NOTABLE DECREASE IN TRANSCRIPTOME CONSERVATION DURING MAMMALIAN AGING

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

ZELIHA GOZDE TURAN

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
BIOLOGY

SEPTEMBER 2016

Aproval of the thesis:

NOTABLE DECREASE IN TRANSCRIPTOME CONSERVATION DURING MAMMALIAN AGING

submitted by **ZELIHA GOZDE TURAN** in partial fulfillment of the requirements for the degree of **Master of Science in Biology Department, Middle East Technical University** by,

Prof. Dr. Gülbin Dural Ünver
Dean, Graduate School of **Natural and Applied Sciences** _____

Prof. Dr. Orhan Adalı
Head of Department, **Biology** _____

Assoc. Prof. Dr. Mehmet Somel
Supervisor, **Biology Dept., METU** _____

Examining Committee Members:

Assoc. Prof. Dr. Çağdaş Devrim Son
Biology Dept., METU _____

Assoc. Prof. Dr. Mehmet Somel
Biology Dept., METU _____

Assoc. Prof. Dr. Sreeparna Banerjee
Biology Dept., METU _____

Assoc. Prof. Dr. Özlen Konu
Dept. of Mol. Biol. and Genet., İhsan Doğramacı Bilkent Uni. _____

Assoc. Prof. Dr. Michelle M. Adams
Psychology Dept., İhsan Doğramacı Bilkent Uni. _____

Date: 30.09.2016

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last name	: ZELIHA GOZDE TURAN
Signature	:

ABSTRACT

NOTABLE DECREASE IN TRANSCRIPTOME CONSERVATION DURING MAMMALIAN AGING

Turan, Zeliha Gözde
M.S., Department of Biology
Supervisor : Assoc. Prof. Dr. Mehmet Somel

September 2016, 88 pages

Aging is a complex process that causes decline in organisms' reproductive capacity and chance of survival. Even though aging tends to reduce fitness, it is not eliminated by natural selection and is observed in many multicellular species, and this leads to an evolutionary paradox. The mutation accumulation theory states that due to the declining force of natural selection with age, old-age-expressed deleterious mutations will not be effectively eliminated, and can contribute to the aging phenotype. A limited number of empirical studies showed effects consistent with the mutation accumulation theory with controversial results, but this theory has not been tested using transcriptomic data. One prediction of mutation accumulation theory would be that genes highly expressed later in life would be less conserved than those expressed early. In this study, I performed a meta-analysis of 35 microarray gene-expression datasets including 8 tissues from 4 mammalian species, and studied the protein sequence conservation of genes expressed at different levels during adulthood. Age-related decrease in transcriptome conservation was detected in brain, liver, and lung, with the contribution of both genes having increased expression with age and low conservation, and genes having decreased expression with age and high conservation. Meanwhile, no such trend was observed in muscle

tissues. To find functional groups associated with decrease in transcriptome conservation with age, I then performed Gene Ontology (GO) analysis. GO analysis revealed that genes showing increased expression and low conservation tend to be associated with apoptosis across different tissues. These results may indicate that genes highly expressed at old age and with low sequence conservation may contribute to the senescence phenotype in different mammalian species, consistent with the mutation accumulation theory.

Keywords: aging, evolution, gene expression, mutation accumulation theory, meta-analysis, protein sequence conservation, dN/dS

ÖZ

MEMELİ YAŞLANMASI ESNASINDA TRANSKRİPTOM KORUNMASINDA MEYDANA GELEN ÖNEMLİ AZALIŞ

Turan, Zeliha Gözde
Yüksek Lisans, Biyoloji Bölümü
Tez Yöneticisi : Doç. Dr. Mehmet Somel

Eylül 2016, 88 sayfa

Yaşlanma, canlının üreme kapasitesinin ve hayatta kalma şansının azalmasına sebep olan oldukça kompleks bir süreçtir. Uyum başarısını azaltma eğiliminde olmasına rağmen, doğal seçilim tarafından elenmemesi ve çok hücreli türlerin çoğunda gözlenmesi evrimsel bir paradoksa sebep olmaktadır. Mutasyon birikimi teorisi, yaşla birlikte doğal seçilimin gücünün zayıflaması nedeniyle, ileri yaşa etkisini gösteren zararlı mutasyonların efektif olarak elenemeyeceğini ve yaşlanma fenotipine katkı sağlayabileceğini ileri sürmektedir. Sınırlı sayıda ampirik çalışma mutasyon birikimi teorisi ile uyumlu olan etkileri tartışmalı sonuçlar ile göstermiştir, ancak bu teori transkriptom verisi kullanılarak test edilmemiştir. Mutasyon birikimi teorisinin bir öngörüsü yaşamın ileri döneminde yüksek seviyede anlatılan genlerin erken dönemde anlatılanlara göre daha düşük seviyede korunmuş olabileceğidir. Bu çalışmada, 8 doku ve 4 memeli türünü kapsayan 35 mikrodizin gen anlatımı verisinin meta-analizini gerçekleştirdim, ve erişkinlik boyunca farklı seviyelerde anlatılan genlerin protein sekans korunmasını çalıştım. Transkriptom korunma seviyesinde yaşa bağlı azalış, hem yaşla beraber artan gen anlatımı ve düşük korunmaya sahip genler hemde yaşla beraber azalan gen anlatımı ve yüksek korunmaya sahip genlerin katkısıyla beyin, akciğer ve karaciğerde tespit edilmiştir. Öte yandan, böyle bir

eğilim kas dokusunda gözlemlenmemiştir. Daha sonra, transkriptom korunma seviyesinde yaşa bağlı azalışa sebep olan fonksiyonel grupları bulmak için Gen Ontoloji (GO) analizi gerçekleştirdim. GO analizi artan gen anlatımı ve düşük korunma gösteren genlerin apoptoz ile bağlantılı olma eğiliminde olduğunu farklı dokularda göstermiştir. Bu sonuç ileri yaşta yüksek seviyede anlatılan ve düşük sekans korunmasına sahip genlerin mutasyon birikimi teorisi ile uyumlu olarak, yaşlanma fenotipine katkı sağlayabileceğine birden fazla memeli türünde işaret etmektedir.

Anahtar Kelimeler: yaşlanma, evrim, gen anlatımı, mutasyon birikimi teorisi, meta-analiz, protein-dizi korunması, dN/dS

To my beloved grandmother Fatma Sür and grandfather Ramazan Sür...

ACKNOWLEDGEMENTS

First and foremost, I want to thank my advisor Mehmet Somel for his support, encouragement and motivation throughout my master. I also thank him for sharing his knowledge and experience with me. I really appreciate his consistent guidance and patience.

I have to thank my lab mates, especially Handan Melike Dönertaş for answering my questions regardless of the time and sharing knowledge with me, and Hamit İzgi for fruitful discussions. Special thanks go to Poorya Parvizi for his suggestions, contributions, and providing an important insight into the study.

I would like to thank my spouse, Anıl Turan, for his positive attitude and encouragement throughout my master. He kept me continue in the times of crisis. Without his constant source of support, I may not achieve to finish this work.

Lastly, I would like to thank my mother Nezekat Şahin, my father Memduh Şahin, and my aunt Gökşen Sür for all their support, love and encouragement. I also thank my mother-in-law Bedriye Turan, and father-in-law Ruhi Turan.

TABLE OF CONTENTS

LIST OF TABLES

TABLES

LIST OF FIGURES

FIGURES

LIST OF ABBREVIATIONS

BA22	Superior temporal cortex
EC	Entorhinal cortex
PCG	Postcentral gyrus
HC	Hippocampus
SG	Superior frontal gyrus
FC	Frontal cortex
BA10	Anterior prefrontal cortex
PFC	Prefrontal cortex
CB	Cerebellar cortex
EOMs	Extraocular muscles
EDL	Extensor digitorum longus muscles
T	Thoracic
BB	Biceps brachii
VL	Vastus lateralis
RA	Rectus abdominis
PC	Precursor cells
NC	Neocortex
WB	Whole brain
Gastro	Gastrocnemius tissue
Other	Various other regions of the anatomy
Hs	*Homo sapiens*
Rm	*Rhesus macaque* & *Macaca fascicularis*
Rn	*Rattus norvegicus*
Mm	*Mus musculus*

CHAPTER 1

INTRODUCTION

1.1 Aging

Aging can be defined as changes that take place in the cells, tissues, and organs with increasing age. Many of these changes are thought to represent either stochastic events or responses to such events. Senescence is defined as the deleterious parts of those changes, which eventually lead to functional impairment. Nearly all the alterations that accumulate with age are harmful to an organism's viability and reproduction, therefore aging is usually used interchangeably with senescence (Longo, Mitteldorf & Skulachev, 2005).

Aging is also accompanied by a wide range of age-related diseases, such as cancer, neurodegenerative and cardiovascular disorders (López-Otín, 2013).Understanding the cause and effect relationship between aging and age-related pathologies, and finding treatment is both demanding and time consuming. Therefore, according to one argument, the primary aim of the aging studies should be to find the underlying driving force of this complex phenotype, which may eventually allow developing approaches that delay the entire aging process, and thus also delay age-related disorders (de Magalhães, 2003).

At the cellular and molecular level, diverse pathways and mechanisms can contribute to aging either directly or indirectly. Each of them has a different role in determining the aging phenotype. A recent study categorized cellular and molecular hallmarks of aging into nine groups, described below, and summarized in

Figure 1.1.

Figure 1.1: The figure illustrates the nine hallmarks described in (López-Otín, 2013).

1. **Genomic Instability:** Accumulation of DNA damage throughout the species lifespan is one of the leading cause of aging (Moskalev et al., 2013). Both nuclear DNA and mitochondrial DNA (mtDNA) are prone to extrinsic and intrinsic damaging agents (Hoeijmakers, 2009). These agents can cause genomic instability in different ways, including point mutations, telomere shortening, and chromosomal abnormalities. Failure to properly overcome such damage in essential genes may result in disruption of homeostasis (Jones & Rando, 2011). Also, DNA repair mechanism has an important role in both mitotic and postmitotic tissues. For example, one study showed that genes having role in DNA repair mechanism showing increased expression with age in brain (Lu et al., 2004). In addition, numerous studies have reported

that there is a relationship between efficiency of DNA repair mechanism and longevity (Promislow, 1994).

2. **Telomere Attrition:** Most somatic cells lack the telomerase enzyme, a condition that causes progressive loss of telomere length with each cell division. A critically short telomere length triggers the activation of DNA repair mechanism that results in cellular senescence or apoptosis. Also, several studies showed that dysfunctional telomerase are associated with accelerated aging phenotypes (Blasco, 2007).

3. **Epigenetic Alterations:** Aging is accompanied by epigenetic alterations, including histone post-translational modifications, changes in DNA methylation, and modification of chromatin architecture (Talens et al., 2012). Several studies in animal models have demonstrated that increased activity of SIRT6, which plays a role in epigenetic regulation, results in prolonged longevity. In addition, causal links between perturbation of epigenetic systems and progeroid syndromes were also found in model organisms (Kanfi et al., 2012; Mostoslavsky et al., 2006).

4. **Loss Of Proteostasis:** Maintenance of protein homeostasis (proteostasis) is crucial to preserving cell function. Chaperones and proteasomes have an important role in securing protein turnover. Chaperones isolate unfolded/ misfolded proteins, and if possible repair them, while proteasomes identify and degrade damaged proteins (Koga, Kaushik & Cuervo, 2011). Impairment in protein homeostasis with increasing age results in accumulation of damaged proteins, and also contributes to age-related disorders such as Alzheimer's disease and cataracts (Powers, Morimoto, Dillin, Kelly & Balch, 2009).

5. **Deregulated Nutrient Sensing:** The rate of aging is controlled by manipulations of anabolic and nutrient signaling pathways. Decreased activity of the nutrient signaling pathways is associated with prolonged longevity, and can be achieved by caloric restriction (CR) or drugs that mimic the effect of CR. In contrast to this, increased activity of anabolic signaling pathways can accelerate the aging rate (Fontana, Partridge & Longo, 2010; Harrison et al., 2009).

6. **Mitochondrial Dysfunction:** Progressive accumulation of mutations within the mtDNA leads to mitochondrial dysfunction and subsequent decline in

energy production. Several studies have demonstrated that mitochondrial deficiency can accelerate aging in mammals (Kujoth et al., 2005; Trifunovic et al., 2004; Vermulst et al., 2008). However, little is known about the effects of improved mitochondrial function on lifespan extension.

7. **Cellular Senescence:** Cellular senescence can be defined as permanent cell cycle arrest. Increased number of senescent cells during aging can contribute to age-related dysfunction. This is because, senescent cells lose their regenerative capacity, and raised number of those cells with age can impair tissue structure and renewal in old species (Campisi & d'Adda di Fagagna, 2007). However, senescence has also an important role in protecting against tumor formation, especially in cells with high proliferative capacity. Several studies showed that both removal of senescent cells and also increased activity of the senescence-inducing tumor suppressor mechanisms can ameliorate age-related deterioration (Baker et al., 2011; Matheu et al., 2009, 2007). These studies provide a link between aging and cellular senescence.

8. **Stem Cell Exhaustion:** Decreased regenerative capacity of stem cells with aging is one of the leading factors that hampers tissue renewal (Sharpless & DePinho, 2007). The end result could be decline in physiological integrity. One recent study showed that rejuvenation of stem cells may improve the aging phenotype at the organismal level (Rando & Chang, 2012).

9. **Altered Intercellular Communication:** Changes in intercellular communication can be observed at different systems including endocrine and neuroendocrine systems. One important age-related change in intercellular communication is low level of chronic inflammation, also called "inflammaging". Several factors, such as defects in autophagy, can cause inflammaging (Salminen, Kaarniranta & Kauppinen, 2012). Furthermore, age-related deterioration in one organ can affect other organs, and harmful signaling in inter-organ communication can also contribute to the aging phenotype.

1.2 Theories of Aging

More than 300 theories have been postulated to date to explain aging (Medvedev, 1990). Besides evolutionary theories of aging, there are an incredibly high number of mechanistic ones. This is probably because of the complexity of the aging

phenotype, which leads to many ideas that each describe different hallmarks as the primary causes of aging. Although most ideas lack direct experimental support, they are called as "theory" instead of "hypothesis" (Kirkwood, 2005). I will also use the term "theory" throughout my thesis to be consistent with the aging literature.

1.2.1 Mechanistic Theories

Mechanistic theories try to explain the proximal molecular or physiological mechanisms driving aging. Most are based on the idea that aging occurs due to harmful byproducts of the cells. One of the prominent mechanistic theories of aging is the "free radical theory", which was proposed in the middle of the 20^{th} century (Harman, 1955). According to this theory, reactive oxygen species cause damage in cellular molecules, and such damaged molecules accumulate throughout the organisms' lifespan. This is especially deleterious for postmitotic cells, which have low protein turnover rate (Kirkwood, 2005).

One of the other mechanistic theories of aging is the "telomere theory," which suggests that telomere shortening triggers persistent activation of the p53 pathway, which in turn halts cellular proliferation and induces cell death. This is detrimental for maintenance of cells with high proliferative capacity, such as blood cells (Kelly, 2011; Lee et al., 1998).

There are many other mechanistic theories that answer the question of "how aging occurs?". However, no single mechanistic theory can explain all the changes that contribute to aging phenotype.

1.2.2 Evolutionary Theories

Natural selection shapes organisms' genomes to increase the probability of reproduction and survival, which are the two main components of Darwinian fitness (Demetrius & Ziehe, 2007). It might be expected that aging –one of the causes of impairment in fitness related traits- is strongly selected against by natural selection. However, aging is a common phenomenon across many multicellular organisms, despite some exceptions (Jones et al., 2014), and this leads to the question of "why aging occurs despite its disadvantages?" (Kirkwood, 2005).

Adaptive and non-adaptive theories answer the question of "why aging happens" taking into account different aspects of the evolutionary process. Adaptive or programmed aging theories suggest that aging is a genetically controlled mechanism as in development and morphogenesis (Austad, 2004). In contrast to the adaptive theories, non-adaptive ones propose that aging is caused by the indirect effects of stochastic evolutionary processes.

Non-adaptive Theories

The declining force of natural selection is the key concept for an evolutionary basis of non-adaptive theories of aging, including three main theories: Mutation accumulation (MA) and antagonistic pleiotropy (AP), which are classified as evolutionary genetic theories, and disposable soma, which involves physiology (Kirkwood, 2005).

Mutation Accumulation Theory

The mutation accumulation (MA) theory, first developed by J.B.S. Haldane (Rose, 1991) and Peter Medawar (Medawar, 1952), is among the keystones of such theories, attempting to explain aging in evolutionary context.

The idea is as follows: First, due to extrinsic mortality such as cold, diseases, and, predators, after a certain period, only a small fraction of a population could survive under a specific environment. Because in a small population the effect of natural selection is weak compared to genetic drift, individuals at advanced age are also under weak negative selection (**Figure 1.2**). Second, due to declining force of natural selection, deleterious mutations that express their harmful effects at later ages cannot be efficiently eliminated. For example, if certain deleterious mutation expresses its harmful effect before reproduction, it has a chance to affect fitness-related traits, such as reproduction or survival ability, and thus will be eliminated by natural selection before passing into next generation. However, if there exist germ-line mutations that exhibit harmful effects only at old age, negative selection against these will be inefficient. In other words, those mutations will already have passed into the next generation by the time their harmful effects become apparent.

Such old-age-expressed harmful mutations can then fix in a population, and contribute to senescent phenotypes (Kirkwood & Austad, 2000).

