Alzheimer disease (AD)

Based on the above evidence for species-shared transcriptional relationships in the *APOE* gene cluster, we examined potential regulatory networks. First, we analyzed gene modules in the consensus weighted gene co-expression network for the brain transcriptome of human and mouse (both sexes of wildtype B6 and *APOE*-TR). Chimpanzee was included as a close human relative that, like wildtype rodents, does not incur AD brain-region specific neurodegeneration in old age [19, 20]. Second, we screened *APOE* cluster genes for transcriptional binding sites in promoters of human and mouse.

The consensus WGCNA identified four modules in association with *APOE* cluster genes in human, chimpanzee, and mouse. The top 50 hub genes of these modules by IPA were enriched for pathways of protein ubiquitination, HIPPO signaling, osteoarthritis signaling, and STAT3 (Fig. 5 A,B). The top upstream regulators of modules ME1-4 included TGF-β3 (transforming growth factor-β3), CLPP (caseinolytic mitochondrial matrix peptidase proteolytic subunit), OGA (O-GlcNAcase), and NFKBIA (NF-κB inhibitor alpha). *APOE* cluster expression showed differential correlations with these four modules. *CLPTM1*, *BCAM*, and *NECTIN2* were positively correlated with all modules, whereas *APOE*, *APOC1*, and *TOMM40* were selectively correlated with different subsets of modules: *APOC1* had strong positive correlation with ME2, but negative correlation with ME1,3,4 (Fig. 5B); *APOE* expression was positively correlated with ME2; *TOMM40* was positively correlated with ME1, 3, and 4, but not ME2. Overall, the four gene modules show conserved co-expression in human, chimpanzee, and mouse.

