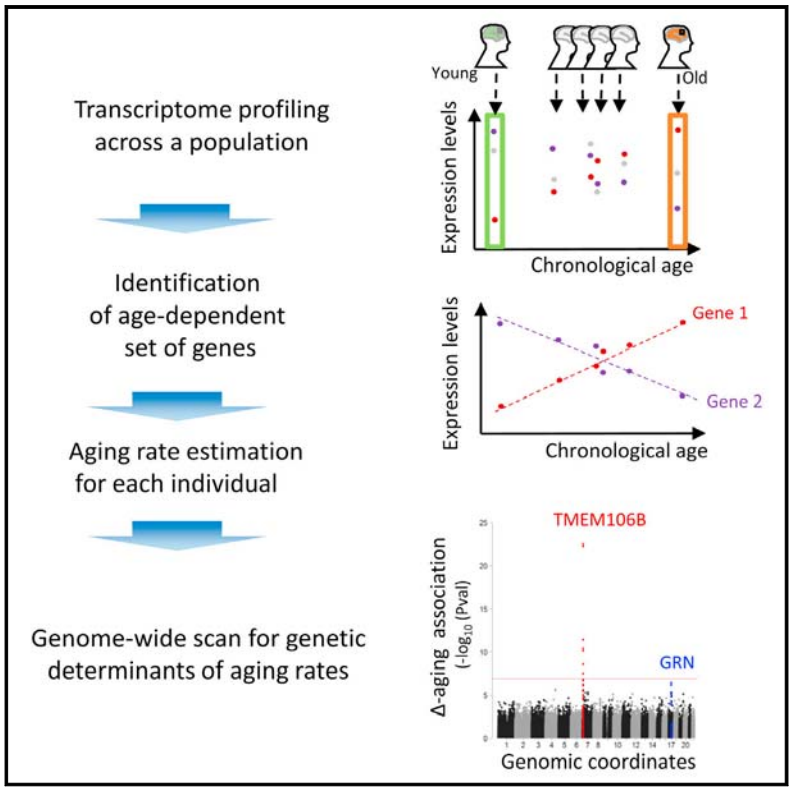
Article

Cell Systems

Differential Aging Analysis in Human Cerebral Cortex Identifies Variants in TMEM106B and GRN that Regulate Aging Phenotypes

Graphical Abstract



Highlights

- Δ-Aging quantifies determinants of individuals appearing younger/older than peers
- TMEM106B and GRN variants interact genetically in the regulation of Δ-aging
- The role of TMEM106B in aging appears CNS region and life-stage selective
- TMEM106B modulates CNS inflammatory and degenerative changes independently of disease

Authors

Herve Rhinn, Asa Abeliovich

Correspondence

hr2239@columbia.edu (H.R.),
aa900@columbia.edu (A.A.)

In Brief

Rhinn et al. describe an integrative genomics approach to quantify aging rate in a tissue of interest. A subsequent GWAS analysis identifies the TMEM106B-progranulin genetic pathway as a key determinant of age-associated manifestations in the human cerebral cortex.

Rhinn & Abeliovich, 2017, Cell Systems *4*, 1–12
April 26, 2017
http://dx.doi.org/10.1016/j.cels.2017.02.009

CellPress

Please cite this article in press as: Rhinn and Abeliovich, Differential Aging Analysis in Human Cerebral Cortex Identifies Variants in TMEM106B and GRN that Regulate Aging Phenotypes, Cell Systems (2017), http://dx.doi.org/10.1016/j.cels.2017.02.009

Cell Systems
Article

Differential Aging Analysis in Human Cerebral Cortex Identifies Variants in TMEM106B and GRN that Regulate Aging Phenotypes

Herve Rhinn[1,2,*] and Asa Abeliovich[1,2,3,*]
[1]Departments of Pathology, Cell Biology, and Neurology
[2]Taub Institute for Alzheimer's Disease and the Aging Brain
Columbia University, New York, NY 10032, USA
[3]Lead Contact
*Correspondence: hr2239@columbia.edu (H.R.), aa900@columbia.edu (A.A.)
http://dx.doi.org/10.1016/j.cels.2017.02.009

SUMMARY

Human age-associated traits, such as cognitive decline, can be highly variable across the population, with some individuals exhibiting traits that are not expected at a given chronological age. Here we present differential aging (Δ-aging), an unbiased method that quantifies individual variability in age-associated phenotypes within a tissue of interest, and apply this approach to the analysis of existing transcriptome-wide cerebral cortex gene expression data from several cohorts totaling 1,904 autopsied human brain samples. We subsequently performed a genome-wide association study and identified the TMEM106B and GRN gene loci, previously associated with frontotemporal dementia, as determinants of Δ-aging in the cerebral cortex with genome-wide significance. TMEM106B risk variants are associated with inflammation, neuronal loss, and cognitive deficits, even in the absence of known brain disease, and their impact is highly selective for the frontal cerebral cortex of older individuals (>65 years). The methodological framework we describe can be broadly applied to the analysis of quantitative traits associated with aging or with other parameters.

INTRODUCTION

The rate at which human age-associated phenomena advance in otherwise healthy individuals, termed healthy biological aging, is highly variable (Deary et al., 2012; Jones et al., 2014; Pitt and Kaeberlein, 2015). This has been hypothesized to be a consequence, in part, of genetic heterogeneity across the population. However, specific genetic factors that determine the rate of normal biological aging remain to be elucidated. Rare Progeria syndromes are caused by single gene mutations, but these disorders are likely to be mechanistically distinct from the common healthy aging process (Burtner and Kennedy, 2010). Prior studies have also identified genetic factors such as apolipoprotein E (APOE) that modify the likelihood of extreme longevity, as with centenarians, using genome-wide association studies (GWAS) (Deelen et al., 2014). But such extreme longevity may reflect a selective reduction in the incidence of some major causes of age-associated mortality, such as atherosclerosis, rather than an altered rate of biological aging.

The relationship between physiological "healthy" aging and aging-associated diseases is complex. Pathological hallmarks of Alzheimer's disease (AD), which is a progressive dementia seen primarily in late life, include neurofibrillary tangles and amyloid plaques in the CNS, but these changes can also be found in the CNS of adults without clinical evidence of dementia, albeit to a lesser degree (Yu et al., 2015). Thus, it has been hypothesized that some aspect of healthy brain aging may represent a prodromal state to disease pathology. However, by functional criteria such as cognitive measures, changes associated with healthy aging appear distinct from those seen in neurodegeneration (Small et al., 2011).

Here, we describe a quantitative approach, termed differential aging (Δ-aging), to evaluate the rate of aging across a cohort of tissue without prior assumption on the nature of the age-associated phenotypic changes, based on transcriptome-wide gene expression analyses. The Δ-aging trait represents the difference between an observed level of an age-associated phenotype, as measured for an individual of a given age, and the level of such a phenotype that would be predicted for the same individual by interpolating from a cohort of age-phenotype correlative data. Transcriptomic or epigenetic analyses have previously been used to identify age-associated phenotypic changes in a hypothesis-free manner (Bocklandt et al., 2011; Colantuoni et al., 2011; de Magalhaes et al., 2009; Glass et al., 2013; Hannum et al., 2013; Kang et al., 2011; Ori et al., 2015; Zahn et al., 2007). Such descriptive results highlight the biological pathways that are most affected by aging. To identify genes with a functional role upstream of the observed age-associated changes and that would be regulators of the rate of aging, we turned to a genetic approach. Specifically, we performed a GWAS to identify common genetic variants in the human genome that are associated with increased or decreased Δ-aging, based on large-scale gene expression datasets. Common variants at two genetic loci, TMEM106B and progranulin, were found to be associated with an increased rate of biological aging. The effect of TMEM106B risk variants on Δ-aging was observed selectively in frontal cortex tissue and most prominently in individuals over 65 years. The

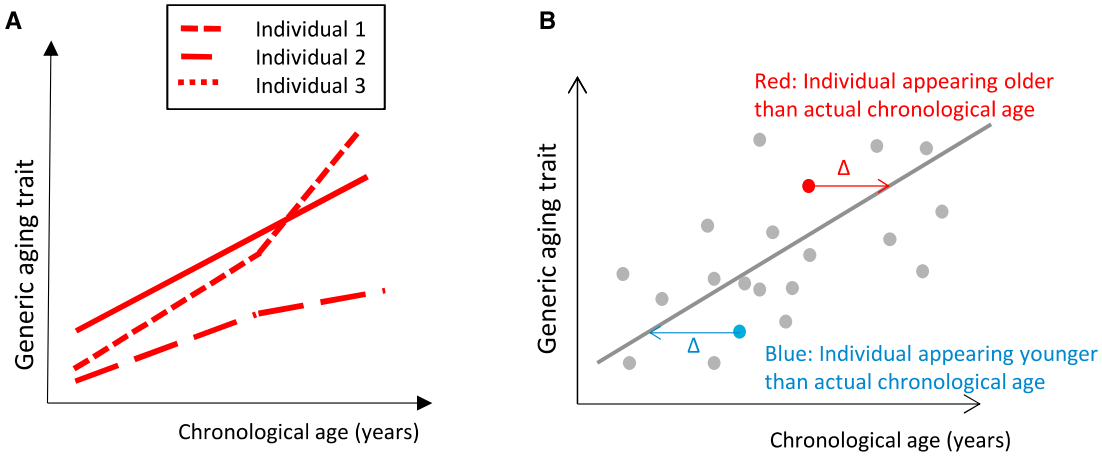


Figure 1. Aging Rates are Heterogeneous across Individuals within a Cohort

(A) Schematic representation of variability in the relative rate and pattern of progression of age-associated traits as a function of time. In this hypothetical example, a generic aging trait progresses more rapidly in individual 2 than individual 1, whereas individual 3 displays a bimodal pattern.

(B) Schematics representation of Δ-aging analysis. Each dot represents, for a single individual, the tissue expression level (x axis) of a hypothetical age-dependent gene as a function of chronological age (y axis). In this example, expression levels are positively correlated with chronological age across the cohort, as shown by the regression line. Individuals that display an expression level higher than predicted for their chronological age, such as the sample highlighted in red, exhibit an estimated biological age higher than their chronological age (positive Δ-aging). In contrast, samples that display an expression level lower than predicted for their chronological age, such as the sample highlighted in blue, would be associated with an estimated biological age lower than the chronological age (negative Δ-aging). Integration across all age-associated genes constitutes the aggregate Δ-aging for that individual.

effect of TMEM106B risk variants was found to be selective to frontal cortex tissue in late life. Further annotation analyses revealed that the presence of TMEM106B risk variants is associated with an increased inflammatory polarization of innate immune markers, and a reduction in neuronal markers. As the pro-inflammatory impact of the risk variants was seen even in the context of isolated innate immune cells, we hypothesize that this represents a proximal effect of the risk variant. Analysis of tissue from individuals with neurodegenerative diseases, including AD, suggested a broader role for TMEM106B in the CNS response to pathological or age-associated insults.

RESULTS

Data-Driven Quantification of Biological Aging

To investigate potential determinants of the rate of biological aging, we defined a quantitative trait (termed Δ-aging) that captures whether an individual displays age-associated phenotypes that are more or less marked than expected for his or her true chronological age (Figure 1A). The Δ-aging trait for a given individual within a cohort is a theoretical value defined as the difference between the age as evaluated on the basis of the phenotypic measurements and the true chronological age of the individual, and thus with the dimension of time (Figure 1B).

We applied this approach using transcriptome-wide gene expression data but note that other biological datasets could similarly be used. Δ-Aging analysis of transcriptomic data is performed in two steps (Figures S1A and S1B, detailed in the Supplemental Information): (1) all transcripts that are correlated in their expression levels with the chronological ages of individuals within a given cohort of samples are identified, and (2) for each individual (or sample) within the cohort, the quantitative trait Δ-aging is defined as the difference between an age predicted from the aggregate expression levels of the age-dependent transcripts, and the actual age of the individual (Figures 1B and S1). To address non-linear aspects of biological aging in a facile manner, cohorts may be subdivided into selected age ranges (such as late life) to be studied independently.