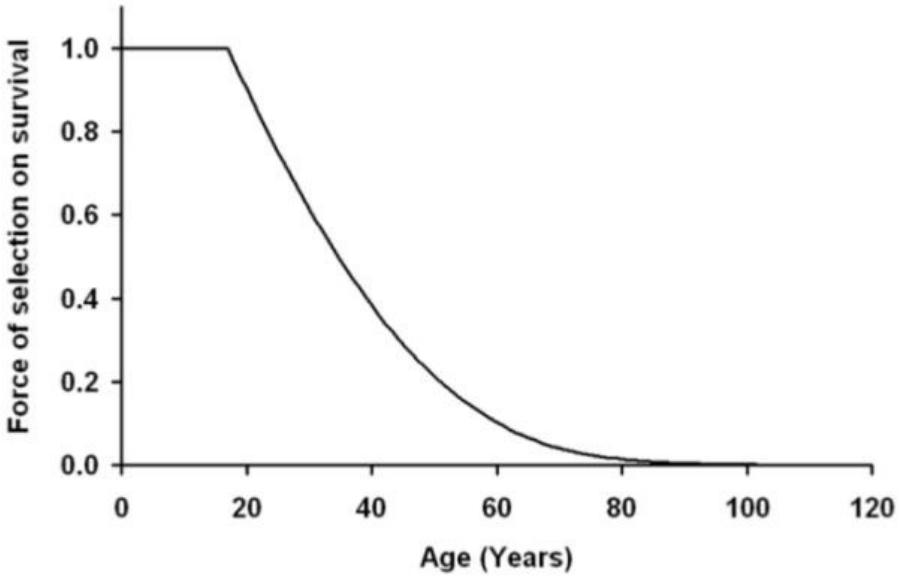

Figure 1.2: Intensity of natural selection on survival (Flatt & Schmidt, 2009).

Evolutionary genetic theories of aging, MA and AP, are based on the idea that aging is a genetically inherited and variable trait which results from germ-line mutations (*i.e.* alleles) that exhibit neutral and positive effects in early period of lifespan, respectively, but have deleterious effects at later ages.

Some predictions of MA are as follows: Genetic variance in fitness-related traits, such as reproductive success or survival, and inbreeding depression will increase with age (Flatt & Schmidt, 2009). This is because, mutations which are initially neutral and exhibit their deleterious effect after reproduction will not be effectively eliminated by natural selection, and will be reflected in phenotypic variation. They will be variable within the population because they become common due to genetic drift.

Consistent with MA, increase in genetic variation in fitness with age was shown in both laboratory populations of *D. melanogaster* (Hughes, Alipaz, Drnevich & Reynolds, 2002), and wild populations of soay sheep (*Ovis aries*) and red deer

(*Cervus elaphus*) (Wilson et al., 2007). A 2007 study, also found that rate of aging is influenced by the genetic differences among individuals within a populations of soay sheep and red deer. In addition, genetic variation changes during aging was studied in natural population of snails (*Physa acuta*). This study showed that inter-breeding between wild populations increase the chance of survival with age (Escobar, Jarne, Charmantier & David, 2008), which implies that deleterious recessive alleles that contribute to aging have fixed in different populations by genetic drift.

Antagonistic Pleiotropy

If some alleles have positive effect on fitness early in life, they would be selected even though they may express a deleterious effect at later ages (**Figure 1.3**). Due to the declining force of negative selection, those with harmful effects will not be eliminated efficiently by natural selection, and they may thus contribute the aging phenotype. Williams, who first proposed the antagonistic pleiotropy theory, acknowledged that it is hard to demonstrate pleiotropic genes that have an opposite effect on fitness at different ages (Williams, 1957). Thus, he exemplified this theory with a hypothetical mutation which has a positive effect on bone calcification early in life, but causes arterial calcification at later ages (Kirkwood, 2005; Williams, 1957).

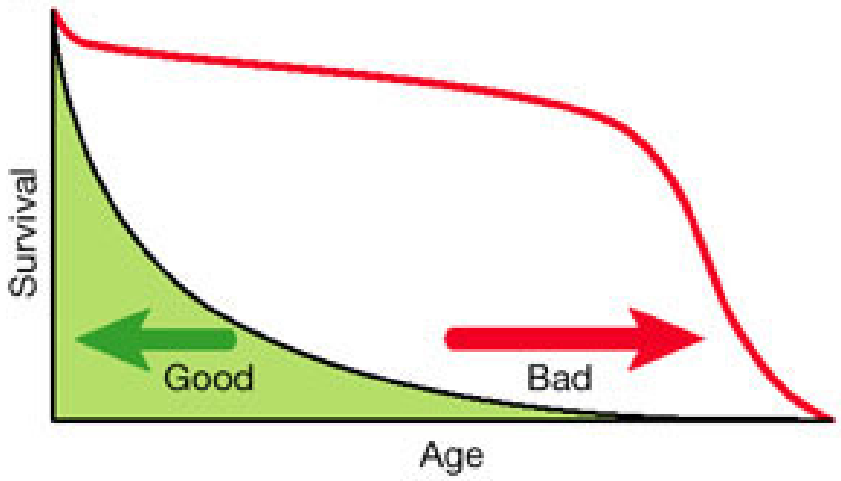

Figure 1.3: Pleiotropic genes that benefit organisms early in life will be favoured by selection even if they have bad effects at later ages (pleiotropy theory) (Kirkwood & Austad, 2000).

Several studies later proposed different candidates for AP genes. It was suggested that p53 fulfills the criteria of being a pletropic gene. p53 is a tumor suppressor which prevents cancer formation by controlling several mechanisms such as apoptosis and cell cycle arrest. Enhanced cancer protection is beneficial early in life, but may result in increase number of senescent cells, and so promotes tissue aging at later ages (Campisi, 2005; Leroi et al., 2005).

On the other hand, AP is not compatible with the observation that life-extending mutations do not cause impairment in growth or fertility. AP proposed that late life benefit of extended lifespan would be paid by early life cost of fitness-related traits, such as impairment in growth or fertility. However, worm *daf-2* mutants, and yeast *RAS2* lived more than twice as long as normal wild-type, and they did not have impaired reproduction or growth (Kenyon et al., 1993; Longo & Finch, 2003). One possible explanation is that impairment in fitness-related traits as in reproduction are hard to detect (Austad, 2004). Consistent with this explanation, for instance, worm *age-1* mutants live longer than controls, without having impairment in fertility in the laboratory, but mutant animals could not tolerate changes in food availability, a condition the species must frequently encounter in nature (Walker, McColl, Jenkins, Harris & Lithgow, 2000).

Disposable Soma

The disposable soma theory was first proposed by August Weismann in the 19^{th} century and then developed by Thomas Kirkwood in the late 20^{th} century. This theory suggests that an organism has limited energy sources, and this energy is allocated between somatic maintenance and reproduction. However, investment in somatic maintenance is a costly process, and it is only necessary for keeping the organism in sound condition to find a chance to reproduce (**Figure 1.4**). Insufficient investment in somatic maintenance, such as in DNA repair, leads to increase in cellular damage, and causes aging (Kirkwood, 2005). For instance, in nature, most of the mice die from cold in their first year. Thus, energy investment in thermogenesis or reproduction would be more advantageous than in repair mechanisms (Berry & Bronson, 1992).

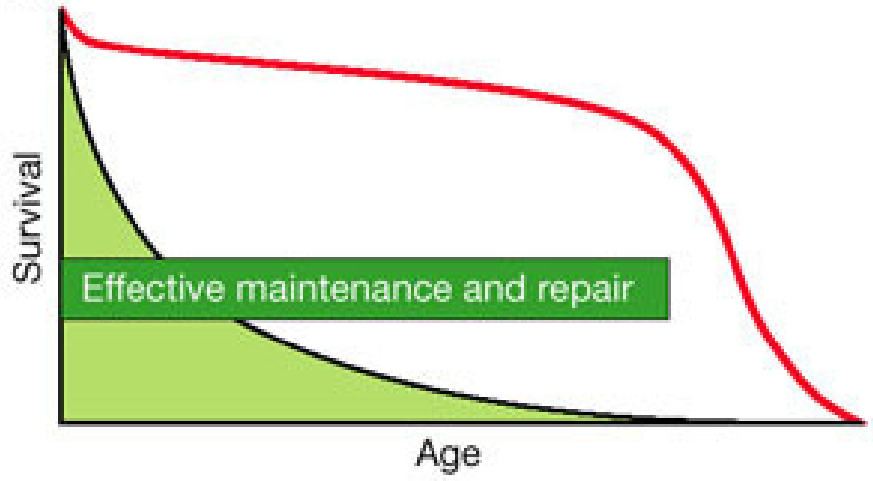

Figure 1.4: Selection pressure to invest metabolic resources in somatic maintenance and repair is limited; all that is required is to keep the organism in sound condition for as long as it might survive in the wild. This is the basis of disposable-soma theory (Kirkwood & Austad, 2000).

Efficiency of genes that regulate repair mechanisms can determine the rate of aging. If there is high investment in those genes, damage accumulates slowly, and may lead to longer lifespan. For example, long-lived rodent species *Peromyscus leucopus* expresses a high amount of antioxidant enzymes. Thus, this species produces low level of reactive oxygen species, and this leads to less oxidatively damaged proteins, than its short-lived relatives such as *Mus musculus* (Sohal, Ku & Agarwal, 1993).

Consistent with the disposable soma theory, increased longevity results in reduced fecundity in some model organisms (Longo, 2003). Furthermore, removal of the germ-line in *C. elegans* causes prolonged longevity (Arantes-Oliveira, Apfeld, Dillin & Kenyon, 2002; Hsin & Kenyon, 1999). This means there is a trade-off between extended lifespan and reproduction.

Reassessment of the disposable soma with advanced molecular techniques also provided more insight into one of its assumptions. For example, Weismann suggested that in organisms having a distinction between germ-line and somatic cells, the main role of the soma is to support and secure the germ-line. In line with this idea, Kirkwood proposed that immortal germ-line may require more energy

than soma. However, one of the recent study showed that the relationship between soma and germ-line is not unidirectional (Douglas & Dillin, 2014; Ermolaeva et al., 2013). In this study, when genotoxic stress was applied to germ-line cells in *C. elegans*, germ-line sent signals to the soma. Once the soma received the signal, it increased the protein turnover rate, and became more resistant to stress.

Adaptive Theories

In addition to non-adaptive theories, a number of ideas have been proposed suggesting that aging has evolved as an adaptation. August Weismann himself proposed the basic programmed theory at the beginning of the 19^{th} century (Weismann, 1889), and it was later revised by others (Longo et al., 2005; Skulachev, 1997). According to this theory, aging is a genetically programmed death mechanism that benefits the species for several reasons. First, it limits the population size by eliminating old individuals. Second, reproductively exhausted individuals sacrifice themselves to secure the turnover of next generations. Finally, aging is directly regulated/controlled by genes which are favored by natural selection. However, arguments for the programmed theory include misconceptions about evolution, as I will discuss below:

1. **"Aging limits the population size" :** The adaptive or programmed theory of aging supports the idea that aging serves as a death mechanism which limits population size by eliminating old individuals. However, in the wild, a relatively low number of individuals can reach "old age" due to extrinsic mortality, such as cold, predation and diseases that affect every individual irrespective of age (Medawar, 1952). Physiologically deteriorated old individuals are hence rarely seen in nature. Thus, these evidences may indicate that aging (*i.e.* intrinsic mortality) in nature is not a fundamental force limiting population size.

2. **"Aging secures the turnover of generations :** The idea here is that elimination of old individuals increases the turnover rate of generations, so species can more easily adapt to the changing environment. Moreover, there would be more resources, such as space and food for their younger kin. However, today it is generally accepted that selection acts on the individual level rather

than the group level (Kirkwood & Melov, 2011), except in some cases (see below). Thus, old individuals sacrificing themselves for the benefit of the group will decrease the probability of any "old-age sacrifice alleles" being passed to the next generation, and such alleles will thus be eliminated.

But there can be exceptions. One theoretical study showed that the possibility of evolution of programmed aging is determined by the range of the population. If the population is spatially distributed, evolution can favor the aging gene that eliminates even reproductively active individuals (Travis, 2004). Thus, in some situations, selection at the group level can be (at least theoretically) more important than the individual level.

The perhaps most striking example of programmed aging is the sudden death of Pacific salmon. Pacific salmon is a semelparous species, *i.e.* it usually reproduces only once in its life cycle. After reproduction, the cortisol level increases in the circulating system, and this leads to organ failure degenerating multiple tissues (McQuillan, Lokman & Young, 2003). According to the programmed theory, the death of the organism has a genetic basis as in development, and may help the offspring. However, this is not necessarily true. Semelparity has evolved in environments that give organisms relatively small chance of reproducing again. Thus, all resources have been allocated to optimize growth and one-time reproduction, and after reproduction, there is no sufficient resource left to maintain homeostasis. This is because, investment in post-reproductive survival has no value (Kirkwood & Cremer, 1982). Thus, the reason for the sudden death of an organism can be seen as the abrupt decline in force of natural selection after reproductive period. The link between reproduction and rapid senescence can be seen in multiple examples. For instance, removal of the gonads increases the lifespan of Pacific salmon (Robertson, 1961). Similar cases can be observed in semelparous plant species. Soybean plants (*Glycine max*) die after the riping of their seeds. When the seeds are removed, extension of the lifespan is observed (Leopold, Niedergang-Kamien & Janick, 1959). Thus, these extreme examples can be readily explained by the disposable soma theory (Kirkwood, 2005).

3. **"Aging is directly regulated/controlled by genes" :** Manipulations in genetic pathways and genes extend lifespan in diverse organisms from yeast to mice (Kenyon, 2010). High similarities between those pathways that regulate/extend lifespan among species is one of the main arguments for existence of programmed aging (Longo et al., 2005). In addition, several single-gene mutations have been identified resulting in significant increase in *C.elegans* lifespan. If certain mutations in certain genes greatly extend the species' lifespan, those genes may be programming aging. However, interpreting these results as evidence for programmed aging may erroneous. First, if only certain genes directly drive the aging process, it would be expected to be able to slow the aging rate or postpone entire aging easily. Thus, complexity of aging phenotype may be explained better by the contribution of multiple genes.

   In addition, single gene manipulations have been conducted in laboratory animals, and there is little known about the survival ability of those animals under severe circumstances (Austad, 2004). For example *age-1* is a gene in *C. elegans* that can cause life extension when mutated, but without causing impairment in fertility, a frequently seen side effect of life-extending interventions. However, it was found that mutant animals could not tolerate changing in food availability as in nature (Walker et al., 2000). So, it is important to carefully examine the possible side-effects of life-extending interventions, before reaching general conclusions and/or applying these to humans.

Overall, arguments supporting programmed aging, including exceptional situations, appear insufficient to explain existence of such programs. Therefore, most biologists today consider aging as caused by non-adaptive evolutionary processes.

1.3 Research Objectives

As discussed earlier, the role of the MA process in aging was previously tested by studying changes in phenotypic (fitness) variation with age and testing for one prediction of MA: higher variance during aging. But the theory awaits testing by new approaches. One such approach could be using transcriptome data. Transcriptome studies of aging have usually focused on identifying functional groups affected by senescence and/or underlying senescence, but such data is not traditionally em-

ployed for testing evolutionary theories.

In previous work, a group of researchers had used pre-frontal cortex transcriptome age-series from humans and tested whether protein sequence conservation varies among genes highly expressed at different ages (Somel et al., 2010). This analysis showed that highly expressed genes in young adults are more highly conserved than those in old individuals. This is consistent with MA theory, which postulates that genes that are relevant for old age should be under less purifying selection, and therefore appear evolutionarily less conserved. However, this work involved limited sample sizes, and only one brain region, and one species.

Here I study the prevalence of MA effects on the aging transcriptome, by conducting a meta-analysis of mammalian aging datasets, including 8 different tissue types and 4 mammalian species (n = 35 total datasets, n = 768 total individuals). I test the hypothesis that MA will be reflected in age-related decrease in conservation of the transcriptome with age (ADICT), in other words, I expect to observe lower evolutionary conservation in protein coding sequence among genes expressed at higher levels in old adults, relative to young adults. This would happen due to slightly harmful mutations fixing in genes more relevant for old adults than for young adults.

CHAPTER 2

MATERIAL AND METHODS

2.1 Datasets Information

To assess the relationship between gene sequence conservation and gene expression changes during aging, I collected 35 published age-series microarray datasets from different brain regions (from humans, macaques, rats, and mice), muscle (from humans, rats, and mice), aorta (from macaques and rats), skin (from humans and mice), and kidney, liver, lung, and spleen from mice.

Across all analyzed datasets, human ages range from 16 to 106 years, macaques ages range from 4 to 28 years, rats ages range from 3 to 30 months, and mice ages range from 8 to 130 weeks.

I limited the study to studies with sample sizes above 10 to ensure sufficient power. I also used only microarray experiment data, as these are most abundant in public repositories and can be readily processed and compared.

The analyzed datasets includes 13 different Affymetrix platforms. Information about the all analyzed datasets is presented in **Table 2.1**.

Table 2.1: Information about the all analyzed datasets.