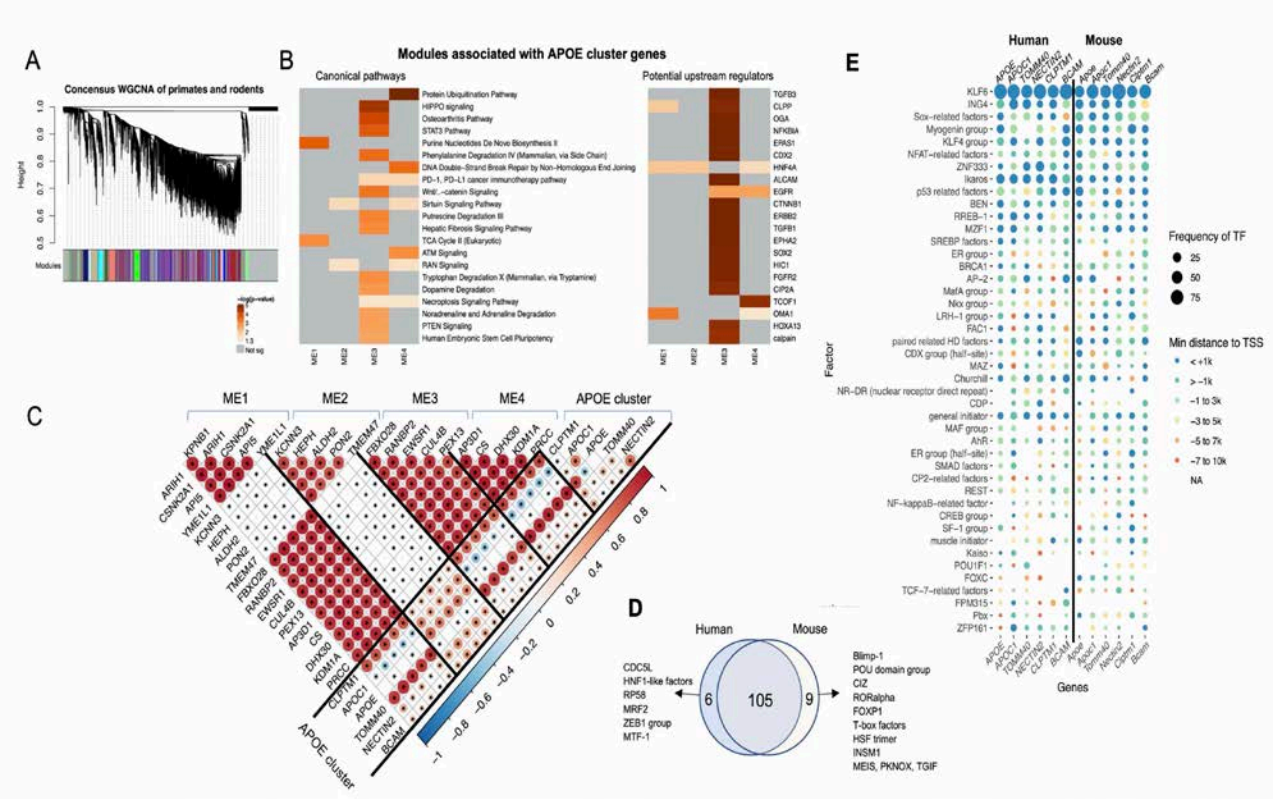


Figure 5. *APOE* cluster genes shows conserved co-expression with four gene-modules in human, chimpanzee, and mouse. Data sources: human, GTEx; chimpanzee, GSE7540; mouse, GSE142066. For hub genes, see Supplementary excel file. A) Consensus WGCNA identified four modules associated *APOE*-cluster in brains of human, chimpanzee, and mouse (B6, *APOE*3 and *APOE*4-TR; both sexes). B) The heatmaps show the top canonical pathways and potential upstream regulators of top 50 hub genes of modules ME1-4, based on IPA analysis. P-values below 10^{-5} were converted to 10^{-5} for better visualization. C) Pearson correlation heatmap of top 5 hub genes of each model and *APOE* cluster genes in GTEx brain data. * $p < 0.05$ after multiple test correction. D) Venn diagram showing transcriptional factor binding motifs of *APOE* cluster genes of human and mouse. E) Heatmap presenting the number of TF binding motifs in the promoter of *APOE* cluster genes in human and mouse based on TRANFAC database for -10,000 to +1000 nucleotides from transcriptional start site. Potential TF binding sites were examined or each gene in the geneXplain and Ensembl databases.

Next we searched for shared transcription factor (TF) binding motifs in the promoters of *APOE* cluster genes in human and mouse. Our target six genes are located on a 0.15 million nucleotide genome region, which has a conserved order but with inverted synteny [78] between human and mouse (Fig. 6). The *APOE* cluster of select other mammals is shown in Supplement (Fig. S8). Potential promoters in proximal and distal regions were examined in the TRANFAC database for -10,000 upstream to +1000 downstream nucleotides from transcriptional starting site (TSS) of each of these six genes [43]. Human and mouse shared 105 potential TF binding sites that comprise 87% of those identified in the *APOE* cluster (Fig. 5D). The highest binding sites included KLF6, ING4, and Sox-related factors, which are proximal to the TSS of most *APOE* cluster genes of mouse and human (Fig. 5E). Genes differed in the extent of shared motifs: for example, the *APOE* promoter did not share the CREB group, and NR-DR in gene neighbors. For details, see Supplementary excel file.

Because human and mouse shared 89% of TF binding motifs (Fig 5A), we compared the promoters of *APOE*, *APOC1*, *BCAM*, *CLPTM1*, and *NECTIN2* of human with neanderthal, chimpanzee, macaque; mouse and rat (rodents); dog and pig (omnivore) (Fig.6). The human and primate promoters were clearly separated from promoters of mouse-rat, and dog-pig, as expected. The human and mouse *APOE* promoters were more conserved than the other four *APOE* cluster genes (Figure 6). In contrast, the *TOMM40* promoter was the least conserved between humans and rodents. The human and mouse *APOE* promoters were more conserved than the other four *APOE* cluster genes examined: human-mouse phylogenetic differences are ranked in ascending order of distance *APOE*, 1.1 (26.6% sequence similarity); *APOC1*, 1.4 (22.5%); *NECTIN2*, 1.4 (17.7%); *CLPTM1*, 1.6 (19.7%); *BCAM*, 1.8 (15.6%); *TOMM40*, 2.1 (19%).