In the most elementary case, Δ-aging for an individual may be quantified based on the analysis of a single gene, the expression level of which is significantly correlated with age within a given cohort (Figures 1B and S1). However, such a limited analysis would most likely reflect gene-specific variations across the cohort, rather than an aspect of aging. Thus, to capture diverse age-associated phenotypes within a tissue of interest, Δ-aging herein represents an aggregated analysis of gene expression across the entire transcriptome of each individual (Figure S1, see the Supplemental Information).

Age-Associated Changes in Gene Expression Patterns in Human Frontal Cortex

We initially sought to characterize the global effect of aging on transcriptome-wide gene expression in post-mortem, autopsied prefrontal cortex tissue samples from individuals free of known neuropathology (Figures 2A–2C). The correlation between chronological age and gene expression levels was queried in transcriptome-wide microarray datasets from four independent cohorts (*Tgen*, Myers et al., 2007; Webster et al., 2009; *BrainEqtl*, Gibbs et al., 2010; *HBTRC*, Zhang et al., 2013; and *BrainCloud*, Colantuoni et al., 2011; detailed in STAR Methods and Table S1). Meta-analysis of the results obtained in the four datasets (n = 716 individuals >25 years) identified 3,329 genes that were significantly correlated in expression with chronological age (false discovery rate < 5% by linear regression, after correction for gender and batch effects, among the 10,474 genes present in all four datasets; see STAR Methods for details; Table S2). Functional

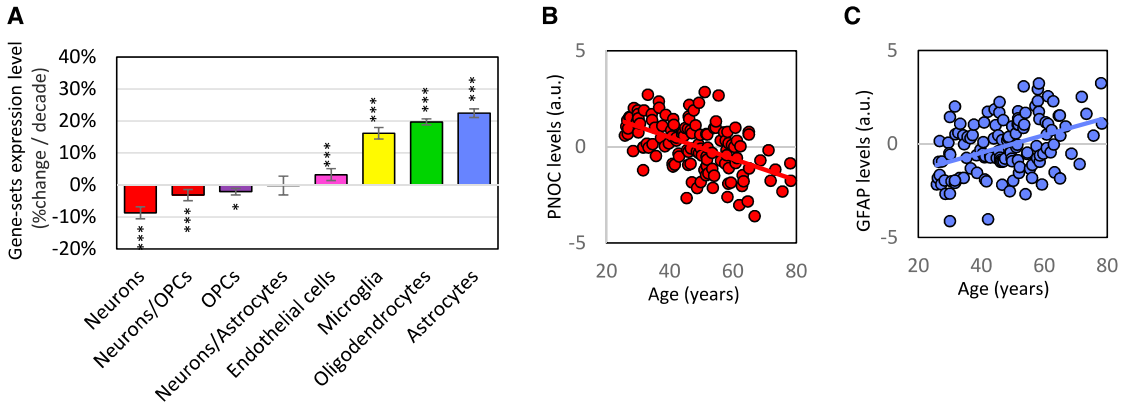


Figure 2. Transcriptome-wide Analysis Identifies Age-Associated Gene Expression Changes in Human Brain

(A) Age-associated changes in the aggregated expression levels of gene sets that typify different human CNS cell types in prefrontal cortex, as labeled. Transcriptome-wide meta-analysis of age-associated gene expression levels was conducted using four prefrontal cortex datasets; n = 716 neurodegenerative disease-free individuals >25 years. Mean age-associated fold-changes across gene sets members are plotted (n = 21 per group). Error bars are SEM. *$p < 0.05$, ***$p < 0.001$ for linear association between age and aggregated expression levels in a meta-analysis across the four expression datasets. Neurons (red): $Z = -3.93$, $p = 8.47 \times 10^{-5}$; neurons/oligodendrocytes precursor cells (OPCs) (burgundy): $Z = -4.24$, $p = 2.24 \times 10^{-5}$; OPCs (purple) $Z = -2.27$, $p = 2.33 \times 10^{-2}$; neurons/astrocytes (orange): $Z = -1.89$, $p = 5.81 \times 10^{-2}$; endothelial cells (pink): $Z = 4.02$, $p = 5.78 \times 10^{-5}$; microglia (yellow): $Z = 2.93$, $p = 3.39 \times 10^{-3}$; oligodendrocytes (green): $Z = 3.97$, $p = 7.31 \times 10^{-5}$; and astrocytes (blue): $Z = 5.29$, $p = 1.25 \times 10^{-7}$.

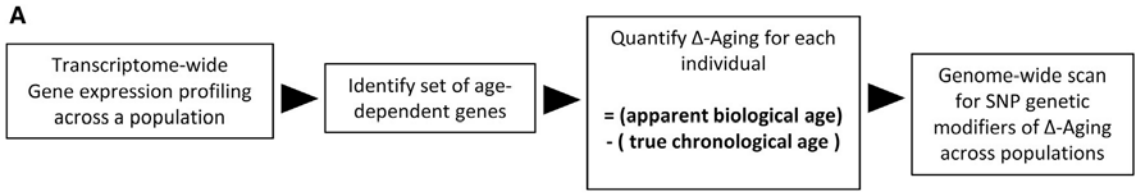
(B and C) Expression levels of prepronociceptin (PNOC), a representative neuronal gene that is decreased in expression with aging (B) (red) and glial fibrillary acidic protein (GFAP), a representative astrocyte gene that is increased in expression with aging (C) (blue) as a function of age. Each dot represents an individual sample; n = 128 individuals of age >25 years from the BrainCloud dataset; regression lines are shown.

annotation revealed an age-associated decrease in the expression of neuronal genes and a parallel age-associated increase in the expression of genes characteristic of astrocytes, microglia, and oligodendrocytes, as defined by the molecular signatures obtained by single-cell RNA sequencing (RNA-seq) from the human brain (Darmanis et al., 2015) (Figure 2A). These observations are consistent with the progressive age-associated loss of neurons and their processes, concomitant with astrogliosis and microglial expansion, as described in neuropathological studies (Beach et al., 1989; Mosher and Wyss-Coray, 2014).

A detailed examination of individual genes the expression levels of which are positively or negatively correlated with chronological age (such as the neuronal gene prepronociceptin or glial fibrillary acidic protein) nonetheless revealed considerable variability between tissue samples obtained from different individuals at any given age (Figures 2B and 2C, Table S2). Consistent with the Δ-aging theoretical model presented above (Figure 1B), we hypothesized that such variability may in part reflect biological diversity in the aging process, due to genetic or non-genetic factors. In this model, for any given age-associated gene, some individual tissue samples may display expression levels that are higher or lower than expected for their chronological age, in part as a consequence of accelerated or decelerated biological aging (Figure 1B).

TMEM106B as a Genetic Determinant of the Rate of Aging in Human Cerebral Cortex

Using Δ-aging as a quantitative trait, we next undertook a GWAS (Figure 3A) in search of genetic modifiers of biological aging, in a meta-analysis across four transcriptome-wide frontal cortex gene expression datasets (N = 716 individuals without known brain pathology; Table S1). A strong association was observed between Δ-aging and the SNP variant rs1990622, at the TMEM106B gene locus in the discovery cohort ($p = 2.77 \times 10^{-7}$, n = 716, Figures 3B and S2), which was replicated with genome-wide significance in the replication cohort ($p = 1.68 \times 10^{-15}$, n = 497, Religious Orders Study and Memory and Aging Project [ROS-MAP] cohort, see details below, Figures 3B and S2), leading to a combined association p value of 1.5×10^{-19} in 1,213 samples (Figure 3B). As this association appeared most prominently within a dataset composed of older adults only (>65 years; *Tgen* dataset; Figure 3A, Table S1), we further refined the entire meta-analysis by stratifying individuals as either younger (<65 years) or older (>65 years) adults within all datasets. This approach revealed an age-dependence in the association between rs1990622 and Δ-aging that reached genome-wide significance in the older but not the younger cohorts ($p = 5.4 \times 10^{-10}$, N = 413 and $p = 8.0 \times 10^{-2}$, N = 303, respectively; Figures 3B, S1, and S2). Analysis of an independent, population-based RNA-seq gene expression dataset of prefrontal cortex samples from older adults, which includes both unaffected individuals and individuals with neurological diseases, replicated the association between rs1990622 and Δ-aging (ROS-MAP; Chan et al., 2015; $p = 1.68 \times 10^{-15}$, N = 497 samples from individuals >65 years; Figure 3B). Joint meta-analysis that included all of the above datasets yielded a combined p value of 2.5×10^{-23} for the association between rs1990622 and Δ-aging (n = 910 samples; Figures 3C–3E). In this meta-analysis, a second association was observed at SNP rs708384 on chromosome 17, which falls at the progranulin gene locus, but this did not quite reach genome-wide statistical significance after correction for multiple testing (GRN; $p = 6.23 \times 10^{-7}$ at rs708384, Figure 3C). Taken together, these findings strongly support the hypothesis that genetic factors account for some of the variability seen in the apparent rate of biological aging traits across the human population.



B

Stage	Dataset Name	Status	In all adults (>25yo) Effect (yrs/allele)	n	p	In younger only (<65yo) Effect (yrs/allele)	n	p	In older only (>65yo) Effect (yrs/allele)	n	p
Discovery	Tgen/HBTRC/BrainEqtl/BrainCloud	Und/Und/Und/Und	-4.50	716	2.77×10^{-7}	-1.62	303	0.08	-6.69	413	5.41×10^{-10}
Replication	RosMap	Long.	-5.54	497	1.68×10^{-15}	NA	0	NA	-5.54	497	1.68×10^{-15}
Discovery +Replication			-4.92	1213	1.50×10^{-19}	-1.62	303	0.08	-6.06	910	2.54×10^{-23}
Disease	Tgen/HBTRC/HBTRC	AD/AD/HD	-3.72	691	1.29×10^{-4}	1.04	145	0.77	-4.75	546	9.61×10^{-6}
All combined			-4.49	1904	1.48×10^{-21}	-0.76	448	0.85	-5.57	1456	1.20×10^{-25}

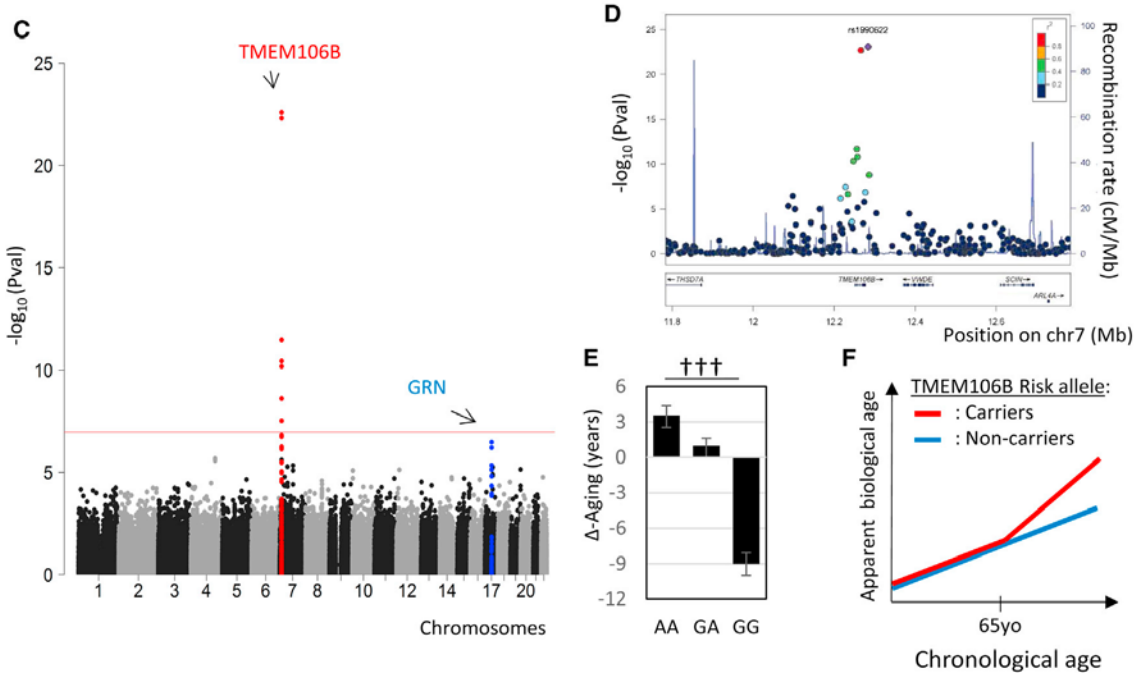


Figure 3. Genome-wide Association Study Identifies a Genetic Determinant of Aging Rate in Human Frontal Cortex at the TMEM106B Locus

(A) Schematic of the genetic analysis of modifiers of Δ-aging in human frontal cortex.