Species & Tissue	Age Range	# of Samples	Platform
Hs_Brain_BA22	25 - 94 yrs	19	HG-U133_Plus_2
Hs_Brain_EC	20 - 97 yrs	35	HG-U133_Plus_2
Hs_Brain_PCG	20 - 99 yrs	39	HG-U133_Plus_2
Hs_Brain_HC	20 - 99 yrs	41	HG-U133_Plus_2
Hs_Brain_SG	20 - 99 yrs	44	HG-U133_Plus_2
Hs_Brain_FC	26 - 106 yrs	29	HG_U95Av2
Hs_Brain_BA10	25 - 94 yrs	23	HG-U133_Plus_2
Hs_Brain_PFC	17 - 98 yrs	12	HuGene-1_0-st
Rm_Brain_PFC	4 - 28 yrs	11	HuGene-1_0-st
Hs_Brain_CB	23 - 98 yrs	11	HuGene-1_0-st
Rm_Brain_CB	4 - 28 yrs	9	HuGene-1_0-st
Rn_Brain_HC	7 - 22 mo	14	Rat230_2
Rn_Muscle_EOMs	6 - 30 mo	12	RG_U34A
Rn_Muscle_EDL	6 - 30 mo	11	RG_U34A
Rn_Aorta_T	3 - 28 mo	16	RG_U34A
Rm_Aorta	6.6 - 21.2 yrs	24	HG-U133A_2
Hs_Muscle_BB	19 - 76 yrs	19	HG-U133_Plus_2
Hs_Muscle_VL1	21 - 75 yrs	15	HG-U133A&B
Hs_Muscle_VL2	20 - 71 yrs	15	HG-U133A&B
Hs_Muscle_Other	16 - 89 yrs	15	HG-U133_Plus_2
Hs_Muscle_RA	16 - 89 yrs	62	HG-U133_Plus_2
Hs_Skin	19 - 86 yrs	98	HuEx-1_0-st
Mm_Muscle_PC	2 - 24 mo	14	Mouse430_2
Mm_Brain_HC	2 - 15 mo	22	MG_U74Av2
Mm_Lung_1	2 - 26 mo	9	Mouse430_2
Mm_Brain_NC	5 - 30 mo	10	Mouse430_2
Mm_Liver_1	8 - 130 wks	13	Mouse430_2
Mm_Skin	5 - 30 mo	20	Mouse430_2
Mm_Muscle_Gastro	5 - 25 mo	10	MOE430A
Mm_Brain_WB	13 - 130 wks	14	Mouse430_2
Mm_Kidney	13 - 130 wks	18	Mouse430_2
Mm_Liver_2	13 - 130 wks	15	Mouse430_2
Mm_Lung_2	13 - 130 wks	18	Mouse430_2
Mm_Spleen	13 - 130 wks	16	Mouse430_2
Mm_Liver_3	6 - 24 mo	15	Mouse430_2

2.2 Normalization

Affymetrix .CEL files from 27 datasets were downloaded from NCBI GEO (www .ncbi.nlm.nih.gov/geo/) (Edgar, Domrachev & Lash, 2002) and EBI Array Express (www.ebi.ac.uk/arrayexpress/) (Kolesnikov et al., 2014). These raw datasets were processed using the Bioconductor “affy” package “expresso” function. The selected options for the “expresso” function were: “rma” for background correction, "quantiles" for normalization, and “medianpolish” for summarization.

Some studies did not provide the raw data so preprocessed values were used. Preprocessed series matrix files (for n = 8 datasets) were downloaded from NCBI GEO and transformed using log2 transformation on the gene expression levels. One of the preprocessed dataset (GSE18876) is normally distributed, so I did not apply any normalization procedure. Normalization methods for the analyzed datasets is presented in **Table 2.2**.

2.3 Probeset Conversion

Affymetrix probe set IDs were converted to Ensembl gene IDs using the Bioconductor "biomaRt" package (Durinck et al., 2005) “useMart” function to select the dataset for the species of interest, and the “getBM” function to retrieve the Ensembl gene IDs corresponding to Affymetrix probe set IDs. I then followed 2 steps: **(a)** if one probe set corresponds to more than one Ensembl gene, I remove them; **(b)** if more than one probe set corresponds to one Ensembl gene, I take the probe set which has a maximum expression value across samples. This is because, that probe set may contain more exons than others, and in that case I consider that probe set the best representative of the gene.

2.4 Outlier Individuals

I next determined whether the datasets include outlier individuals (**Table 2.2**) using principal component analysis (PCA) in each dataset. Principal component analysis (PCA) was conducted using "prcomp" function “scale=T” parameter which is implemented in the “stats” package in R. If samples in the datasets are clustered to-

gether, and there is a distance between a particular sample and the rest, that sample was determined as an outlier. When an outlier was detected, I repeated the preprocessing procedure described above after discarding the outlier sample.

Table 2.2: Normalization methods and outlier information about the analyzed datasets. Method 1, including RMA, log2 transformation and quantile normalization, was applied to the raw datasets; Method 2, including only log2 transformation, was applied to the preprocessed datasets.

GEO/Array Acc	Norm. Methods	Outliers
GSE21935	Method 1	-
GSE11882 (EC)	Method 1	GSM300192, GSM300196, GSM300228, GSM300300
GSE11882 (PCG)	Method 1	GSM300198, GSM300212, GSM300287, GSM300326
GSE11882 (HC)	Method 1	GSM300255, GSM300301
GSE11882 (SG)	Method 1	GSM300213, GSM300250, GSM300288, GSM318840
GSE1572	Method 1	X48F.CEL
GSE17612	Method 1	-
GSE22521 (Hs)	Method 1	-
GSE22521 (Rm)	Method 1	-
GSE22569 (Hs)	Method 1	-
GSE22569 (Rm)	Method 1	GSM560170
GSE20219	Method 1	GSM506979
GSE3309 (EOMs)	Method 2	-
GSE3309 (EDL)	Method 2	GSM74440
GSE7281	Method 1	-
GSE6599	Method 1	-
GSE38718	Method 1	GSM948631, GSM948637, GSM948640
GSE362	Method 2	-
GSE674	Method 2	-
GSE5086 (Other)	Method 2	GSM114704, GSM114706, GSM114719, GSM114720
GSE5086 (RA)	Method 2	-
GSE18876	No prepro.	-
GSE50821	Method 1	-
GSE5078	Method 2	GSM114424
GSE6591	Method 1	-
GSE8150	Method 1	-
E-MEXP-839	Method 1	X33.CEL
GSE35322	Method 1	-
GSE6323	Method 1	-
GSE34378 (Brain)	Method 1	GSM847851, GSM847853, GSM847857, GSM847865
GSE34378 (Kidney)	Method 1	-
GSE34378 (Liver)	Method 1	GSM847798, GSM847810, GSM847812
GSE34378 (Lung)	Method 1	-
GSE34378 (Spleen)	Method 1	GSM847828, GSM847830
GSE21716	Method 1	GSM541754

2.5 Age test

Genes showing age-related changes in expression levels was identified using the Spearman correlation test, and p-values were corrected for multiple testing using the "p.adjust" function with the "Benjamini-Hochberg (BH)" method in R, yielding q-values, a measure of false discovery rate. I used a q-value cutoff $q < 0.10$, which is a commonly used threshold (Hartmann et al., 2009). To avoid statistical power reduction due to type II error in datasets with low number of age-related genes (n$<$ 50), those datasets were discarded from the analyses. The number of the age-related and all detected genes is presented in **Table 2.3**.

To overcome the problems related to conducting meta-analysis -each dataset displaying unique and possibly non-normal distributions, outliers having large effects on analyses, etc.- nonparametric Spearman correlation test was preferred. Variables, *i.e.* gene expression level and age, are converted to ranks in the Spearman correlation test, so that the test is not affected by outliers or the exact shape of the data distribution, unlike parametric tests.

2.6 Direction of the age-related changes in expression level

To find direction of gene-expression changes with age, I first chose genes showing age-related changes (method= "BH", $q - value < 0.10$). I then separated these genes into 2 categories using Spearman correlation coefficient (rho): **(a)** genes showing age-related increase: $q < 10$ and $rho > 0.1$; **(b)** genes showing age-related decrease: $q < 0.10$ and $rho < -0.1$. Categorizing genes depending only on the q-value would be affected by the sample size (power) of each dataset, so using the correlation coefficient (rho) as cutoff I aimed to avoid such possible bias. Genes having $q - value > 0.10$ were selected as having no change in expression level with age. Genes with $q < 0.10$ and $|rho| < 0.1$ were discarded from the analysis.

Table 2.3: The number of age-related and all detected genes calculated for the 35 datasets.

Species & Tissue	**# of age-related genes**	**# of all genes**
Hs_Brain_BA22	28	21323
Hs_Brain_EC	488	21323
Hs_Brain_PCG	5026	21323
Hs_Brain_HC	2144	21323
Hs_Brain_SG	7799	21323
Hs_Brain_FC	720	8258
Hs_Brain_BA10	2354	21323
Hs_Brain_PFC	213	1200
Rm_Brain_PFC	4	12002
Hs_Brain_CB	4	12467
Rm_Brain_CB	0	12467
Rn_Brain_HC	216	11555
Rn_Muscle_EOMs	1	3906
Rn_Muscle_EDL	0	3906
Rn_Aorta_T	347	3905
Rm_Aorta	24	12187
Hs_Muscle_BB	3879	21323
Hs_Muscle_VL1	5	18517
Hs_Muscle_VL2	1	18517
Hs_Muscle_Other	37	21323
Hs_Muscle_RA	6	21323
Hs_Skin	91	14356
Mm_Muscle_PC	616	1829
Mm_Brain_HC	48	7775
Mm_Lung_1	318	18299
Mm_Brain_NC	7152	18299
Mm_Liver_1	449	18369
Mm_Skin	2201	18299
Mm_Muscle_Gastro	1046	12330
Mm_Brain_WB	244	18299
Mm_Kidney	657	18299
Mm_Liver_2	956	18299
Mm_Lung_2	403	18299
Mm_Spleen	383	18299
Mm_Liver_3	850	18299

2.7 Conservation ratio

I used different types of metrics to determine the negative selection pressure on the protein coding sequence. First, I calculated conservation in protein coding regions between pairs of species with different evolutionary distance between them, using dN (nonsynonymous substitution) and dS (synonymous substitution) statistics downloaded from Ensembl Biomart (v.83). Here I used "one-to-one orthologs" between human-mouse, human-elephant, and human-cow, in order to identify whether evolutionary distance between species affect the conservation in protein coding sequence. Note that dN/dS ratios can measure both the strength of negative and of positive selection.

To exclude the genes possibly under positive selection, I repeated the analysis only using genes with $dN/dS < 1$. In addition to this, GO categories and subcategories related to immune system genes ("GO:0008150") were selected using the R "get" function. I then repeated the analysis after discarding these genes.

Second, I used ω_0, which is the protein sequence conservation statistic calculated for each Ensembl gene, and downloaded from the Selectome database (Moretti et al., 2014). The statistic is based on coding sequence alignments across mammalian species and is estimated for the Homininae branch for human and Murinae branch for mouse, using the branch-site model (Nielsen et al., 2005). The branch-site model is one of the methods to estimate different dN/dS among branches and among sites. A branch of interest is called the "foreground branch". Other branches in the phylogenetic tree are called "background" branches, which have the same distribution of dN/dS value among sites. On the other hand, different values can apply to the foreground branches. The branch-site model thus estimates selective pressure -positive or negative selection- on a protein coding gene sequence, and I used here dN/dS ratio calculated for the sites determined to be under negative selection. Thus, ω_0 is expected to be a measure of the strength of negative selection on a gene.

This measure of ω_0 can vary among genes due to multiple factors that are not the focus of this study. To disentangle the effects of such factors from ω_0, I used information on GC content, CDS length, intron length, intron number, mean expression,

median expression, maximum expression, tissue specificity, network connectivity, phyletic age, number of paralogs, which were directly obtained from the Supplemental Material of Kryuchkova-Mostacci & Robinson-Rechavi, 2015. To remove the effect of these variables from ω_0, I used residuals from a multiple regression model where ω_0 is the response variable. Here I used the "lm" and "residuals" functions and in the R "stats" package. I call the resulting protein conservation statistic, which are the residual from this model, as ω_0*.

The ω_0* statistic was calculated separately for human and for mouse ω_0 values. I used the human ω_0* data in analyses involving primate transcriptome datasets, and the mouse ω_0* data in analyses involving rodent transcriptome datasets.

2.8 Bootstrapping

Bootstrapping is a nonparametric method to calculate confidence interval and standard errors. Bootstrapping was performed with the "sample" function with "replacement=TRUE" option in R. I used separately bootstrapping to calculate 95% confidence intervals for the mean ω_0* among genes showing **(a)** age-related increase in expression level, **(b)** age-related decrease in expression level, and **(c)** no age-related change in expression level. For each case I resampled genes for 1000 times, and calculated the mean. To visually compare the ω_0* among datasets, I then subtracted the median for genes showing no age-related change, from genes showing age-related increase or age-related decrease. The upper and lower 2.5% quantiles were plotted in **Figure 3.3**.

2.9 Consistency among datasets

To determine consistency in gene expression level distributions, I first chose common genes (n = 1119) detected among all 35 datasets. The aim of the consistency analyses is to check whether the same type of biological samples (tissues and species) exhibit similar expression patterns and cluster together. If not, this can indicate artifacts in the experimental data.

For this, all rodent genes were mapped to human genes using "one-to-one orthologs" as defined by Ensembl (v.83). I then merged datasets and calculated the youngest 5 individuals' mean gene expression level. After reprocessing the merge

file using “quantile” normalization, I plotted the PCA using "prcomp" function “scale=T” parameter (**Figure 2.1**).

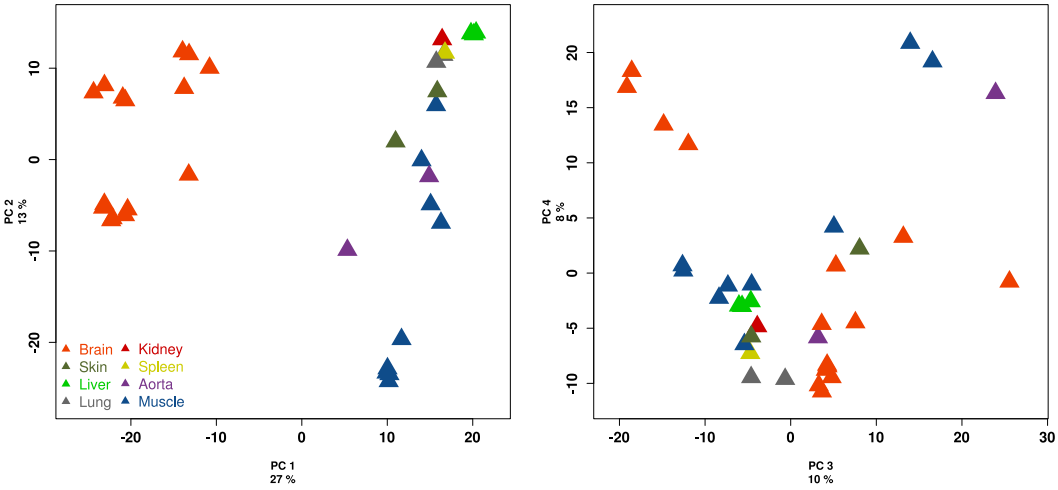

Figure 2.1: Mean gene expression level of the youngest 5 individual from each of the datasets.

2.10 Gene Ontology Analysis

I developed a statistic, z, which can capture the relative expression level and relative conservation of a gene simultaneously:

$$z = x^2 - y^2,$$

where x stands for the rank of a gene’s expression level across all detected genes in a dataset, and y stands for the rank of ω_0*. When z is high, this indicates genes having relatively high expression and low conservation, and when z is low this indicates genes having relatively low expression and high conservation. After sorting z values, the first 10% of genes were selected as increasing expression and low conservation (IELC) and last 10% were selected as decreasing expression and high conservation (DEHC).

To find functional groups driving the ADICT pattern in tissues, I conducted Gene Ontology (GO) analyses for 3 GO domains - Biological Process (BP), Cellular

Component (CC), and Molecular Function (MF) - performing Fisher's exact test using the "fisher.test" function in the R "stats" package (GO data provided by Handan Melike Dönertaş). The GO groups having higher than 1.5 odds ratio were selected as the enriched groups. Here, 1.5 is a random threshold, and the significance of the enrichment (how many GO groups have odds ratio > 1.5) was calculated using permutations. Specifically, I randomized ages of individuals in each dataset by conducting 1000 permutations, calculated expression correlations with age, and repeated the GO analysis using these correlation values.

CHAPTER 3

RESULTS

3.1 Age-related decrease in conservation of the transcriptome (ADICT)

To test the ADICT hypothesis, I first studied the correlation between protein sequence conservation and gene expression levels for each individual in the analysis. In each dataset, I used two non-exclusive gene sets: **(a)** genes showing significant age-related change in expression levels (at Spearman correlation test $q - value < 0.10$), and **(b)** all expressed genes. I conducted analyses using all expressed genes in order to avoid statistical power reduction due to type II error in datasets with low sample sizes, and further to determine the overall trend across the transcriptome (**Table 2.3**). Note that in 12 of 35 datasets, I could not identify a set of significant age-related gene set at $q < 0.10$ (see **Section 2.5**).

To measure conservation, I used protein sequence conservation data (ω_0) from the Selectome database (Moretti et al., 2014), based on coding sequence alignments across mammalian species and estimated for the human branch or the mouse branch using the branch-site model (Zhang, Nielsen & Yang, 2005) (see **Section 2.7**). This is the dN/dS ratio calculated for the sites determined to be under purifying selection, and thus is expected to be a direct measure of the strength of purifying selection on a gene. I further calculated a corrected protein conservation statistic (ω_0*) for each gene, disentangling conservation measurements from the effects GC content, CDS length, intron length, intron number, mean expression, median expression, maximum expression, tissue specificity, network connectivity, phyletic age, and number of paralogs, using a multiple regression model following

(Kryuchkova-Mostacci & Robinson-Rechavi, 2015).

I then calculated the Spearman correlation coefficient between gene expression level and ω_0* per individual in each dataset (**Figure 3.1a** & **Figure 3.1b**). A positive correlation indicates that genes more highly expressed in an individual's transcriptome tend to be more conserved in their protein sequence. Although a positive correlation between expression and sequence conservation is generally expected (Pál, Papp & Lercher, 2006), the level may vary among individual transcriptomes. Hence I asked whether the expression - ω_0* correlation is dependent on individual age, again using the Spearman correlation test. **Figure 3.1c** provides an example of such a correlation in one brain aging dataset (Lu et al., 2004), and **Figure 3.2** shows the results across all datasets.

