EVOLUTIONARILY SELECTED AND DIFFERENTIALLY EXPRESSED *CIS*-REGULATORY ELEMENTS AND THEIR RELATION TO HUMAN AGING

An Honors Thesis Presented
by

LILY HILTEBEITEL

Approved By:

Committee Chair: Dr. Courtney Babbitt

Committee Member: Dr. Margaret Riley

May 2022
Department of Biology
University of Massachusetts Amherst

ACKNOWLEDGEMENTS

Throughout this thesis, I have received a great deal of support and assistance! I'd like to thank my supervisor and principal investigator, Dr. Courtney Babbitt, for her guidance and feedback. This project would not have been possible without her expertise. In addition, thank you to members of the Babbitt lab for support. I would also like to thank my committee member, Dr. Margaret Riley, for her support and feedback.

I would also like to acknowledge the iCons Program, particularly Dr. Scott Auerbach and my peers, for support throughout the process. Feedback received was invaluable, and this thesis benefitted from our in-depth and insightful discussions.

TABLE OF CONTENTS

ABSTRACT

Noncoding *cis*-regulatory elements (CREs) of our genome are hypothesized to be a heavy contributor to evolved human phenotypes. However, the complex interplay of CRE functions across genes remains elusive, particularly in regards to aging of the human brain. We have not been able to attribute all evolutionary changes in the brain to coding mutations alone, so regulatory elements are a prime candidate of investigation for causes of evolutionary changes. In this study, datasets of two types of CREs were examined: those with high activity in the human brain, and those that have experienced recent positive selection in the genome when compared to non-human primates. In addition, known brain aging-related genes were analyzed against these CREs in order to detect signs of CREs regulating these aging-related genes. If evolutionarily selected CREs regulate aging genes, then these CREs and their methods of gene regulation may be unique to humans. We identified AKT3, CDH13, CDKAL1, CSMD1, UNC5D, and WWOX as genes regulated by selected and/or differentially expressed CREs. These genes are orthologous in non-human primates, so while the genes have been selectively retained in all primates studied, the regulation appears to be species specific. This provides insight into how our noncoding regulatory elements, such as enhancers or promoters, are linked to uniquely human aging processes. Because age is a major risk factor for neurodegenerative diseases, such as Alzheimer's disease, we also gain a better understanding of regulatory factors that could be contributing to neurodegeneration.

INTRODUCTION

While the neural phenotypes of humans and chimpanzees clearly differ, their genomes are highly similar in sequence. Because they are so similar, it is likely that many of the more dramatic phenotypic differences are caused by changes in gene expression rather than changes in protein-coding regions. This was first hypothesized decades ago (King & Wilson, 1975), but with the -increased availability of whole genome data across thousands of humans and numerous chimps, it is now more possible than ever to analyze regulatory noncoding elements, such as promoters and enhancers, and connect them to human-specific phenotypic changes. Not only are regulatory regions a larger part of the genome than coding regions (Franchini & Pollard, 2017), but noncoding regulatory regions tend to have more modulatory functions compared to coding regions (Carroll, 2008). The 'omnigenic model' of genetics hypothesizes that most phenotypes arise from the combination of thousands of gene regulatory networks, rather than a change in a particular coding region (Boyle et al., 2017). Based on these ideas, it is likely that noncoding regulatory regions have been a target of evolution and that changes in these regions have helped drive unique human phenotype (Franchini & Pollard, 2017).

The lineages leading to humans and chimpanzees diverged approximately 7 million years ago (Amster & Sella, 2016). Many obvious phenotypic differences now separate these two species, but those related to neurological functions are most prominent, and thought to be driven by neocortical expansion (Rilling, 2014). Brain expansion in humans, in addition to a higher number of cortical neurons are correlated with higher cognitive abilities (Roth & Dicke, 2005).

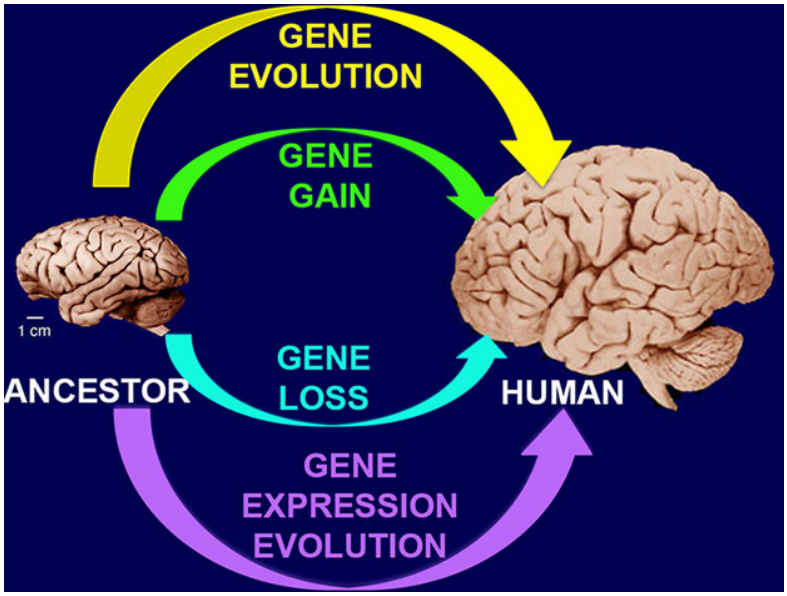


Figure 1. Chimpanzee brain (left) and human brain (right) with possible genetic evolution explanations (Franchini & Pollard, 2017).

A recent study in the Babbitt Lab compared the characteristics of *cis*-regulatory elements (CREs) in both humans and chimpanzee genomes (Pizzollo et al., 2019). A CRE is a noncoding sequence that regulates expression of a gene by binding to proteins that initiate and modulate transcription, such as transcription factors (Hernandez-Garcia & Finer, 2014). In the study, two types of sequences were focused on in both humans and chimps: highly conserved noncoding regions that have high substitution rates throughout evolutionary history, and enhancer sequences that are active in human brain development. This resulted in the determination of the number of CREs experiencing positive selection, how many are differentially expressed in humans and chimpanzees, and the overlap of CREs that have both of these characteristics (Pizzollo et al., 2019). By investigating these two categories of sequences, we can better understand which active, brain-related regulatory regions have been selected for throughout evolution. This thesis project furthers this investigation with a focus on brain aging.

Aging is a major risk factor for neurodegenerative disorders, so studying genomic components of aging also facilitates further understanding of neurodegenerative diseases. However, investigating the role of regulatory non-coding sequences in aging and resulting neurological disorders poses challenges. High-throughput genomic sequencing technology has resulted in many recent revelations linking genetic variants to various cognitive diseases, such as Alzheimer's disease. However, another discovery from the genomic sequencing revolution is that these disorders are genetically complex and cannot be traced to single mutations (Berto et al., 2020; Schaub et al., 2012). Mimicking neurological diseases in model systems is also difficult because of this (Berto et al., 2020). Genome-wide association studies (GWAS) show correlation between neurological disorders and genes associated with cognitive function, inspiring a hypothesis that these disorders could be an underlying consequence to human brain evolution (Berto et al., 2020). In addition, the complex relationship between *cis*-regulatory elements and hallmarks of human aging is not yet unraveled. While it is known that they are intertwined, we do not yet have a detailed map of which CRE changes result in which neural-related, aging phenotypes.

This thesis project builds off of this knowledge to investigate the cause-effect relationship between aging-related genes and the CREs identified by the Babbitt lab. By doing this, we can further understand the role of evolution of noncoding regions in human aging. In the following review, which precedes the methods and results of the thesis, I will describe selected CREs, differentially expressed CREs, and how the human genome and aging are related.

LITERATURE REVIEW

Conserved and Selected *Cis*-Regulatory Elements

Multiple studies have characterized selected, human accelerated noncoding regions. These regions are considered evolutionarily selected due to a couple of parameters. First, noncoding sequences that are conserved among vertebrates across millions of years are focused on because this implies that, generally, they are kept because they are functional (Pizzollo et al., 2019). Second, these regions have high substitution rates in the human genome when compared to the genomes of other primate species (Haygood et al., 2007; Lindblad-Toh et al., 2011). Combined, these parameters help sift through vast amounts of non-coding sequences to find ones that are most likely to be both functional and accelerated. Another perspective of this is to look for areas of the genome that have experienced negative selection across vertebrates, then recent positive selection when comparing the human genome with other primate species (Pollard, Salama, King, et al., 2006). It is important to note that the results of the following characterization studies may be slightly different due to different methods of statistical parameters. When considered together, we can begin to construct a map of the human genome and its regulatory elements.

An early study identified 202 elements, mostly noncoding, that were both conserved and contained high human acceleration rates (Pollard, Salama, King, et al., 2006). In order to discover this, researchers first recognized any increase in substitution rate, then used likelihood ratio tests (LRTs) on these regions. LRTs for a genomic sequence require a molecular evolutionary model of that sequence of interest, constructed using known data. Then, a ratio is calculated of the likelihood of the model with acceleration against the model with no substitution

acceleration (Pollard, Salama, King, et al., 2006). Using this approach, they found 202 significant human accelerated regions (HARs), majority of which enriched outside coding genes (66.3% intergenic regions, 31.7% intronic regions) and were in areas enriched with transcription factors. The top five of the most significant HARs had significant bias for AT to GC substitutions and overwhelmingly were in high recombination areas (Pollard, Salama, King, et al., 2006). The most significant HAR, referred to as HAR1, was found to be part of a gene that is expressed in Cajal-Retzius neurons in a critical period of neocortical development, migration and differentiation (Pollard, Salama, Lambert, et al., 2006). One of the first major studies of accelerated regions using comparative genomics, this study provided evidence to help confirm King and Wilson's 1975 hypothesis of *noncoding evolution driving human phenotypes*.

Another study in the same year also examined accelerated evolution of conserved noncoding regions. 992 conserved noncoding sequences (CNSs) were identified by looking for an excess of human specific mutations, with a p value < 0.005 (Prabhakar et al., 2006). Using gene ontology annotations of other primates, the CNSs were significantly enriched near neuronal cell adhesion genes and other structural neuron genes. When doing the same analysis using the chimpanzee genome, a very similar neuronal cell adhesion enrichment pattern was observed, but only 34 of the 1050 CNSs in chimpanzees overlapped with the human CNSs (Prabhakar et al., 2006). When analyzing the mouse genome, enrichment near neuronal cell adhesion genes was not found. This surprising discovery suggests that humans evolved their neuronal phenotypes independently and that evolution of these CNSs were likely major contributors of unique human brain development and function (Prabhakar et al., 2006).

A 2011 study, rather than focusing on only larger elements (>100 bp) like the previous studies, also took into account smaller conserved elements using 29 mammal species for

comparison (Lindblad-Toh et al., 2011). Using a SiPhy measure of overall substitution rate (Garber et al., 2009), they approximated 5.3-5.5% of the human genome to be under selection, and significantly in noncoding areas. Only around 60% of these regions could be annotated with known coding or regulatory roles (Lindblad-Toh et al., 2011). Much of the CREs showed overlap with known disease-related genetic variants (Lindblad-Toh et al., 2011), marking one of the first noncoding evolution studies that also investigated a role in disease.

In preparation for a study comparing both conserved/selected and differentially expressed neural enhancers, researchers of the Babbitt lab (Pizzollo et al., 2019) used 1579 known evolutionarily accelerated CREs (Lindblad-Toh et al., 2011; Pollard, Salama, King, et al., 2006; Prabhakar et al., 2006) in their library. They performed their own positive selection test, which calculated the rate of nucleotide substitution of the non-coding sequence (aligned between human, chimpanzee, and a macaque outgroup) compared to a neutral substitution rate (Haygood et al., 2007; Pizzollo et al., 2019). This resulted in their own characterization of 722 positively selected CREs. All 1579 literature-identified CREs were then analyzed along with a set of differentially expressed CREs (particularly enhancers) in order to gain insight into which CREs have both experienced recent evolution and have high brain activity.

Differentially Expressed CREs and Their Relationship to Selected CREs

The higher cognition of humans is thought to be largely due to changes in gene expression regulation rather than just changes in protein-coding genes. *Cis*-regulatory elements, such as enhancers or promoters, are defined as the noncoding regions that *trans*-factors, such as transcription factors, bind to. Enhancers typically regulate transcription in regions independent of distance to the target gene (Pizzollo et al., 2019), which separates them from other classes of

CREs. While transcription factors are vital for regulation, enhancers are largely responsible for specifically where and when gene expression occurs (Long et al., 2016). When the local chromatin state of enhancer sequences (or other CREs) is loosened or restricted, gene expression rate can increase or decrease. These epigenetic modifications are critical during brain development (Vermunt et al., 2014). Due to recent large-scale sequencing, CREs have been relatively straightforward to find in our genome because they have specific footprints to look for, such as occupying open chromatin (Vermunt & Creyghton, 2016). However, the web of CREs and its role in what, where, and when gene activity occurs in our brain is extremely complex. Untangling it will give insights into how our genomes lead to human brain function.