(B) Tabular presentation of the associations observed between rs1990622 genotype at TMEM106B and Δ-aging through stages of the GWAS. Effects are expressed in terms of years per minor allele load. See STAR Methods for details on the statistical analyses. NA, not applicable.

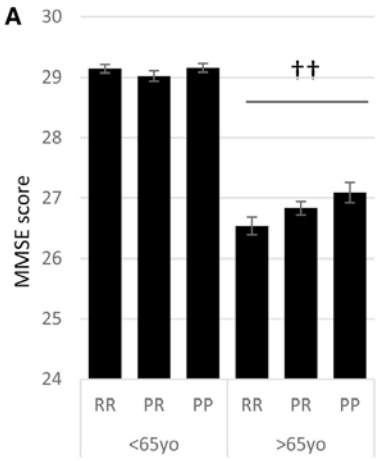
(C) Manhattan plot representing the association between Δ-aging, as quantified in frontal cortex tissue samples from older adults, and each of 468,129 common SNP variants (meta-analysis of 5 cohorts; Discovery + Replication as in (B) and Table S1; n = 910). The red line corresponds to a threshold ($p < 1.06 \times 10^{-7}$) for genome-wide significance after Bonferroni correction for the multiple SNPs tested. Highlighted in red are the SNPs in the region of interest (chr7:11,783,787-12,783,787) surrounding the association peaks at the TMEM106B (rs1990622) and GRN loci.

(D) Local Manhattan plot in the region of interest of the TMEM106B gene locus, representing the genome-wide association p value between common genetic variants and Δ-aging in older adults frontal cortex samples in a meta-analysis of 5 cohorts (Discovery + Replication, n = 910) after local imputation.

(E) Effect of rs1990622 allele load on Δ-aging value in a meta-analysis of 910 tissue samples from older neurodegenerative-free individuals. Homozygosity for the minor protective (PP) allele is associated with a 12-year decrease in Δ-aging, relative to homozygosity of the risk (RR) allele. N = 182, 438, and 290 for the PP, PR, and RR genotypes. Mean values are presented. Error bars are SEM. $^{†††}p = 2.74 \times 10^{-22}$ for the effect of "R" allelic load on Δ-aging by Kruskal-Wallis test, chi-square = 99.30 with 1 degree of freedom.

(F) Schematic representation of the effect of TMEM106B rs1990622 genotype on brain aging trajectories. Risk allele carriers display an accelerated aging phenotype in late life (>65 years).

The TMEM106B SNP most strongly associated with increased Δ-aging in the meta-analysis, rs1990622, was previously also associated with frontotemporal dementia (FTD) (Cruchaga et al., 2011; Finch et al., 2011; Van Deerlin et al., 2010). Specifically, the major allele at rs1990622 (A, "Risk"; with a 59.2% allelic frequency in population of European ascent), which in the meta-analysis above is associated to an increase in Δ-aging (Figure 3A), had previously been shown to increase the risk of FTD (Van Deerlin



B

	MMSE , LongLife Study			
	< 65yo Individuals (n=1731)		> 65yo Individuals (n=2154)	
Factor	Effect	p	Effect	p
Age	-0.01/year	0.13	-0.17/year	$2.2x10^{-96}$
TMEM106B Risk allele load	-0.03/G allele	0.68	0.26/G allele	$7.7x10^{-3}$

C

	Memory Score , Health and Retirement Study			
	50-65yo Individuals (n=6605)		65-80yo Individuals (n=5606)	
Factor	Effect	p	Effect	p
Age	-0.08/year	$1.2.x10^{-35}$	-0.16/year	$9.8x10^{-58}$
TMEM106B Risk allele load	0.007/G allele	0.88	0.13/ G allele	$2.7x10^{-2}$

Figure 4. TMEM106B and Age Affect Cognitive Function in Elderly Individuals

(A and B) Carriers of the TMEM106B Δ-aging risk-associated haplotype display age-associated cognitive deficits, as assessed by Mini Mental State Examination (MMSE) in a cross-sectional cohort (Long Life Study) of individuals. Consistent with the Δ-aging findings, the deficits are seen in older (65–80 years) but not younger (50–65 years) individuals. rs1060700 was used a proxy for rs1990622 (pairwise linkage disequilibrium: R2 = 1, D' = 1 in the CEU 1000 Genomes datasets). N = 262, 744, and 541, N = 192, 429 and 321 for the homozygous protective-allele (PP), heterozygous (PR), and homozygous risk (RR) genotypes in the 50–65- and 65–80-year-old age groups, respectively. Mean values are presented. Error bars are SEM. $^{††}p < 0.01$ by linear regression for the additive allelic load, correcting for age, gender, and three principal eigenvectors for population stratification.

(C) Carriers of the TMEM106B Δ-aging risk-associated haplotype display age-associated cognitive deficits, as assessed by a memory recall score in the Health and Retirement Study. The Δ-aging-associated genotype was queried using rs1060700 as a proxy for rs1990622 (pairwise linkage disequilibrium: R2 = 1, D' = 1 in the CEU 1000 Genomes datasets).

et al., 2010), as well to reduce the age of onset comparably (Cruchaga et al., 2011; Finch et al., 2011). Similarly, local imputation at the GRN locus revealed that the strongest association with Δ-aging was observed for a variant (rs5848, $p = 1.85 \times 10^{-8}$; Figure S3) previously identified as associated with FTD as well as other neurodegenerative disorders (Chen et al., 2015; Rademakers et al., 2008; van Blitterswijk et al., 2014), with the disease-associated risk allele being associated with an increase in Δ-aging. Thus, the same genetic determinants that modify the rate of apparent biological aging in the frontal cortex of otherwise healthy individuals also appear to play a significant role in a rare neurodegenerative disorder of the frontal cortex.

We next sought to determine whether the effect of TMEM106B genetic variation on the Δ-aging endophenotype (Figures 3F and S4) might be reflected in functional measures of brain aging. We thus queried the relationship between TMEM106B haplotypes and cognitive function, as assessed by the Mini Mental State Examination in a large cross-sectional cohort of genotyped individuals from the National Institute on Aging (NIA) Long Life Family Study (N = 4,953 individuals; Newman et al., 2011; Barral et al., 2012). Remarkably, in individuals without diagnosed dementia, the TMEM106B haplotype associated with a decreased rate of aging selectively in individuals >65 years in the analysis above was also associated with better cognitive scores, specifically in such older individuals ($p = 2.8 \times 10^{-3}$, n = 774; Figures 4A and 4B, rs1060700 was used a proxy for rs1990622 as these are co-inherited, see STAR Methods for details). Similar results were obtained in a second independent cross-sectional cohort (Health and Retirement Study; N = 12,507 genotyped individuals; Juster and Suzman, 1995) in which the TMEM106B haplotype associated with a decrease rate of aging in our analysis above was again associated with improved scores in a memory test specifically among individuals >65 years, but not younger individuals ($p = 1.9 \times 10^{-2}$ in the >65 years age group, n = 5,604; p = 0.48, in the younger individuals, n = 4432; Figure 4C). These analyses of cognitive measures support the relevance of our transcriptomic-based Δ-aging studies of human frontal cortex tissue.

The TMEM106B Genetic Variant Modulates Innate Immune Activation and Neuronal Loss Markers

To assess the impact of the TMEM106B rs1990622 risk allele load (defined as 0, 1, or 2 risk allele copies, as determined by SNP genotyping) on cerebral cortex gene expression in detail, we next compared the transcriptome-wide gene expression changes observed in the context of increasing risk allele load (termed the TMEM106B risk-associated transcriptomic signature of change) to the gene expression changes associated with chronological aging (termed the age-associated transcriptomic signature of change) in four independent gene expression datasets from prefrontal cortex tissue of individuals free of known neuropathology (*Tgen*, *BrainEqtl*, *HBTRC*, and *BrainCloud*; Table S1). A meta-analysis demonstrated that the TMEM106B risk-associated transcriptomic signature of change was broadly correlated with the age-associated transcriptomic signature of change. Furthermore, this correlation appeared selective for tissue from older adults (>65 years), relative to tissue from younger adults (Figures 5A–5C, Tables S2 and S3).

To gain more insight into the biological pathways that are affected by the TMEM106B risk variant, we next applied a gene co-expression network approach, termed whole-genome co-expression network analysis (WGCNA) (Fuller et al., 2007; Langfelder and Horvath, 2008), that can functionally probe

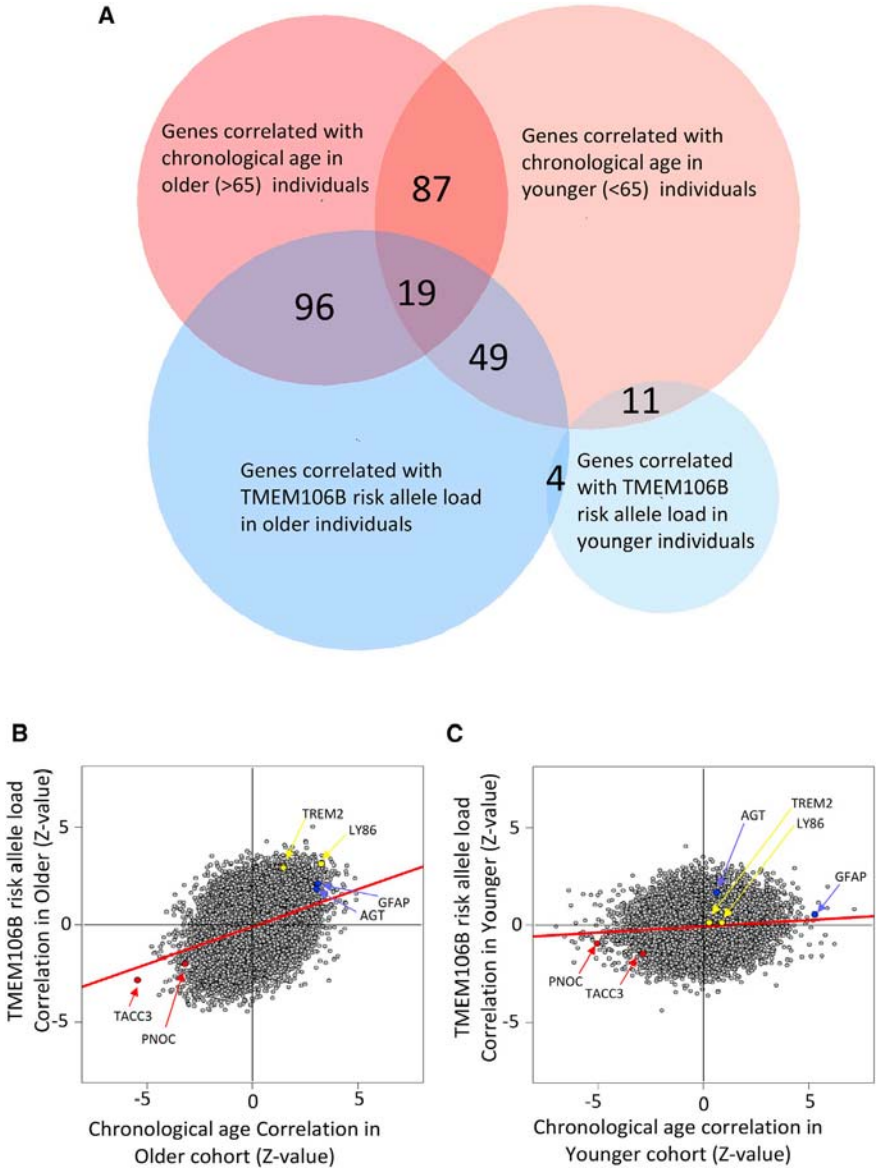


Figure 5. The Effect of TMEM106B Genetic Variant on the Transcriptome Appears Similar to the Effect of Age Exclusively in Elderly Individuals

(A) Venn diagram representing the number of genes, the levels of which correlate with age or TMEM106B risk allele load in cohorts of older (n = 413) or younger (n = 303) adults from the "Discovery" datasets (with an association p value of <0.01 for each factor with same direction of changes, Figure 3B, Tables S1, S2, and S3). The effect of the TMEM106B risk allele appears very different in younger versus older individuals: its impact on the transcriptome is potentiated in individuals over 65 years, in whom its global signature resembles the one associated to chronological age. By contrast, the impact of chronological age appears marked in both individuals below or above 65 years.