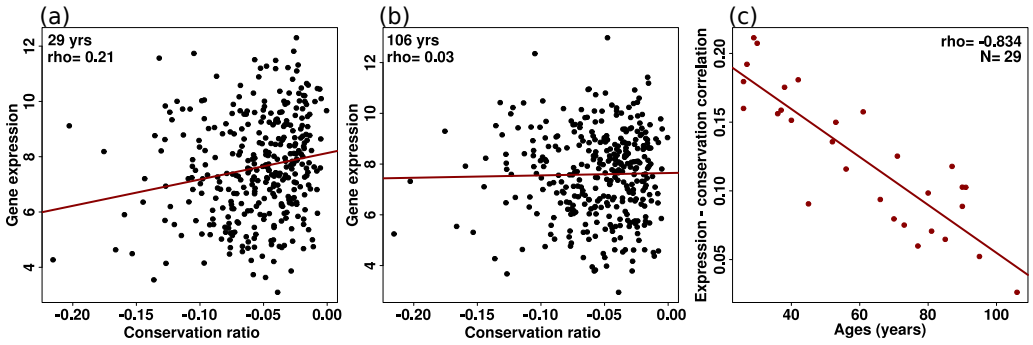

Figure 3.1: Relationship between gene expression level and protein conservation. Examples of gene expression level – protein conservation statistic (ω_0*) correlations **(a)** for a 29 year-old human, and **(b)** for a 106-year old human, in the cerebral frontal cortex (Lu et al., 2004). The analysis includes only age-related genes detected in this dataset ($q < 0.10$). Each point represents a gene (n = 345). The x-axis shows the conservation statistic, ω_0*. The y-axis shows gene expression level. The expression-conservation (ω_0*) Spearman correlation coefficients are indicated in the inset. **(c)** Age-dependent change in expression-conservation (ω_0*) correlation in the human frontal cortex, based on age-related genes in the same dataset as panels (a) and (b). The y-axis shows expression-conservation Spearman correlation coefficient calculated for each individual in this dataset (n = 29). The x-axis shows the age of an individual. The Spearman correlation coefficient between age and expression-conservation correlation is indicated in the inset.

In the brain, I found negative correlation between ω_0* and age among all significant

results (Spearman correlation test, $p < 0.05$) involving age-related genes. Specifically, among the 10 brain datasets where I had identified a significant age-related gene set, 9 showed significant expression-conservation correlation with age, and all were negative. Repeating this analysis with all expressed genes, 11 out of 15 datasets showed significant correlation, and all were again negative, indicating a general trend of ADICT in the brain.

In contrast to the brain, across the multiple datasets of various muscle types (n = 11), I found no common ADICT pattern. In mouse gastrocnemius muscle and rat thoracic aorta I even observed the opposite of what I had hypothesized: a significant age-related increase in transcriptome conservation. I also note that in 8/11 muscle datasets, I did not detect significant age-related expression change (*i.e.* n $<$ 50 at $q < 0.10$; **Appendix B**).

Finally, in different datasets for mouse liver (n = 3) and lung (n = 2) I found the same significant ADICT trend, as in brain, whereas skin datasets (n = 2) showed no significant trend. Thus, in brain and possibly in some other tissues, such as liver and lung, I find a consistent ADICT trend, whereas in other tissues, such as muscle, ADICT may not exist.

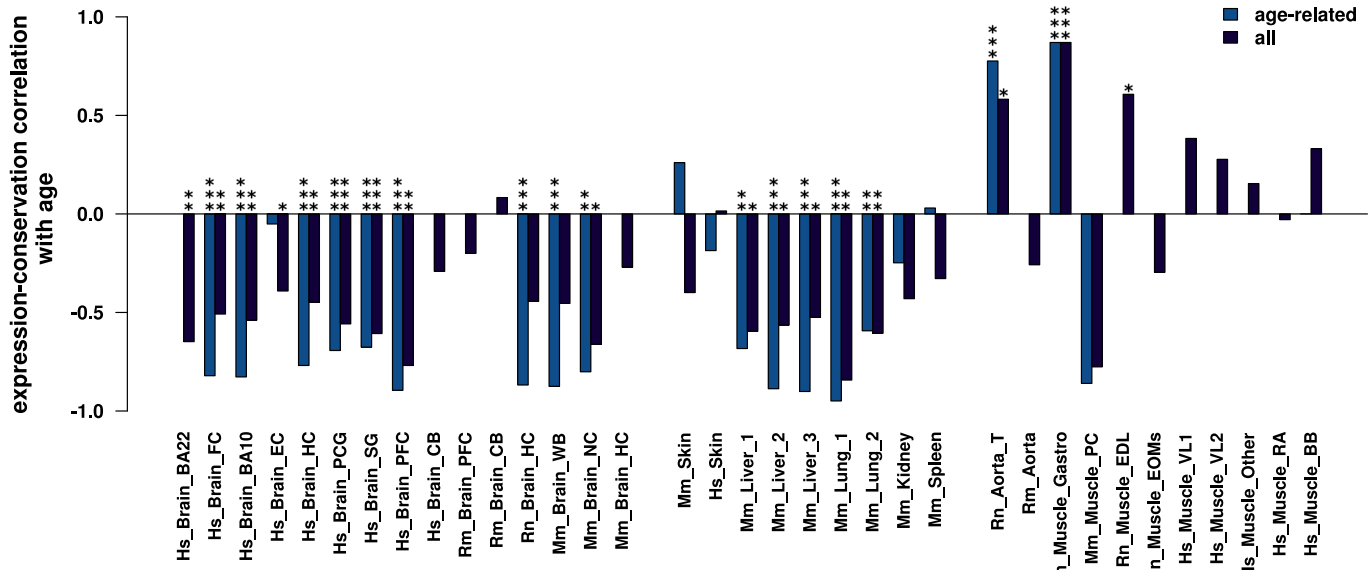

Figure 3.2: Age-dependent changes in transcriptome conservation (BA22: Superior temporal cortex; EC: Entorhinal cortex; PCG: Postcentral gyrus; HC: Hippocampus; SG: Superior frontal gyrus; FC: Frontal cortex; BA10: Anterior prefrontal cortex; PFC: Prefrontal cortex; CB: Cerebellar cortex; EOMs: Extraocular muscles; EDL: Extensor digitorum longus muscles; T: Thoracic; BB: Biceps brachii; VL: Vastus lateralis; RA: Rectus abdominis; PC: Precursor cells; NC: Neocortex; WB: Whole brain; Gastro: Gastrocnemius tissue; Other: Various other regions of the anatomy; Hs: *Homo sapiens*; Rm: *Rhesus macaque* & *Macaca fascicularis*; Rn: *Rattus norvegicus*; Mm: *Mus musculus*).

To determine the robustness of ADICT with respect to the protein-coding sequence conservation metric chosen, I repeated the analysis using ω_0 values without multiple regression, as well as ω (dN/dS) values obtained from Ensembl for 3 species pairs (see **Section 2.7**). I further tested whether the trend holds when I exclude **(a)** putatively positively selected genes (with $\omega > 1$ in my data), **(b)** immune system genes known to be generally fast evolving (Nielsen et al., 2005), and **(c)** genes showing age-related expression down-regulation in a dataset (expression-age rho $<$ -0.1). ADICT was consistently detected across datasets the same brain, liver and lung datasets, irrespective of the metric used and the gene sets involved (**Appendix A-H**).

3.2 Distinct processes contribute to ADICT

I next studied two non-exclusive scenarios that could lead to ADICT: **(a)** genes showing age-related increase in expression levels could have low ω_0*, consistent with MA, **(b)** genes showing age-related decrease in expression levels could have high ω_0*, relative to genes showing no change in expression. The latter could be expected if a set of highly conserved genes (e.g. synaptic genes) show down-regulation during postnatal lifespan, as reported earlier (Somel et al., 2010).

To test which of these scenarios underlie ADICT, I compared the mean ω_0* among **(a)** genes showing increase in expression with age, **(b)** genes showing decrease in expression with age, with **(c)** genes showing no change in expression level, as control. This analysis was repeated across the 14 datasets where I had detected significant ADICT using age-related genes. I found results consistent with both scenarios (**Figure 3.3**): genes showing increase in expression with age had on average lower conservation in nearly all cases (n = 13/14, 8 of these with significant with bootstrap support $>$95%). The only exception was one mouse neocortex dataset. Meanwhile, genes showing decrease in expression with age also had on average higher conservation in most cases (n = 12/14, 6 of these significant with bootstrap support $>$95%) than genes with no change. The exceptions to higher conservation in genes showing decreased expression with age were two mouse liver datasets. Thus, I conclude that both processes could be contributing to ADICT.

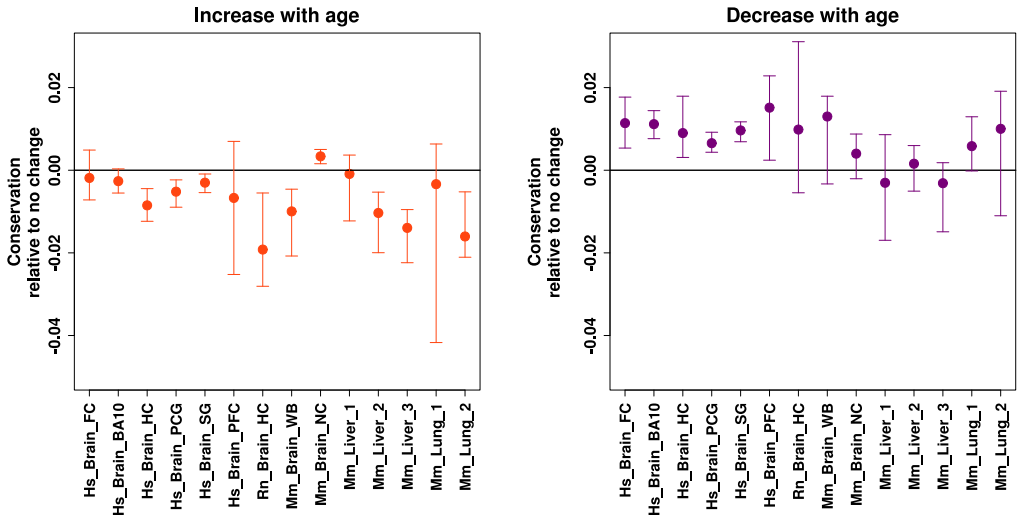

Figure 3.3: Comparison of ω_0* among gene sets showing different age-related expression level change patterns. The plots show mean ω_0* for genes showing age-related increase (left) and age-related decrease (right) in expression level, compared to mean ω_0* among genes showing no significant age-related change in expression level. The error bars indicate 95% confidence intervals calculated by 1,000 bootstraps (FC: Frontal cortex; BA10: Anterior prefrontal cortex; HC: Hippocampus; PCG: Postcentral gyrus; SG: Superior frontal gyrus; PFC: Prefrontal cortex; NC: Neocortex; WB: Whole brain; Hs: *Homo sapiens*; Rn: *Rattus norvegicus*; Mm: *Mus musculus*).

3.3 Functional analysis of ADICT

To find functional groups driving the ADICT pattern in brain, lung, and liver, I conducted Gene Ontology (GO) analysis for the 3 GO domains. Given the earlier result, I separately analyzed **(a)** genes showing increased expression with age and with low conservation (IELC), and **(b)** genes showing decreased expression with age and with high conservation (DEHC). I sought for common GO categories enriched in IELC genes and in DEHC genes (with odds ratio > 1.5) in all 14 datasets that showed significant ADICT. I further sought for common GO categories across different datasets for the same type of tissue (**Table 3.1**, **Table 3.2**, **Table 3.3**). The significance of the results were assessed using random permutations of individual ages (see **Section 2.10**).

Table 3.1: Biological Process

	IELC	**IELC**	**DEHC**	**DEHC**
Tissue	**# GO groups**	p	**# GO groups**	p
All	2	< 0.001	3	n.s.
Brain	11	n.s.	11	n.s.
Liver	238	0.028	79	n.s.
Lung	117	n.s.	213	n.s.

Table 3.2: Cellular Component

	IELC	**IELC**	**DEHC**	**DEHC**
Tissue	**# GO groups**	p	**# GO groups**	p
All	1	< 0.001	-	-
Brain	12	0.012	15	< 0.001
Liver	18	n.s.	11	n.s.
Lung	12	n.s.	8	n.s.

Table 3.3: Molecular Function

	IELC	**IELC**	**DEHC**	**DEHC**
Tissue	**# GO groups**	p	**# GO groups**	p
All	-	-	1	n.s.
Brain	-	-	11	0.016
Liver	-	-	21	n.s.
Lung	-	-	37	0.001

I found 2 GO Biological Process categories, "apoptotic signaling pathway" and "extrinsic apoptotic signaling pathway", enriched for IELC across all 14 datasets (permutation test $p < 0.001$) (**Appendix I**).

In the 3 mouse liver datasets, 238 GO categories ($p = 0.028$) were enriched for

IELC, mainly including immune system and metabolism-related genes (**Appendix J**). **Figure 3.4** contains a summary of these results as provided by REVIGO (Supek, Bošnjak, Škunca & Šmuc, 2011). For other tissues, or for DEHC I did not find a common significant enrichment for Biological Process categories based on the permutation test.

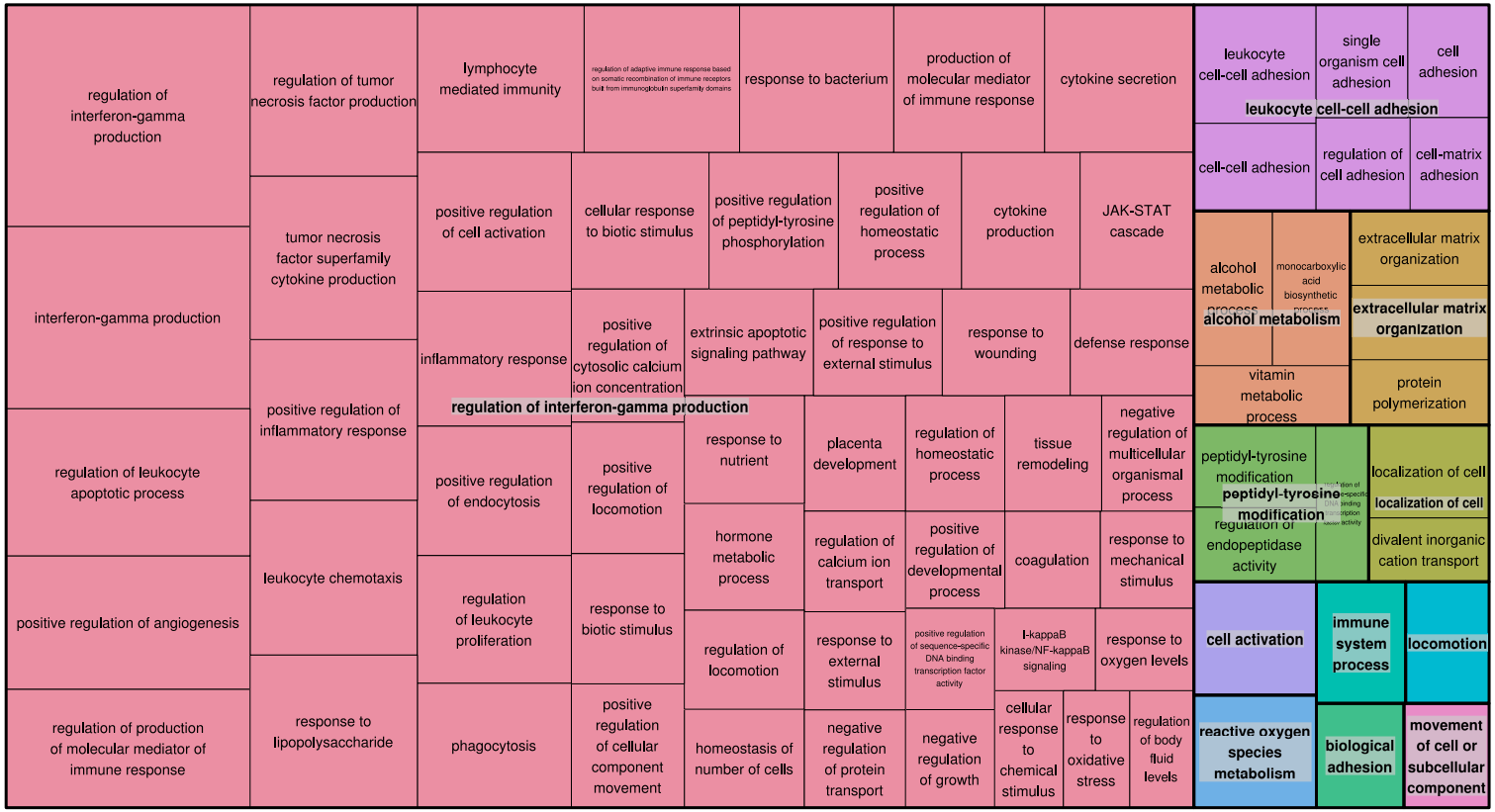

Figure 3.4: GO Biological Process enrichment result for genes showing IELC pattern in liver datasets, summarized by REVIGO.

In Cellular Component, "external side of plasma membrane" was enriched for IELC across all 14 datasets ($p < 0.001$) (**Appendix K**).

In different datasets for primate and rodent brains (n = 9), 12 GO categories ($p = 0.012$) were enriched for IELC, including "vesicle", "external membrane", and related categories (**Figure 3.5** and **Appendix L**). Meanwhile, among the 9 brain datasets, 15 GO categories ($p < 0.001$) were enriched for DEHC including "synapse", "transporter complex", and related functions (**Figure 3.6** and **Appendix M**).