The recent project in the Babbitt lab (Pizzollo et al., 2019) studied human differentially expressed (DE) CRE sequences with specific chromatin marks, against their chimpanzee orthologs. These chromatin marks are H3K27ac and H3K4me2, previously annotated to be associated with ChIP-seq-defined high activity in the human brain (Reilly et al., 2015; Vermunt et al., 2014). Chromatin immunoprecipitation followed by sequencing (ChIP-seq) is generally the standard for measuring regulatory region interactions and activity. Chromatin immunoprecipitation (ChIP) identifies the DNA fragments that are attached to proteins of interest, such as a transcription factor or acetylated histones, using an antibody and a specific cell type (Jiang & Mortazavi, 2018). Because the proteins of interest will typically bind if chromatin is open, results of ChIP give insight into which DNA regions are expressed in brain cells (Pizzollo et al., 2019). After ChIP, sequencing is done in order to identify the genetic regions.

The Babbitt lab's library of 695 CREs was narrowed down to 179 CREs that have differential activity in the human brain compared to chimpanzee orthologs. This was done using a lentivirus-based massively parallel reporter assay (lentiMPRA) in neurons and induced

pluripotent stem cell-derived neural progenitor cells (iPSC-derived NPCs). MPRA is a broad method that can analyze a large amount of regulatory elements in just one assay and identify their activity by seeing which elements activate transcription in the NPCs and neurons (Inoue et al., 2017; Pizzollo et al., 2019). Then, using the gene expression software package limma, differential expression was able to be compared against orthologous CREs, and the 179 differentially expressed CREs were identified (Pizzollo et al., 2019).

Of the 722 positive selection CREs and the 179 CREs with differential activity, only 34 overlapped. In addition, the features of these genomic regions differed. On average, selection CREs had a higher amount of transcription factor gains, and differentially expressed (DE) CREs had a more neutral balance of gains and losses. In addition, selection CREs were further from genes and other enhancers than DE CREs, and they are generally shorter and have less GC content. Despite being shorter, selection CREs also tended to be a part of large enhancer regions, more so than differentially expressed CREs are (Pizzollo et al., 2019). Because the human brain has such higher cognition when compared to other primates, some may expect selection CREs and DE CREs to have a large overlap. This analysis tells us that the two sets actually have quite a few different characteristics and therefore may function differently. It also suggests that human-specific selected CRE evolution contains small, fine-tuned changes—however, it doesn't change large enhancer regions that have important and preserved ancestral functions. Because neurally differentially expressed CREs did not appear to be as highly conserved, differentially expressed CRE evolution may be less constrained and more recently evolved in history than selected CREs (Pizzollo et al., 2019).

Aging-Related Cognitive Decline and Neural Phenotypes

In humans, neural structural changes such as volume loss or neuronal loss correlate tightly with cognitive decline that comes with aging. However, it is not a clear cut, causal relationship. Sometimes, changes in gene expression occur before these structural changes, and possibly underlie them or increase their chances of materializing (Ham & Lee, 2020). For this reason, transcriptome analysis of gene expression and association of regulatory regions with aging-related genes is critical. It can help us further understand causes of structural changes in the brain, cognitive decline, and neurodegenerative diseases.

Executive function describes a set of higher-order cognitive processes that underlie brain functions that we often take for granted: planning, working memory, focus, among others. A decline in these processes comes with aging, and is associated with structural changes in the brain cortices (Lacreuse et al., 2020). To some extent, executive function worsens with age in all primate species. However, in humans it is still unclear what the nature of the relationship is between executive function decline and cortical alterations during aging. These two changes are also correlated with higher risk of neurodegenerative diseases, such as Alzheimer's disease and Parkinson's disease, but the underlying genomic variables of this are also unknown (Lacreuse et al., 2020). *Could higher risk of neurodegenerative diseases be a genetic consequence of the evolution of higher cognition in humans?*

In chimpanzees, the closest relative of humans, executive function has been shown to decline with age as it does in humans (Lacreuse et al., 2020; Manrique & Call, 2015). However, studies of neural changes in chimpanzees are less conclusive. Multiple studies have found little evidence for significant white or gray matter volume loss in chimpanzee brains, while in humans, the relationship between aging and brain volume loss is much more linear (Autrey et al., 2014).

While current research has proven that chimpanzees experience executive function decline with aging, this decline cannot be as directly linked to structural neural changes as it is in humans. These observations provide further evidence that the significant structural aging-related changes in the human brain could be correlated with the genomic evolution.

The role of regulatory elements and gene expression in aging and neurodegenerative diseases is complex. Many transcriptome studies focus on aging-related differential expression patterns in different age groups and brain regions (Işıldak et al., 2020; Mori, 2020). Overall, gene expression heterogeneity (or, shared expression patterns) seems to increase with age (Işıldak et al., 2020). Evidence also shows that neurodegeneration can result from lowered transcription and translation accuracy, particularly in the REST transcription factor and NRSF which are master regulators in neural development (Mori, 2020).

In addition to characterizing DE and selected CREs, the Babbitt lab also connected the CREs they identified to neural phenotypes (Pizzollo et al., 2019). Genome-wide association studies (GWAS) of single nucleotide polymorphisms (SNPs) often are located in non-coding genomic regions, and these SNPs are often in areas where transcription factors would normally bind. The Babbitt lab found that annotated neuropsychiatric SNPs were enriched in their differentially expressed CREs. These neuropsychiatric SNPs are associated with Alzheimer's disease, schizophrenia, depression, ADHD, autism, and bipolar disorder; these diseases all involve cognitive alterations (Buniello et al., 2019). A GWAS locus? is defined as the genetic region around the gene that shares at least one regulatory element and one SNP. Pizzollo et al. (2019) found that 69 of their CREs were associated with a GWAS SNP. The majority of these were differentially expressed CREs rather than selected CREs. One aging-related locus of note contains the FAT1 gene, CRE.2063, and an Alzheimer's disease associated SNP rs75718659

(Pizzollo et al., 2019). This discovery suggests that the differentially expressed CRE.2063 could be associated with Alzheimer's disease, which aging is a monumental risk factor for. Overall, these findings show that mutated noncoding regions can have an impact on neural phenotype, despite the enhancers being distal from the target genes. If CREs were impacted in human evolution, then it is possible that gene expression was also impacted. This likely has an effect on both neural structural changes that are hallmarks of aging, as well as likelihood of neurodegenerative diseases.

Current State of the Field

Humans have many unique phenotypic features compared to other primates, including neocortical expansion and higher cognitive abilities. While it is hypothesized that many of these features have evolved due to changes in non-coding regulatory regions rather than protein-coding regions, the exact role of regulatory regions remains unclear. Identifying the roles of *cis*-regulatory elements has two primary components: how these elements evolved, and where they act in human processes. These components can be investigated by (1) Analyzing human and other primate genomes to see which CREs are both conserved across primates and have undergone positive selection in recent history and (2) Using experimental techniques to see which CREs have differential activity in neural cells. Linking the impact of CREs to their target genes is difficult, due to enhancers acting from distal regions and often having many targets (Pizzollo et al., 2019). However, successfully linking them through further data analysis has many implications. If the functions of specific CRE regions are identified, then we will have a more in depth understanding of what their causal roles are in neural phenotypes, aging, and neurodegenerative disease risk. More broadly: we can understand genomically what makes us

human, and why we develop aging related diseases. My thesis research will aim to answer the following questions building off of work by the Babbitt lab (Pizzollo et al., 2019): Where do aging-related genes lie in datasets of evolutionarily selected and differentially expressed *cis*-regulatory elements? Do aging-related genes appear to be more selected in humans or more differentially expressed, or both? Are CREs close or distal to these genes? What links do the CREs have with genes linked to Alzheimer's disease or other neurodegenerative diseases? Are orthologous aging-related genes found in chimpanzees or other primates? By further analyzing selected and DE CREs with these questions in mind, a further understanding can be developed of how human-specific aging has evolved and how it is manifested in the human brain.

METHODS

Cis-Regulatory Element (CRE) Data

CRE data were collected from a previous project in the Babbitt lab that investigated both selected and differentially expressed neural CREs (Pizzollo et al., 2019). Researchers began with 1579 known evolutionarily accelerated CREs (Haygood et al., 2007; Lindblad-Toh et al., 2011; Pollard, Salama, King, et al., 2006; Prabhakar et al., 2006) conserved throughout vertebrate evolution, then tested for the signature of positive selection. This test calculated the rate of nucleotide substitution of the noncoding sequence (aligned between human, chimpanzee, and a macaque outgroup) compared to a neutral substitution rate. From this test, they characterized 722 positively selected CREs.

In addition, they also characterized a set of CREs that had high brain activity in humans. Similarly to the selected CRE data, they began with 695 CREs that were previously annotated to

be associated with ChIP-seq-defined high brain activity (Reilly et al., 2015; Vermunt et al., 2014). A lentivirus-based massively parallel reporter assay (lentiMPRA) was then done using neurons and iPSC-derived neural progenitor cells (NPCs) in order to characterize their own CRE set with differential brain activity. The R package limma was then used to compare differential expression data against orthologous CREs. 179 differentially expressed CREs were identified. 34 CREs overlapped between the positive selection dataset and the differential expression dataset. An overview of this selection process can be seen in Figure 2.

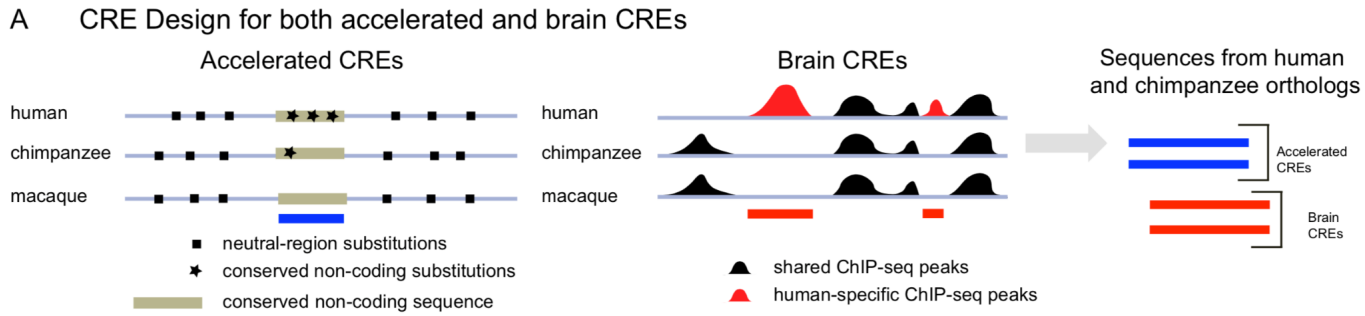


Figure 2. Visual overview of the sources of both evolutionarily accelerated and brain-active CREs. Figure taken from Pizzollo et al., 2019. Accelerated CREs were identified by computationally comparing substitution rates in conserved noncoding regions to neutral surrounding regions. Brain CREs were identified by comparing human-specific ChIP-seq peaks to ChIP-seq peaks in chimpanzees and macaques (Pizzollo et al., 2019).

Aging Genes

Aging genes were chosen through review of the literature and are represented below in Table 1.

Table 1. Chosen aging genes and a description of their function and link to aging. Many aging genes were found using GenAge Database (Tacutu et al., 2018).

Gene ID	**Aging Link**
ADCY5	Inhibition of ADCY5 experimentally results in longer lifespan, with suspected mechanism of oxidative stress protection (Vatner et al., 2015).
ADIPOQ	Evidenced as a core frailty biomarker, particularly in cytoskeleton and hormone function (Cardoso et al., 2018).
AKT1	Human longevity is shown to be associated with genetic variation, based on GWAS analysis (Deelen et al., 2013).
AKT3	Human longevity is shown to be associated with genetic variation, based on GWAS analysis (Deelen et al., 2013).
APOC1	Recent studies have directly tied APOC1 variants to memory impairment and increased risk in k of Alzheimer's disease (AD) (Kulminski et al., 2022).
APOC3	APOC3 is shown to be tied to longevity via involvement in lipoprotein metabolism (Atzmon et al., 2006).
APOE	Locus is widely established as a contributor to longevity, with variants being associated with increased AD risk (Benitez et al., 2013; Deelen et al., 2013; Kulminski et al., 2022)**.**
APP	Amyloid precursor proteins (APP) are directly tied to AD and age-related cognitive impairment (Burrinha et al., 2021).
ARHGAP1	As a regulator of CDC42, mutations result in shorter average lifespans (L. Wang et al., 2007).
ATR	Involved in DNA repair as a checkpoint regulator, particularly regulation of WRN and TP53; deletion leads to loss of stem cells and premature aging related phenotypes (Ruzankina et al., 2007).
BCL2	Tied to longevity of dendritic cells; inhibition of BCL2 promotes longevity (Fernández et al., 2018).
BUB3	With mitotic spindle checkpoint functions, BUB3 insufficiency is tied to early aging and reduced lifespan (Baker et al., 2006, p. 3).
CDH13	Associated with prevention of age-related tissue loss (Yang et al., 2020), as well as implicated in causation of age-related disorders (ARDs) based on GWAS and additional pathway analysis.