(B and C) Dot plots display, for each of 15,139 individual genes (each represented by a single dot), the degree to which gene expression is correlated with TMEM106B risk allele load (y axis) versus the degree to which gene expression is correlated with chronological age (x axis). Separate analyses are presented of older (B) (n = 413) or younger (C) (n = 303) adults from the "Discovery" datasets (see see Figure 3B, Table S1). *Z* scores represent the statistical significance and direction of the age- or genotype-associated correlations. Z values of 1.96, 2.56, and 3.29 correspond to p values of 0.05, 0.01, and 0.001, respectively. Regression lines (in red) show that in older individuals (B), those genes that are more highly correlated (positively or negatively) in expression with TMEM106b risk allele load (y axis) are also more highly correlated with age (x axis). Individual genes characteristic of neurons (red), astrocytes (blue), and microglia (yellow) are highlighted as examples.

transcriptomic patterns of change. WGCNA is a computational tool that clusters genes in an unsupervised (hypothesis-free) manner based on their correlated co-expression, and thus defines biologically relevant groups of genes that typically correspond to specific cell types or processes (Langfelder and Horvath, 2008). WGCNA analysis of a frontal cortical gene expression dataset from older adults defined five gene clusters (Figure S5A; *Tgen* dataset; 179 individuals) that could be categorized in relation to major CNS cell types: microglia, astroglial, oligodendroglia, and two different neuron-associated groups. Assessment of the effect size of the rs1990622 genotype on the expression level of each of these gene clusters revealed the greatest impact to be on the microglia-associated gene cluster, which overall is significantly increased in expression with increased risk allele load. A similar pattern of transcriptome-wide gene expression change was observed with chronological aging, as expected. In contrast, other factors (such as gender or post-mortem interval to time of autopsy) did not significantly affect these gene expression categories in a consistent manner. In a complementary approach to WGCNA, we functionally annotated the TMEM106B risk-associated transcriptomic signature, as well as the age-associated transcriptomic signature, using previously described human CNS single-cell RNA-seq data (Darmanis et al., 2015). This analysis confirmed the microglia gene set to be most increased in expression in the context of the rs1990622 risk variant; again, the effect was seen selectively in the older adult cohort (Figure S5B, Table S3; N = 716 unaffected individuals from *Tgen*, *BrainEqtl*, *HBTRC*, and *BrainCloud* cohorts).

Prior studies have associated pathological aging with an inappropriate polarization of microglia and other innate immune myeloid cells toward an increasingly pro-inflammatory state (Deelen et al., 2014; Gabuzda and Yankner, 2013; Hu et al., 2015; Mosher and Wyss-Coray, 2014; Salminen et al., 2012). Given the altered expression of microglia-associated genes in the context of the TMEM106B risk haplotype, we hypothesized that this may be associated with a selective change in the expression of microglial inflammatory state-associated ("M1") or anti-inflammatory/repair state-associated ("M2") genes (Figure 6A).

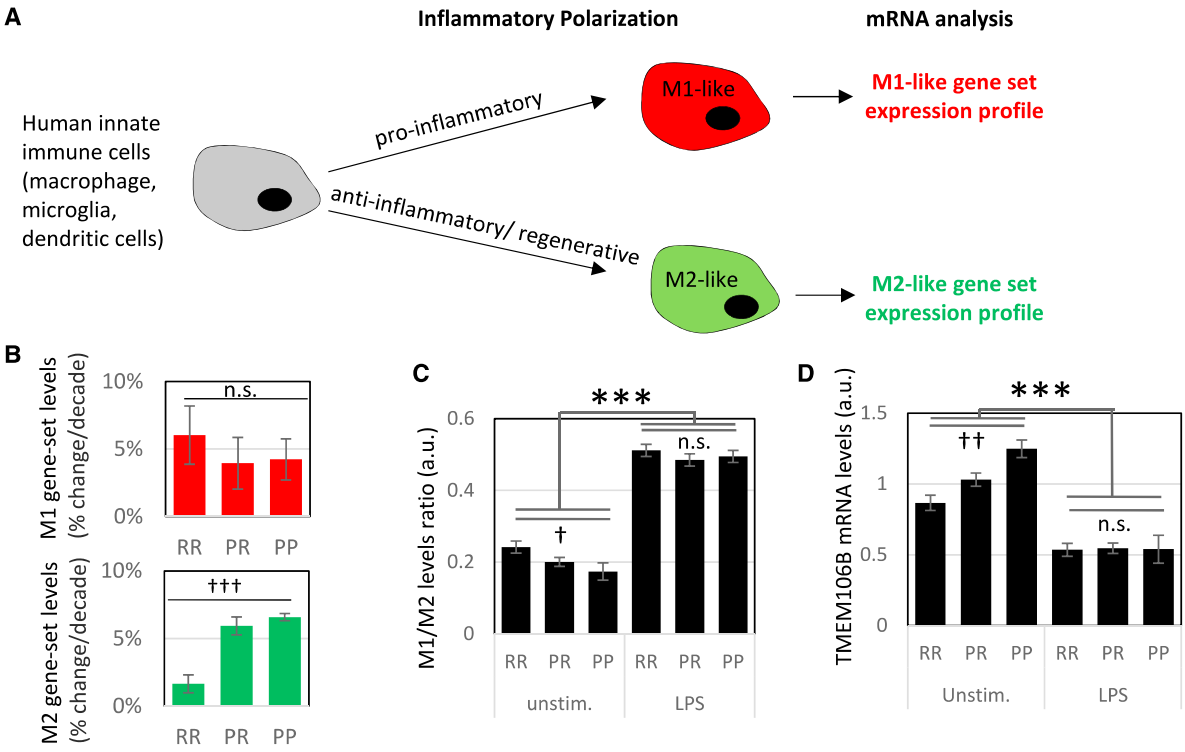


Figure 6. TMEM106B Modulates Innate Immune Cell Inflammatory Polarization

(A) Schematic of the inflammatory polarization of myeloid cells by pro- or anti-inflammatory factors.

(B) Aggregate cerebral cortical expression levels of gene sets associated with M1-type (red, n = 39 genes) or M2-type (green n = 39 genes) microglia were quantified for neurodegenerative disease-free individuals homozygous for the risk (RR) or protective (PP) alleles, or heterozygous (PR) for these alleles, at the TMEM106B locus. Expression was found to increase steadily in all groups over time (presented as percent change per decade); for the M2 gene set, the increase was significantly less rapid in homozygous carriers of the risk allele, relative to carriers of the protective allele. For each group, median values across the datasets are presented (n = 4 datasets from the Discovery stage, with 716 individuals in total; see Table S1). Error bars are SEM. $^{\dagger\dagger\dagger}p < 0.001$, n.s., $p > 0.05$ for the effect of rs1990622 risk allele load by ANOVA, correcting for dataset-specific effects ($p = 9.41 \times 10^{-4}$, $F = 21.41$ for 1 degree of freedom and $p = 0.737$, $F = 0.119$ for 1 degree of freedom for M2 and M1, respectively).

(C) Effect of rs1990622 genotype or LPS treatment on the aggregate expression levels of M1- and M2-selective genes in human monocyte-derived dendritic cells, as quantified by Nanostring (GEO: GSE53165). Data are presented as a ratio of the M1/M2 aggregate levels; higher values are thus associated with a pro-inflammatory state. Mean values are presented. Error bars are SEM. N = 29/82/49 and 13/51/23 for the RR/PR/PP genotypes in the unstimulated and LPS-treated groups, respectively. $^{***}p < 0.001$ treatment effect, $^{\dagger}p < 0.05$ allelic load association by Kruskal-Wallis test ($p = 1.08 \times 10^{-32}$, chi-square = 141.78 for 1 degree of freedom and $p = 1.09 \times 10^{-2}$, chi-square = 9.02 for 1 degree of freedom, respectively).

(D) Effects of rs1990622 genotype or LPS treatment on TMEM106B expression levels in human monocyte-derived dendritic cells, as measured by Affymetrix Human Gene 1.0 ST Array (GSE53166). Mean values are presented. Error bars are SEM. N = 2/14/7 and 2/14/8 for the RR/PR/PP genotypes in the unstimulated and LPS-treated groups respectively. $^{***}p < 0.001$ for treatment effect by ANOVA, $^{\dagger\dagger}p < 0.01$ for allelic load association in unstimulated by ANOVA ($p = 5.0 \times 10^{-3}$, $F = 9.85$ for 1 degree of freedom and $p = 1.13 \times 10^{-12}$, $F = 96.624$ for 1 degree of freedom).

Overall, both the M1- and M2-associated gene sets (detailed in Table S4), as defined by RNA-seq in polarized mouse microglia (Butovsky et al., 2014), were generally increased in expression over the course of chronological aging (Figure 6B). TMEM106B rs1990622 risk allele carriers showed a significantly muted age-associated increase in the expression of the M2 gene set (Figure 6B, Table S5). M1 genes showed a trend toward a potentiated age-associated increase in expression that did not reach statistical significance in protective allele carriers. Taken together, these data suggest an "inflammaging" (Franceschi et al., 2007; Giunta et al., 2008) mechanism in which the TMEM106B risk allele polarizes myeloid cells toward a pro-inflammatory stage in the cerebral cortex of older individuals.

We extended this analysis of inflammatory gene expression by additionally querying the effect of the TMEM106B rs1990622 genotype on another human myeloid cell lineage, dendritic cells. In a gene expression dataset from isolated, unstimulated human monocyte-derived dendritic cells from genotyped individuals without known pathology (GEO: GSE53165; Lee et al., 2014), the presence of the TMEM106B Δ-aging risk allele was again associated with an increased pro-inflammatory M1-like gene expression signature. Furthermore, the effect of the TMEM106B risk allele appeared non-additive with classical pro-inflammatory stimuli such as lipopolysaccharide (LPS) treatment (Figures 6C and S6).

An unexpected observation in the analysis of this dendritic cell dataset (Lee et al., 2014) was that LPS exposure led to a significant reduction in the gene expression level of TMEM106B (Figure 6D). We further note that the TMEM106B rs1990622 risk allele was itself similarly associated with decreased expression of TMEM106B mRNA (Figure 6D), suggesting a potential direct molecular mechanism for the effect of the TMEM106B risk

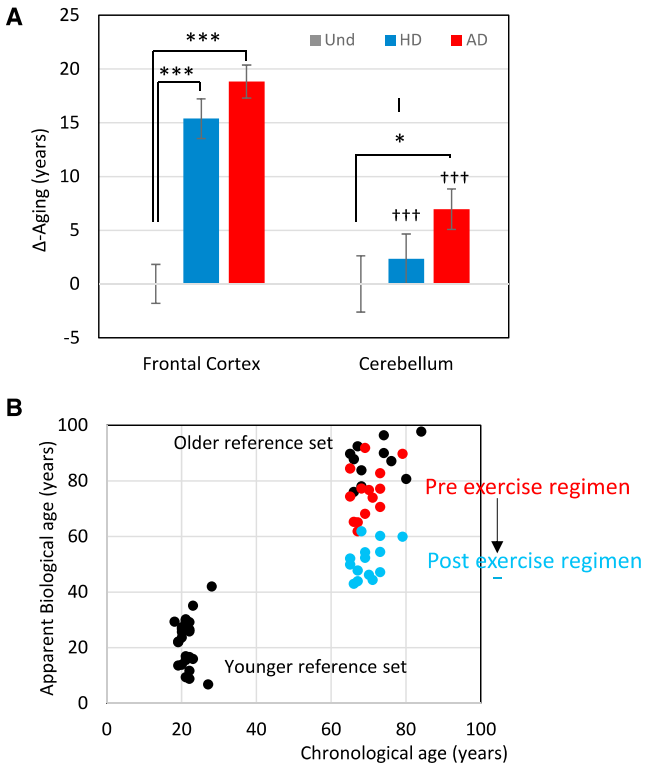


Figure 7. Non-genetic Factors Can Modify Δ-Aging

(A) Δ-Aging values observed in frontal cortex or cerebellum tissue samples derived from neurological disease-free, AD, or HD cohorts (Harvard Brain Bank). Δ-Aging values are presented relative to the values observed in unaffected individuals for each brain region. Mean values are presented. Error bars are SEM. N = 154, 345, 170, 124, 269, and 140; ***$p < 0.001$, *$p < 0.05$ by Kruskal-Wallis test followed by Dunn's multiple comparisons test with the undiagnosed (Und) group in the same tissue; $^{\dagger\dagger\dagger}p < 0.001$ by Kruskal-Wallis followed by Dunn's multiple comparisons test for comparison with frontal cortex tissue in the same disease category ($p = 3.91 \times 10^{-18}$ chi-square = 91.12 with 5 degrees of freedom for Kruskal-Wallis, with Dunn's test $p = 3.72 \times 10^{-13}$ and 9.36×10^{-8} for AD versus Und and HD versus Und, respectively in the frontal cortex, $p = 1.74 \times 10^{-2}$ for AD versus Und in cerebellum and $p = 2.71 \times 10^{-7}$ and 1.14×10^{-4} for cerebellum versus frontal cortex in AD and HD, respectively).