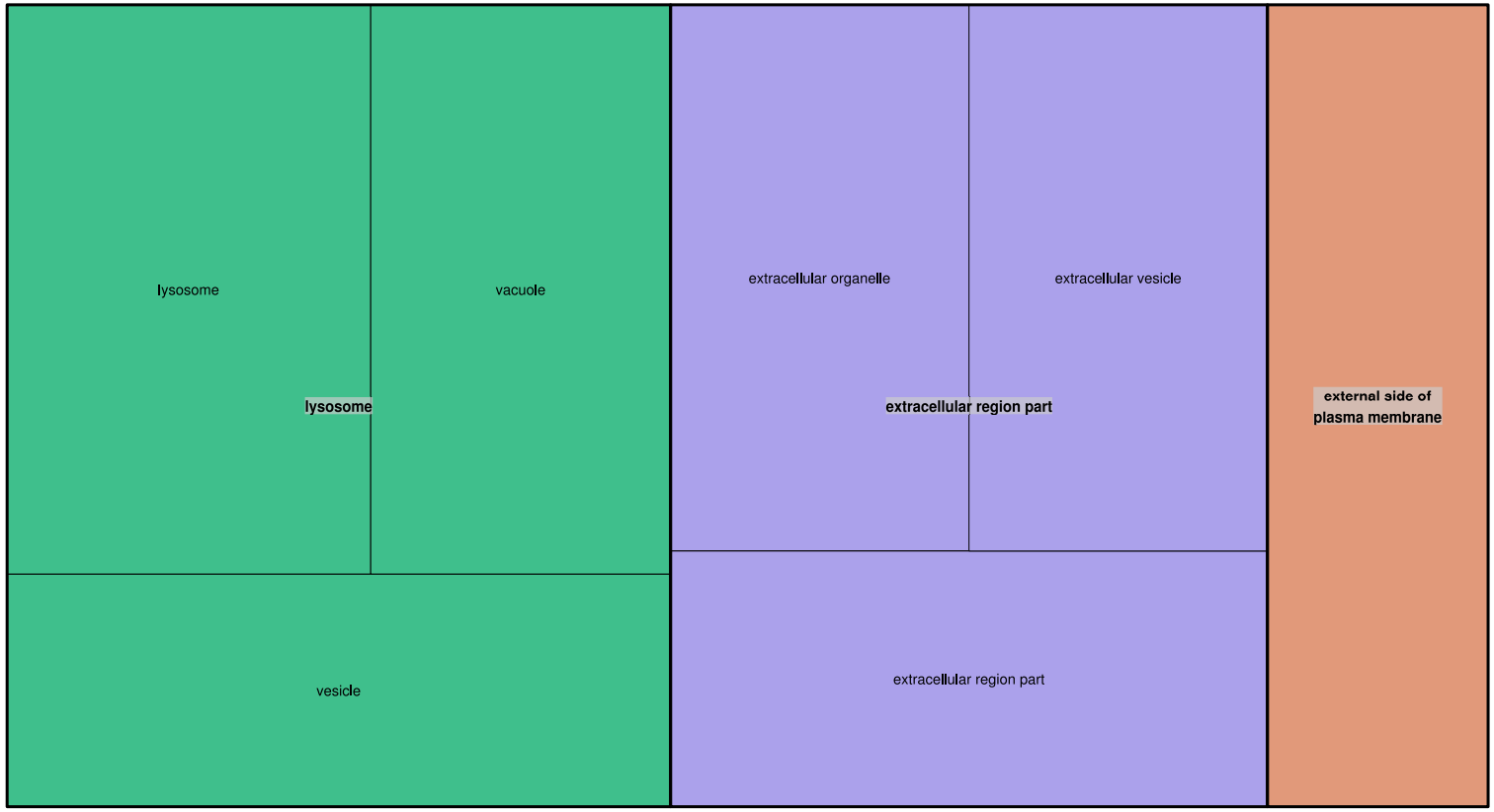

Figure 3.5: GO Cellular Component enrichment result for genes showing IELC pattern in brain datasets, summarized by REVIGO.

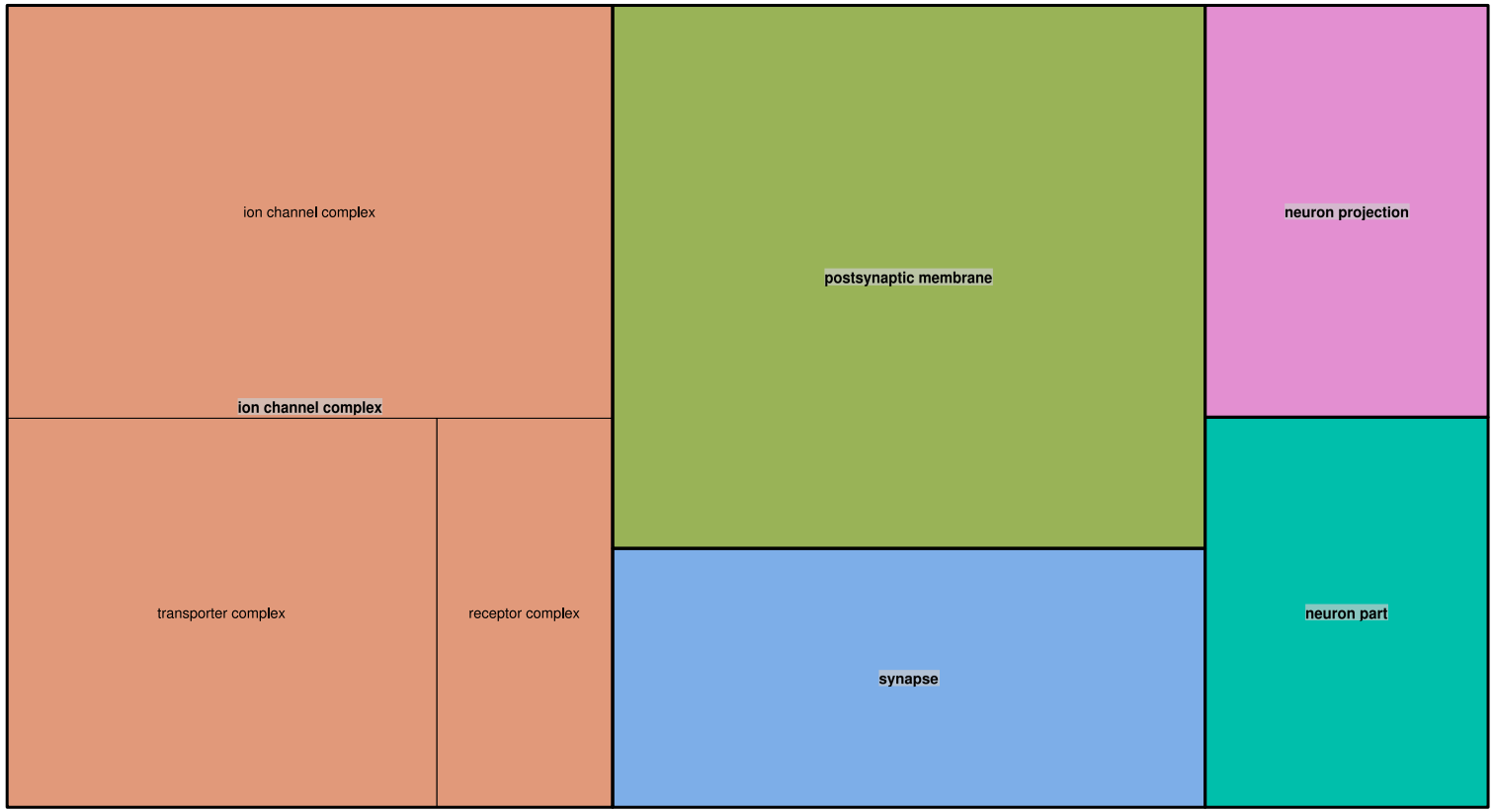

Figure 3.6: GO Cellular Component enrichment result for genes showing DEHC pattern in brain datasets, summarized by REVIGO.

Finally I repeated the analysis with GO Molecular Function. Among the 9 brain datasets, 11 common categories were enriched for DEHC ($p = 0.016$), including "ligand-gated ion channel activity" (see the REVIGO summary in **Figure 3.7** and **Appendix N**).

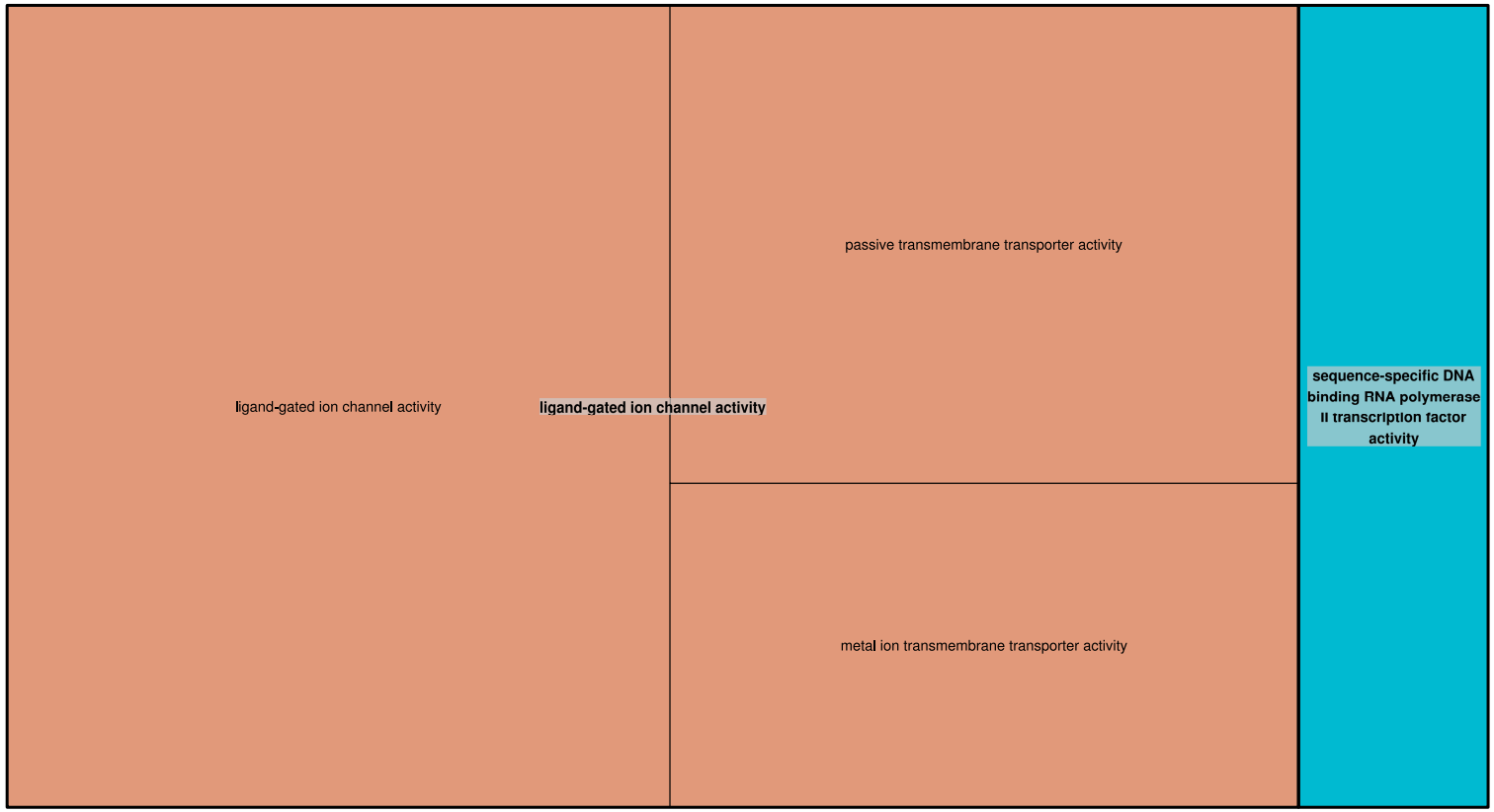

Figure 3.7: GO Molecular Process enrichment result for genes showing DEHC pattern in brain datasets, summarized by REVIGO.

In 2 different datasets for mouse lung, 37 common GO categories ($p = 0.001$) were also enriched for DEHC, including ”ubiquitin-protein transferase activity” and “transcription regulatory region DNA binding” (**Figure 3.8** and **Appendix O**). I did not find common GO categories among all 14 datasets for IELC in Molecular Function.

Figure 3.8: GO Molecular Process enrichment result for genes showing DEHC pattern in lung datasets, summarized by REVIGO.

CHAPTER 4

DISCUSSION

The MA theory predicts that mutational load increases with age due to the declining force of negative selection, in turn caused by extrinsic mortality (Medawar, 1952). This could mean that genes expressed at higher levels at old age are less conserved than those expressed at younger ages. Although proposed many decades ago, there have been few tests of the theory. To test the prediction, I analyzed transcriptome datasets of various tissues from primates and murids. In 3 tissues in which I could clearly determine changes in gene expression level during aging, in the brain, liver, and lung, I found consistent ADICT patterns. I could further attribute these patterns, at least partially, to lower conservation of genes with increased expression during aging, as predicted by MA. The ADICT propensity could not be explained by up-regulation of fast-evolving immune-related genes, or of positively selected genes. To my knowledge, this is the first molecular-based indication for the widespread presence of MA in mammals.

ADICT was not universally detected among all 35 datasets I analyzed. Notably, no consistent trend could be detected for kidney, skin, spleen, and muscle. Discrepancy among datasets, even in the same type of the tissues, could be interpreted as false negative results due to artifacts related to experimental design, platform differences of the analyzed datasets and heterogeneity of aging process within species, between species and even the different types of cells in the same tissue (Bahar et al., 2006; Kenyon, 2005; Somel, Khaitovich, Bahn, Pääbo & Lachmann, 2006). For this reason, conducting meta-analyses is of high importance especially in aging

research, and I could find consistent trends in multiple tissues, and especially the brain.

Intriguingly, however, while nearly all brain datasets showed ADICT, in muscle, I observed no consistent trend across the 11 datasets. Why do brain and muscle behave so differently in age-dependent change in transcriptome conservation? Brain aging has previously been reported to differ from aging in other tissues in a mouse experiment (Zahn et al., 2007); similarly, I find that brain datasets showed the most divergent age-related expression trends relative to those in other tissues (**Figure 2.1**). This could be related to the RNA composition the of brain being mostly dominated by neurons, which are mainly post-mitotic, and thus subject to different aging dynamics than other tissues with high cell turnover rates (Kirkwood, 2005; Yu & He, 2016). However, muscle cells are also post-mitotic and the tissue has low turnover rates as in brain (Richardson, Allan & Le, 2014). In addition, liver and lung also display a significant ADICT trend, and these are tissues which undergo high cell turnover (Tomasetti & Vogelstein, 2015). Also, brain and skin arise from exoderm, lung and liver from the endoderm, and muscle and kidney from mesoderm (Pansky, 1982); this grouping clearly does not correspond to ADICT patterns shown in **Figure 3.2**. Thus, the difference among tissues in ADICT propensity does not appear to be related to their general age-related transcriptome change trends, to their mitotic versus post-mitotic properties, or to their ontogenic origins.

Another possible explanation for the discrepancy among tissues is that a fraction of age-dependent changes in gene expression level reflect responses to stochastic environmental factors, rather than internal processes linked to age-related dysfunction, and this fraction differs among tissues. Thus, I may not expect to find an MA signature in tissues where the transcriptome is highly influenced by the environment, and muscle may be such a tissue. Indeed, in 8 / 11 muscle datasets, no age-related genes could be found at the cutoff $q < 0.10$, implying that convergent expression changes across individuals in muscle are limited. Also interestingly, recent studies have shown that muscle transcriptome profiles of older individual could return to younger levels by exercise (Melov, Tarnopolsky, Beckman, Felkey & Hubbard, 2007). I thus propose that ADICT propensity, and the putative MA effect in mammalian aging, may not be visible when the transcriptome is strongly affected by external noise. Notably tissues such as skin and liver are also affected

by external factors, but perhaps these may not show rejuvenation potential like muscle.

Does the ADICT signature I detect contribute to physiological decline in organismal senescence? The trend I observe likely involves the age-related expression of nearly-neutral (Ohta, 2002), *i.e.* very slightly harmful mutations, rather than strongly penetrant mutations; it also involves hundreds to thousands of genes, rather than a few genes driving senescence. Directly associating ADICT with physiological senescence will therefore remain a challenge. Still, the functional enrichment analysis provides a number of clues. Among genes with age-dependent increased expression and low conservation, the main culprits of senescence under MA, I find enrichment in two categories related to apoptosis and cell signaling among all 14 datasets with significant ADICT signal ("apoptotic signaling pathway" and "external side of plasma membrane". Both categories are related to the hallmarks of aging (López-Otín, 2013). In particular, apoptosis is crucial for eliminating senescent cells during healthy aging, and disruptions in apoptosis could lead to accumulating dysfunctional cells over time (Childs, Durik, Baker & van Deursen, 2015). Conversely, apoptosis is also thought to have role in the Alzeimer's disease etiology by causing loss of neurons (Currais, Hortobágyi & Soriano, 2009). Overall, relatively weak purifying selection on apoptosis-related genes may contribute to suboptimal regulation of the apoptosis process during aging, in mitotic as well as non-mitotic tissues, and contribute to senescent phenotypes.

Previous work has noted difficulties in distinguishing MA and AP based on studies of fitness variation changes with age (Moorad & Promislow, 2009). ADICT could be compatible not only with MA, but also with the AP model, if low conservation of the genes with high expression at late age is caused by their selection-driven rapid evolution due to their effects in earlier life phases. Genes involved in immune response or spermatogenesis are generally known to be fast evolving under positive selection (Consortium, 2005; Nielsen et al., 2005). Notably, apoptosis, the category I find enriched in high expression and low conservation genes, also shows large overlap (42%) with the immune system GO category (only 6% with spermatogenesis). On the other hand, the ADICT signal is robust to removing genes with a positive selection signal ($dN/dS > 1$) or immune genes. Therefore, ADICT at least partly represents an MA-specific molecular signature in mammals.

4.1 Limitations and Possible Improvements

1. I used kidney, liver, lung, and spleen datasets only in mice. The age-series microarray datasets including those tissues from different mammalian species could be added to more effectively interpret the effects of ADICT in mammalian aging. In addition to that, I used only microarray datasets in the analyses. Another possible improvement would be to add RNA-sequencing datasets.

2. Female and male samples were analyzed together. Whether the ADICT signal is different between sexes is another question waiting to be answered.

3. I would like to determine the contribution of the ADICT signature to physiological decline in organismal senescence and age-related diseases, using different methods. One approach would be to analyze the expression changes of proteins having role in Huntington's or Alzheimer's diseases, and repeat containing proteins which can cause protein aggregation by sticking to other proteins that eventually contribute to impairment in function. If those genes would be categorized into increased expression and low conservation, such a result could be considered as indication that the ADICT pattern is contributing to both age-related neurodegenerative diseases and the aging phenotype.

4. In the data I used, different samples were taken from the same individuals, especially in the brain datasets. I could improve the permutation tests to include individual information.

5. Discrepancy among datasets could be assessed using different methods such as surrogate variable analysis (SVA) (Leek & Storey, 2007), which removes the technical noise from unknown sources after discarding the effects of age on gene expression level.

6. Conservation of the proteins that are specific to early evolved brain regions, *i.e.* thalamus, could be studied to determine whether evolutionary history of the tissues have an effect on protein conservation.