CDK7	Maintains homeostasis of tissues; inactivation leads to premature aging, reduced lifespan, and telomere shortening (Ganuza et al., 2012).
CDKN2A	GWAS analysis reveals a thorough association with aging phenotypes and lifespan length (Jeck et al., 2012).
CDKN2B	GWAS analysis reveals a thorough association with aging phenotypes and lifespan length (Jeck et al., 2012).
CDKAL1	Implicated in causation of ARDs based on GWAS and additional pathway analysis (Srivastava et al., 2015).
CETP	Associated with longer human lifespans and is linked to AD susceptibility (Yu et al., 2012).
CSMD1	Associated with aging-related memory and cognitive functions through synaptic plasticity (Athanasiu et al., 2017)**.**
DGAT1	Codes for a lipid synthesis enzyme that is associated with longevity, particularly longevity extension (Streeper et al., 2012).
EP300	Transcriptional coactivator which stimulates WRN, mutation of which causes Werner syndrome, a disorder that causes premature aging phenotypes (Blander et al., 2002).
ERCC1	Encodes an important DNA repair protein and is tied to aging of the immune system; evidence shows that an aged immune system has a causal role in systemic aging (Yousefzadeh et al., 2021).
ERCC2	Also primarily involved in DNA repair as well as transcription, studies have shown that its protein affects human lifespan and longevity (Savina et al., 2018).
ERCC8	Involved in transcription and in repair of damaged DNA; when mutated, results in premature aging phenotype (Paccosi & Proietti-De-Santis, 2021)**.**
FOXO1A	GWAS studies show that SNPs of FOXO1A are linked to longevity (Li et al., 2009).
FOXO3A	GWAS studies show that SNPs of FOXO3A are linked to longevity (Li et al., 2009).
GDF11	Factor that is secreted in the TGFß cytokine family, and secretion declines with age. Suspected to be involved in aging of skeletal muscle and skin (Egerman & Glass, 2019).
GPX1	SNPs are associated with aging-related memory decline and ARD development (da Rocha et al., 2018).

IL6	Linked to AD and aging-related inflammation (Cardoso et al., 2018; Greene & Loeser, 2015)**.**
JUND	Part of AP-1 redox-regulated transcription process, playing a role in age-related changes. Associated with premature aging phenotypes and aging of human fibroblasts (Paneni et al., 2013).
KCNA3	Potassium voltage-gated channel that is involved in ARDs (Ota et al., 2011).
KL	Evidenced as a core frailty biomarker, particularly in cytoskeleton and hormones; alleles are implicated in longevity (Cardoso et al., 2018)
LMNA	Mutation causes Hutchinson-Gilford's progeroid syndrome, with an accelerated aging phenotype; in many studies, mutants result in premature aging (Gonzalo et al., 2017)**.**
MAT1A	SNPs show association with human longevity (Luciano et al., 2012).
MAT2A	SNPs show association with human longevity (Luciano et al., 2012).
MTOR	Widely established as a regulator of longevity, lifespan, and aging-related processes like metabolism and cellular senescence (Weichhart, 2018).
NFKB1	Associated with lifespan and age-related inflammatory changes (Lorenzini et al., 2020).
NUDT1	Degrades oxidative damage to both RNA and DNA, linked to longevity and lifespan (Garre et al., 2011).
PAPPA	Has an important role in fetal development and pregnancy, but is also evidenced to play a role in longevity and risk development of ARD (Conover, 2013).
PIK3CB	Linked to both longevity and plasma IGF1 levels in humans (Elzinga et al., 2019).
PLAU	Evidenced to extend longevity through metabolic functions; variants linked to ARDs in humans (Cardoso et al., 2018; Gaweda-Walerych et al., 2021).
PLD3	Regulates late-stage neurogenesis and processing of amyloid precursor proteins; variants increase risk of ARD in humans, particularly AD (J. Wang et al., 2015).
POLG	Deficiency results in premature aging phenotype, associated with ARD risk (Cohen et al., 1993).
PON1	Affects longevity through suspected atheroprotective effect (Lescai et al., 2009).

POU1F1	Associated with lifespan, regulates pituitary development and hormone expression (Jadhav et al., 2021).
PPARGC1A	Associated with ARDs, evidenced to accelerate vascular aging and atherosclerosis through metabolic and inflammatory regulation (Liang & Ward, 2006).
RAE1	Along with BUB3, insufficiency is tied to early aging and reduced lifespan (Baker et al., 2006, p. 3).
SHC1	A splice variant of SHC1, p66, is associated with lifespan, cognitive performance, and vascular aging (Paneni & Cosentino, 2012).
SIRT6	Strongly evidenced to promote longevity and protect against ARDs through functions in chromatin signaling and genome maintenance (Tasselli et al., 2017).
SOD2	Overexpression associated with increased lifespan and longevity; polymorphisms associated with longevity and ARD development (El Assar et al., 2013).
TERC	Critical in cellular senescence and telomere elongation; mutations resulting in telomere shortening may cause accelerated aging (Grill & Nandakumar, 2021).
TERT	Critical in telomere elongation; mutations resulting in telomere shortening may cause accelerated aging (Grill & Nandakumar, 2021).
TNFA	An immune response cytokine, TNFA, is associated with inflammatory aspects of aging (Clarke et al., 2018).
TP53	Through involvement in DNA repair, cell cycle regulation, and apoptosis, TP3 is associated with lifespan and longevity (Wu & Prives, 2018).
TP63	Regulates proliferation and differentiation, and is associated with premature aging and lifespan (Holder-Espinasse et al., 2007).
TREM2	Receptor expressed on myeloid cells that regulates inflammation, cell death, and synaptic pruning, associated with ARDs (Mecca et al., 2018).
UNC5D	Promotes neuronal cell survival, as well as cell migration and cell-cell adhesion (Takemoto et al., 2011).
WWOX	Plays a role in apoptosis and connected to ARD risk, particularly dementia development (Dugan et al., 2022; Kunkle et al., 2019).
XRCC5	Involved in DNA repair and chromatin remodeling that has been associated with premature aging (Gennaccaro et al., 2021).

Functional Profiling of Aging Genes

Functional enrichment analysis of the 59 aging-related genes was performed using g:Profiler's g:GOSt package in RStudio (Kolberg et al., 2020). Using statistical enrichment analysis, GOSt finds significant functional information for genes based on known Gene Ontology terms, biological pathways, regulatory DNA elements, human disease gene annotations, and protein-protein interaction networks (Raudvere et al., 2019). In addition to functionally profiling the full list of 59 aging-related genes, functional profiling was also performed for the 6 aging-related genes (that are a part of the full list) that are regulated by selected and/or differentially expressed CREs. For both lists, Benjamini-Hochberg FDR correction method, with a p-value threshold of 0.05, was used in order to decrease the chance of false positive results occurring.

Finding Intersection Between CREs and Aging Genes

RStudio and UCSC Genome Browser were the primary tools used to identify if any of the 901 total CREs regulate any of the 59 total aging-related genes. The CRE dataset was imported into R, containing the following information: CRE ID, chromosome, start coordinate location, stop coordinate location, and the gene that it is annotated to regulate. CRE-gene assignment was performed using GREAT (McLean et al. 2010) with the "basal plus extension" method (Pizzollo et al., 2019), which defines regulatory domains for genes and their CRE annotations automatically.

To attain the genomic locations of the aging-related genes, we entered the gene IDs into USCS Genome Table Browser. Resulting output was a BED (Browser Extensible Data) file,

which includes the following information: chromosome, start coordinate location, stop coordinate location, name, score (read depth), strand orientation, thick start coordinate (typically where the exon starts), thick end coordinate, Rgb value, number of blocks (such as exons), sizes of these blocks, and start coordinates of these blocks. The BED file was imported into RStudio, aligned with the CRE dataset, then cleaned up using the Tidyverse package. In order to calculate the distance between the CRE and the gene that it regulates, a new column ("distance") was created by subtracting the CRE start coordinate location from the gene start coordinate location and taking the absolute value.

Interested readers can view the RStudio code at https://github.com/lilyhilt.

Transcription Factor Enrichment

Cis-regulatory element activity often occurs when transcription factors bind to motifs, causing more transcriptional machinery to bind to promoters and induce gene transcription (Pizzollo et al., 2019). Because of this interaction, we investigated motif enrichment of the three selected aging genes (CDKAL1: 2 sequences; CSMD1: 4 sequences; UNC5D: 6 sequences). This was performed using Meme Suite's Simple Enrichment Analysis (SEA) package (Bailey & Grant, 2021) by attaining each splicing exon variant sequence using NCBI, inputting the sequences into SEA, and scanning for motifs based off of their JASPAR database of 1227 vertebrate motifs with E less than or equal to 10.

Orthologous Gene Search

We utilized NCBI (National Center for Biotechnology Information) to search for orthologous genes of the 6 aging-related genes regulated by selected/DE CREs. Each of the 6

genes were searched in the ortholog database, then narrowed down from jawed vertebrates to placental primates in order to compare with the closest human ancestors. NCBI calculates these ortholog gene groups using their Eukaryotic Genome Annotation Pipeline. This calculation uses a combination of protein sequence similarity and local synteny, which is measured by homology of neighboring genes. Specifically among placental primates, *Homo sapiens* (humans), *Pan troglodytes* (chimpanzees), and *Macaca mulatta* (Rhesus monkey) were examined. Presence of the gene, amino acid length, sequence alignment, and gene architecture were noted.

RESULTS & DISCUSSION

Six aging genes are regulated by CREs with either positive selection and/or high brain activity

Of the 59 human aging-related related genes, *6 were identified as being regulated by cis-regulatory elements that are either evolutionarily selected for, differentially expressed in the human brain, or both*. A summary of these genes, brief description of their function, chromosome, and CRE information are shown in Table 2. AKT3, CDH13, and WWOX are regulated by CREs that were differentially expressed in neurons and iPSC-derived neural progenitor cells. CSMD1 and UNC5D are regulated by CREs that have experienced significant evolutionary positive selection compared to a standard nucleotide substitution rate. CDKAL1 is regulated by a CRE that falls under both of these categories, making it a strong candidate for further investigation.

Table 2. Aging Genes Regulated by Selected and/or Brain DE CREs

Gene ID	**Function**	**CRE**	**Chromosome**	**Selected, DE, or both?**

AKT3	Codes for serine/threonine kinase Linked to longevity Regulation of proliferation and survival via PI3K/Akt/mTOR pathway Mediator of insulin effects	Cre.1641	Chr1	Differentially expressed
CDH13	Codes for T-cadherin Involved in cognitive ability, prevents age-related bone loss, implicated in age-related disorders	Cre.1825	Chr16	Differentially expressed
CDKAL1	Codes for CDK5 regulatory subunit-associated protein 1-like 1 Implicated in age-related disorders	Cre.1210	Chr6	Both
CSMD1	Membrane component Associated with aging-related memory and cognitive issues	Cre.1368	Chr8	Selected
UNC5D	Netrin receptor activity Cortical cell survival	Cre.1378	Chr8	Selected
WWOX	Correlated with dementia Tumor suppressor	Cre.518	Chr16	Differentially expressed

AKT3

AKT3 has an extremely wide range of involvement in biological processes. It is perhaps most well known for its role in proliferation, tumorigenesis, and apoptosis due to its part in the PI3K/Akt pathway, which regulates cell survival during conditions of stress. The pathway is commonly mutated in cancerous cell conditions, resulting in lack of control over cell proliferation (Porta et al., 2014). Because of this, AKT3, along with the entire PI3K/Akt pathway, is a heavily researched target of cancer mechanisms and therapies. In addition, with high expression throughout the hippocampus, AKT3 has also been found to be essential for determining brain size. Regulation and coordination of organ tissue growth must be extremely precise - particularly in the brain, where an organ size that is either too large or too small can

severely affect the function of the organism. In a study examining AKT3 deficiency, mice with a knockout of AKT3 showed a 20% decrease in brain size compared to control mice due to brain cells that were both reduced in count and reduced in size (Easton et al., 2005). The mechanism of this is likely to be involvement with insulin-like growth factor 1 (IGF1), which has high interaction with AKT3 and is also important in regulating brain size. It is extremely possible that aging-related brain size decline can be partly attributed to altered regulation of AKT3.