(B) Longitudinal analysis of Δ-aging in serial muscle tissue biopsies from elderly individuals before and after a 6-month-long vigorous exercise routine program (GEO: GSE8479). Black dots represent muscle samples from individuals that did not perform an exercise routine; these data were used to define the pattern of age-associated gene expression changes within the cohort. Red dots represent muscle samples taken from elderly individuals before the exercise regimen, whereas blue dots represent muscle samples from the same individuals after training. The mean Δ-aging shift observed after exercise training is −24.43 ± 1.94 years (SEM, n = 14 individuals, $p = 1.14 \times 10^{-8}$ by paired two-tailed t test).

haplotype on TMEM106B activity. Taken together, these findings implicate a regulatory circuitry, where TMEM106B activity impacts myeloid cell inflammatory status, and conversely where myeloid cell inflammatory status impacts TMEM106B activity.

TMEM106B, Inflammaging, and Neurodegenerative Disorders

To further explore the relationships between Δ-aging and age-associated neurodegenerative diseases, we next analyzed frontal cortex gene expression datasets from individuals with a diagnosis of AD or Huntington's disease (HD). Frontal cortex tissue from such individuals demonstrated significantly increased Δ-aging relative to unaffected individuals (Figure 7A, plus 18.8 and 15.4 years, respectively; HBTRC datasets), and thus appeared significantly older than expected in terms of their transcriptomic profiles. As this dataset also includes tissue samples from the cerebellum, we could further extend the analysis of Δ-aging to this second brain region. The effect of either AD or HD on Δ-aging appeared selective for frontal cortex, relative to cerebellar tissue, consistent with the neuropathological regional patterns that typify these disorders (Figure 7A).

Further utilizing this cerebellar gene expression dataset, we extended the analysis of the association between the TMEM106B risk allele and Δ-aging to this brain region. The impact of the TMEM106B allele on Δ-aging was significantly less robust in cerebellar cortex tissue relative to frontal cortex, and a GWAS analysis in these datasets failed to identify an association of the TMEM106B alleles with Δ-aging in cerebellar tissue, in contrast to frontal cortex tissue (Figure S7; in older individuals from the same cohorts).

We note that the effect of the TMEM106B risk allele appeared additive with the effect of AD on Δ-aging (Figures 7A and S2), rather than occlusive, suggesting that the mechanisms are distinct. Consistent with this, previous studies have established a lack of association between TMEM106B and AD risk (association p values for rs1990622 of 0.26 and 0.87 in stage 1 and 2, respectively, of the Alzheimer Disease Genetics Consortium GWAS in more than 22,000 individuals; Harold et al., 2009). Furthermore, APOE alleles (which are major genetic determinants of AD risk) were not associated with an alteration in Δ-aging (Table S6).

A broader analysis of 70 neurodegenerative disease-associated genetic variants (Jun et al., 2016; Lambert et al., 2013; McMillan et al., 2014; Nalls et al., 2014; Rollinson et al., 2011; van Es et al., 2009) did not reveal any significant additional associations with Δ-aging, beyond TMEM106B and GRN (Table S6). TMEM106B and GRN share a number of common attributes: both have previously been associated with risk of FTD (Cruchaga et al., 2011; Finch et al., 2011; Van Deerlin et al., 2010), with primary hippocampal sclerosis (Aoki et al., 2015), and with TDP-43 neuropathology in the absence of a clinical neurological diagnosis (Dickson et al., 2015; Yu et al., 2015). Furthermore, both genes have been implicated together in the regulation of lysosomal function (Schwenk et al., 2014; Stagi et al., 2014), and TMEM106B has been reported to regulate progranulin protein accumulation (Chen-Plotkin et al., 2012). Consistent with a common pathway of action for these genes, the risk-associated genetic variants at these two loci showed a significant genetic interaction in their modulation of Δ-aging, in that the effect of GRN rs5848 variants on Δ-aging was observed only in carriers of the TMEM106B risk allele, where it reached genome-wide statistical significance ($p = 1.91 \times 10^{-9}$ and 0.48 with N = 689 and 187 in rs1990622 risk allele carriers and non-carriers, respectively, in Discovery + Replication cohorts, $p = 6.42 \times 10^{-10}$ and 0.97 with N = 1,137 and 296 in Discovery + Replication Disease cohorts, as defined in Figure 3A and Table S1).

The fact that the TMEM106B genetic variant is not associated with AD risk and that AD genetic risk factors such as APOE4 do not appear associated with Δ-aging would suggest distinct phenomena. It moreover argues strongly against the possibility that we are merely observing a prodromal AD phenotype in individuals with high Δ-aging values. The effect of the TMEM106B risk variant on Δ-aging in AD patients is significant, but as these patients display markers of accelerated aging, including inflammation and neuronal loss, even independent of TMEM106B, the pathological relevance of TMEM106B is likely to be limited in this context.

Non-genetic Modifiers of Aging

To extend the application of our approach to aging beyond the study of genetic determinants, we further hypothesized that interventions previously associated with healthy aging, such as exercise, may decrease the apparent biological age, as quantified by Δ-aging, in contrast to the increased Δ-aging seen in pathological contexts such as AD (de Cabo et al., 2014). As a proof of principle, we analyzed existing datasets from a longitudinal gene expression study of serial human muscle tissue biopsies, sampled before and after vigorous exercise. Δ-Aging analysis revealed that muscle tissue from individuals who undertook a vigorous exercise regimen for 6 months displayed a significant reduction in Δ-aging, and thus appeared significantly younger than their chronological age (Figure 7B). Δ-Aging may thus serve as a useful biomarker for the evaluation of anti-aging interventions.

DISCUSSION

Δ-Aging analysis allows for an unbiased quantification of an individual's apparent (biological) age, relative to other individuals within the same cohort. Δ-Aging differs qualitatively from other aging analysis frameworks such as the "epigenetic clock" (Bocklandt et al., 2011; Lu et al., 2016), which assumes that aging impacts the expression of the same genes, and the same cellular processes, through all stages of life and in every tissue or context. By contrast, Δ-aging analysis allows a context-dependent identification of age-associated genes, enabling tissue- or age-range-specific processes to be detected and taken into account. The lack of an association between TMEM106B or GRN genetic variants and pleiotropic age-associated markers such as the "methylation clock" or telomere length suggests that TMEM106B and GRN impact the rate of healthy aging in prefrontal cortex independently of such factors. While in the present work we primarily used transcriptome-wide expression data to study aging in brain, subsequent studies may apply this approach to other tissues, other data types, or more systematically query the overlap with other aging-associated markers and the tissue specificity of the effect of TMEM106B and GRN on healthy aging.

Our GWAS identified two genetic loci, at TMEM106B and progranulin, which function cooperatively to modify Δ-aging in the cerebral cortex of older individuals. Genetic variants previously associated with extreme longevity, such as APOE (Deelen et al., 2011, 2014), were not associated with the rate of healthy aging or with age-associated inflammation in the present study. Our findings underscore the genetic differences between mechanisms that govern healthy aging in a given tissue and those that underlie extreme longevity.

The relationship between normal aging and neurodegenerative disorders appears complex. The two gene loci we identified as modifying Δ-aging have previously been associated with FTD risk. It appears extremely unlikely that our results are merely a reflection of "early onset FTD" in the cohorts we analyze for several reasons: first, FTD is an extremely rare disease (its occurrence rate is 10 per 100,000 individuals in individuals aged 60–69 years; Knopman et al., 2004; Knopman and Roberts, 2011); the prevalence of contaminating case of FTD (or even of prodromal FTD) in any of the cohorts we studied are thus most likely very low, while the association between TMEM106B genotype and Δ-aging is seen strongly in all of those cohorts. Furthermore, the "Unaffected individuals" in these cohorts were neuropathologically assessed; an FTD case would very likely have been identified and filtered out.

A potential explanation for the dual association of those loci with both FTD risk and Δ-aging is that the proximate effect of these risk variants is to cause the tissue substrate for FTD (frontal cortex) to appear older, and thus secondarily to increase the prevalence of FTD (as prevalence of FTD is highly age dependent). Although there are pathological criteria that differentiate healthy aging from disorders such as AD or FTD, certain processes are common, such as inflammation. Furthermore, neurodegenerative disease hallmarks, such as TDP-43 aggregates, are seen even in apparently healthy individuals, albeit to a limited extent (Beecham et al., 2014; Crary et al., 2014; Yu et al., 2015). Indeed, TMEM106B risk variants have been associated with increased TDP-43 aggregates in neuropathology-based association studies of apparently healthy older individuals (Dickson et al., 2015; Yu et al., 2015). We hypothesize that the selective role of TMEM106B in the aging frontal cortex may reflect unique stressors present in this tissue late in life, such as the accumulation of inflammatory cell debris or protein aggregates. The TMEM106B-progranulin pathway may modulate the response to such stressors both during healthy aging and in the context of neurodegenerative disease (Figure 6C) (Martens et al., 2012; Tanaka et al., 2013; Yin et al., 2010). Further studies in model systems may help to unravel underlying cellular mechanisms. The identification of TMEM106B risk variant-dependent phenotypic changes in peripheral circulating myeloid cells (Figures 6A, 6C, 6D, and S5) are of particular interest as they open the perspective of peripheral biomarkers of brain aging in monocytes or macrophages. Finally, although our analysis focused on the role of TMEM106B in inflammation, we do not exclude a function in neurons (Brady et al., 2013; Chen-Plotkin et al., 2012; Schwenk et al., 2014).

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- METHOD DETAILS
  - Gene Expression Analysis in Human Brain
  - Δ-Aging Analysis
  - Theory and Calculation of Delta-Age

- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Genotype Association Analysis
  - Gene Co-expression Analysis
  - Human Dendritic Cell Analysis
- DATA AND SOFTWARE AVAILABILITY
  - Custom Dependent Functions

SUPPLEMENTAL INFORMATION

Supplemental Information includes eight figures and six tables and can be found with this article online at http://dx.doi.org/10.1016/j.cels.2017.02.009.

AUTHOR CONTRIBUTIONS

A.A. supervised the research, conceived and designed the analysis and wrote the manuscript. H.R. conceived, designed, and performed the analysis and wrote the manuscript. A.A. is a co-founder and consultant for Alector. H.R. is a consultor for Alector.

ACKNOWLEDGMENTS

We thank Dr. Philip L. De Jager for sharing genotypes for the human monocyte-derived dendritic cells with us. We thank Jessie Chang and Ralitsa Petrova for comments on the manuscript. We are extremely grateful to the teams who made their data available through the GEO, Synapse, the AMP-AD Partnership, the database of Genotypes and Phenotypes (dbGaP), or the NIA Genetics of Alzheimer's Disease Data Storage Site (NIAGADS). The work was supported by grants from the NIA (AG042317), NINDS, and the Michael J. Fox Foundation. A.A. is a co-founder and consultant for Alector. Herve Rhinn is a consultant for Alector.

Received: November 9, 2016
Revised: January 14, 2017
Accepted: February 3, 2017
Published: March 15, 2017

REFERENCES

Aoki, N., Murray, M.E., Ogaki, K., Fujioka, S., Rutherford, N.J., Rademakers, R., Ross, O.A., and Dickson, D.W. (2015). Hippocampal sclerosis in Lewy body disease is a TDP-43 proteinopathy similar to FTLD-TDP type A. Acta Neuropathol. *129*, 53–64.