7. ADICT was only observed using conservation of protein coding sequences, but not using conservation of regulatory regions (Poorya Parvizi, unpublished

results). The possible reasons for the difference between conservation of protein coding sequences and regulatory regions need to be addressed.

8. To improve to accuracy of conservation ratio (dN/dS), different situations can be considered. For example, synonymous mutations (dS) do not take into account the codon bias, and each type of the non-synonymous mutations treated as the same. Therefore, dS would not be completely neutral, and dN would not be totally harmful across all possible changes.

CHAPTER 5

CONCLUSION

Although recent advances in molecular techniques provide a better understanding of the genetics of aging, its evolutionary basis remains little understood. To test evolutionary theories of aging using new approaches is important to gain further insights into the underlying reasons for this complex process, and also age-related diseases.

For the aforementioned reasons, I tested one prediction of mutation accumulation theory, that is genes expressed later in life (after initiation of adulthood) would have lower evolutionary conservation than genes expressed early in life (at early adulthood). This I did by analyzing 35 microarray datasets that include a total of 768 samples from diverse mammalian species and tissues. My findings are as follows:

- Age-related decrease in conservation of the transcriptome (ADICT) was found in different brain regions in mice, rats, macaques and humans. In addition, mice liver and lung show similar trends as in brain. However, I did not find any consistent ADICT trend in kidney, skin, spleen and muscle.

- Both genes showing increased expression with age and low conservation, and genes showing decreased expression with age and high conservation, could be contributing to the ADICT pattern.

- Among GO groups having decreased expression with age and high conservation, functional categories showed enrichment across the same type of tissues displaying DEHC. This may indicate highly conserved genes are tissue-

specific, and that is the reason why I did not find any common GO categories that were enriched in decreased expression and high conservation among different types of the tissues.

- The presence of genes with increased expression with age and low conservation in tissues showing ADICT is compatible with the mutation accumulation theory's prediction, that genes expressed at old ages will be less conserved. This pattern is observed in 4 different mammalian species, mice, rats, macaques and humans, and 3 different tissues, brain, liver and lung. This may indicate the universality of ADICT among mammals, at least in some tissues.

- Among genes showing increased expression with age and low conservation, two functional categories showed enrichment across all tissues displaying ADICT: apoptosis and cell signaling. These groups may provide a link between age-related decline and diseases.

REFERENCES

Arantes-Oliveira, N., Apfeld, J., Dillin, A. & Kenyon, C. (2002). Regulation of life-span by germ-line stem cells in Caenorhabditis elegans. *Science (80-.).*, *295*(5554), 502–505.

Austad, S. N. (2004). Is aging programed? *Aging Cell*, *3*(5), 249–251. Retrieved from http://www.blackwell-synergy.com/doi/abs/10.1111/j.1474-9728.2004.00112.xpapers://d2952c50-9509-4ba2-9a03-22fbc04267d4/Paper/p999 doi: 10.1111/j.1474-9728.2004.00112.x

Bahar, R., Hartmann, C. H., Rodriguez, K. A., Denny, A. D., Busuttil, R. A., Dolle, M. E. T., ... Vijg, J. (2006, jun). Increased cell-to-cell variation in gene expression in ageing mouse heart. *Nature*, *441*(7096), 1011–1014. Retrieved from http://dx.doi.org/10.1038/nature04844http://www.nature.com/nature/journal/v441/n7096/suppinfo/nature04844{_}S1.html

Baker, D. J., Wijshake, T., Tchkonia, T., LeBrasseur, N. K., Childs, B. G., Van De Sluis, B., ... van Deursen, J. M. (2011). Clearance of p16Ink4a-positive senescent cells delays ageing-associated disorders. *Nature*, *479*(7372), 232–236.

Berry, R. J. & Bronson, F. H. (1992). Life history and bioeconomy of the house mouse. *Biol. Rev.*, *67*(4), 519–550.

Blasco, M. A. (2007). Telomere length, stem cells and aging. *Nat. Chem. Biol.*, *3*(10), 640–649.

Campisi, J. (2005). Aging, tumor suppression and cancer: high wire-act! *Mech. Ageing Dev.*, *126*(1), 51–58.

Campisi, J. & d'Adda di Fagagna, F. (2007). Cellular senescence: when bad things happen to good cells. *Nat. Rev. Mol. cell Biol.*, *8*(9), 729–740.

Childs, B. G., Durik, M., Baker, D. J. & van Deursen, J. M. (2015). Cellular senescence in aging and age-related disease: from mechanisms to therapy.

Nat. Med., *21*(12), 1424–1435.

Consortium, A. (2005). Initial sequence of the chimpanzee genome and comparison with the human genome. *Nature*, *437*(7055), 69.

Currais, A., Hortobágyi, T. & Soriano, S. (2009, apr). The neuronal cell cycle as a mechanism of pathogenesis in Alzheimer's disease. *Aging (Albany. NY).*, *1*(4), 363–371. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2806021/

de Magalhães, J. P. (2003). Is mammalian aging genetically controlled? *Biogerontology*, *4*(2), 119–120.

Demetrius, L. & Ziehe, M. (2007). Darwinian fitness. *Theor. Popul. Biol.*, *72*(3), 323–345. doi: 10.1016/j.tpb.2007.05.004

Douglas, P. M. & Dillin, A. (2014). The disposable soma theory of aging in reverse. *Cell Res.*, *24*(1), 7–8. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3879701{\protect\T1\textbraceleft}{&}{\protect\T1\textbraceright}tool=pmcentrez{\protect\T1\textbraceleft}{&}{\protect\T1\textbraceright}rendertype=abstract doi: 10.1038/cr.2013.148

Durinck, S., Moreau, Y., Kasprzyk, A., Davis, S., De Moor, B., Brazma, A. & Huber, W. (2005). BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis. *Bioinformatics*, *21*(16), 3439–3440.

Edgar, R., Domrachev, M. & Lash, A. E. (2002). Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.*, *30*(1), 207–210.

Ermolaeva, M. A., Segref, A., Dakhovnik, A., Ou, H.-L., Schneider, J. I., Utermöhlen, O., … Schumacher, B. (2013). DNA damage in germ cells induces an innate immune response that triggers systemic stress resistance. *Nature*, *501*(7467), 416–420. Retrieved from http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed{\protect\T1\textbraceleft}{&}{\protect\T1\textbraceright}id=23975097{\protect\T1\textbraceleft}{&}{\protect\T1\textbraceright}retmode=ref{\protect\T1\textbraceleft}{&}{\protect\T1\textbraceright}cmd=prlinks{\protect\T1\textdollar}\delimiter"026E30F{\protect\T1\textdollar}npapers3://publication/doi/10.1038/nature12452 doi: 10.1038/nature12452

Escobar, J. S., Jarne, P., Charmantier, A. & David, P. (2008, sep). Outbreeding Alleviates Senescence in Hermaphroditic Snails as Expected from the Mutation-

Accumulation Theory. *Curr. Biol.*, *18*(12), 906–910. Retrieved from http://dx.doi.org/10.1016/j.cub.2008.04.070 doi: 10.1016/j.cub.2008.04.070

Flatt, T. & Schmidt, P. S. (2009). Integrating Evolutionary and Molecular Genetics of Aging. *Biochim. Biophys. Acta*, *1790*(10), 951–962. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2972575/ doi: 10.1016/j.bbagen.2009.07.010

Fontana, L., Partridge, L. & Longo, V. D. (2010). Extending healthy life span—from yeast to humans. *Science (80-.).*, *328*(5976), 321–326.

Harman, D. (1955). Aging: a theory based on free radical and radiation chemistry.

Harrison, D. E., Strong, R., Sharp, Z. D., Nelson, J. F., Astle, C. M., Flurkey, K., ... Others (2009). Rapamycin fed late in life extends lifespan in genetically heterogeneous mice. *Nature*, *460*(7253), 392–395.

Hartmann, A., Nürnberg, G., Repsilber, D., Janczyk, P., Walz, C., Ponsuksili, S., ... Others (2009). Effects of threshold choice on the results of gene expression profiling, using microarray analysis, in a model feeding experiment with rats. *Arch. Tierzucht*, *52*, 65–78.

Hoeijmakers, J. H. J. (2009). DNA damage, aging, and cancer. *N. Engl. J. Med.*, *361*(15), 1475–1485.

Hsin, H. & Kenyon, C. (1999). Signals from the reproductive system regulate the lifespan of C. elegans. *Nature*, *399*(6734), 362–366.

Hughes, K. A., Alipaz, J. A., Drnevich, J. M. & Reynolds, R. M. (2002). Wilson. *Proc. Natl. Acad. Sci.*, *99*(22), 14286–14291.

Jones & Rando, T. A. (2011). Emerging models and paradigms for stem cell ageing. *Nat. Cell Biol.*, *13*(5), 506–512.

Jones, Scheuerlein, A., Salguero-Gómez, R., Camarda, C. G., Schaible, R., Casper, B. B., ... Vaupel, J. W. (2014). Diversity of ageing across the tree of life. *Nature*, *505*(7482), 169–73. Retrieved from http://www.nature.com/nature/journal/v505/n7482/nature12789/metrics doi: 10.1038/nature12789

Kanfi, Y., Naiman, S., Amir, G., Peshti, V., Zinman, G., Nahum, L., ... Cohen, H. Y. (2012). The sirtuin SIRT6 regulates lifespan in male mice. *Nature*, *483*(7388), 218–221.

Kelly, D. P. (2011). Cell biology: ageing theories unified. *Nature*, *470*(7334), 342–343.

Kenyon. (2005, feb). The Plasticity of Aging: Insights from Long-Lived Mutants. *Cell*, *120*(4), 449–460. Retrieved from http://www.sciencedirect.com/

science/article/pii/S0092867405001108 doi: http://dx.doi.org/10.1016/j.cell.2005.02.002

Kenyon. (2010). The genetics of ageing. *Nature*, *464*(7288), 504–512.

Kenyon, Chang, J., Gensch, E., Rudner, A., Tabtiang, R. & Others. (1993). A C. elegans mutant that lives twice as long as wild type. *Nature*, *366*(6454), 461–464.

Kirkwood, T. B. L. (2005, feb). Understanding the Odd Science of Aging. *Cell*, *120*(4), 437–447. Retrieved from http://www.sciencedirect.com/science/article/pii/S0092867405001017 doi: http://dx.doi.org/10.1016/j.cell.2005.01.027

Kirkwood, T. B. L. & Austad, S. N. (2000). Why do we age? *Nature*, *408*(6809), 233–238.

Kirkwood, T. B. L. & Cremer, T. (1982). Cytogerontology Since 1881: A Reappraisal of August Weismann and a Review of Modern Progress. *Hum Genet*, *60*, 101–121.

Kirkwood, T. B. L. & Melov, S. (2011). On the programmed/non-programmed nature of ageing within the life history. *Curr. Biol.*, *21*(18), R701—-R707. Retrieved from http://dx.doi.org/10.1016/j.cub.2011.07.020 doi: 10.1016/j.cub.2011.07.020

Koga, H., Kaushik, S. & Cuervo, A. M. (2011). Protein homeostasis and aging: The importance of exquisite quality control. *Ageing Res. Rev.*, *10*(2), 205–215.

Kolesnikov, N., Hastings, E., Keays, M., Melnichuk, O., Tang, Y. A., Williams, E., … Others (2014). ArrayExpress update—simplifying data submissions. *Nucleic Acids Res.*, gku1057.

Kryuchkova-Mostacci, N. & Robinson-Rechavi, M. (2015, jun). Tissue-Specific Evolution of Protein Coding Genes in Human and Mouse. *PLoS One*, *10*(6), e0131673. Retrieved from http://dx.doi.org/10.1371{%}2Fjournal.pone.0131673

Kujoth, G. C., Hiona, A., Pugh, T. D., Someya, S., Panzer, K., Wohlgemuth, S. E., … Others (2005). Mitochondrial DNA mutations, oxidative stress, and apoptosis in mammalian aging. *Science (80-.).*, *309*(5733), 481–484.

Lee, H.-W., Blasco, M. A., Gottlieb, G. J., Horner, J. W., Greider, C. W. & DePinho, R. A. (1998). Essential role of mouse telomerase in highly proliferative organs. *Nature*, *392*(6676), 569–574.

Leek, J. T. & Storey, J. D. (2007). Capturing heterogeneity in gene expression studies by surrogate variable analysis. *PLoS Genet*, *3*(9), e161.

Leopold, A. C., Niedergang-Kamien, E. & Janick, J. (1959). Experimental Modification of Plant Senescence. *Plant Physiol.*, *34*(5), 570.

Leroi, A. M., Bartke, A., De Benedictis, G., Franceschi, C., Gartner, A., Gonos, E., ... Others (2005). What evidence is there for the existence of individual genes with antagonistic pleiotropic effects? *Mech. Ageing Dev.*, *126*(3), 421–429.

Longo, V. D. & Finch, C. E. (2003). Evolutionary medicine: from dwarf model systems to healthy centenarians? *Science (80-.).*, *299*(5611), 1342–1346.

Longo, V. D., Mitteldorf, J. & Skulachev, V. P. (2005). Programmed and altruistic ageing. *Nat. Rev. Genet.*, *6*(11), 866–872. doi: 10.1038/nrg1706

López-Otín, C. (2013). *The Hallmarks of Aging* (Vol. 153) (No. 6). doi: 10.1016/j.cell.2013.05.039

Lu, T., Pan, Y., Kao, S.-Y., Li, C., Kohane, I., Chan, J. & Yankner, B. A. (2004, jun). Gene regulation and DNA damage in the ageing human brain. *Nature*, *429*(6994), 883–891. Retrieved from http://dx.doi.org/10.1038/nature02661http://www.nature.com/nature/journal/v429/n6994/suppinfo/nature02661{_}S1.html

Matheu, A., Maraver, A., Collado, M., Garcia-Cao, I., Cañamero, M., Borras, C., ... Serrano, M. (2009). Anti-aging activity of the Ink4/Arf locus. *Aging Cell*, *8*(2), 152–161.

Matheu, A., Maraver, A., Klatt, P., Flores, I., Garcia-Cao, I., Borras, C., ... Serrano, M. (2007). Delayed ageing through damage protection by the Arf/p53 pathway. *Nature*, *448*(7151), 375–379.

McQuillan, H. J., Lokman, P. M. & Young, G. (2003). Effects of sex steroids, sex, and sexual maturity on cortisol production: an in vitro comparison of chinook salmon and rainbow trout interrenals. *Gen. Comp. Endocrinol.*, *133*(1), 154–163.

Medawar, P. B. (1952). *An Unsolved Problem of Biology: An Inaugural Lecture Delivered at University College, London, 6 December, 1951*. H.K. Lewis and Company. Retrieved from https://books.google.com.tr/books?id=ozEmPQAACAAJ

Medvedev, Z. A. (1990). An attempt at a rational classification of theories of ageing. *Biol. Rev.*, *65*(3), 375–398.

Melov, S., Tarnopolsky, M. A., Beckman, K., Felkey, K. & Hubbard, A. (2007, may). Resistance Exercise Reverses Aging in Human Skeletal Muscle. *PLoS One*, *2*(5), e465. Retrieved from http://dx.plos.org/10.1371/journal.pone.0000465

Moorad, J. A. & Promislow, D. E. L. (2009). What can genetic variation tell us about the evolution of senescence? *Proc. R. Soc. London B Biol. Sci.*, *276*(1665), 2271–2278.

Moretti, S., Laurenczy, B., Gharib, W. H., Castella, B., Kuzniar, A., Schabauer, H., ... Robinson-Rechavi, M. (2014, jan). Selectome update: quality control and computational improvements to a database of positive selection. *Nucleic Acids Res.*, *42*(Database issue), D917–D921. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3964977/ doi: 10.1093/nar/gkt1065

Moskalev, A. A., Shaposhnikov, M. V., Plyusnina, E. N., Zhavoronkov, A., Budovsky, A., Yanai, H. & Fraifeld, V. E. (2013). The role of DNA damage and repair in aging through the prism of Koch-like criteria. *Ageing Res. Rev.*, *12*(2), 661–684.

Mostoslavsky, R., Chua, K. F., Lombard, D. B., Pang, W. W., Fischer, M. R., Gellon, L., ... Others (2006). Genomic instability and aging-like phenotype in the absence of mammalian SIRT6. *Cell*, *124*(2), 315–329.

Nielsen, R., Bustamante, C., Clark, A. G., Glanowski, S., Sackton, T. B., Hubisz, M. J., ... Cargill, M. (2005, may). A Scan for Positively Selected Genes in the Genomes of Humans and Chimpanzees. *PLoS Biol*, *3*(6), e170. Retrieved from http://dx.doi.org/10.1371{%}2Fjournal.pbio.0030170

Ohta, T. (2002, dec). Near-neutrality in evolution of genes and gene regulation. *Proc. Natl. Acad. Sci.*, *99*(25), 16134–16137. Retrieved from http://www.pnas.org/content/99/25/16134.abstract doi: 10.1073/pnas.252626899

Pál, C., Papp, B. & Lercher, M. J. (2006). An integrated view of protein evolution. *Nat. Rev. Genet.*, *7*(5), 337–348.

Pansky, B. (1982). *Review of medical embryology*. Macmillan New York.

Powers, E. T., Morimoto, R. I., Dillin, A., Kelly, J. W. & Balch, W. E. (2009). Biological and chemical approaches to diseases of proteostasis deficiency. *Annu. Rev. Biochem.*, *78*, 959–991.

Promislow, D. E. L. (1994). DNA repair and the evolution of longevity: a critical analysis. *J. Theor. Biol.*, *170*(3), 291–300.

Rando, T. A. & Chang, H. Y. (2012). Aging, rejuvenation, and epigenetic repro-

gramming: resetting the aging clock. *Cell*, *148*(1), 46–57.

Richardson, R. B., Allan, D. S. & Le, Y. (2014, jul). Greater organ involution in highly proliferative tissues associated with the early onset and acceleration of ageing in humans. *Exp. Gerontol.*, *55*, 80–91. Retrieved from http://www.sciencedirect.com/science/article/pii/S0531556514000874 doi: http://dx.doi.org/10.1016/j.exger.2014.03.015

Robertson, O. H. (1961). Prolongation of the life span of kokanee salmon (Oncorhynchus nerka kennerlyi) by castration before beginning of gonad development. *Proc. Natl. Acad. Sci.*, *47*(4), 609–621.