In this project, AKT3 was found to be regulated by a CRE with significant differential expression in neural cells. This aligns with the logic of previous findings that AKT3 itself has high expression patterns in neural tissues. While this project did not find evidence that AKT3 is regulated by evolutionarily selected CREs, other research has found evidence of adaptive evolution in mammary gland and immune cell AKT3 expression (Farmanullah et al., 2020).

CDH13

Like AKT3, CDH13 (Cadherin-13) also plays a role in cancer development, but is often downregulated in cancer rather than upregulated. The protein it codes for, T-cadherin, is an anchored membrane receptor that can alter cell properties and phenotype through its unique signaling properties (Andreeva & Kutuzov, 2010). While cadherins have an established role in cancer, CDH13 is a novel candidate gene for neuropsychiatric disorders (particularly ADHD) and age-related disorders. The link was initially discovered through genome-wide association studies, and recent and current studies are investigating mechanisms. During neural differentiation, CDH13 negatively regulates the growth of axons. It's involved in control of axon growth and guidance while neurons differentiate and develop. As such, it is hypothesized that altered axon growth could underlie a wide variety of neurological disorders, including

age-related neurological disorders (Rivero et al., 2013). CDH13 expression also appears to resist development of atherosclerosis, a disease common with increased age (Takeuchi et al., 2007).

Experimentally, CDH13 has high expression through the brain, particularly the prefrontal cortex, medulla, and thalamus (Rivero et al., 2013). This aligns with this project's finding that CDH13 is regulated by a CRE with significant differential neural cell expression.

CDKAL1

CDKAL1 is most well known for having a polymorphism that is associated with type 2 diabetes by regulating insulin production and adipose tissue mitochondrial function (Palmer et al., 2017). Based on GWAS, it is also associated with development of age-related disorders (Srivastava et al., 2015). With high neural expression, CDKAL1 is suspected to contribute to normal brain function due to its involvement with insulin (Srivastava et al., 2015). However, aside from the hypotheses stated, the function of CDKAL1 is largely unknown. CDKAL1 is thought to help compose CDK5, which is a key regulator of several neuronal functions after being activated by p35. Because of this, it is likely that CDKAL1's high neuronal expression is due to regulatory involvement in processes such as neuron migration, growth, and neurodegeneration (Takasugi et al., 2016). This project's finding that CDKAL1 is regulated by a CRE with high neuronal differential expression aligns with these hypotheses. However, it is also extremely interesting that the CRE which regulates CDKAL1 has been evolutionarily selected as well. This suggests that expression of CDKAL1 in the brain is altered in humans than other primates, and could be contributing to unique human neural features and age-related neurological diseases.

CSMD1

CSMD1 is known for acting in several cognitive processes, particularly learning and memory, partly due to its role in the complement cascade. The complement cascade is a part of the immune system that enhances the innate immune system and consists of approximately 50 proteins that activate each other in a cascade. However, the complement cascade's role extends beyond immunity. It also regulates synaptic pruning and synaptic plasticity, which are the processes of eliminating and changing the junctions of neurons when needed (Athanasiu et al., 2017). In neurodegenerative diseases, the complement cascade plays a part in synapse loss (Druart & Le Magueresse, 2019). In summary, CSMD1 has been associated with cognitive and memory issues, and its role in the complement cascade is a likely mechanism for this.

Little is known about the evolutionary history of CSMD1. It is conserved in our close ancestors, which is discussed further in the results section. This project identifies that CSMD1 is regulated by a CRE that has experienced positive selection. This finding suggests that regulation and expression of CSMD1 could be novel in humans compared to other primates.

UNC5D

UNC5D codes for a class of netrin receptors. Netrins are a class of proteins involved in axon guidance, and as a receptor for that, UNC5D is involved with axon guidance, cell migration, cell-cell adhesion, and cell survival in the human brain (Takemoto et al., 2011). A recent study using whole genome sequencing to analyze brain imaging phenotypes found that UNC5D influences total cerebral brain volume, hippocampal volume, and white matter hyperintensity (Sarnowski et al., 2018). All three of these aspects are often decreased in the aging brain and tend to accompany memory impairment (Peters, 2006), suggesting a role of

UNC5D in these cognitive declines. As discussed in the literature review, it is suspected that aging-related cognitive decline follows brain volume loss, but this relationship is nonlinear and the cause-effect nature of the relationship is not confirmed. In addition, a related gene of the same family, UNC5C, is associated with Alzheimer's disease and middle temporal volume (Sarnowski et al., 2018).

This project finds that UNC5D is regulated by a CRE that have experienced positive selection. Interestingly, under the scope of this project, UNC5D was not found to be regulated by CREs with differential brain activity, despite the gene itself typically being strongly expressed in human neocortical cells (Sarnowski et al., 2018). However, because its CREs are evolutionarily selected for, there is evidence that regulation of UNC5D and therefore brain volume attributes are evolutionarily important.

WWOX

Originally discovered as a tumor suppressor that promotes apoptosis (Aldaz & Hussain, 2020), research into WWOX has revealed that downregulation is a major risk factor for neurodegeneration as well. This is due to its role in the development and function of the central nervous system (CNS). WWOX is involved in key CNS development and neural differentiation signaling pathways involving Wnt and Hedgehog that help not only to develop, but maintain the adult human brain (Kośla et al., 2020). Among neurodegeneration, WWOX is particularly involved in Alzheimer's disease (AD) based on both experimental research and genome-wide association studies (Aldaz & Hussain, 2020; Dugan et al., 2022; Hsu et al., 2021). WWOX's protein, a WW domain-containing oxidoreductase, limits AD development by binding to Tau (aggregation of which is a key hallmark of AD) and preventing its aggregation. It is hypothesized

that in the aging human brain, gradual loss of WWOX induces increased levels of tau tangles (Hsu et al., 2021). This proposed mechanism would help explain WWOX as an age-related risk gene for Alzheimer's disease in GWA studies. A summary of WWOX's known neural functions is shown in Figure 3.

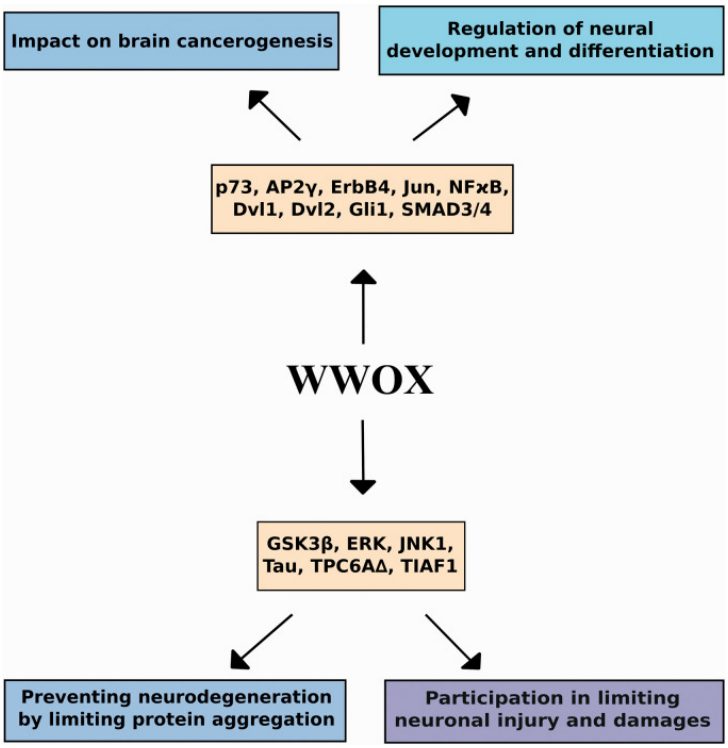


Figure 3. Summary of pleiotropic roles of WWOX in the central nervous system along with the proteins that it is involved with. Taken from (Kośla et al., 2020).

Previous gene expression data has identified that expression of WWOX tends to be higher in the cerebellar cortex than in other human CNS structures (Aldaz & Hussain, 2020). This, while considering the function of WWOX, aligns with the findings of this project that WWOX is regulated by a CRE with high differential expression in neurons and iPSC-derived

neural progenitor cells. This provides further evidence that WWOX is a significant candidate to investigate when researching aging and age-related neurodegeneration.

CREs tend to be closer to the aging genes that they regulate

Among the 6 significant aging-related genes, distances (in base pairs) between the genes and their CREs were calculated, and visually represented in Figure 4. The reason that there are over 100 total counts, rather than 6 counts, is because of different protein-splicing variants of each gene. As shown, CREs tended to be closer to the genes that they regulate rather than distal, with a median distance of approximately 250,000 base pairs away. The original CRE dataset includes both enhancers and promoters and are undistinguished. Because of this, it is possible that the 6 genes tend to be regulated by promoters rather than enhancers, since promoters regulate from close to the gene, and enhancers regulate regardless of distance. However, with the exception of one peak of distance counts (at 250,000 base pairs), distribution of the distances is not extreme. In these genes, regulation is still occurring from up to 2 million base pairs away from the gene start location. So, while regulation tends to occur closer to the gene that it is regulating, distance can vary.

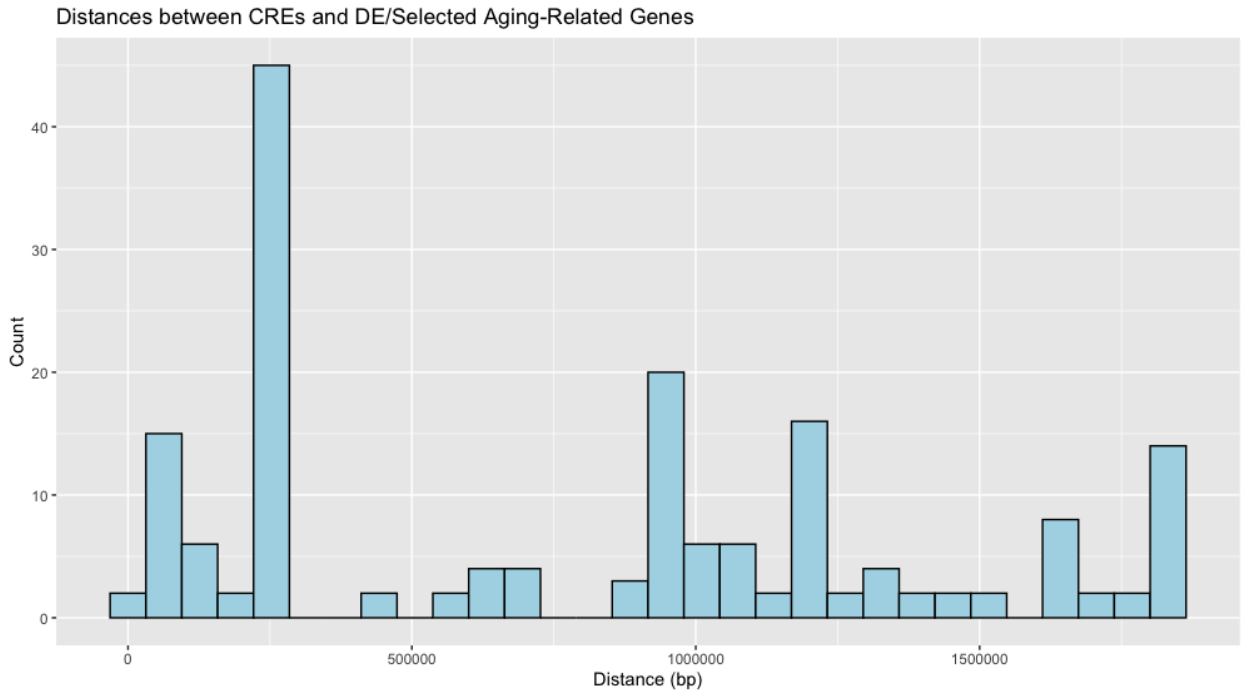


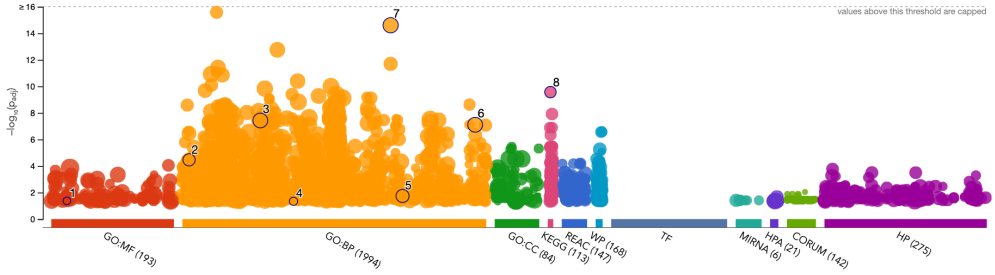
Figure 4. The distances between CREs and the protein-coding splicing variants of the six significant aging genes. Distances were calculated by subtracting the start base pair of the CREs from the start base pair of the genes.