Barral, S., Cosentino, S., Costa, R., Matteini, A., Christensen, K., Andersen, S.L., Glynn, N.W., Newman, A.B., and Mayeux, R. (2012). Cognitive function in families with exceptional survival. Neurobiol. Aging *33*, 619.e1-7.

Beach, T.G., Walker, R., and McGeer, E.G. (1989). Patterns of gliosis in Alzheimer's disease and aging cerebrum. Glia *2*, 420–436.

Beecham, G.W., Hamilton, K., Naj, A.C., Martin, E.R., Huentelman, M., Myers, A.J., Corneveaux, J.J., Hardy, J., Vonsattel, J.P., Younkin, S.G., et al. (2014). Genome-wide association meta-analysis of neuropathologic features of Alzheimer's disease and related dementias. PLoS Genet. *10*, e1004606.

Bocklandt, S., Lin, W., Sehl, M.E., Sanchez, F.J., Sinsheimer, J.S., Horvath, S., and Vilain, E. (2011). Epigenetic predictor of age. PLoS One *6*, e14821.

Brady, O.A., Zheng, Y., Murphy, K., Huang, M., and Hu, F. (2013). The frontotemporal lobar degeneration risk factor, TMEM106B, regulates lysosomal morphology and function. Hum. Mol. Genet. *22*, 685–695.

Burtner, C.R., and Kennedy, B.K. (2010). Progeria syndromes and ageing: what is the connection? Nat. Rev. Mol. Cell Biol. *11*, 567–578.

Butovsky, O., Jedrychowski, M.P., Moore, C.S., Cialic, R., Lanser, A.J., Gabriely, G., Koeglsperger, T., Dake, B., Wu, P.M., Doykan, C.E., et al. (2014). Identification of a unique TGF-beta-dependent molecular and functional signature in microglia. Nat. Neurosci. *17*, 131–143.

Chan, G., White, C.C., Winn, P.A., Cimpean, M., Replogle, J.M., Glick, L.R., Cuerdon, N.E., Ryan, K.J., Johnson, K.A., Schneider, J.A., et al. (2015). CD33 modulates TREM2: convergence of Alzheimer loci. Nat. Neurosci. *18*, 1556–1558.

Chang, C.C., Chow, C.C., Tellier, L.C., Vattikuti, S., Purcell, S.M., and Lee, J.J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience *4*, 7.

Chen-Plotkin, A.S., Unger, T.L., Gallagher, M.D., Bill, E., Kwong, L.K., Volpicelli-Daley, L., Busch, J.I., Akle, S., Grossman, M., Van Deerlin, V., et al. (2012). TMEM106B, the risk gene for frontotemporal dementia, is regulated by the microRNA-132/212 cluster and affects progranulin pathways. J. Neurosci. *32*, 11213–11227.

Chen, Y., Li, S., Su, L., Sheng, J., Lv, W., Chen, G., and Xu, Z. (2015). Association of progranulin polymorphism rs5848 with neurodegenerative diseases: a meta-analysis. J. Neurol. *262*, 814–822.

Colantuoni, C., Lipska, B.K., Ye, T., Hyde, T.M., Tao, R., Leek, J.T., Colantuoni, E.A., Elkahloun, A.G., Herman, M.M., Weinberger, D.R., et al. (2011). Temporal dynamics and genetic control of transcription in the human prefrontal cortex. Nature *478*, 519–523.

Crary, J.F., Trojanowski, J.Q., Schneider, J.A., Abisambra, J.F., Abner, E.L., Alafuzoff, I., Arnold, S.E., Attems, J., Beach, T.G., Bigio, E.H., et al. (2014). Primary age-related tauopathy (PART): a common pathology associated with human aging. Acta Neuropathol. *128*, 755–766.

Cruchaga, C., Graff, C., Chiang, H.H., Wang, J., Hinrichs, A.L., Spiegel, N., Bertelsen, S., Mayo, K., Norton, J.B., Morris, J.C., et al. (2011). Association of TMEM106B gene polymorphism with age at onset in granulin mutation carriers and plasma granulin protein levels. Arch. Neurol. *68*, 581–586.

Darmanis, S., Sloan, S.A., Zhang, Y., Enge, M., Caneda, C., Shuer, L.M., Hayden Gephart, M.G., Barres, B.A., and Quake, S.R. (2015). A survey of human brain transcriptome diversity at the single cell level. Proc. Natl. Acad. Sci. U S A *112*, 7285–7290.

de Cabo, R., Carmona-Gutierrez, D., Bernier, M., Hall, M.N., and Madeo, F. (2014). The search for antiaging interventions: from elixirs to fasting regimens. Cell *157*, 1515–1526.

de Magalhaes, J.P., Curado, J., and Church, G.M. (2009). Meta-analysis of age-related gene expression profiles identifies common signatures of aging. Bioinformatics *25*, 875–881.

Deary, I.J., Yang, J., Davies, G., Harris, S.E., Tenesa, A., Liewald, D., Luciano, M., Lopez, L.M., Gow, A.J., Corley, J., et al. (2012). Genetic contributions to stability and change in intelligence from childhood to old age. Nature *482*, 212–215.

Deelen, J., Beekman, M., Uh, H.W., Helmer, Q., Kuningas, M., Christiansen, L., Kremer, D., van der Breggen, R., Suchiman, H.E., Lakenberg, N., et al. (2011). Genome-wide association study identifies a single major locus contributing to survival into old age; the APOE locus revisited. Aging Cell *10*, 686–698.

Deelen, J., Beekman, M., Uh, H.W., Broer, L., Ayers, K.L., Tan, Q., Kamatani, Y., Bennet, A.M., Tamm, R., Trompet, S., et al. (2014). Genome-wide association meta-analysis of human longevity identifies a novel locus conferring survival beyond 90 years of age. Hum. Mol. Genet. *23*, 4420–4432.

Dickson, D.W., Rademakers, R., Nicholson, A.M., Schneider, J.A., Yu, L., and Bennett, D.A. (2015). The TMEM106B locus and TDP-43 pathology in older persons without FTLD. Neurology *85*, 1354–1355.

Finch, N., Carrasquillo, M.M., Baker, M., Rutherford, N.J., Coppola, G., Dejesus-Hernandez, M., Crook, R., Hunter, T., Ghidoni, R., Benussi, L., et al. (2011). TMEM106B regulates progranulin levels and the penetrance of FTLD in GRN mutation carriers. Neurology *76*, 467–474.

Franceschi, C., Capri, M., Monti, D., Giunta, S., Olivieri, F., Sevini, F., Panourgia, M.P., Invidia, L., Celani, L., Scurti, M., et al. (2007). Inflammaging and anti-inflammaging: a systemic perspective on aging and longevity emerged from studies in humans. Mech. Ageing Dev. *128*, 92–105.

Fuller, T.F., Ghazalpour, A., Aten, J.E., Drake, T.A., Lusis, A.J., and Horvath, S. (2007). Weighted gene coexpression network analysis strategies applied to mouse weight. Mamm. Genome *18*, 463–472.

Gabuzda, D., and Yankner, B.A. (2013). Physiology: inflammation links ageing to the brain. Nature *497*, 197–198.

Gibbs, J.R., van der Brug, M.P., Hernandez, D.G., Traynor, B.J., Nalls, M.A., Lai, S.L., Arepalli, S., Dillman, A., Rafferty, I.P., Troncoso, J., et al. (2010). Abundant quantitative trait loci exist for DNA methylation and gene expression in human brain. PLoS Genet. *6*, e1000952.

Giunta, B., Fernandez, F., Nikolic, W.V., Obregon, D., Rrapo, E., Town, T., and Tan, J. (2008). Inflammaging as a prodrome to Alzheimer's disease. J. Neuroinflammation *5*, 51.

Glass, D., Vinuela, A., Davies, M.N., Ramasamy, A., Parts, L., Knowles, D., Brown, A.A., Hedman, A.K., Small, K.S., Buil, A., et al. (2013). Gene expression changes with age in skin, adipose tissue, blood and brain. Genome Biol. *14*, R75.

Hannum, G., Guinney, J., Zhao, L., Zhang, L., Hughes, G., Sadda, S., Klotzle, B., Bibikova, M., Fan, J.B., Gao, Y., et al. (2013). Genome-wide methylation profiles reveal quantitative views of human aging rates. Mol. Cell *49*, 359–367.

Harold, D., Abraham, R., Hollingworth, P., Sims, R., Gerrish, A., Hamshere, M.L., Pahwa, J.S., Moskvina, V., Dowzell, K., Williams, A., et al. (2009). Genome-wide association study identifies variants at CLU and PICALM associated with Alzheimer's disease. Nat. Genet. *41*, 1088–1093.

Howie, B.N., Donnelly, P., and Marchini, J. (2009). A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. PLoS Genet. *5*, e1000529.

Hu, X., Leak, R.K., Shi, Y., Suenaga, J., Gao, Y., Zheng, P., and Chen, J. (2015). Microglial and macrophage polarization-new prospects for brain repair. Nat. Rev. Neurol. *11*, 56–64.

Jones, O.R., Scheuerlein, A., Salguero-Gomez, R., Camarda, C.G., Schaible, R., Casper, B.B., Dahlgren, J.P., Ehrlen, J., Garcia, M.B., Menges, E.S., et al. (2014). Diversity of ageing across the tree of life. Nature *505*, 169–173.

Jun, G., Ibrahim-Verbaas, C.A., Vronskaya, M., Lambert, J.C., Chung, J., Naj, A.C., Kunkle, B.W., Wang, L.S., Bis, J.C., Bellenguez, C., et al. (2016). A novel Alzheimer disease locus located near the gene encoding tau protein. Mol. Psychiatry *21*, 108–117.

Juster, F.T., and Suzman, R. (1995). An overview of the health and retirement study. J. Hum. Resour. *30*, S7.

Kang, H.J., Kawasawa, Y.I., Cheng, F., Zhu, Y., Xu, X., Li, M., Sousa, A.M., Pletikos, M., Meyer, K.A., Sedmak, G., et al. (2011). Spatio-temporal transcriptome of the human brain. Nature *478*, 483–489.

Knopman, D.S., and Roberts, R.O. (2011). Estimating the number of persons with frontotemporal lobar degeneration in the US population. J. Mol. Neurosci. *45*, 330–335.

Knopman, D.S., Petersen, R.C., Edland, S.D., Cha, R.H., and Rocca, W.A. (2004). The incidence of frontotemporal lobar degeneration in Rochester, Minnesota, 1990 through 1994. Neurology *62*, 506–508.

Lambert, J.C., Ibrahim-Verbaas, C.A., Harold, D., Naj, A.C., Sims, R., Bellenguez, C., DeStafano, A.L., Bis, J.C., Beecham, G.W., Grenier-Boley, B., et al. (2013). Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. Nat. Genet. *45*, 1452–1458.

Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics *9*, 559.

Lee, M.N., Ye, C., Villani, A.C., Raj, T., Li, W., Eisenhaure, T.M., Imboywa, S.H., Chipendo, P.I., Ran, F.A., Slowikowski, K., et al. (2014). Common genetic variants modulate pathogen-sensing responses in human dendritic cells. Science *343*, 1246980.

Lu, A.T., Hannon, E., Levine, M.E., Hao, K., Crimmins, E.M., Lunnon, K., Kozlenkov, A., Mill, J., Dracheva, S., and Horvath, S. (2016). Genetic variants near MLST8 and DHX57 affect the epigenetic age of the cerebellum. Nat. Commun. *7*, 10561.

Martens, L.H., Zhang, J., Barmada, S.J., Zhou, P., Kamiya, S., Sun, B., Min, S.W., Gan, L., Finkbeiner, S., Huang, E.J., et al. (2012). Progranulin deficiency promotes neuroinflammation and neuron loss following toxin-induced injury. J. Clin. Invest. *122*, 3955–3959.

McMillan, C.T., Toledo, J.B., Avants, B.B., Cook, P.A., Wood, E.M., Suh, E., Irwin, D.J., Powers, J., Olm, C., Elman, L., et al. (2014). Genetic and neuroanatomic associations in sporadic frontotemporal lobar degeneration. Neurobiol. Aging *35*, 1473–1482.