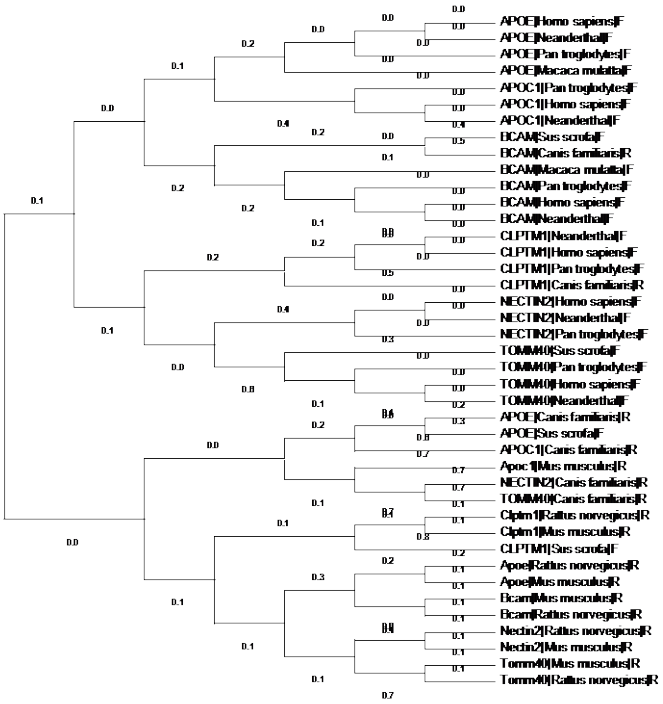


Figure 6. Phylogenetic comparison of the six *APOE* cluster promoters by multiple sequence alignments for human, neanderthal, chimpanzee, macaque monkey (primate lineage); mouse and rat (rodent); dog, and pig (omnivore). Figure S8 shows the *APOE* clusters of these species. Lengths of tree branches represent the relative sequence similarity, calculated by neighbor-joining algorithm. The promoters of human, neanderthal, chimpanzee, and macaque were clearly separated from promoters of mouse-rat, and dog-pig, as expected. The human and mouse *APOE* promoters were more conserved than the other four *APOE* cluster genes examined. The human and mouse *APOE* promoters were more conserved than the other four *APOE* cluster genes we examined: human-mouse distances rank in ascending order of distance *APOE*, 1.1; *APOC1*, 1.4; *NECTIN2*, 1.4; *BCAM*, 1.8; *TOMM40*, 2.1, *CLPTM1*, 1.6.

The shared co-regulation of these six genes may have been a factor in the synteny that persisted in multiple lineages that diverged at least 100 million years ago, as observed for other co-regulated gene clusters [18]. Because topologically associated domains (TAD) of chromatin are conserved particularly for syntenic regions [44], the *APOE* gene cluster is likely to be within a TAD.

The top regulatory candidates sought by screening for mRNA differences of TF genes that were identified by binding motifs in GTEx brain data (Fig. 7B). Genes were ranked by their correlation with *APOE* cluster brain PC1 in GTEx data. The analysis included all brain regions, and was then restricted to hippocampus, or amygdala for relevance to AD. The top correlated TF genes included *POU3F4* (r=0.22), *SOX2* (r=0.17), *MLX* (r= -0.17), *ATF1* (r= -0.17), MLX (r= -0.17), and BHLHE40 (r= -0.16) (Table 1). Hippocampal genes with the strongest correlations were *SOX1* (r= 0.2), and *DMBX1* (r = -0.3). In amygdala, *DBP* (r = 0.3) and *DMBX1* (r = -0.3) were highly correlated with *APOE* PC1. A subset was enriched in upstream regulators of ME3 from consensus WGCNA of total brain: *SOX2*, *NEUROG2*, *SOX3*, *DBP,* and *KLF1*. Using a relaxed criteria of lower

correlation, several genes from the MAF family showed inverse correlations with PC1: *NFE2* (whole brain r= -0.07; amygdala r = -0.21), *MAFG* (brain r= -0.09), *BACH1* (brain r= -0.14), *NFE2L3* (brain r= -0.12), and *MAFB* (brain r = -0.12) (Table S1). These genes are associated with NRF2 mediated phase II gene responses to oxidative stress involving several genes that respond to air pollution: *Nhlh2* (-30%) and *Mafg* (+30%). These genes also showed correlated *APOE* cluster expression in human brain for PC1: *NHLH2* (amygdala r = 0.22); *MAFG* (brain r = -0.09).

Candidate regulatory genes were then screened for variants associated with AD risk, using GWAS of the International Genomics of Alzheimer's project (IGAP). At stringent criteria for genome wide significance ($p < 5 \times 10^{-8}$), there were no SNPs for our genes of interest. A relaxed significance ($p<0.003$) showed 10 SNPs for AD risk in *ELK4*, and 2 SNPs in *SOX1*. *SOX1* was also positively correlated with *APOE* cluster PC1 in hippocampus. For additional gene candidates, see Supplement excel file.