Functional profiling of aging-related genes

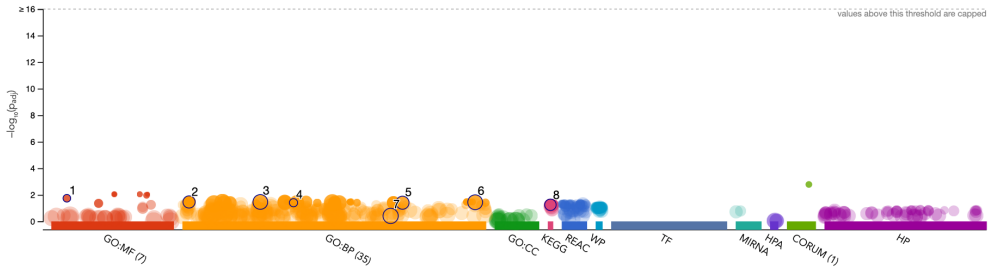
The 59 aging-related genes were chosen based on review of the literature. In order to visualize annotated functions of the genes using established gene ontology (GO) categories, we utilized gProfiler g:GOSt R package for functional enrichment analysis. A summary of the results, along with explanations of the GO categories, is shown in Figure 5. The most prevalent biological processes (BP) results include cellular stress response, cell population proliferation, cell aging, response to abiotic stimulus, response to oxidative stress, and regulation of cell migration/motility. The most prevalent KEGG pathway results include longevity regulating pathways, cellular senescence (deterioration of cells with age), and cancers such as T-cell leukemia, prostate cancer, and gastric cancer. The most prevalent reactome (REAC) terms

include regulation of p53 (a well known tumor suppressor), BH3-only protein activation (apoptotic proteins that often activate in response to stress), and many DNA repair and cell stress pathways. Notably, these genes are also enriched for netrin receptor activity. Identification of enriched GO categories is important because while the 59 aging-related genes were chosen based on GWAS and/or experimental evidence, many of their mechanistic involvements with human aging are unknown. Aging is a complex process that is not directly controlled by known genes, but it is likely that biological processes such as stress response regulation, proliferation and apoptosis regulation, cellular senescence, and netrin activity are key contributors of why and how humans age.

> full aging genes



> cre aging genes



ID	Source	Term ID	Term Name	p_{adj} (full aging g...	p_{adj} (cre aging g...
1	GO:MF	GO:0005042	netrin receptor activity	4.254×10^{-2}	1.814×10^{-2}
2	GO:BP	GO:0001935	endothelial cell proliferation	3.441×10^{-5}	3.445×10^{-2}
3	GO:BP	GO:0030334	regulation of cell migration	3.724×10^{-8}	3.445×10^{-2}
4	GO:BP	GO:0038007	netrin-activated signaling pathway	4.586×10^{-2}	3.932×10^{-2}
5	GO:BP	GO:0098742	cell-cell adhesion via plasma-membrane adhes...	1.815×10^{-2}	3.932×10^{-2}
6	GO:BP	GO:2000145	regulation of cell motility	7.860×10^{-8}	3.630×10^{-2}
7	GO:BP	GO:0080134	regulation of response to stress	2.498×10^{-15}	3.875×10^{-1}
8	KEGG	KEGG:04211	Longevity regulating pathway	2.720×10^{-10}	6.643×10^{-2}

version	e105_eg52_p16_e84549f
date	4/27/2022, 6:58:20 PM
organism	hsapiens

g:Profiler

Figure 5. Functional profiling of all 59 aging-related genes (top) and 6 aging-related genes (bottom) regulated by the discussed CREs, performed using the g:GOSt gProfiler open source package (Raudvere et al., 2019). In this Manhattan plot, each circle represents a gene ontology (GO) term that fall under eleven "source" categories based on previous functional annotation: molecular function (MF), biological processes (BP), cellular component (CC), KEGG pathways (KEGG), reactome (REAC), Wiki pathways (WP), transcription binding sites (TF), genes targeted by miRNAs (MIRNA), Human Protein Atlas (HPA), CORUM protein complexes (CORUM), and human phenotype ontology (HP). Adjusted p-values for particular functions of interest are also shown.

Along with functional profiling of all 59 aging-related genes, we also performed functional enrichment of the 6 aging-related genes regulated by selected and/or differentially expressed CREs: AKT3, CDH13, CDKAL1, CSMD1, UNC5D, and WWOX. Because these 6 genes are included in the full list of 59 genes, their results (shown in the bottom plot of Figure 5) are also included in the results of the full list (shown in the top plot of Figure 5). Notable functions include netrin receptor activity, endothelial cell proliferation, cell migration regulation, cell-cell adhesion, stress response, and enrichment in a longevity regulating pathway. While these results had a larger p-value than many of the results for the full list, they are still significant. The larger p-value is likely due to some of the 6 genes being less researched and therefore less annotated in terms of gene ontology. Overall, the functional enrichments of the 6 aging-related genes represent similar biological processes as the full list of 59 genes. Because some genes are under-researched, particularly CDKAL1 and CSMD1, many gene ontology functions may have not been identified yet.

Aging genes are orthologous in our closest ancestors

NCBI (National Center for Biotechnology Information) was used to search for orthologous genes in our closest ancestors for AKT3, CDH13, CDKAL1, CSMD1, UNC5D, and WWOX. All 6 genes are present and conserved in our closest ancestors. There are generally no duplications, losses, or novel genes present, and the amino acid sequences are generally the same. The only slight exceptions are in CDH13 and WWOX. In CDH13, the most common splicing variant is slightly shorter in humans than in chimpanzees (713 amino acids vs 760 amino acids), In WWOX, the most common splicing variant is also slightly shorter in chimpanzees (436 amino acids vs 414 amino acids). However, this orthologous gene search provides further evidence that these genes are not novel. *Any evolutionary change in these genes is likely due to regulation, possibly from CREs, rather than any mutations in the coding regions themselves*.

Motif Enrichment in Selected Aging Genes

Due to particular interest in the aging genes that are regulated by evolutionarily selected aging-related genes, motif enrichments for CDKAL1, CSMD1,and UNC5D were examined using SEA (Bailey & Grant, 2021). Using the JASPAR Vertebrate database, no motif sequences (with E less than or equal to 10) were identified in CDKAL1 or CSMD1. However, 254 motifs were enriched in the 6 UNC5D alternative splicing sequences. The top 5 of these motifs are identified and described in Table 3 below.

Table 3. Top 5 transcription factors enriched in UNC5D sequences. Obtained with SEA; results limited to JASPAR Vertebrate database.

Name and Matrix ID	**Class**	**Description**
Mecom MA0029.1	C2H2 zinc finger factors	Association to promoter expression in various human samples and cell lines (*JASPAR Motif:MA0029.1 - Resource_browser*, n.d.).
PBX1 MA0070.1	Homeo domain factors	Involvement with pre-B-cell leukemia homeobox 1 (*JASPAR Motif:MA0070.1 - Resource_browser*, n.d.).
REST MA0138.2	C2H2 zinc finger factors	Repressor element-1 silencing transcription factor (REST) is involved in neuronal death. Dysregulation of this transcription factor is involved in neurodegenerative disease, heart disease, and cancer (Noh et al., 2012).
HNF1B MA0153.2	Homeo domain factors	Involvement with insulin release, beta cells in pancreas, and kidneys; association to promoter expression in various human samples and cell lines (*JASPAR Motif:MA0153.1 - Resource_browser*, n.d.).
ESRRB MA0141.3	Nuclear receptors with C4 zinc fingers	Involvement in self-renewal and pluripotency maintenance through signaling pathways such as Wnt and FGF (Zhou et al., 2007).

In UNC5D, the average motif density is ~26 motifs/100 base pairs. While motif abundance is an indicator of transcription factor and CRE activity, the presence of a motif does not guarantee that transcription factors will bind at that location. In addition, because cofactors can be recruited to sequences that do not match predicted motifs, the absence of transcription factor motifs does not mean that there is no enhancer activity occurring (Grossman et al., 2017; Pizzollo et al., 2019). However, overall, these results suggest that UNC5D has high levels of enhancer activity occurring. Based on the top five enriched transcription factors identified (Table 3), it is likely that UNC5D has a wide range of participation in functions such as neuronal death, insulin release, self-renewal, and pluripotency maintenance.

CONCLUSIONS

While mutations in human coding regions presumably contribute to a portion of our neurological differences compared to non-human primates (Suntsova & Buzdin, 2020), many of the phenotypic differences are likely attributable to changes in gene expression. Previous research in the Babbitt lab identified a library of *cis*-regulatory elements (CREs), including enhancers and promoters, which modulate gene expression. These CREs fell under one of two categories (or, in the case of 34 CREs, overlapped): evolutionarily accelerated in humans, or differentially expressed in the human brain (Pizzollo et al., 2019).

This honors thesis identifies 6 aging-related genes, narrowed down from a list of 59, that are regulated by these evolutionarily accelerated and/or brain active *cis*-regulatory elements. *These genes (AKT3, CDH13, CDKAL1, CSMD1, UNC5D, and WWOX) and their respective CREs are excellent candidates to investigate in future research regarding how the human brain ages and how humans develop age-related neurodegenerative disorders.* In addition, this project identified that all 6 aging-related genes are orthologous in recent ancestors of humans, and that their CREs tend to be closer (in bp) to the genes that they regulate. Motif enrichment of genes regulated by evolutionarily selected CREs was also performed, resulting in identification of several significant transcription factors within UNC5D's sequence. Finally, functional profiling of the aging genes was performed and gave insight into the genes' wide range of functions, such as stress response regulation and netrin receptor activity.

As a field with massive amounts of data, further research into regulatory genomics is limitless. In future studies investigating the role of CREs in human aging, there are many steps to take next. This project investigated 59 aging-related genes in humans chosen from literature

search and the GenAge Database (Tacutu et al., 2018), but further research could expand this list to include other human aging-related genes that were not included in this project's search. Other entries in the GenAge Database include genes that are linked to aging in other mammalian organisms, non-mammalian organisms, or cell line models. In addition, because the chosen genes are suspected to be linked to aging based on association and/or experimental studies, there are likely many more genes that contribute to human aging and degeneration that have not been definitively identified yet. The library of CREs could also be expanded in future research. Evolutionarily accelerated CREs were identified based on computational tests comparing region substitution rate to a neutral substitution rate, but other statistical tests and parameters exist that could potentially alter or expand the database of CREs that have undergone positive selection. Similarly, the library of brain active CREs could be altered or expanded based on further experimental research in neuronal cells.

The scope of this project primarily focuses on aging in the brain. Further research could investigate genes that affect aging in other organ systems, such as the immune or musculoskeletal system. In addition, many of the 59 aging-related genes overlap with neurodegenerative disease risk, but further studies could focus specifically on risk genes for neurodegeneration. Another avenue for further research is investigation of motif enrichment. Motif density could be investigated for all of the aging-related genes, then compared to positive and negative controls using a permutation test. This project focused on motif enrichment of genes regulated by evolutionarily selected CREs, but motifs could also be examined for genes regulated by the differentially expressed CREs as well.

Cis-regulatory elements are not the only factors that regulate expression of genes. Of particular recent interest in the field is long noncoding RNAs (lncRNAs), which have a wide

range of regulatory methods and localizations. Various lncRNAs can interfere with signaling pathways, regulate chromatin function, affect stability and translation of cytoplasmic mRNAs—all of which regulate rates of gene expression (Statello et al., 2021). Regulatory lncRNAs are a popular recent area of research and their role in human brain aging would be fascinating to investigate.

As described, this area of research has many paths left to explore. AKT3, CDH13, CDKAL1, CSMD1, UNC5D, WWOX, and their CREs are strong candidates for aging research. In particular, CDKAL1 is regulated by a CRE that has both experienced recent positive selection and is active in the brain. Because this gene is modulated by an active regulatory region that is unique in humans, it is particularly critical for further investigation. Medical applications, such as targeting these genes when treating age-related disorders, have strong potential.