Mosher, K.I., and Wyss-Coray, T. (2014). Microglial dysfunction in brain aging and Alzheimer's disease. Biochem. Pharmacol. *88*, 594–604.

Myers, A.J., Gibbs, J.R., Webster, J.A., Rohrer, K., Zhao, A., Marlowe, L., Kaleem, M., Leung, D., Bryden, L., Nath, P., et al. (2007). A survey of genetic human cortical gene expression. Nat. Genet. *39*, 1494–1499.

Nalls, M.A., Pankratz, N., Lill, C.M., Do, C.B., Hernandez, D.G., Saad, M., DeStefano, A.L., Kara, E., Bras, J., Sharma, M., et al. (2014). Large-scale meta-analysis of genome-wide association data identifies six new risk loci for Parkinson's disease. Nat. Genet. *46*, 989–993.

Newman, A.B., Glynn, N.W., Taylor, C.A., Sebastiani, P., Perls, T.T., Mayeux, R., Christensen, K., Zmuda, J.M., Barral, S., Lee, J.H., et al. (2011). Health and function of participants in the Long Life Family Study: a comparison with other cohorts. Aging (Albany NY) *3*, 63–76.

Ori, A., Toyama, B.H., Harris, M.S., Bock, T., Iskar, M., Bork, P., Ingolia, N.T., Hetzer, M.W., and Beck, M. (2015). Integrated transcriptome and proteome analyses reveal organ-specific proteome deterioration in old rats. Cell Syst. *1*, 224–237.

Pitt, J.N., and Kaeberlein, M. (2015). Why is aging conserved and what can we do about it? PLoS Biol. *13*, e1002131.

Pruim, R.J., Welch, R.P., Sanna, S., Teslovich, T.M., Chines, P.S., Gliedt, T.P., Boehnke, M., Abecasis, G.R., and Willer, C.J. (2010). LocusZoom: regional visualization of genome-wide association scan results. Bioinformatics *26*, 2336–2337.

Rademakers, R., Eriksen, J.L., Baker, M., Robinson, T., Ahmed, Z., Lincoln, S.J., Finch, N., Rutherford, N.J., Crook, R.J., Josephs, K.A., et al. (2008). Common variation in the miR-659 binding-site of GRN is a major risk factor for TDP43-positive frontotemporal dementia. Hum. Mol. Genet. *17*, 3631–3642.

Rollinson, S., Mead, S., Snowden, J., Richardson, A., Rohrer, J., Halliwell, N., Usher, S., Neary, D., Mann, D., Hardy, J., et al. (2011). Frontotemporal lobar degeneration genome wide association study replication confirms a risk locus shared with amyotrophic lateral sclerosis. Neurobiol. Aging *32*, 758.e1-7.

Salminen, A., Kaarniranta, K., and Kauppinen, A. (2012). Inflammaging: disturbed interplay between autophagy and inflammasomes. Aging (Albany NY) *4*, 166–175.

Schwenk, B.M., Lang, C.M., Hogl, S., Tahirovic, S., Orozco, D., Rentzsch, K., Lichtenthaler, S.F., Hoogenraad, C.C., Capell, A., Haass, C., et al. (2014). The FTLD risk factor TMEM106B and MAP6 control dendritic trafficking of lysosomes. EMBO J. *33*, 450–467.

Small, S.A., Schobel, S.A., Buxton, R.B., Witter, M.P., and Barnes, C.A. (2011). A pathophysiological framework of hippocampal dysfunction in ageing and disease. Nat. Rev. Neurosci. *12*, 585–601.

Stagi, M., Klein, Z.A., Gould, T.J., Bewersdorf, J., and Strittmatter, S.M. (2014). Lysosome size, motility and stress response regulated by fronto-temporal dementia modifier TMEM106B. Mol. Cell Neurosci. *61*, 226–240.

Tanaka, Y., Matsuwaki, T., Yamanouchi, K., and Nishihara, M. (2013). Exacerbated inflammatory responses related to activated microglia after traumatic brain injury in progranulin-deficient mice. Neuroscience *231*, 49–60.

van Blitterswijk, M., Mullen, B., Wojtas, A., Heckman, M.G., Diehl, N.N., Baker, M.C., DeJesus-Hernandez, M., Brown, P.H., Murray, M.E., Hsiung, G.Y., et al. (2014). Genetic modifiers in carriers of repeat expansions in the C9ORF72 gene. Mol. Neurodegener. *9*, 38.

Van Deerlin, V.M., Sleiman, P.M., Martinez-Lage, M., Chen-Plotkin, A., Wang, L.S., Graff-Radford, N.R., Dickson, D.W., Rademakers, R., Boeve, B.F., Grossman, M., et al. (2010). Common variants at 7p21 are associated with frontotemporal lobar degeneration with TDP-43 inclusions. Nat. Genet. *42*, 234–239.

van Es, M.A., Veldink, J.H., Saris, C.G., Blauw, H.M., van Vught, P.W., Birve, A., Lemmens, R., Schelhaas, H.J., Groen, E.J., Huisman, M.H., et al. (2009). Genome-wide association study identifies 19p13.3 (UNC13A) and 9p21.2 as susceptibility loci for sporadic amyotrophic lateral sclerosis. Nat. Genet. *41*, 1083–1087.

Webster, J.A., Gibbs, J.R., Clarke, J., Ray, M., Zhang, W., Holmans, P., Rohrer, K., Zhao, A., Marlowe, L., Kaleem, M., et al. (2009). Genetic control of human brain transcript expression in Alzheimer disease. Am. J. Hum. Genet. *84*, 445–458.

Willer, C.J., Li, Y., and Abecasis, G.R. (2010). METAL: fast and efficient meta-analysis of genome-wide association scans. Bioinformatics *26*, 2190–2191.

Yin, F., Banerjee, R., Thomas, B., Zhou, P., Qian, L., Jia, T., Ma, X., Ma, Y., Iadecola, C., Beal, M.F., et al. (2010). Exaggerated inflammation, impaired host defense, and neuropathology in progranulin-deficient mice. J. Exp. Med. *207*, 117–128.

Yu, L., De Jager, P.L., Yang, J., Trojanowski, J.Q., Bennett, D.A., and Schneider, J.A. (2015). The TMEM106B locus and TDP-43 pathology in older persons without FTLD. Neurology *84*, 927–934.

Zahn, J.M., Poosala, S., Owen, A.B., Ingram, D.K., Lustig, A., Carter, A., Weeraratna, A.T., Taub, D.D., Gorospe, M., Mazan-Mamczarz, K., et al. (2007). AGEMAP: a gene expression database for aging in mice. PLoS Genet. *3*, e201.

Zhang, B., Gaiteri, C., Bodea, L.G., Wang, Z., McElwee, J., Podtelezhnikov, A.A., Zhang, C., Xie, T., Tran, L., Dobrin, R., et al. (2013). Integrated systems approach identifies genetic nodes and networks in late-onset Alzheimer's disease. Cell *153*, 707–720.

Please cite this article in press as: Rhinn and Abeliovich, Differential Aging Analysis in Human Cerebral Cortex Identifies Variants in TMEM106B and GRN that Regulate Aging Phenotypes, Cell Systems (2017), http://dx.doi.org/10.1016/j.cels.2017.02.009

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited Data		
Tgen Expression	(Myers et al., 2007; Webster et al., 2009)	GEO: GSE15222
BrainEqtl Expression	(Gibbs et al., 2010)	GEO: GSE30272
BrainCloud Expression	(Colantuoni et al., 2011)	GEO: GSE30272
HTRC Expression and Genotypes	(Zhang et al., 2013)	Synapse: syn4505
ROS-MAP Expression	(Chan et al., 2015)	Synapse: syn3388564
Tgen Genotypes	(Myers et al., 2007; Webster et al., 2009)	NIAGADS: NG0028
ROS-MAP Genotypes	(Chan et al., 2015)	NIAGADS: NG0029
BrainEqtl Genotypes	(Gibbs et al., 2010)	dbGap: phs000249
BrainCloud Genotypes	(Colantuoni et al., 2011)	dbGap: phs000417
Dendritic cells expression	(Lee et al., 2014)	GEO: GSE57542
Software and Algorithms		
PLINK 1.9	(Chang et al., 2015)	https://www.cog-genomics.org/plink2
Metal	(Willer et al., 2010)	http://csg.sph.umich.edu/abecasis/metal/
LocusZoom	(Pruim et al., 2010)	http://locuszoom.org/
Impute 2	(Howie et al., 2009)	https://mathgen.stats.ox.ac.uk/impute/impute_v2.html
R WGCNA package	(Langfelder and Horvath, 2008)	https://cran.r-project.org/web/packages/WGCNA/index.html
Delta-aging R function	This paper	STAR Methods

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources should be directed to and will be fulfilled by the Lead Contact, Asa Abeliovich (aa900@columbia.edu).

METHOD DETAILS

Gene Expression Analysis in Human Brain

Studies described in this manuscript used de-identified data and were reviewed by the Columbia University Medical Center IRB. All data manipulations and analyses were done using R statistical software. Within each dataset independently of the others, expression level matrixes were log-centered before further manipulation for homogenization. The effect of chronological age on each probe/gene was assessed using R *lm()* function, with gender, batch and post-mortem interval as correlates. The effect of rs1990622 genotype was similarly studied, with age, gender, batch and post-mortem interval as correlates. The effects are studied in either all individuals (age >25 years old), in older adults (age >65 years old for all datasets except BrainCloud for which individuals with age >50 years old are included, as only n = 9 individuals are aged of more than 65 years) and in younger adults (age >25 years old and <65 years old for all datasets except BrainCloud for which individuals with age >25 years old and <50 years old are included), as indicated in legends. Meta-analysis across different datasets were carried at a gene-level using *Stouffer's* weighted Z-score method.

Δ-Aging Analysis

For each individual sample an associated Δ-aging numerical value, expressed in time units is evaluated. It corresponds to the aggregation of Δ-aging values evaluated for each gene found to be significantly correlated with age in the studied dataset (with FDR = 5%). For such a given gene G, the gene-specific Δ-aging value in a sample from individual I corresponds to the difference between the age as it would be imputed on the sole basis of gene G expression level in the studied sample, and the actual chronological age of I. Formally, it is expressed as $\Delta_{G,I} = \frac{\sigma_{G,I}}{a_G}$, where a_G is the coefficient of the linear regression analysis of gene G expression levels in function of age across samples of the studied dataset and $\sigma_{G,I}$ the residual expression level of gene G in individual I in the context of the above-mentioned linear regression. Integrating over the N genes with expression levels significantly correlated with age in the dataset of interest, the Δ-aging value for individual I is expressed as $\Delta_I = \frac{1}{N}\sum_{G=1}^{N}\frac{\sigma_{G,I}}{a_{G,I}}$. Detailed explanations are provided below.

Theory and Calculation of Delta-Age

Principle

At the level of a gene whose expression level is positively correlated with age within a given cohort, we define the Delta Age of a given individual as the difference between the actual (chronological) age and the biological age as is would be imputed for this individual on the basis of gene G expression level data across the entire cohort.

Figure S8 presents a theoretical case for illustration. Individuals, represented as dots, are plotted as a function of their chronological age (X-axis) and their measured expression level for gene G (Y-axis). The dotted line corresponds to the regression line of Gene G expression levels as a function of chronological age across the entire cohort. A graphical interpretation shows that dots that are above this regression line (in red) correspond to individuals with expression levels of G higher than expected for their chronological age, while those below (in blue) to individual with G levels lower than expected for their chronological age.

In Figure S8, the biological age imputed on the basis of gene G expression level is presented for 2 individuals by green arrows, corresponding to the projection on the age axis through the regression line. The yellow lines correspond to the chronological age, and the Delta-age represents the difference.

Formal Calculation of Delta-Age for a Given Individual and a Single Gene

The linear regression across individuals of the expression level of a gene G ($Expr_G$) in function of chronological age (*ChrAge*) yields a regression line defined by the following equation:

$$Expr_G = a_G.ChrAge + b_G \quad \text{(Equation 1)}$$

Where a_G and b_G are respectively the factor and constant associated to gene G expression levels in the linear regression in function of age. As graphically illustrated on Figure S1, for a given individual *I* the expression level of gene G, $Expr_{G,I}$ can be expressed by both expression:

$$Expr_{G,I} = a_G.ChrAge_I + b_G + \sigma_{G,I} \quad \text{(Equation 2a)}$$

where $\sigma_{G,I}$ is the residual expression level of gene G in individual I after the above-mentioned linear regression

and

$$Expr_{G,I} = a_G.AppAge_{G,I} + b_G \quad \text{(Equation 2b)}$$

where $AppAge_{G,I}$ is the apparent age of individual I for gene G.