Acknowledgement: We appreciate comments from Christian Pike (USC), Hussein Yassine (USC), Derek Wildman (University of South Florida), Alexander Kulminski (Duke University). The evolutionary framework draws on discussions with members of the Center for Academic Research and Training in Anthropogeny (CARTA).

Author contributions: Conceptualization, A.H., C.E.F.; Statistical analysis: A.H.; Protein analysis: A.H., M.T.; Writing, A.H., M.T., T.E.M., C.E.F.; Supervision, Project Administration, and Funding Acquisition, T.E.M., and C.E.F.

Funding: This work was supported by the Cure Alzheimer's Fund (C.E.F.) and the National Institutes on Aging: C.E.F. (R01-AG051521, P50-AG005142, P01-AG055367); A.H. (PI: Kelvin Davis, T32- AG052374; Nelson Freimer, 5T32NS048004-15); M.T. (PI: Eileen Crimmins, T32-AG000037).

Data availability: All data used in this paper are publicly available.

Conflict of interest: The authors have no conflict of interest to declare.

References

[1] Cacciottolo M, Wang X, Driscoll I, Woodward N, Saffari A, Reyes J, et al. Particulate air pollutants, APOE alleles and their contributions to cognitive impairment in older women and to amyloidogenesis in experimental models. Transl Psychiatry. 2017;7:e1022.

[2] Kulick ER, Elkind MSV, Boehme AK, Joyce NR, Schupf N, Kaufman JD, et al. Long-term exposure to ambient air pollution, APOE-epsilon4 status, and cognitive decline in a cohort of older adults in northern Manhattan. Environ Int. 2020;136:105440.

[3] Roses AD, Lutz MW, Amrine-Madsen H, Saunders AM, Crenshaw DG, Sundseth SS, et al. A TOMM40 variable-length polymorphism predicts the age of late-onset Alzheimer's disease. Pharmacogenomics J. 2010;10:375-84.

[4] Kulminski AM, Shu L, Loika Y, Nazarian A, Arbeev K, Ukraintseva S, et al. APOE region molecular signatures of Alzheimer's disease across races/ethnicities. Neurobiol Aging. 2020;87:141 e1- e8.

[5] Zhou X, Chen Y, Mok KY, Kwok TCY, Mok VCT, Guo Q, et al. Non-coding variability at the APOE locus contributes to the Alzheimer's risk. Nat Commun. 2019;10:3310.

[6] Lutz MW, Sundseth SS, Burns DK, Saunders AM, Hayden KM, Burke JR, et al. A Genetics-based Biomarker Risk Algorithm for Predicting Risk of Alzheimer's Disease. Alzheimers Dement (N Y). 2016;2:30-44.

[7] Davignon J, Roederer G. [Phenotype of apolipoprotein E, hyperlipidemia and atherosclerosis]. Union Med Can. 1988;117:56-61.

[8] Shi J, Liu Y, Liu Y, Li Y, Qiu S, Bai Y, et al. Association between ApoE polymorphism and hypertension: A meta-analysis of 28 studies including 5898 cases and 7518 controls. Gene. 2018;675:197-207.

[9] Kim IY, Grodstein F, Kraft P, Curhan GC, Hughes KC, Huang H, et al. Interaction between apolipoprotein E genotype and hypertension on cognitive function in older women in the Nurses' Health Study. PLoS One. 2019;14:e0224975.

[10] Mooyaart AL, Valk EJ, van Es LA, Bruijn JA, de Heer E, Freedman BI, et al. Genetic associations in diabetic nephropathy: a meta-analysis. Diabetologia. 2011;54:544-53.

[11] Davignon J, Gregg RE, Sing CF. Apolipoprotein E polymorphism and atherosclerosis. Arteriosclerosis. 1988;8:1-21.

[12] Deelen J, Evans DS, Arking DE, Tesi N, Nygaard M, Liu X, et al. A meta-analysis of genome-wide association studies identifies multiple longevity genes. Nat Commun. 2019;10:3669.

[13] Lu F, Guan H, Gong B, Liu X, Zhu R, Wang Y, et al. Genetic variants in PVRL2-TOMM40-APOE region are associated with human longevity in a Han Chinese population. PLoS One. 2014;9:e99580.