REFERENCES

1. Aldaz, C. M., & Hussain, T. (2020). WWOX Loss of Function in Neurodevelopmental and Neurodegenerative Disorders. *International Journal of Molecular Sciences*, *21*(23), E8922. https://doi.org/10.3390/ijms21238922
2. Amster, G., & Sella, G. (2016). Life history effects on the molecular clock of autosomes and sex chromosomes. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(6), 1588–1593. https://doi.org/10.1073/pnas.1515798113
3. Andreeva, A. V., & Kutuzov, M. A. (2010). Cadherin 13 in cancer. *Genes, Chromosomes & Cancer*, *49*(9), 775–790. https://doi.org/10.1002/gcc.20787
4. Athanasiu, L., Giddaluru, S., Fernandes, C., Christoforou, A., Reinvang, I., Lundervold, A. J., Nilsson, L.-G., Kauppi, K., Adolfsson, R., Eriksson, E., Sundet, K., Djurovic, S., Espeseth, T., Nyberg, L., Steen, V. M., Andreassen, O. A., & Le Hellard, S. (2017). A genetic association study of CSMD1 and CSMD2 with cognitive function. *Brain, Behavior, and Immunity*, *61*, 209–216. https://doi.org/10.1016/j.bbi.2016.11.026
5. Atzmon, G., Rincon, M., Schechter, C. B., Shuldiner, A. R., Lipton, R. B., Bergman, A., & Barzilai, N. (2006). Lipoprotein genotype and conserved pathway for exceptional longevity in humans. *PLoS Biology*, *4*(4), e113. https://doi.org/10.1371/journal.pbio.0040113
6. Autrey, M. M., Reamer, L. A., Mareno, M. C., Sherwood, C. C., Herndon, J. G., Preuss, T., Schapiro, S. J., & Hopkins, W. D. (2014). Age-related effects in the neocortical organization of chimpanzees: Gray and white matter volume, cortical thickness, and gyrification. *NeuroImage*, *101*, 59–67. https://doi.org/10.1016/j.neuroimage.2014.06.053
7. Bailey, T., & Grant, C. (2021). SEA: Simple Enrichment Analysis of Motifs. *BioRxiv*.
8. Baker, D. J., Jeganathan, K. B., Malureanu, L., Perez-Terzic, C., Terzic, A., & van Deursen, J. M. A. (2006). Early aging-associated phenotypes in Bub3/Rae1 haploinsufficient mice. *The Journal*

of Cell Biology, *172*(4), 529–540. https://doi.org/10.1083/jcb.200507081

9. Benitez, B. A., Karch, C. M., Cai, Y., Jin, S. C., Cooper, B., Carrell, D., Bertelsen, S., Chibnik, L., Schneider, J. A., Bennett, D. A., Alzheimer's Disease Neuroimaging Initiative, Genetic and Environmental Risk for Alzheimer's Disease Consortium GERAD, Fagan, A. M., Holtzman, D., Morris, J. C., Goate, A. M., & Cruchaga, C. (2013). The PSEN1, p.E318G variant increases the risk of Alzheimer's disease in APOE-ε4 carriers. *PLoS Genetics*, *9*(8), e1003685. https://doi.org/10.1371/journal.pgen.1003685
10. Berto, S., Liu, Y., & Konopka, G. (2020). Genomics at cellular resolution: Insights into cognitive disorders and their evolution. *Human Molecular Genetics*, *29*(R1), R1–R9. https://doi.org/10.1093/hmg/ddaa117
11. Blander, G., Zalle, N., Daniely, Y., Taplick, J., Gray, M. D., & Oren, M. (2002). DNA damage-induced translocation of the Werner helicase is regulated by acetylation. *The Journal of Biological Chemistry*, *277*(52), 50934–50940. https://doi.org/10.1074/jbc.M210479200
12. Boyle, E. A., Li, Y. I., & Pritchard, J. K. (2017). An expanded view of complex traits: From polygenic to omnigenic. *Cell*, *169*(7), 1177–1186. https://doi.org/10.1016/j.cell.2017.05.038
13. Buniello, A., MacArthur, J. A. L., Cerezo, M., Harris, L. W., Hayhurst, J., Malangone, C., McMahon, A., Morales, J., Mountjoy, E., Sollis, E., Suveges, D., Vrousgou, O., Whetzel, P. L., Amode, R., Guillen, J. A., Riat, H. S., Trevanion, S. J., Hall, P., Junkins, H., … Parkinson, H. (2019). The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Research*, *47*(D1), D1005–D1012. https://doi.org/10.1093/nar/gky1120
14. Burrinha, T., Martinsson, I., Gomes, R., Terrasso, A. P., Gouras, G. K., & Almeida, C. G. (2021). Upregulation of APP endocytosis by neuronal aging drives amyloid-dependent synapse loss. *Journal of Cell Science*, *134*(9), jcs255752. https://doi.org/10.1242/jcs.255752
15. Cardoso, A. L., Fernandes, A., Aguilar-Pimentel, J. A., de Angelis, M. H., Guedes, J. R., Brito,

M. A., Ortolano, S., Pani, G., Athanasopoulou, S., Gonos, E. S., Schosserer, M., Grillari, J., Peterson, P., Tuna, B. G., Dogan, S., Meyer, A., van Os, R., & Trendelenburg, A.-U. (2018). Towards frailty biomarkers: Candidates from genes and pathways regulated in aging and age-related diseases. *Ageing Research Reviews*, *47*, 214–277. https://doi.org/10.1016/j.arr.2018.07.004

16. Carroll, S. B. (2008). Evo-devo and an expanding evolutionary synthesis: A genetic theory of morphological evolution. *Cell*, *134*(1), 25–36. https://doi.org/10.1016/j.cell.2008.06.030
17. Clarke, L. E., Liddelow, S. A., Chakraborty, C., Münch, A. E., Heiman, M., & Barres, B. A. (2018). Normal aging induces A1-like astrocyte reactivity. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(8), E1896–E1905. https://doi.org/10.1073/pnas.1800165115
18. Cohen, B. H., Chinnery, P. F., & Copeland, W. C. (1993). POLG-Related Disorders. In M. P. Adam, H. H. Ardinger, R. A. Pagon, S. E. Wallace, L. J. Bean, K. W. Gripp, G. M. Mirzaa, & A. Amemiya (Eds.), *GeneReviews*®. University of Washington, Seattle. http://www.ncbi.nlm.nih.gov/books/NBK26471/
19. Conover, C. A. (2013). Role of PAPP-A in aging and age-related disease. *Experimental Gerontology*, *48*(7), 612–613. https://doi.org/10.1016/j.exger.2012.06.017
20. da Rocha, T. J., Silva Alves, M., Guisso, C. C., de Andrade, F. M., Camozzato, A., de Oliveira, A. A., & Fiegenbaum, M. (2018). Association of GPX1 and GPX4 polymorphisms with episodic memory and Alzheimer's disease. *Neuroscience Letters*, *666*, 32–37. https://doi.org/10.1016/j.neulet.2017.12.026
21. Deelen, J., Uh, H.-W., Monajemi, R., van Heemst, D., Thijssen, P. E., Böhringer, S., van den Akker, E. B., de Craen, A. J. M., Rivadeneira, F., Uitterlinden, A. G., Westendorp, R. G. J., Goeman, J. J., Slagboom, P. E., Houwing-Duistermaat, J. J., & Beekman, M. (2013). Gene set analysis of GWAS data for human longevity highlights the relevance of the insulin/IGF-1

signaling and telomere maintenance pathways. *Age (Dordrecht, Netherlands)*, *35*(1), 235–249. https://doi.org/10.1007/s11357-011-9340-3

22. Druart, M., & Le Magueresse, C. (2019). Emerging Roles of Complement in Psychiatric Disorders. *Frontiers in Psychiatry*, *10*. https://www.frontiersin.org/article/10.3389/fpsyt.2019.00573
23. Dugan, A. J., Nelson, P. T., Katsumata, Y., Shade, L. M. P., Teylan, M. A., Boehme, K. L., Mukherjee, S., Kauwe, J. S. K., Hohman, T. J., Schneider, J. A., Fardo, D. W., & Alzheimer's Disease Genetics Consortium. (2022). Association between WWOX/MAF variants and dementia-related neuropathologic endophenotypes. *Neurobiology of Aging*, *111*, 95–106. https://doi.org/10.1016/j.neurobiolaging.2021.10.011
24. Easton, R. M., Cho, H., Roovers, K., Shineman, D. W., Mizrahi, M., Forman, M. S., Lee, V. M.-Y., Szabolcs, M., de Jong, R., Oltersdorf, T., Ludwig, T., Efstratiadis, A., & Birnbaum, M. J. (2005). Role for Akt3/Protein Kinase Bγ in Attainment of Normal Brain Size. *Molecular and Cellular Biology*, *25*(5), 1869–1878. https://doi.org/10.1128/MCB.25.5.1869-1878.2005
25. Egerman, M. A., & Glass, D. J. (2019). The role of GDF11 in aging and skeletal muscle, cardiac and bone homeostasis. *Critical Reviews in Biochemistry and Molecular Biology*, *54*(2), 174–183. https://doi.org/10.1080/10409238.2019.1610722
26. El Assar, M., Angulo, J., & Rodríguez-Mañas, L. (2013). Oxidative stress and vascular inflammation in aging. *Free Radical Biology & Medicine*, *65*, 380–401. https://doi.org/10.1016/j.freeradbiomed.2013.07.003
27. Elzinga, S., Murdock, B. J., Guo, K., Hayes, J. M., Tabbey, M. A., Hur, J., & Feldman, E. L. (2019). Toll-like receptors and inflammation in metabolic neuropathy; a role in early versus late disease? *Experimental Neurology*, *320*, 112967. https://doi.org/10.1016/j.expneurol.2019.112967
28. Farmanullah, null, Hosseini, S. M., Liang, A., Hua, G., Rehman, Z. U., Talpur, H. S., Salim, M., Ahmad, S., Abulaiti, A., Khan, M., Safdar, M., Kakar, I. U., Ahmad, Z., Ahmad, M. Z., Tingzhu,

Y., Schreurs, N. M., Bano, I., & Yang, L. (2020). Adaptive Molecular Evolution of AKT3 Gene for Positive Diversifying Selection in Mammals. *BioMed Research International*, *2020*, 2584627. https://doi.org/10.1155/2020/2584627

29. Fernández, Á. F., Sebti, S., Wei, Y., Zou, Z., Shi, M., McMillan, K. L., He, C., Ting, T., Liu, Y., Chiang, W.-C., Marciano, D. K., Schiattarella, G. G., Bhagat, G., Moe, O. W., Hu, M. C., & Levine, B. (2018). Disruption of the beclin 1-BCL2 autophagy regulatory complex promotes longevity in mice. *Nature*, *558*(7708), 136–140. https://doi.org/10.1038/s41586-018-0162-7
30. Franchini, L. F., & Pollard, K. S. (2017). Human evolution: The non-coding revolution. *BMC Biology*, *15*(1), 89. https://doi.org/10.1186/s12915-017-0428-9
31. Ganuza, M., Sáiz-Ladera, C., Cañamero, M., Gómez, G., Schneider, R., Blasco, M. A., Pisano, D., Paramio, J. M., Santamaría, D., & Barbacid, M. (2012). Genetic inactivation of Cdk7 leads to cell cycle arrest and induces premature aging due to adult stem cell exhaustion. *The EMBO Journal*, *31*(11), 2498–2510. https://doi.org/10.1038/emboj.2012.94
32. Garber, M., Guttman, M., Clamp, M., Zody, M. C., Friedman, N., & Xie, X. (2009). Identifying novel constrained elements by exploiting biased substitution patterns. *Bioinformatics*, *25*(12), i54–i62. https://doi.org/10.1093/bioinformatics/btp190
33. Garre, P., Briceño, V., Xicola, R. M., Doyle, B. J., de la Hoya, M., Sanz, J., Llovet, P., Pescador, P., Puente, J., Díaz-Rubio, E., Llor, X., & Caldés, T. (2011). Analysis of the oxidative damage repair genes NUDT1, OGG1, and MUTYH in patients from mismatch repair proficient HNPCC families (MSS-HNPCC). *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research*, *17*(7), 1701–1712. https://doi.org/10.1158/1078-0432.CCR-10-2491
34. Gaweda-Walerych, K., Sitek, E. J., Borczyk, M., Berdyński, M., Narożańska, E., Brockhuis, B., Korostyński, M., Sławek, J., & Zekanowski, C. (2021). Two Rare Variants in PLAU and BACE1 Genes-Do They Contribute to Semantic Dementia Clinical Phenotype? *Genes*, *12*(11), 1806.