As a consequence:

$$\Delta_{G,I} = \text{AppAge}_{G,I} - ChrAge_I = \frac{\sigma_{G,I}}{a_G} \quad \text{(Equation 3)}$$

The Delta-Age for an individual I for a given gene G is expressed as the ratio between the residual value for the individual and the coefficient obtained by linear regression of the expression level of gene G in function of Age across individuals.

Formal Calculation of Delta-Age across Multiple Genes

For a given individual I, the global Delta-age Δ_I is obtained by integration of all the gene-specific Delta-Age $\Delta_{G,I}$ over all genes which expression levels are found to be correlated with chronological age during the original linear regression:

$$\Delta_I = \frac{1}{N}\sum_{G=1}^{N}\Delta_{G,I} = \frac{1}{N}\sum_{G=1}^{N}\frac{\sigma_{G,I}}{a_{G,I}} \quad \text{(Equation 4)}$$

QUANTIFICATION AND STATISTICAL ANALYSIS

Genotype Association Analysis

Genotypes datasets were downloaded from dbGap (phs000249 and GSE30272 for *BrainEqtl* and *BrainCloud* respectively), NIAGADS (NG00029 and NG0028 for *ROS-MAP* and *Tgen*) or Synapse (phs000417 for *HBTRC*). All subsequent data manipulations and analyses were done using PLINK 1.9 software (https://www.cog-genomics.org/plink2, (Chang et al., 2015)). Meta-analyses were carried using Metal software (Willer et al., 2010). Manhattan plots were drawn using R qqman package or LocusZoom (Pruim et al., 2010). Genotype imputation at targeted loci was performed using Impute2 software (Howie et al., 2009) and 1000 Genomes Phase I integrated haplotypes. For cognition association analysis, genotypes and cognitive assessment phenotypes datasets were downloaded from dbGap (phs000397 and phs000428 for Long Life Family Study and Health and Retirement Study respectively). Association between genotype and cognitive scores were tested using plink with the following covariates: age, gender and 3 population eigenvectors as defined by PCA.

Gene Co-expression Analysis

Unsupervised clustering was carried using R *WGCNA* package (Langfelder and Horvath, 2008) with the following settings: power = 8, TOMType = "signed", minModuleSize = 40, reassignThreshold = 0.05, mergeCutHeight = 0.25. Enrichment analysis for the identified clusters was done using WGCNA R package built-in *userListEnrichment* function and associated brain cell types categories.

Human Dendritic Cell Analysis

Only individuals described as from European ascent were included in the analysis. The effects of treatment and rs1990622 on M1- or M2- related genes or on TMEM106B mRNA levels were studied by R by Kruskal-Wallis test or ANOVA, depending on the normality of the distribution and the similarity of variance between the groups as assessed using Shapiro and Levene tests respectively.

DATA AND SOFTWARE AVAILABILITY

The R annotated code for Δ-aging calculation and the custom dependent functions are provided below, with the R code being italicized:

##Delta Calculation using a gene expression matrix *ExprNumLog* with expression probes as rows, samples as columns and expression levels in a log-scale.
Age and Gender (or additional covariates) are provided as vectors corresponding to the organization of the samples in the expression matrix columns.
Coord_Temp defines which samples –defined by their column coordinates - are to be included in the analysis. By default, all samples are included:
Coord_Temp<-1:length(ExprNumLog[1,])
##The expression matrix is normed and centered. Its values correspond to $Expr_{G,I}$ in Equations 2a and 2b above.
Mtemp<-CentreNormeDoubleGenesSdInd(ExprNumLog[,Coord_Temp])
##A temporary matrix *Genes_ResidualsTemp* is created to store the residual values for each probe/sample pair after linear regression analysis for age and gender (and other potential experimental cofactors such as Pmi, Batch, …). This matrix has the same size and structure as *ExprNumLog.* Its values corresponds to the $\sigma_{G,I}$ in Equations 2a, 3 and 4 above.
Genes_ResidualsTemp<-matrix(0,length(Mtemp[,1]),length(Coord_Temp))
##3 matrices are created to store for each probe the effect of age and gender. Factor_AgeGender, Stat_AgeGender and Pval_AgeGender will respectively store in columns 1/2 the estimated coefficient, t-statistic and corresponding p-value of the association with age/gender for the expression level of each probe, as determned by R's *lm()* function summary. Those matrices have the same number of rows as *ExprNumLog* (1 per probe) and 2 columns (Age/Gender effects). The values stored in Factor_AgeGender correspond to the a_G in Equations 1, 2a, 2b, 3 and 4 above.
Pval_AgeGender<-matrix(NA,length(Mtemp[,1]),2)
Factor_AgeGender<-matrix(NA,length(Mtemp[,1]),2)
Stat_AgeGender<-matrix(NA,length(Mtemp[,1]),2)
##Each probe/row of the centered-normed expression matrix *Mtemp* is queried for the linear association of its levels with Age and Gender using R's *lm()* function. The results are stored in *Genes_ResidualsTemp*, *Pval_AgeGender*, *Stat_AgeGender* and *Factor_AgeGender.*
for (i in 1:length(Mtemp[,1]))
{
##The linear regression below corresponds to the one described in Equation 1 above, with the addition of Gender as a covariate. Other covariates could be added at this stage (Batch, Pmi…).
y<-lm(Mtemp[i,]~Age[Coord_Temp]+Gender[Coord_Temp],na.action=na.exclude)
Genes_ResidualsTemp[i,]<-residuals(y)
Pval_AgeGender[i,1]<-t(summary(y)[[4]][2:3,4])
Stat_AgeGender[i,1]<-t(summary(y)[[4]][2:3,3])
Factor_AgeGender[i,1]<-t(summary(y)[[4]][2:3,1])
}
##For each sample, the Deta-Age value is calculated using the residual expression levels after linear regression for age and gender. This corresponds to Equation 4 above. The genes included in the calculation are the one found to be significantly associated with age in the regression analysis. 3 false-discovery rate cut-off thresholds for the inclusion of genes are considered: fdr = 1%, 5% or 10%. Deta-Age values are calculated independently for each of them.
The expression *SeuilPos(1-p.adjust(Pval_AgeGender[,1], method = "fdr"),0.99)* is used to select the genes that are associated with age with fdr = 1%: it returns a vector which length equals the number of probes in the original expression matrix, with binary values of 1 if the probe's level is linearly correlated with age or 0 if not, for the considered significance level.
DeltaAge_Div_Age_Factors_FDR1pc<-as.numeric(t(t(Genes_ResidualsTemp)%*%(SeuilPos(1-p.adjust(Pval_AgeGender[,1], method = "fdr"),0.99)/Factor_Age[,1])/sum(SeuilPos(1-p.adjust(Pval_AgeGender[,1], method = "fdr"),0.99))))
DeltaAge_Div_Age_Factors_FDR5pc<-as.numeric(t(t(Genes_ResidualsTemp)%*%(SeuilPos(1-p.adjust(Pval_AgeGender [,1], method = "fdr"),0.95)/Factor_AgeGender[,1])/sum(SeuilPos(1-p.adjust(Pval_AgeGender[,1], method = "fdr"),0.95))))
DeltaAge_Div_Age_Factors_FDR10pc<-as.numeric(t(t(Genes_ResidualsTemp)%*%(SeuilPos(1-p.adjust(Pval_AgeGender[,1], method = "fdr"),0.9)/Factor_AgeGender[,1])/sum(SeuilPos(1-p.adjust(Pval_AgeGender[,1], method = "fdr"),0.9))))

Object such as *DeltaAge_Div_Age_Factors_FDR5pc* are the final output, being vector of length equal to the number of samples included in the analysis, containing the Differential-Aging values for each individual. Such values are later used as quantitative trait in genetic analysis.
##Delta Calculation from a gene expression matrix *ExprNumLin* with expression probes as rows, samples as columns and expression levels in a linear-scale. The process is similar to the one described into detail above, except for the creation of Mtemp at line #2, for which the function *CentreNormeDoubleGenesLogSdInd* is used instead of *CentreNormeDoubleGenesSdInd*, with the consequence to first log-transforming of the expression matrix before norming/centering it.
##NOTE: due to the log-transformation step, linear expression level matrix can contain only strictly positive values. Among the strategies to enforce this: 1)Rows/probes containing zero/negative values can be filtered out, 2)The whole expression matrix can be offset by its minimal value +1 (*ExprNumLin<- ExprNumLin+min(ExprNumLin)+1*), …

```
Coord_Temp<-1:length(ExprNumLin[1,])
Mtemp<-CentreNormeDoubleGenesLogSdInd(ExprNumLin[,Coord_Temp])
Genes_ResidualsTemp<-matrix(0,length(Mtemp[,1]),length(Coord_Temp))
Pval_AgeGender<-matrix(NA,length(Mtemp[,1]),2)
Factor_AgeGender<-matrix(NA,length(Mtemp[,1]),2)
Stat_AgeGender<-matrix(NA,length(Mtemp[,1]),2)
for (i in 1:length(Mtemp[,1]))
{
y<-lm(Mtemp[i,]~Age[Coord_Temp],na.action=na.exclude)
Genes_ResidualsTemp[i,]<-residuals(y)
Pval_AgeGender[i,1]<-t(summary(y)[[4]][2,4])Stat_AgeGender[i,1]<-t(summary(y)[[4]][2,3])
Factor_AgeGender[i,1]<-t(summary(y)[[4]][2,1])
}
DeltaAge_Div_Age_Factors_FDR1pc<-as.numeric(t(t(Genes_ResidualsTemp)%*%(SeuilPos(1-p.adjust(Pval_AgeGender[,1],
method = "fdr"),0.99)/Factor_AgeGender[,1])/sum(SeuilPos(1-p.adjust(Pval_AgeGender[,1], method = "fdr"),0.99))))
DeltaAge_Div_Age_Factors_FDR5pc<-as.numeric(t(t(Genes_ResidualsTemp)%*%(SeuilPos(1-p.adjust(Pval_AgeGender[,1],
method = "fdr"),0.95)/Factor_AgeGender[,1])/sum(SeuilPos(1-p.adjust(Pval_AgeGender[,1], method = "fdr"),0.95))))
DeltaAge_Div_Age_Factors_FDR10pc<-as.numeric(t(t(Genes_ResidualsTemp)%*%(SeuilPos(1-p.adjust(Pval_AgeGender[,1],
method = "fdr"),0.9)/Factor_AgeGender[,1])/sum(SeuilPos(1-p.adjust(Pval_AgeGender[,1], method = "fdr"),0.9))))
```

Custom Dependent Functions

```
## CentreNormeDoubleGenesSdInd
CentreNormeDoubleGenesSdInd<-function (Matrix)
{
Mout <- Matrix
for (i in 1:length(Matrix[, 1])) {
A <- sum(Mout[i, ])/length(Mout[i, ])
Mout[i, ] <- (Mout[i, ] - A)
}
for (j in 1:length(Matrix[1, ])) {
B <- sum(Mout[,j])/length(Mout[,j])
Mout[,j] <- (Mout[,j] - B)/sd(Mout[,j])
}
Mout
}
## CentreNormeDoubleGenesLogSdInd
CentreNormeDoubleGenesLogSdInd<-function (Matrix)
{
Mout <- Matrix
for (i in 1:length(Matrix[, 1])) {
Mout[i, ] <- log((Matrix[i, ]), 2)
A <- sum(Mout[i, ])/length(Mout[i, ])
Mout[i, ] <- (Mout[i, ] - A)
}
for (j in 1:length(Matrix[1, ])) {
B <- sum(Mout[,j])/length(Mout[,j])
Mout[,j] <- (Mout[,j] - B)/sd(Mout[,j])
}
```

```
Mout
}
##SeuilPos
SeuilPos<-function(x,s)
{m<-max(x,na.rm=TRUE)
Pos(ceiling((x-s)/m))}
##Pos
Pos<-function(x)
{(x+abs(x))/2}
```