[14] Babenko VN, Smagin DA, Kudryavtseva NN. RNA-Seq Mouse Brain Regions Expression Data Analysis: Focus on ApoE Functional Network. J Integr Bioinform. 2017;14:356-16.

[15] Allan CM, Taylor S, Taylor JM. Two hepatic enhancers, HCR.1 and HCR.2, coordinate the liver expression of the entire human apolipoprotein E/C-I/C-IV/C-II gene cluster. J Biol Chem. 1997;272:29113-9.

[16] Subramanian S, Gottschalk WK, Kim SY, Roses AD, Chiba-Falek O. The effects of PPARgamma on the regulation of the TOMM40-APOE-C1 genes cluster. Biochim Biophys Acta Mol Basis Dis. 2017;1863:810-6.

689 [17] Mijalski T, Harder A, Halder T, Kersten M, Horsch M, Strom TM, et al. Identification
690 of coexpressed gene clusters in a comparative analysis of transcriptome and proteome
691 in mouse tissues. Proc Natl Acad Sci U S A. 2005;102:8621-6.
692 [18] Ghanbarian AT, Hurst LD. Neighboring Genes Show Correlated Evolution in Gene
693 Expression. Mol Biol Evol. 2015;32:1748-66.
694 [19] Finch CE, Austad SN. Commentary: is Alzheimer's disease uniquely human?
695 Neurobiol Aging. 2015;36:553-5.
696 [20] Rapoport SI. Hypothesis: Alzheimer's disease is a phylogenetic disease. Medical
697 Hypotheses. 1989;29:147-50.
698 [21] Haghani A, Johnson R, Safi N, Zhang H, Thorwald M, Mousavi A, et al. Toxicity of
699 urban air pollution particulate matter in developing and adult mouse brain: Comparison of
700 total and filter-eluted nanoparticles. Environ Int. 2020;136:105510.
701 [22] Haghani A, Cacciottolo M, Doty KR, D'Agostino C, Thorwald M, Safi N, et al. Mouse
702 brain transcriptome responses to inhaled nanoparticulate matter differed by sex and
703 APOE in Nrf2-Nfkb interactions. Elife. 2020;9.
704 [23] Woodward NC, Pakbin P, Saffari A, Shirmohammadi F, Haghani A, Sioutas C, et al.
705 Traffic-related air pollution impact on mouse brain accelerates myelin and neuritic aging
706 changes with specificity for CA1 neurons. Neurobiol Aging. 2017;53:48-58.
707 [24] Pomatto LCD, Cline M, Woodward N, Pakbin P, Sioutas C, Morgan TE, et al. Aging
708 attenuates redox adaptive homeostasis and proteostasis in female mice exposed to
709 traffic-derived nanoparticles ('vehicular smog'). Free Radic Biol Med. 2018;121:86-97.
710 [25] Mumaw CL, Levesque S, McGraw C, Robertson S, Lucas S, Stafflinger JE, et al.
711 Microglial priming through the lung-brain axis: the role of air pollution-induced circulating
712 factors. FASEB J. 2016;30:1880-91.
713 [26] Forman HJ, Finch CE. A critical review of assays for hazardous components of air
714 pollution. Free Radic Biol Med. 2018;117:202-17.
715 [27] Aragon MJ, Topper L, Tyler CR, Sanchez B, Zychowski K, Young T, et al. Serum-
716 borne bioactivity caused by pulmonary multiwalled carbon nanotubes induces
717 neuroinflammation via blood-brain barrier impairment. Proc Natl Acad Sci U S A.
718 2017;114:E1968-E76.
719 [28] Morgan TE, Davis DA, Iwata N, Tanner JA, Snyder D, Ning Z, et al. Glutamatergic
720 neurons in rodent models respond to nanoscale particulate urban air pollutants in vivo
721 and in vitro. Environ Health Perspect. 2011;119:1003-9.
722 [29] Woodward NC, Levine MC, Haghani A, Shirmohammadi F, Saffari A, Sioutas C, et
723 al. Toll-like receptor 4 in glial inflammatory responses to air pollution in vitro and in vivo.
724 J Neuroinflammation. 2017;14:84.
725 [30] Maloney B, Ge YW, Alley GM, Lahiri DK. Important differences between human and
726 mouse APOE gene promoters: limitation of mouse APOE model in studying Alzheimer's
727 disease. J Neurochem. 2007;103:1237-57.
728 [31] Rowan-Carroll A, Halappanavar S, Williams A, Somers CM, Yauk CL. Mice exposed
729 in situ to urban air pollution exhibit pulmonary alterations in gene expression in the lipid
730 droplet synthesis pathways. Environ Mol Mutagen. 2013;54:240-9.
731 [32] Labib S, Williams A, Kuo B, Yauk CL, White PA, Halappanavar S. A framework for
732 the use of single-chemical transcriptomics data in predicting the hazards associated with
733 complex mixtures of polycyclic aromatic hydrocarbons. Arch Toxicol. 2017;91:2599-616.