https://doi.org/10.3390/genes12111806

35. Gennaccaro, L., Fuchs, C., Loi, M., Pizzo, R., Alvente, S., Berteotti, C., Lupori, L., Sagona, G., Galvani, G., Gurgone, A., Raspanti, A., Medici, G., Tassinari, M., Trazzi, S., Ren, E., Rimondini, R., Pizzorusso, T., Giovanna, Z., Maurizio, G., & Elisabetta, C. (2021). Age-Related Cognitive and Motor Decline in a Mouse Model of CDKL5 Deficiency Disorder is Associated with Increased Neuronal Senescence and Death. *Aging and Disease*, *12*(3), 764–785. https://doi.org/10.14336/AD.2020.0827
36. Gonzalo, S., Kreienkamp, R., & Askjaer, P. (2017). Hutchinson-Gilford Progeria Syndrome: A premature aging disease caused by LMNA gene mutations. *Ageing Research Reviews*, *33*, 18–29. https://doi.org/10.1016/j.arr.2016.06.007
37. Greene, M. A., & Loeser, R. F. (2015). Aging-related inflammation in osteoarthritis. *Osteoarthritis and Cartilage*, *23*(11), 1966–1971. https://doi.org/10.1016/j.joca.2015.01.008
38. Grill, S., & Nandakumar, J. (2021). Molecular mechanisms of telomere biology disorders. *The Journal of Biological Chemistry*, *296*, 100064. https://doi.org/10.1074/jbc.REV120.014017
39. Grossman, S. R., Zhang, X., Wang, L., Engreitz, J., Melnikov, A., Rogov, P., Tewhey, R., Isakova, A., Deplancke, B., Bernstein, B. E., Mikkelsen, T. S., & Lander, E. S. (2017). Systematic dissection of genomic features determining transcription factor binding and enhancer function. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(7), E1291–E1300. https://doi.org/10.1073/pnas.1621150114
40. Ham, S., & Lee, S.-J. V. (2020). Advances in transcriptome analysis of human brain aging. *Experimental & Molecular Medicine*, *52*(11), 1787–1797. https://doi.org/10.1038/s12276-020-00522-6
41. Haygood, R., Fedrigo, O., Hanson, B., Yokoyama, K.-D., & Wray, G. (2007). Promoter regions of many neural- and nutrition-related genes have experienced positive selection during human evolution. *Nature Precedings*, 1–1. https://doi.org/10.1038/npre.2007.69.1

42. Hernandez-Garcia, C. M., & Finer, J. J. (2014). Identification and validation of promoters and cis-acting regulatory elements. *Plant Science*, *217–218*, 109–119. https://doi.org/10.1016/j.plantsci.2013.12.007

43. Holder-Espinasse, M., Martin-Coignard, D., Escande, F., & Manouvrier-Hanu, S. (2007). A new mutation in TP63 is associated with age-related pathology. *European Journal of Human Genetics: EJHG*, *15*(11), 1115–1120. https://doi.org/10.1038/sj.ejhg.5201888

44. Hsu, C.-Y., Lee, K.-T., Sun, T.-Y., Sze, C.-I., Huang, S.-S., Hsu, L.-J., & Chang, N.-S. (2021). WWOX and Its Binding Proteins in Neurodegeneration. *Cells*, *10*(7), 1781. https://doi.org/10.3390/cells10071781

45. Inoue, F., Kircher, M., Martin, B., Cooper, G. M., Witten, D. M., McManus, M. T., Ahituv, N., & Shendure, J. (2017). A systematic comparison reveals substantial differences in chromosomal versus episomal encoding of enhancer activity. *Genome Research*, *27*(1), 38–52. https://doi.org/10.1101/gr.212092.116

46. Işıldak, U., Somel, M., Thornton, J. M., & Dönertaş, H. M. (2020). Temporal changes in the gene expression heterogeneity during brain development and aging. *Scientific Reports*, *10*, 4080. https://doi.org/10.1038/s41598-020-60998-0

47. Jadhav, S., Diwaker, C., Lila, A. R., Gada, J. V., Kale, S., Sarathi, V., Thadani, P. M., Arya, S., Patil, V. A., Shah, N. S., & Bandgar, T. R. (2021). POU1F1 mutations in combined pituitary hormone deficiency: Differing spectrum of mutations in a Western-Indian cohort and systematic analysis of world literature. *Pituitary*, *24*(5), 657–669. https://doi.org/10.1007/s11102-021-01140-9

48. *JASPAR motif:MA0029.1—Resource_browser*. (n.d.). Retrieved April 28, 2022, from https://fantom.gsc.riken.jp/5/sstar/JASPAR_motif:MA0029.1

49. *JASPAR motif:MA0070.1—Resource_browser*. (n.d.). Retrieved April 28, 2022, from https://fantom.gsc.riken.jp/5/sstar/JASPAR_motif:MA0070.1

50. *JASPAR motif:MA0153.1—Resource_browser*. (n.d.). Retrieved April 28, 2022, from https://fantom.gsc.riken.jp/5/sstar/JASPAR_motif:MA0153.1

51. Jeck, W. R., Siebold, A. P., & Sharpless, N. E. (2012). Review: A meta-analysis of GWAS and age-associated diseases. *Aging Cell*, *11*(5), 727–731. https://doi.org/10.1111/j.1474-9726.2012.00871.x

52. Jiang, S., & Mortazavi, A. (2018). Integrating ChIP-seq with other functional genomics data. *Briefings in Functional Genomics*, *17*(2), 104–115. https://doi.org/10.1093/bfgp/ely002

53. King, M. C., & Wilson, A. C. (1975). Evolution at two levels in humans and chimpanzees. *Science (New York, N.Y.)*, *188*(4184), 107–116. https://doi.org/10.1126/science.1090005

54. Kolberg, L., Raudvere, U., Kuzmin, I., Vilo, J., & Peterson, H. (2020). gprofiler2-an R package for gene list functional enrichment analysis and namespace conversion toolset g: Profiler. *F1000Research*, *9*(ELIXIR), 709. https://doi.org/doi:10.12688/f1000research.24956.2

55. Kośla, K., Kałuzińska, Ż., & Bednarek, A. K. (2020). The WWOX gene in brain development and pathology. *Experimental Biology and Medicine (Maywood, N.J.)*, *245*(13), 1122–1129. https://doi.org/10.1177/1535370220924618

56. Kulminski, A. M., Philipp, I., Shu, L., & Culminskaya, I. (2022). Definitive roles of TOMM40-APOE-APOC1 variants in the Alzheimer's risk. *Neurobiology of Aging*, *110*, 122–131. https://doi.org/10.1016/j.neurobiolaging.2021.09.009

57. Kunkle, B. W., Grenier-Boley, B., Sims, R., Bis, J. C., Damotte, V., Naj, A. C., Boland, A., Vronskaya, M., van der Lee, S. J., Amlie-Wolf, A., Bellenguez, C., Frizatti, A., Chouraki, V., Martin, E. R., Sleegers, K., Badarinarayan, N., Jakobsdottir, J., Hamilton-Nelson, K. L., Moreno-Grau, S., … Genetic and Environmental Risk in AD/Defining Genetic, Polygenic and Environmental Risk for Alzheimer's Disease Consortium (GERAD/PERADES),. (2019). Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates Aβ, tau, immunity and lipid processing. *Nature Genetics*, *51*(3), 414–430.

https://doi.org/10.1038/s41588-019-0358-2

58. Lacreuse, A., Raz, N., Schmidtke, D., Hopkins, W. D., & Herndon, J. G. (2020). Age-related decline in executive function as a hallmark of cognitive ageing in primates: An overview of cognitive and neurobiological studies. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *375*(1811), 20190618. https://doi.org/10.1098/rstb.2019.0618
59. Lescai, F., Marchegiani, F., & Franceschi, C. (2009). PON1 is a longevity gene: Results of a meta-analysis. *Ageing Research Reviews*, *8*(4), 277–284. https://doi.org/10.1016/j.arr.2009.04.001
60. Li, Y., Wang, W.-J., Cao, H., Lu, J., Wu, C., Hu, F.-Y., Guo, J., Zhao, L., Yang, F., Zhang, Y.-X., Li, W., Zheng, G.-Y., Cui, H., Chen, X., Zhu, Z., He, H., Dong, B., Mo, X., Zeng, Y., & Tian, X.-L. (2009). Genetic association of FOXO1A and FOXO3A with longevity trait in Han Chinese populations. *Human Molecular Genetics*, *18*(24), 4897–4904. https://doi.org/10.1093/hmg/ddp459
61. Liang, H., & Ward, W. F. (2006). PGC-1alpha: A key regulator of energy metabolism. *Advances in Physiology Education*, *30*(4), 145–151. https://doi.org/10.1152/advan.00052.2006
62. Lindblad-Toh, K., Garber, M., Zuk, O., Lin, M. F., Parker, B. J., Washietl, S., Kheradpour, P., Ernst, J., Jordan, G., Mauceli, E., Ward, L. D., Lowe, C. B., Holloway, A. K., Clamp, M., Gnerre, S., Alföldi, J., Beal, K., Chang, J., Clawson, H., … Kellis, M. (2011). A high-resolution map of human evolutionary constraint using 29 mammals. *Nature*, *478*(7370), 476–482. https://doi.org/10.1038/nature10530
63. Long, H. K., Prescott, S. L., & Wysocka, J. (2016). Ever-changing landscapes: Transcriptional enhancers in development and evolution. *Cell*, *167*(5), 1170–1187. https://doi.org/10.1016/j.cell.2016.09.018
64. Lorenzini, T., Fliegauf, M., Klammer, N., Frede, N., Proietti, M., Bulashevska, A., Camacho-Ordonez, N., Varjosalo, M., Kinnunen, M., de Vries, E., van der Meer, J. W. M., Ameratunga, R., Roifman, C. M., Schejter, Y. D., Kobbe, R., Hautala, T., Atschekzei, F., Schmidt,

R. E., Schröder, C., … NIHR BioResource. (2020). Characterization of the clinical and immunologic phenotype and management of 157 individuals with 56 distinct heterozygous NFKB1 mutations. *The Journal of Allergy and Clinical Immunology*, *146*(4), 901–911. https://doi.org/10.1016/j.jaci.2019.11.051

65. Luciano, M., Lopez, L. M., de Moor, M. H. M., Harris, S. E., Davies, G., Nutile, T., Krueger, R. F., Esko, T., Schlessinger, D., Toshiko, T., Derringer, J. L., Realo, A., Hansell, N. K., Pergadia, M. L., Pesonen, A.-K., Sanna, S., Terracciano, A., Madden, P. A. F., Penninx, B., … Deary, I. J. (2012). Longevity candidate genes and their association with personality traits in the elderly. *American Journal of Medical Genetics. Part B, Neuropsychiatric Genetics: The Official Publication of the International Society of Psychiatric Genetics*, *159B*(2), 192–200. https://doi.org/10.1002/ajmg.b.32013
66. Manrique, H. M., & Call, J. (2015). Age-dependent cognitive inflexibility in great apes. *Animal Behaviour*, *102*, 1–6. https://doi.org/10.1016/j.anbehav.2015.01.002
67. Mecca, C., Giambanco, I., Donato, R., & Arcuri, C. (2018). Microglia and Aging: The Role of the TREM2-DAP12 and CX3CL1-CX3CR1 Axes. *International Journal of Molecular Sciences*, *19*(1), E318. https://doi.org/10.3390/ijms19010318
68. Mori, N. (2020). [Brain and Neuronal Aging: Aged Brain Controls via Gene Expression Fidelity and Master Regulatory Factors]. *Yakugaku Zasshi: Journal of the Pharmaceutical Society of Japan*, *140*(3), 395–404. https://doi.org/10.1248/yakushi.19-00193-4
69. Noh, K.-M., Hwang, J.-Y., Follenzi, A., Athanasiadou, R., Miyawaki, T., Greally, J. M., Bennett, M. V. L., & Zukin, R. S. (2012). Repressor element-1 silencing transcription factor (REST)-dependent epigenetic remodeling is critical to ischemia-induced neuronal death. *Proceedings of the National Academy of Sciences*, *109*(16), E962–E971. https://doi.org/10.1073/pnas.1121568109
70. Ota, M., Ito, T., Umemura, T., Katsuyama, Y., Yoshizawa, K., Hamano, H., & Kawa, S. (2011).