Rose, M. R. (1991). *Evolutionary biology of aging*. Oxford University Press. Retrieved from https://books.google.com.tr/books?id=TaHwAAAAMAAJ

Salminen, A., Kaarniranta, K. & Kauppinen, A. (2012). Inflammaging: disturbed interplay between autophagy and inflammasomes. *Aging (Albany NY)*, *4*(3), 166–175.

Sharpless, N. E. & DePinho, R. A. (2007). How stem cells age and why this makes us grow old. *Nat. Rev. Mol. Cell Biol.*, *8*(9), 703–713.

Skulachev, V. P. (1997). Aging is a specific biological function rather than the result of a disorder in complex living systems: biochemical evidence in support of Weismann's hypothesis. *Biochem. York-English Transl. Biokhimiya*, *62*(11), 1191–1195.

Sohal, R. S., Ku, H.-H. & Agarwal, S. (1993). Biochemical correlates of longevity in two closely related rodent species. *Biochem. Biophys. Res. Commun.*, *196*(1), 7–11.

Somel, M., Guo, S., Fu, N., Yan, Z., Hu, H. Y., Xu, Y., ... Khaitovich, P. (2010, sep). MicroRNA, mRNA, and protein expression link development and aging in human and macaque brain. *Genome Res.*, *20*(9), 1207–1218. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2928499/ doi: 10.1101/gr.106849.110

Somel, M., Khaitovich, P., Bahn, S., Pääbo, S. & Lachmann, M. (2006). Gene expression becomes heterogeneous with age. *Curr. Biol.*, *16*(10), R359—-R360.

Supek, F., Bošnjak, M., Škunca, N. & Šmuc, T. (2011, jul). REVIGO Summarizes and Visualizes Long Lists of Gene Ontology Terms. *PLoS One*, *6*(7), e21800. Retrieved from http://dx.doi.org/10.1371{%}2Fjournal.pone.0021800

Talens, R. P., Christensen, K., Putter, H., Willemsen, G., Christiansen, L., Kremer,

D., ... Heijmans, B. T. (2012). Epigenetic variation during the adult lifespan: cross-sectional and longitudinal data on monozygotic twin pairs. *Aging Cell*, *11*(4), 694–703.

Tomasetti, C. & Vogelstein, B. (2015, jan). Variation in cancer risk among tissues can be explained by the number of stem cell divisions. *Science (80-.).*, *347*(6217), 78–81. Retrieved from http://science.sciencemag.org/content/347/6217/78.abstract

Travis, J. M. J. (2004). The Evolution of Programmed Death in a Spatially Structured Population. *J. Gerontol.*, *59*(4), 301–305.

Trifunovic, A., Wredenberg, A., Falkenberg, M., Spelbrink, J. N., Rovio, A. T., Bruder, C. E., ... Others (2004). Premature ageing in mice expressing defective mitochondrial DNA polymerase. *Nature*, *429*(6990), 417–423.

Vermulst, M., Wanagat, J., Kujoth, G. C., Bielas, J. H., Rabinovitch, P. S., Prolla, T. A. & Loeb, L. A. (2008). DNA deletions and clonal mutations drive premature aging in mitochondrial mutator mice. *Nat. Genet.*, *40*(4), 392–394.

Walker, D. W., McColl, G., Jenkins, N. L., Harris, J. & Lithgow, G. J. (2000). Natural selection: evolution of lifespan in C. elegans. *Nature*, *405*(6784), 296–297.

Weismann, A. (1889). *Essays upon heredity and kindred biological problems* (Vol. 1). Clarendon press.

Williams, G. C. (1957). PLEIOTROPY, NATURAL SELECTION, AND THE EVOLUTION OF SENESCENCE1. *Evolution (N. Y).*, *11*(4), 398–411.

Wilson, A. J., Nussey, D. H., Pemberton, J. M., Pilkington, J. G., Morris, A., Pelletier, F., ... Kruuk, L. E. B. (2007). Evidence for a genetic basis of aging in two wild vertebrate populations. *Curr. Biol.*, *17*(24), 2136–2142.

Yu, Q. & He, Z. (2016, jul). Cell type composition is the primary - but far from the only - power shaping temporal transcriptome of human brains. *bioRxiv*. Retrieved from http://biorxiv.org/content/early/2016/07/26/065292.abstract

Zahn, J. M., Poosala, S., Owen, A. B., Ingram, D. K., Lustig, A., Carter, A., ... Becker, K. G. (2007, nov). AGEMAP: A Gene Expression Database for Aging in Mice. *PLoS Genet.*, *3*(11), e201. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2098796/ doi: 10.1371/journal.pgen.0030201

Zhang, J., Nielsen, R. & Yang, Z. (2005, dec). Evaluation of an Improved

Branch-Site Likelihood Method for Detecting Positive Selection at the Molecular Level. *Mol. Biol. Evol.*, *22*(12), 2472–2479. Retrieved from http://mbe.oxfordjournals.org/content/22/12/2472.abstract doi: 10.1093/molbev/msi237

APPENDIX A

RESULTS FOR THE CHANGES IN EXPRESSION - ω_0 CORRELATION WITH AGE

Table A.1: Results – number of genes, rho, p values - for the changes in expression - ω_0 correlation with age both all genes and age-related genes.

Dataset Name	**# all**	**rho**	p	**# age-rel.**	**rho**	p
Hs_Brain_BA22	8744	−0.51	0.026	-	-	-
Hs_Brain_FC	3966	−0.443	0.016	345	−0.834	0
Hs_Brain_BA10	8744	−0.387	0.068	1136	−0.819	0
Hs_Brain_EC	8744	−0.35	0.039	220	−0.518	0.001
Hs_Brain_HC	8744	−0.256	0.106	950	−0.771	0
Hs_Brain_PCG	8744	−0.593	0	2345	−0.635	0
Hs_Brain_SG	8744	−0.555	0	3650	−0.658	0
Hs_Brain_PFC	6347	−0.587	0.049	103	−0.93	0
Hs_Brain_CB	6406	−0.018	0.968	-	-	-
Rm_Brain_PFC	6347	−0.009	0.989	-	-	-
Rm_Brain_CB	6406	−0.017	0.982	-	-	-
Rn_Brain_HC	6657	−0.372	0.19	113	−0.868	0
Mm_Brain_WB	8639	−0.189	0.517	104	−0.873	0
Mm_Brain_NC	8639	0.522	0.122	3641	0.731	0.016
Mm_Brain_HC	4009	−0.299	0.177	-	-	-
Mm_Skin	8639	0.017	0.942	1053	0.642	0.002
Hs_Skin	7377	0.011	0.915	39	−0.444	0
Mm_Liver_3	8639	−0.182	0.515	404	−0.772	0.001
Mm_Liver_2	8639	−0.442	0.099	496	−0.918	0
Mm_Liver_1	8632	−0.672	0.012	188	−0.746	0.003

Mm_Lung_2	8639	−0.285	0.251	189	−0.724	0.001
Mm_Lung_1	8639	−0.738	0.023	153	−0.949	0
Mm_Kidney	8639	−0.179	0.478	253	−0.712	0.001
Mm_Spleen	8639	−0.263	0.326	180	0.746	0.001
Rn_Aorta_T	2016	0.437	0.091	192	0.776	0
Rm_Aorta	5973	−0.226	0.288	-	-	-
Mm_Muscle_Gastro	6394	0.522	0.122	573	0.731	0.016
Mm_Muscle_P C	8639	−0.503	0.067	280	−0.881	0
Rn_Muscle_EDL	2017	0.559	0.074	-	-	-
Rn_Muscle_EOMs	2017	0.118	0.714	-	-	-
Hs_Muscle_VL2	8307	0.345	0.208	-	-	-
Hs_Muscle_VL1	8307	0.349	0.203	-	-	-
Hs_Muscle_Other	8744	0.1	0.724	-	-	-
Hs_Muscle_RA	8744	0.077	0.552	-	-	-
Hs_Muscle_BB	8744	0.021	0.933	1582	−0.579	0.009

APPENDIX B

RESULTS FOR THE CHANGES IN EXPRESSION - ω_0* CORRELATION WITH AGE

Table B.1: Results – number of genes, rho, p values - for the changes in expression - ω_0* correlation with age both all genes and age-related genes.

Dataset Name	**# all**	**rho**	p	**# age-rel.**	**rho**	p
Hs_Brain_BA22	6221	-0.648	0.003	-	-	-
Hs_Brain_FC	3048	-0.508	0.005	246	-0.821	0
Hs_Brain_BA10	6221	-0.539	0.008	835	-0.827	0
Hs_Brain_EC	6221	-0.391	0.02	155	-0.051	0.772
Hs_Brain_HC	6221	-0.448	0.003	688	-0.769	0
Hs_Brain_PCG	6221	-0.558	0	1688	-0.693	0
Hs_Brain_SG	6221	-0.607	0	2633	-0.676	0
Hs_Brain_PFC	4579	-0.769	0.005	80	-0.895	0
Hs_Brain_CB	4602	-0.291	0.386	-	-	-
Rm_Brain_PFC	4579	-0.2	0.558	-	-	-
Rm_Brain_CB	4602	0.083	0.843	-	-	-
Rn_Brain_HC	4414	-0.443	0.113	71	-0.868	0
Mm_Brain_WB	5532	-0.454	0.103	70	-0.875	0
Mm_Brain_NC	5532	-0.661	0.037	2321	-0.801	0.005
Mm_Brain_HC	2960	-0.27	0.225	-	-	-
Mm_Skin	5532	-0.399	0.081	656	0.26	0.268
Hs_Skin	5437	0.015	0.882	32	-0.186	0.066
Mm_Liver_3	5532	-0.525	0.044	265	-0.901	0
Mm_Liver_2	5532	-0.564	0.029	341	-0.887	0
Mm_Liver_1	5530	-0.595	0.032	110	-0.683	0.01

Mm_Lung_2	5532	−0.605	0.008	129	−0.593	0.01
Mm_Lung_1	5532	−0.843	0.004	100	−0.949	0
Mm_Kidney	5532	−0.43	0.075	176	−0.248	0.322
Mm_Spleen	5532	−0.328	0.214	114	0.03	0.913
Rn_Aorta_T	1479	0.582	0.018	136	0.776	0
Rm_Aorta	4559	−0.258	0.223	-	-	-
Mm_Muscle_Gastro	4461	0.87	0.001	400	0.87	0.001
Mm_Muscle_P C	5532	−0.776	0.001	181	−0.86	0
Rn_Muscle_EDL	1480	0.607	0.048	-	-	-
Rn_Muscle_EOMs	1480	−0.296	0.351	-	-	-
Hs_Muscle_VL2	5992	0.277	0.317	-	-	-
Hs_Muscle_VL1	5992	0.383	0.159	-	-	-
Hs_Muscle_Other	6221	0.154	0.584	-	-	-
Hs_Muscle_RA	6221	−0.029	0.825	-	-	-
Hs_Muscle_BB	6221	0.331	0.167	1140	0	1

APPENDIX C

RESULTS FOR THE CHANGES IN EXPRESSION - ω ("ONE-TO-ONE ORTHOLOGS" BETWEEN HUMAN-MOUSE) CORRELATION WITH AGE

Table C.1: Results – number of genes, rho, p values - for the changes in expression - ω ("one-to-one orthologs" between human-mouse) correlation with age both all genes and age-related genes.

Dataset Name	**# all**	**rho**	p	**# age-rel.**	**rho**	p
Hs_Brain_BA22	20219	-0.5	0.029	-	-	-
Hs_Brain_FC	7815	-0.482	0.008	644	-0.835	0
Hs_Brain_BA10	20219	-0.493	0.017	1963	-0.814	0
Hs_Brain_EC	20219	-0.395	0.019	388	-0.628	0
Hs_Brain_HC	20219	-0.237	0.136	1723	-0.728	0
Hs_Brain_PCG	20219	-0.613	0	3950	-0.625	0
Hs_Brain_SG	20219	-0.613	0	6117	-0.683	0
Hs_Brain_PFC	11453	-0.587	0.049	189	-0.965	0
Hs_Brain_CB	11883	-0.182	0.595	-	-	-
Rm_Brain_PFC	11453	-0.109	0.755	-	-	-
Rm_Brain_CB	11883	-0.017	0.982	-	-	-
Rn_Brain_HC	10747	-0.408	0.148	195	-0.868	0
Mm_Brain_WB	14172	-0.105	0.722	211	-0.944	0
Mm_Brain_NC	14172	0.592	0.071	5834	0.87	0.001
Mm_Brain_HC	6824	-0.342	0.119	-	-	-
Mm_Skin	14172	-0.277	0.236	1837	-0.173	0.465
Hs_Skin	13461	0.013	0.9	83	-0.511	0
Mm_Liver_3	14172	-0.317	0.25	667	-0.733	0.002

Mm_Liver_2	14172	−0.219	0.432	804	−0.747	0.001
Mm_Liver_1	14182	−0.598	0.031	341	−0.507	0.077
Mm_Lung_2	14172	−0.348	0.157	327	−0.756	0
Mm_Lung_1	14172	−0.527	0.145	275	−0.791	0.011
Mm_Kidney	14172	−0.26	0.297	524	−0.749	0
Mm_Spleen	14171	−0.251	0.349	317	0.522	0.038
Rn_Aorta_T	3584	0.412	0.113	320	0.534	0.033
Rm_Aorta	11535	−0.242	0.254	-	-	-
Mm_Muscle_Gastro	10613	0.453	0.189	926	0.731	0.016
Mm_Muscle_P C	14172	−0.126	0.668	541	−0.629	0.016
Rn_Muscle_EDL	3584	0.66	0.027	-	-	-
Rn_Muscle_EOMs	3584	0.118	0.714	-	-	-
Hs_Muscle_VL2	17557	0.12	0.67	-	-	-
Hs_Muscle_VL1	17557	0.188	0.503	-	-	-
Hs_Muscle_Other	20219	0.15	0.593	-	-	-
Hs_Muscle_RA	20219	0.094	0.467	-	-	-
Hs_Muscle_BB	20219	−0.041	0.867	2747	−0.579	0.009

APPENDIX D

RESULTS FOR THE CHANGES IN EXPRESSION - ω ("ONE-TO-ONE ORTHOLOGS" BETWEEN HUMAN-ELEPHANT) CORRELATION WITH AGE

Table D.1: Results – number of genes, rho, p values - for the changes in expression - ω ("one-to-one orthologs" between human-elephant) correlation with age both all genes and age-related genes.

Dataset Name	**# all**	**rho**	p	**# age-rel.**	**rho**	p
Hs_Brain_BA22	13930	-0.574	0.01	-	-	-
Hs_Brain_FC	6943	-0.475	0.009	630	-0.866	0
Hs_Brain_BA10	13930	-0.377	0.076	1917	-0.836	0
Hs_Brain_EC	13930	-0.39	0.021	375	-0.633	0
Hs_Brain_HC	13930	-0.251	0.114	1670	-0.738	0
Hs_Brain_PCG	13930	-0.646	0	3844	-0.644	0
Hs_Brain_SG	13930	-0.582	0	5931	-0.68	0
Hs_Brain_PFC	9998	-0.587	0.049	188	-0.909	0
Hs_Brain_CB	10090	-0.2	0.558	-	-	-
Rm_Brain_PFC	9998	-0.055	0.881	-	-	-
Rm_Brain_CB	10090	-0.017	0.982	-	-	-
Rn_Brain_HC	10000	-0.443	0.113	193	-0.868	0
Mm_Brain_WB	13086	-0.16	0.584	220	-0.862	0
Mm_Brain_NC	13086	0.592	0.071	5409	0.731	0.016
Mm_Brain_HC	6299	-0.386	0.076	-	-	-
Mm_Skin	13086	0.017	0.942	1994	0.017	0.942
Hs_Skin	11837	0.002	0.981	79	-0.483	0
Mm_Liver_3	13086	-0.31	0.262	643	-0.514	0.05

Mm_Liver_2	13086	−0.375	0.168	825	−0.675	0.006
Mm_Liver_1	13077	−0.646	0.017	280	0.048	0.875
Mm_Lung_2	13086	−0.43	0.075	331	−0.756	0
Mm_Lung_1	13086	−0.632	0.068	263	−0.949	0
Mm_Kidney	13086	−0.248	0.322	484	−0.655	0.003
Mm_Spleen	13086	−0.397	0.128	313	−0.069	0.801
Rn_Aorta_T	3315	0.473	0.064	307	−0.133	0.622
Rm_Aorta	9925	−0.264	0.213	-	-	-
Mm_Muscle_Gastro	9835	0.313	0.378	873	0.383	0.275
Mm_Muscle_P C	13086	−0.252	0.386	599	−0.776	0.001
Rn_Muscle_EDL	3315	0.588	0.057	-	-	-
Rn_Muscle_EOMs	3315	0	1	-	-	-
Hs_Muscle_VL2	13361	0.182	0.515	-	-	-
Hs_Muscle_VL1	13361	0.263	0.344	-	-	-
Hs_Muscle_Other	13930	−0.021	0.944	-	-	-
Hs_Muscle_RA	13930	0.094	0.469	-	-	-
Hs_Muscle_BB	13930	0.227	0.349	2624	−0.351	0.14

APPENDIX E

RESULTS FOR THE CHANGES IN EXPRESSION - ω ("ONE-TO-ONE ORTHOLOGS" BETWEEN HUMAN-COW) CORRELATION WITH AGE

Table E.1: Results – number of genes, rho, p values - for the changes in expression - ω ("one-to-one orthologs" between human-elephant) correlation with age both all genes and age-related genes.