[33] Zhang H, Liu H, Davies KJ, Sioutas C, Finch CE, Morgan TE, et al. Nrf2-regulated phase II enzymes are induced by chronic ambient nanoparticle exposure in young mice with age-related impairments. Free Radic Biol Med. 2012;52:2038-46.
[34] Kulminski AM, Huang J, Wang J, He L, Loika Y, Culminskaya I. Apolipoprotein E region molecular signatures of Alzheimer's disease. Aging Cell. 2018;17:e12779.
[35] Zhang H, Haghani A, Mousavi AH, Cacciottolo M, D'Agostino C, Safi N, et al. Cell-based assays that predict in vivo neurotoxicity of urban ambient nano-sized particulate matter. Free Radic Biol Med. 2019;145:33-41.
[36] Malek-Ahmadi M, Perez SE, Chen K, Mufson EJ. Braak Stage, Cerebral Amyloid Angiopathy, and Cognitive Decline in Early Alzheimer's Disease. J Alzheimers Dis. 2020;74:189-97.
[37] Duyckaerts C, Braak H, Brion JP, Buee L, Del Tredici K, Goedert M, et al. PART is part of Alzheimer disease. Acta Neuropathol. 2015;129:749-56.
[38] Guo T, Landau SM, Jagust WJ, Alzheimer's Disease Neuroimaging I. Detecting earlier stages of amyloid deposition using PET in cognitively normal elderly adults. Neurology. 2020;94:e1512-e24.
[39] Morgan TE, Xie Z, Goldsmith S, Yoshida T, Lanzrein AS, Stone D, et al. The mosaic of brain glial hyperactivity during normal ageing and its attenuation by food restriction. Neuroscience. 1999;89:687-99.
[40] Nichols NR, Day JR, Laping NJ, Johnson SA, Finch CE. GFAP mRNA increases with age in rat and human brain. Neurobiology of Aging. 1993;14:421-9.
[41] Akram A, Schmeidler J, Katsel P, Hof PR, Haroutunian V. Association of ApoE and LRP mRNA levels with dementia and AD neuropathology. Neurobiol Aging. 2012;33:628 e1- e14.
[42] Lee EG, Tulloch J, Chen S, Leong L, Saxton AD, Kraemer B, et al. Redefining transcriptional regulation of the APOE gene and its association with Alzheimer's disease. PLoS One. 2020;15:e0227667.
[43] Zeng J, Zhu S, Yan H. Towards accurate human promoter recognition: a review of currently used sequence features and classification methods. Brief Bioinform. 2009;10:498-508.
[44] Liao Y, Zhang X, Chakraborty M, Emerson JJ. Topologically associating domains and their role in the evolution of genome structure and function in Drosophila. bioRxiv. 2020:2020.05.13.094516.
[45] Trumble BC, Finch CE. The Exposome in Human Evolution: From Dust to Diesel. Q Rev Biol. 2019;94:333-94.
[46] Lerner SP, Finch CE. The major histocompatibility complex and reproductive functions. Endocr Rev. 1991;12:78-90.
[47] Finch CE, Rose MR. Hormones and the physiological architecture of life history evolution. Q Rev Biol. 1995;70:1-52.
[48] Qiao L, Luo GG. Human apolipoprotein E promotes hepatitis B virus infection and production. PLoS Pathog. 2019;15:e1007874.
[49] Linard M, Letenneur L, Garrigue I, Doize A, Dartigues JF, Helmer C. Interaction between APOE4 and herpes simplex virus type 1 in Alzheimer's disease. Alzheimers Dement. 2020;16:200-8.