Polymorphism in the KCNA3 gene is associated with susceptibility to autoimmune pancreatitis in the Japanese population. *Disease Markers*, *31*(4), 223–229. https://doi.org/10.3233/DMA-2011-0820

71. Paccosi, E., & Proietti-De-Santis, L. (2021). The emerging role of Cockayne group A and B proteins in ubiquitin/proteasome-directed protein degradation. *Mechanisms of Ageing and Development*, *195*, 111466. https://doi.org/10.1016/j.mad.2021.111466
72. Palmer, C. J., Bruckner, R. J., Paulo, J. A., Kazak, L., Long, J. Z., Mina, A. I., Deng, Z., LeClair, K. B., Hall, J. A., Hong, S., Zushin, P.-J. H., Smith, K. L., Gygi, S. P., Hagen, S., Cohen, D. E., & Banks, A. S. (2017). Cdkal1, a type 2 diabetes susceptibility gene, regulates mitochondrial function in adipose tissue. *Molecular Metabolism*, *6*(10), 1212–1225. https://doi.org/10.1016/j.molmet.2017.07.013
73. Paneni, F., & Cosentino, F. (2012). P66 Shc as the engine of vascular aging. *Current Vascular Pharmacology*, *10*(6), 697–699. https://doi.org/10.2174/157016112803520747
74. Paneni, F., Osto, E., Costantino, S., Mateescu, B., Briand, S., Coppolino, G., Perna, E., Mocharla, P., Akhmedov, A., Kubant, R., Rohrer, L., Malinski, T., Camici, G. G., Matter, C. M., Mechta-Grigoriou, F., Volpe, M., Lüscher, T. F., & Cosentino, F. (2013). Deletion of the activated protein-1 transcription factor JunD induces oxidative stress and accelerates age-related endothelial dysfunction. *Circulation*, *127*(11), 1229–1240, e1-21. https://doi.org/10.1161/CIRCULATIONAHA.112.000826
75. Peters, R. (2006). Ageing and the brain. *Postgraduate Medical Journal*, *82*(964), 84–88. https://doi.org/10.1136/pgmj.2005.036665
76. Pizzollo, J., Zintel, T., & Babbitt, C. (2019). *Differentially active and conserved neural enhancers define two forms of adaptive non-coding evolution in humans*. University of Massachusetts Amherst.
77. Pollard, K. S., Salama, S. R., King, B., Kern, A. D., Dreszer, T., Katzman, S., Siepel, A.,

Pedersen, J. S., Bejerano, G., Baertsch, R., Rosenbloom, K. R., Kent, J., & Haussler, D. (2006). Forces Shaping the Fastest Evolving Regions in the Human Genome. *PLoS Genetics*, *2*(10), e168. https://doi.org/10.1371/journal.pgen.0020168

78. Pollard, K. S., Salama, S. R., Lambert, N., Lambot, M.-A., Coppens, S., Pedersen, J. S., Katzman, S., King, B., Onodera, C., Siepel, A., Kern, A. D., Dehay, C., Igel, H., Ares, M., Vanderhaeghen, P., & Haussler, D. (2006). An RNA gene expressed during cortical development evolved rapidly in humans. *Nature*, *443*(7108), 167–172. https://doi.org/10.1038/nature05113
79. Porta, C., Paglino, C., & Mosca, A. (2014). Targeting PI3K/Akt/mTOR Signaling in Cancer. *Frontiers in Oncology*, *4*. https://www.frontiersin.org/article/10.3389/fonc.2014.00064
80. Prabhakar, S., Noonan, J. P., Pääbo, S., & Rubin, E. M. (2006). Accelerated Evolution of Conserved Noncoding Sequences in Humans. *Science*, *314*(5800), 786–786. https://doi.org/10.1126/science.1130738
81. Raudvere, U., Kolberg, L., Kuzmin, I., Arak, T., Adler, P., Peterson, H., & Vilo, J. (2019). g:Profiler: A web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Research*.
82. Reilly, S. K., Yin, J., Ayoub, A. E., Emera, D., Leng, J., Cotney, J., Sarro, R., Rakic, P., & Noonan, J. P. (2015). Evolutionary genomics. Evolutionary changes in promoter and enhancer activity during human corticogenesis. *Science (New York, N.Y.)*, *347*(6226), 1155–1159. https://doi.org/10.1126/science.1260943
83. Rilling, J. K. (2014). Comparative primate neuroimaging: Insights into human brain evolution. *Trends in Cognitive Sciences*, *18*(1), 46–55. https://doi.org/10.1016/j.tics.2013.09.013
84. Rivero, O., Sich, S., Popp, S., Schmitt, A., Franke, B., & Lesch, K.-P. (2013). Impact of the ADHD-susceptibility gene CDH13 on development and function of brain networks. *European Neuropsychopharmacology: The Journal of the European College of Neuropsychopharmacology*, *23*(6), 492–507. https://doi.org/10.1016/j.euroneuro.2012.06.009

85. Roth, G., & Dicke, U. (2005). Evolution of the brain and intelligence. *Trends in Cognitive Sciences*, *9*(5), 250–257. https://doi.org/10.1016/j.tics.2005.03.005

86. Ruzankina, Y., Pinzon-Guzman, C., Asare, A., Ong, T., Pontano, L., Cotsarelis, G., Zediak, V. P., Velez, M., Bhandoola, A., & Brown, E. J. (2007). Deletion of the developmentally essential gene ATR in adult mice leads to age-related phenotypes and stem cell loss. *Cell Stem Cell*, *1*(1), 113–126. https://doi.org/10.1016/j.stem.2007.03.002

87. Sarnowski, C., Satizabal, C. L., DeCarli, C., Pitsillides, A. N., Cupples, L. A., Vasan, R. S., Wilson, J. G., Bis, J. C., Fornage, M., Beiser, A. S., DeStefano, A. L., Dupuis, J., Seshadri, S., NHLBI Trans-Omics for Precision Medicine (TOPMed) Consortium, & TOPMed Neurocognitive Working Group. (2018). Whole genome sequence analyses of brain imaging measures in the Framingham Study. *Neurology*, *90*(3), e188–e196. https://doi.org/10.1212/WNL.0000000000004820

88. Savina, N. V., Nikitchenko, N. V., Kuzhir, T. D., Rolevich, A. I., Krasny, S. A., & Goncharova, R. I. (2018). The Involvement of ERCC2/XPD and ERCC6/CSB Wild Type Alleles in Protection Against Aging and Cancer. *Current Aging Science*, *11*(1), 45–54. https://doi.org/10.2174/1874609810666170707101548

89. Schaub, M. A., Boyle, A. P., Kundaje, A., Batzoglou, S., & Snyder, M. (2012). Linking disease associations with regulatory information in the human genome. *Genome Research*, *22*(9), 1748–1759. https://doi.org/10.1101/gr.136127.111

90. Srivastava, I., Thukral, N., & Hasija, Y. (2015). GENETICS OF HUMAN AGE RELATED DISORDERS. *Advances in Gerontology = Uspekhi Gerontologii*, *28*(2), 228–247.

91. Statello, L., Guo, C.-J., Chen, L.-L., & Huarte, M. (2021). Gene regulation by long non-coding RNAs and its biological functions. *Nature Reviews Molecular Cell Biology*, *22*(2), 96–118. https://doi.org/10.1038/s41580-020-00315-9

92. Streeper, R. S., Grueter, C. A., Salomonis, N., Cases, S., Levin, M. C., Koliwad, S. K., Zhou, P.,

Hirschey, M. D., Verdin, E., & Farese, R. V. (2012). Deficiency of the lipid synthesis enzyme, DGAT1, extends longevity in mice. *Aging*, *4*(1), 13–27. https://doi.org/10.18632/aging.100424

93. Suntsova, M. V., & Buzdin, A. A. (2020). Differences between human and chimpanzee genomes and their implications in gene expression, protein functions and biochemical properties of the two species. *BMC Genomics*, *21*(7), 535. https://doi.org/10.1186/s12864-020-06962-8
94. Tacutu, R., Thornton, D., Johnson, E., Budovsky, A., Barardo, T., Diana, E., Lehmann, G., Toren, D., Wang, J., Fraifeld, V., & Magalhaes, J. (2018). Human Ageing Genomic Resources: New and updated databases. *Nucleic Acids Research*, *46*(D1), D1083–D1090.
95. Takasugi, T., Minegishi, S., Asada, A., Saito, T., Kawahara, H., & Hisanaga, S. (2016). Two Degradation Pathways of the p35 Cdk5 (Cyclin-dependent Kinase) Activation Subunit, Dependent and Independent of Ubiquitination*. *Journal of Biological Chemistry*, *291*(9), 4649–4657. https://doi.org/10.1074/jbc.M115.692871
96. Takemoto, M., Hattori, Y., Zhao, H., Sato, H., Tamada, A., Sasaki, S., Nakajima, K., & Yamamoto, N. (2011). Laminar and areal expression of unc5d and its role in cortical cell survival. *Cerebral Cortex (New York, N.Y.: 1991)*, *21*(8), 1925–1934. https://doi.org/10.1093/cercor/bhq265
97. Takeuchi, T., Adachi, Y., Ohtsuki, Y., & Furihata, M. (2007). Adiponectin receptors, with special focus on the role of the third receptor, T-cadherin, in vascular disease. *Medical Molecular Morphology*, *40*(3), 115–120. https://doi.org/10.1007/s00795-007-0364-9
98. Tasselli, L., Zheng, W., & Chua, K. F. (2017). SIRT6: Novel Mechanisms and Links to Aging and Disease. *Trends in Endocrinology and Metabolism: TEM*, *28*(3), 168–185. https://doi.org/10.1016/j.tem.2016.10.002
99. Vatner, S. F., Pachon, R. E., & Vatner, D. E. (2015). Inhibition of adenylyl cyclase type 5 increases longevity and healthful aging through oxidative stress protection. *Oxidative Medicine and Cellular Longevity*, *2015*, 250310. https://doi.org/10.1155/2015/250310

100. Vermunt, M. W., & Creyghton, M. P. (2016). Transcriptional Dynamics at Brain Enhancers: From Functional Specialization to Neurodegeneration. *Current Neurology and Neuroscience Reports*, *16*(10), 94. https://doi.org/10.1007/s11910-016-0689-7

101. Vermunt, M. W., Reinink, P., Korving, J., de Bruijn, E., Creyghton, P. M., Basak, O., Geeven, G., Toonen, P. W., Lansu, N., Meunier, C., van Heesch, S., Netherlands Brain Bank, Clevers, H., de Laat, W., Cuppen, E., & Creyghton, M. P. (2014). Large-scale identification of coregulated enhancer networks in the adult human brain. *Cell Reports*, *9*(2), 767–779. https://doi.org/10.1016/j.celrep.2014.09.023

102. Wang, J., Yu, J.-T., & Tan, L. (2015). PLD3 in Alzheimer's disease. *Molecular Neurobiology*, *51*(2), 480–486. https://doi.org/10.1007/s12035-014-8779-5

103. Wang, L., Yang, L., Debidda, M., Witte, D., & Zheng, Y. (2007). Cdc42 GTPase-activating protein deficiency promotes genomic instability and premature aging-like phenotypes. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(4), 1248–1253. https://doi.org/10.1073/pnas.0609149104

104. Weichhart, T. (2018). mTOR as Regulator of Lifespan, Aging, and Cellular Senescence: A Mini-Review. *Gerontology*, *64*(2), 127–134. https://doi.org/10.1159/000484629

105. Wu, D., & Prives, C. (2018). Relevance of the p53-MDM2 axis to aging. *Cell Death and Differentiation*, *25*(1), 169–179. https://doi.org/10.1038/cdd.2017.187

106. Yang, Y. R., Kabir, M. H., Park, J. H., Park, J.-I., Kang, J. S., Ju, S., Shin, Y. J., Lee, S. M., Lee, J., Kim, S., Lee, K.-P., Lee, S. Y., Lee, C., & Kwon, K.-S. (2020). Plasma proteomic profiling of young and old mice reveals cadherin-13 prevents age-related bone loss. *Aging*, *12*(9), 8652–8668. https://doi.org/10.18632/aging.103184

107. Yousefzadeh, M. J., Flores, R. R., Zhu, Y., Schmiechen, Z. C., Brooks, R. W., Trussoni, C. E., Cui, Y., Angelini, L., Lee, K.-A., McGowan, S. J., Burrack, A. L., Wang, D., Dong, Q., Lu, A., Sano, T., O'Kelly, R. D., McGuckian, C. A., Kato, J. I., Bank, M. P., … Niedernhofer, L. J.

(2021). An aged immune system drives senescence and ageing of solid organs. *Nature*, *594*(7861), 100–105. https://doi.org/10.1038/s41586-021-03547-7

108. Yu, L., Shulman, J. M., Chibnik, L., Leurgans, S., Schneider, J. A., De Jager, P. L., & Bennett, D. A. (2012). The CETP I405V polymorphism is associated with an increased risk of Alzheimer's disease. *Aging Cell*, *11*(2), 228–233. https://doi.org/10.1111/j.1474-9726.2011.00777.x

109. Zhou, W., Lo, S.-C., Liu, J.-H., Hannink, M., & Lubahn, D. B. (2007). ERRβ: A potent inhibitor of Nrf2 transcriptional activity. *Molecular and Cellular Endocrinology*, *278*(1–2), 52–62. https://doi.org/10.1016/j.mce.2007.08.011