[50] Kuo C-L, Pilling LC, Atkins JL, Masoli JAH, Delgado J, Kuchel GA, et al. APOE e4 genotype predicts severe COVID-19 in the UK Biobank community cohort. The Journals of Gerontology: Series A. 2020.
[51] Kulminski AM, Loika Y, Culminskaya I, Huang J, Arbeev KG, Bagley O, et al. Independent associations of TOMM40 and APOE variants with body mass index. Aging Cell. 2019;18:e12869.
[52] Setti L, Passarini F, De Gennaro G, Baribieri P, Perrone MG, Borelli M, et al. SARS-Cov-2 RNA Found on Particulate Matter of Bergamo in Northern Italy: First Preliminary Evidence. medRxiv. 2020:2020.04.15.20065995.
[53] Sweeney MD, Montagne A, Sagare AP, Nation DA, Schneider LS, Chui HC, et al. Vascular dysfunction-The disregarded partner of Alzheimer's disease. Alzheimers Dement. 2019;15:158-67.
[54] Hachinski V, Einhaupl K, Ganten D, Alladi S, Brayne C, Stephan BCM, et al. Preventing dementia by preventing stroke: The Berlin Manifesto. Alzheimers Dement. 2019;15:961-84.
[55] Finch CE, Shams S. Apolipoprotein E and Sex Bias in Cerebrovascular Aging of Men and Mice. Trends Neurosci. 2016;39:625-37.
[56] Rius-Perez S, Tormos AM, Perez S, Talens-Visconti R. Vascular pathology: Cause or effect in Alzheimer disease? Neurologia. 2018;33:112-20.
[57] Kaufman JD, Adar SD, Barr RG, Budoff M, Burke GL, Curl CL, et al. Association between air pollution and coronary artery calcification within six metropolitan areas in the USA (the Multi-Ethnic Study of Atherosclerosis and Air Pollution): a longitudinal cohort study. Lancet. 2016;388:696-704.
[58] Schraufnagel DE, Balmes JR, Cowl CT, De Matteis S, Jung SH, Mortimer K, et al. Air Pollution and Noncommunicable Diseases: A Review by the Forum of International Respiratory Societies' Environmental Committee, Part 1: The Damaging Effects of Air Pollution. Chest. 2019;155:409-16.
[59] Button EB, Robert J, Caffrey TM, Fan J, Zhao W, Wellington CL. HDL from an Alzheimer's disease perspective. Curr Opin Lipidol. 2019;30:224-34.
[60] Ungvari Z, Tarantini S, Nyul-Toth A, Kiss T, Yabluchanskiy A, Csipo T, et al. Nrf2 dysfunction and impaired cellular resilience to oxidative stressors in the aged vasculature: from increased cellular senescence to the pathogenesis of age-related vascular diseases. Geroscience. 2019;41:727-38.
[61] Guzik TJ, Touyz RM. Oxidative Stress, Inflammation, and Vascular Aging in Hypertension. Hypertension. 2017;70:660-7.
[62] Nakashima A, Kawamoto T, Noshiro M, Ueno T, Doi S, Honda K, et al. Dec1 and CLOCK Regulate Na(+)/K(+)-ATPase beta1 Subunit Expression and Blood Pressure. Hypertension. 2018;72:746-54.
[63] Finch CE, Kulminski AM. The Alzheimer's Disease Exposome. Alzheimers Dement. 2019;15:1123-32.
[64] Veitch DP, Weiner MW, Aisen PS, Beckett LA, Cairns NJ, Green RC, et al. Understanding disease progression and improving Alzheimer's disease clinical trials: Recent highlights from the Alzheimer's Disease Neuroimaging Initiative. Alzheimers Dement. 2019;15:106-52.

[65] Durazzo TC, Mattsson N, Weiner MW, Alzheimer's Disease Neuroimaging I. Smoking and increased Alzheimer's disease risk: a review of potential mechanisms. Alzheimers Dement. 2014;10:S122-45.
[66] Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, et al. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc. 2012;7:562-78.
[67] Robinson MD, Oshlack A. A scaling normalization method for differential expression analysis of RNA-seq data. Genome Biol. 2010;11:R25.
[68] Consortium GT. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. Science. 2015;348:648-60.
[69] Aguet F, Barbeira AN, Bonazzola R, Brown A, Castel SE, Jo B, et al. The GTEx Consortium atlas of genetic regulatory effects across human tissues. bioRxiv. 2019.
[70] Berchtold NC, Coleman PD, Cribbs DH, Rogers J, Gillen DL, Cotman CW. Synaptic genes are extensively downregulated across multiple brain regions in normal human aging and Alzheimer's disease. Neurobiol Aging. 2013;34:1653-61.
[71] Caceres M, Lachuer J, Zapala MA, Redmond JC, Kudo L, Geschwind DH, et al. Elevated gene expression levels distinguish human from non-human primate brains. Proc Natl Acad Sci U S A. 2003;100:13030-5.
[72] Lambert JC, Ibrahim-Verbaas CA, Harold D, Naj AC, Sims R, Bellenguez C, et al. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. Nat Genet. 2013;45:1452-8.
[73] Langfelder P, Horvath S. Tutorials for the WGCNA Package. UCLA; 2014.
[74] Matys V, Kel-Margoulis OV, Fricke E, Liebich I, Land S, Barre-Dirrie A, et al. TRANSFAC and its module TRANSCompel: transcriptional gene regulation in eukaryotes. Nucleic Acids Res. 2006;34:D108-10.
[75] Wingender E. The TRANSFAC project as an example of framework technology that supports the analysis of genomic regulation. Brief Bioinform. 2008;9:326-32.
[76] Katoh K, Rozewicki J, Yamada KD. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. Brief Bioinform. 2019;20:1160-6.
[77] Hunt SE, McLaren W, Gil L, Thormann A, Schuilenburg H, Sheppard D, et al. Ensembl variation resources. Database (Oxford). 2018;2018.
[78] Saunders AM, Seldin MF. A molecular genetic linkage map of mouse chromosome 7. Genomics. 1990;8:525-35.

Table 1. Potential transcriptional regulators of APOE-cluster genes. The top ranked candidates were selected based on their transcriptional factor group, correlation with *APOE* PC1 in human brain, response to nPM in mouse cerebral cortex, identification as potential upstream regulator in WGCNA modules by IPA, and genetic variants associated with AD risk.

			Correlation with *APOE* region PC1 in human (r)			Responses to air pollution in the cerebral cortex of adult mice		upstream of modules	SNPs associated with risk of AD in IGAP meta-analysis			
SYMBOL	Factor	Chr	brain	hip	amyg	Log2 fold change	pval		ID	Location	Beta	pval
POU3F4	Dlx group	X	0.22									
SOX2	Sox10	3	0.17		0.21			ME3				
SIX3	paired related HD factors	2	0.17									
HMX1	Nkx group	2	0.16									
SOX1	Sox10	4	0.15	0.23					rs571564	3'UTR	0.05	0.0071
									rs12429920	promoter	-0.05	0.0073
NEUROG2	Myogenin group	13	0.14					ME3				
HES5	Ebox	4	0.14		0.30							
SOX3	Sox10	1	0.12		0.24			ME3				
NHLH2	Myogenin group	X	0.09		0.22	-0.29	0.02					
DBP	ATF-2 group	X	0.06		0.33			ME3				
NFE2	MAF group; AP-1	1	-0.07		-0.21							
POU4F3	Dlx group	19	-0.08	0.24								
CHURC1	Churchill	12	-0.08			-0.24	0.02					
DMBX1	paired related HD factors	12	-0.08	-0.31	-0.30							
NKX3-1	Nkx group	5	-0.08	-0.16								
GLIS2	GLI group	14	-0.09			0.36	0.01					
MAFG	MAF group	1	-0.09			0.30	0.00					
RREB1	RREB-1	1	-0.15			0.49	0.01					
BHLHE40	Ebox; E2A group	8	-0.16									
CEBPD	C/EBP group	16	-0.16		-0.21							
ELK4	Ets-related factors	17	-0.16						rs41304251	Intron	-0.33	0.0001
									rs59990545	Intron	-0.28	0.0003
									rs57847589	Intron	-0.28	0.0003
									rs72750941	Intron	-0.28	0.0004
									rs55736919	Intron	-0.28	0.0004
									rs66742105	Intron	-0.27	0.0005
									rs56927917	Intron	-0.26	0.0005
									rs7521095	Intron	-0.27	0.0006
									rs55781355	Intron	-0.26	0.0008
									rs75638049	Intron	-0.12	0.0024
KLF11	Sp1 group	6	-0.16					ME3				
ATF1	CREB group	3	-0.17									
MLX	Ebox	3	-0.17									