Dataset Name	**# all**	**rho**	p	**# age-rel.**	**rho**	p
Hs_Brain_BA22	14168	-0.5	0.029	-	-	-
Hs_Brain_FC	6997	-0.483	0.008	628	-0.84	0
Hs_Brain_BA10	14168	-0.351	0.101	1920	-0.833	0
Hs_Brain_EC	14168	-0.366	0.031	376	-0.651	0
Hs_Brain_HC	14168	-0.238	0.135	1705	-0.725	0
Hs_Brain_PCG	14168	-0.602	0	3858	-0.64	0
Hs_Brain_SG	14168	-0.546	0	5951	-0.645	0
Hs_Brain_PFC	10067	-0.573	0.055	188	-0.944	0
Hs_Brain_CB	10171	-0.182	0.595	-	-	-
Rm_Brain_PFC	10067	-0.027	0.946	-	-	-
Rm_Brain_CB	10171	-0.017	0.982	-	-	-
Rn_Brain_HC	10068	-0.372	0.19	190	-0.868	0
Mm_Brain_WB	13255	-0.116	0.693	222	-0.931	0
Mm_Brain_NC	13255	0.661	0.037	5456	0.87	0.001
Mm_Brain_HC	6319	-0.357	0.103	-	-	-
Mm_Skin	13255	-0.139	0.56	2008	0.468	0.037
Hs_Skin	12017	0.015	0.881	85	-0.43	0
Mm_Liver_3	13255	-0.45	0.093	642	-0.459	0.085

Mm_Liver_2	13255	−0.368	0.177	836	−0.647	0.009
Mm_Liver_1	13241	−0.629	0.021	288	0.048	0.875
Mm_Lung_2	13255	−0.329	0.182	332	−0.737	0
Mm_Lung_1	13255	−0.58	0.102	265	−0.896	0.001
Mm_Kidney	13255	−0.223	0.375	494	−0.705	0.001
Mm_Spleen	13256	−0.287	0.282	320	0.197	0.465
Rn_Aorta_T	3332	0.521	0.038	304	−0.121	0.655
Rm_Aorta	10012	−0.258	0.223	-	-	-
Mm_Muscle_Gastro	9910	0.313	0.378	879	0.661	0.037
Mm_Muscle_P C	13255	−0.189	0.518	601	−0.755	0.002
Rn_Muscle_EDL	3332	0.66	0.027	-	-	-
Rn_Muscle_EOMs	3332	0.089	0.784	-	-	-
Hs_Muscle_VL2	13566	0.247	0.375	-	-	-
Hs_Muscle_VL1	13566	0.331	0.228	-	-	-
Hs_Muscle_Other	14168	0.004	0.995	-	-	-
Hs_Muscle_RA	14168	0.089	0.49	-	-	-
Hs_Muscle_BB	14168	0.31	0.196	2689	−0.207	0.396

APPENDIX F

RESULTS FOR THE CHANGES IN EXPRESSION - ω - WID ("ONE-TO-ONE ORTHOLOGS" BETWEEN HUMAN-MOUSE) CORRELATION WITH AGE

Table F.1: Results – number of genes, rho, p values - for the changes in expression - ω ("one-to-one orthologs" between human-mouse) correlation after excluding positively selected genes from ω, immun system-related genes and gene showing age-related decrease trend from gene-expression level (as indicated with the "WID" suffix).

Dataset Name	**# all**	**rho**	p	**# age-rel.**	**rho**	p
Hs_Brain_BA22	7424	−0.011	0.963	-	-	-
Hs_Brain_FC	3714	−0.128	0.509	213	−0.425	0.022
Hs_Brain_BA10	7983	−0.081	0.713	986	−0.507	0.013
Hs_Brain_EC	7228	−0.228	0.187	-	−	-
Hs_Brain_HC	8126	−0.029	0.858	950	−0.385	0.013
Hs_Brain_PCG	7351	−0.317	0.049	716	−0.271	0.095
Hs_Brain_SG	7042	−0.382	0.011	2282	−0.197	0.201
Hs_Brain_PFC	5803	−0.392	0.21	112	−0.629	0.032
Hs_Brain_CB	5195	−0.182	0.595	-	-	-
Rm_Brain_PFC	6078	−0.109	0.755	-	-	-
Rm_Brain_CB	5765	−0.083	0.843	-	-	-
Rn_Brain_HC	6426	−0.266	0.358	160	−0.797	0.001
Mm_Brain_WB	8172	0.06	0.838	102	−0.782	0.001
Mm_Brain_NC	9492	0.522	0.122	4524	0.313	0.378
Mm_Brain_HC	3950	−0.255	0.252	-	-	-
Mm_Skin	8247	−0.382	0.097	217	−0.711	0

Hs_Skin	9205	−0.004	0.966	71	−0.404	0
Mm_Liver_3	8001	−0.617	0.014	511	−0.719	0.003
Mm_Liver_2	8865	−0.713	0.003	415	−0.627	0.012
Mm_Liver_1	8967	−0.584	0.036	288	−0.666	0.013
Mm_Lung_2	8677	−0.266	0.285	214	−0.498	0.035
Mm_Lung_1	7682	−0.264	0.493	88	−0.949	0
Mm_Kidney	8589	−0.392	0.108	444	−0.674	0.002
Mm_Spleen	8014	−0.567	0.022	137	−0.719	0.002
Rn_Aorta_T	2125	0.036	0.894	262	−0.158	0.56
Rm_Aorta	5239	−0.194	0.364	-	-	-
Mm_Muscle_Gastro	4784	0.522	0.122	435	−0.313	0.378
Mm_Muscle_P C	7355	−0.314	0.273	392	−0.524	0.054
Rn_Muscle_EDL	2283	0.337	0.311	-	-	-
Rn_Muscle_EOMs	2373	0.059	0.855	-	-	-
Hs_Muscle_VL2	7898	0.045	0.874	-	-	-
Hs_Muscle_VL1	7124	−0.02	0.945	-	-	-
Hs_Muscle_Other	7523	0.046	0.873	-	-	-
Hs_Muscle_RA	8645	0.032	0.806	-	-	-
Hs_Muscle_BB	7411	−0.207	0.396	1044	−0.062	-

APPENDIX G

RESULTS FOR THE CHANGES IN EXPRESSION - ω - WID ("ONE-TO-ONE ORTHOLOGS" BETWEEN HUMAN-ELEPHANT) CORRELATION WITH AGE

Table G.1: Results – number of genes, rho, p values - for the changes in expression - ω ("one-to-one orthologs" between human-elephant) correlation after excluding positively selected genes from ω, immun system-related genes and gene showing age-related decrease trend from gene-expression level (as indicated with the "WID" suffix).

Dataset Name	**# all**	**rho**	p	**# age-rel.**	**rho**	p
Hs_Brain_BA22	7299	-0.211	0.386	-	-	-
Hs_Brain_FC	3706	-0.235	0.22	-	-	-
Hs_Brain_BA10	7868	-0.103	0.639	-	-	-
Hs_Brain_EC	7190	-0.274	0.111	-	-	-
Hs_Brain_HC	8028	-0.046	0.775	940	-0.458	0.003
Hs_Brain_PCG	7212	-0.382	0.016	755	-0.658	0
Hs_Brain_SG	7020	-0.354	0.018	2349	-0.393	0.008
Hs_Brain_PFC	5722	-0.378	0.227	-	-	-
Hs_Brain_CB	5171	-0.2	0.558	-	-	-
Rm_Brain_PFC	5995	-0.055	0.881	-	-	-
Rm_Brain_CB	5692	0	1	-	-	-
Rn_Brain_HC	5217	-0.266	0.358	102	-0.393	0
Mm_Brain_WB	6463	0.165	0.573	-	-	-
Mm_Brain_NC	7661	0.801	0.005	-	-	-
Mm_Brain_HC	2969	-0.313	0.156	-	-	-
Mm_Skin	6732	-0.364	0.114	-	-	-

Hs_Skin	8967	−0.011	0.916	69	−0.591	0.001
Mm_Liver_3	6338	−0.451	0.091	-	-	-
Mm_Liver_2	6957	−0.622	0.013	243	−0.622	0.035
Mm_Liver_1	6982	−0.666	0.013	-	-	-
Mm_Lung_2	6913	−0.21	0.403	-	-	-
Mm_Lung_1	6155	−0.527	0.145	-	-	-
Mm_Kidney	6781	0.034	0.892	254	−0.558	0.006
Mm_Spleen	6447	−0.585	0.017	-	-	-
Rn_Aorta_T	1609	0.255	0.341	-	-	-
Rm_Aorta	5223	−0.258	0.223	-	-	-
Mm_Muscle_Gastro	3803	0.035	0.924	332	−0.269	0.266
Mm_Muscle_P C	5950	−0.147	0.617	-	-	-
Rn_Muscle_EDL	1749	0.554	0.077	-	-	-
Rn_Muscle_EOMs	1822	−0.089	0.784	-	-	-
Hs_Muscle_VL2	7705	0.125	0.657	-	-	-
Hs_Muscle_VL1	7039	0.168	0.549	-	-	-
Hs_Muscle_Other	7422	−0.082	0.773	-	-	-
Hs_Muscle_RA	8449	0.04	0.755	-	-	-
Hs_Muscle_BB	7271	−0.186	0.446	-	-	-

APPENDIX H

RESULTS FOR THE CHANGES IN EXPRESSION - ω - WID ("ONE-TO-ONE ORTHOLOGS" BETWEEN HUMAN-COW) CORRELATION WITH AGE

Table H.1: Results – number of genes, rho, p values - for the changes in expression - ω ("one-to-one orthologs" between human-cow) correlation after excluding positively selected genes from ω, immun system-related genes and gene showing age-related decrease trend from gene-expression level (as indicated with the "WID" suffix).

Dataset Name	**# all**	**rho**	p	**# age-rel.**	**rho**	p
Hs_Brain_BA22	7499	-0.05	0.839	-	-	-
Hs_Brain_FC	3719	-0.144	0.456	-	-	-
Hs_Brain_BA10	8027	-0.019	0.93	-	-	-
Hs_Brain_EC	7328	-0.2	0.249	-	-	-
Hs_Brain_HC	8176	-0.004	0.979	956	-0.146	0.363
Hs_Brain_PCG	7421	-0.247	0.129	752	-0.58	0
Hs_Brain_SG	7173	-0.264	0.083	2348	-0.299	0.049
Hs_Brain_PFC	5770	-0.322	0.308	-	-	-
Hs_Brain_CB	5251	-0.073	0.839	-	-	-
Rm_Brain_PFC	6056	-0.055	0.881	-	-	-
Rm_Brain_CB	5762	-0.017	0.982	-	-	-
Rn_Brain_HC	5274	-0.23	0.428	100	-0.356	0
Mm_Brain_WB	6558	0.087	0.768	-	-	-
Mm_Brain_NC	7739	0.801	0.005	-	-	-
Mm_Brain_HC	2993	-0.197	0.38	-	-	-
Mm_Skin	6837	-0.33	0.156	-	-	-

Hs_Skin	9071	0.013	0.903	72	−0.387	0.038
Mm_Liver_3	6427	−0.501	0.057	-	-	-
Mm_Liver_2	7073	−0.555	0.032	247	−0.734	0.009
Mm_Liver_1	7104	−0.598	0.031	-	-	-
Mm_Lung_2	7011	−0.141	0.577	-	-	-
Mm_Lung_1	6257	−0.158	0.685	-	-	-
Mm_Kidney	6867	−0.06	0.814	258	−0.576	0.004
Mm_Spleen	6531	−0.603	0.013	-	-	-
Rn_Aorta_T	1607	0.097	0.721	-	-	-
Rm_Aorta	5274	−0.162	0.451	-	-	-
Mm_Muscle_Gastro	3836	0.453	0.189	345	−0.103	0.674
Mm_Muscle_P C	5992	0.021	0.943	-	-	-
Rn_Muscle_EDL	1754	0.409	0.211	-	-	-
Rn_Muscle_EOMs	1833	0	1	-	-	-
Hs_Muscle_VL2	7870	0.15	0.593	-	-	-
Hs_Muscle_VL1	7179	0.309	0.262	-	-	-
Hs_Muscle_Other	7533	0.025	0.934	-	-	-
Hs_Muscle_RA	8623	0.038	0.767	-	-	-
Hs_Muscle_BB	7415	0	1	-	-	-

APPENDIX I

GO BIOLOGICAL PROCESS CATEGORIES COMMON AMONG ALL DATASETS ENRICHED FOR IELC

Table I.1: GO Biological Process categories common among all datasets (n=14) enriched for IELC.

GO ID	GO Term
GO:0097190	apoptotic signaling pathway
GO:0097191	extrinsic apoptotic signaling pathway

APPENDIX J

GO BIOLOGICAL PROCESS CATEGORIES COMMON AMONG LIVER DATASETS ENRICHED FOR IELC

Table J.1: GO Biological Process categories common among liver datasets (n=3) enriched for IELC.

GO ID	GO Term
GO:0000302	response to reactive oxygen species
GO:0001525	angiogenesis
GO:0001568	blood vessel development
GO:0001775	cell activation
GO:0001816	cytokine production
GO:0001817	regulation of cytokine production
GO:0001818	negative regulation of cytokine production
GO:0001819	positive regulation of cytokine production
GO:0001890	placenta development
GO:0001944	vasculature development
GO:0002218	activation of innate immune response
GO:0002221	pattern recognition receptor signaling pathway
GO:0002224	toll-like receptor signaling pathway
GO:0002237	response to molecule of bacterial origin
GO:0002250	adaptive immune response
GO:0002263	cell activation involved in immune response
GO:0002366	leukocyte activation involved in immune response
GO:0002376	immune system process
GO:0002443	leukocyte mediated immunity
GO:0002460	adaptive immune response based on somatic

	recombination of immune receptors built from immunoglobulin superfamily domains
GO:0002520	immune system development
GO:0002521	leukocyte differentiation
GO:0002573	myeloid leukocyte differentiation
GO:0002682	regulation of immune system process
GO:0002683	negative regulation of immune system process
GO:0002684	positive regulation of immune system process
GO:0002685	regulation of leukocyte migration
GO:0002687	positive regulation of leukocyte migration
GO:0002757	immune response-activating signal transduction
GO:0002758	innate immune response-activating signal transduction
GO:0002764	immune response-regulating signaling pathway
GO:0002819	regulation of adaptive immune response
GO:0006066	alcohol metabolic process
GO:0006766	vitamin metabolic process
GO:0006873	cellular ion homeostasis
GO:0006874	cellular calcium ion homeostasis
GO:0006875	cellular metal ion homeostasis
GO:0006897	endocytosis
GO:0006898	receptor-mediated endocytosis
GO:0006909	phagocytosis
GO:0006928	movement of cell or subcellular component
GO:0006952	defense response
GO:0006954	inflammatory response
GO:0006955	immune response
GO:0006959	humoral immune response
GO:0006979	response to oxidative stress
GO:0007155	cell adhesion
GO:0007159	leukocyte cell-cell adhesion
GO:0007160	cell-matrix adhesion
GO:0007204	positive regulation of cytosolic calcium ion concentration
GO:0007249	I-kappaB kinase/NF-kappaB signaling
GO:0007259	JAK-STAT cascade

APPENDIX K

GO CELLULAR COMPONENT CATEGORIES COMMON AMONG ALL DATASETS ENRICHED FOR IELC

Table K.1: GO Cellular Component categories common among all datasets (n=14) enriched for IELC.

GO ID	GO Term
GO:0009897	external side of plasma membrane

APPENDIX L

GO CELLULAR COMPONENT CATEGORIES COMMON AMONG BRAIN DATASETS ENRICHED FOR IELC

Table L.1: GO Cellular Component categories common among brain datasets (n=9) enriched for IELC.

GO ID	GO Term
GO:0005764	lysosome
GO:0005773	vacuole
GO:0009897	external side of plasma membrane
GO:0031982	vesicle
GO:0031988	membrane-bounded vesicle
GO:0065010	extracellular membrane-bounded organelle
GO:0070062	extracellular exosome
GO:1903561	extracellular vesicle

APPENDIX M

GO CELLULAR COMPONENT CATEGORIES COMMON AMONG BRAIN DATASETS ENRICHED FOR DEHC

Table M.1: GO Cellular Component categories common among brain datasets (n=9) enriched for IELC.

GO ID	GO Term
GO:0030425	dendrite
GO:0034702	ion channel complex
GO:0036477	somatodendritic compartment
GO:0043005	neuron projection
GO:0043025	neuronal cell body
GO:0043235	receptor complex
GO:0045202	synapse
GO:0045211	postsynaptic membrane
GO:0097060	synaptic membrane
GO:0098794	postsynapse
GO:0098797	plasma membrane protein complex
GO:1902495	transmembrane transporter complex

APPENDIX N

GO MOLECULAR FUNCTION CATEGORIES COMMON AMONG BRAIN DATASETS ENRICHED FOR DEHC

Table N.1: GO Molecular Function categories common among brain datasets (n=9) enriched for IELC.

GO ID	GO Term
GO:0000981	RNA polymerase II transcription factor activity sequence-specific DNA binding
GO:0005216	ion channel activity
GO:0005261	cation channel activity
GO:0008324	cation transmembrane transporter activity
GO:0015267	channel activity
GO:0015276	ligand-gated ion channel activity
GO:0022803	passive transmembrane transporter activity
GO:0046873	metal ion transmembrane transporter activity

APPENDIX O

GO MOLECULAR FUNCTION CATEGORIES COMMON AMONG LUNG DATASETS ENRICHED FOR DEHC

Table O.1: GO Molecular Function categories common among lung datasets (n=2) enriched for IELC.

GO ID	GO Term
GO:0000975	regulatory region DNA binding
GO:0000976	transcription regulatory region sequence-specific DNA binding
GO:0000977	RNA polymerase II regulatory region sequence-specific DNA binding
GO:0000978	RNA polymerase II core promoter proximal region sequence-specific DNA binding
GO:0000981	RNA polymerase II transcription factor activity sequence-specific DNA binding
GO:0000982	transcription factor activity, RNA polymerase II core promoter proximal region sequence-specific binding
GO:0000987	core promoter proximal region sequence-specific DNA binding
GO:0000988	transcription factor activity, protein binding
GO:0000989	transcription factor activity, transcription factor binding
GO:0001012	RNA polymerase II regulatory region DNA binding
GO:0001047	core promoter binding
GO:0001071	nucleic acid binding transcription factor activity
GO:0001076	transcription factor activity, RNA polymerase II transcription factor binding
GO:0001077	transcriptional activator activity, RNA polymerase II

	core promoter proximal region sequence-specific binding
GO:0001159	core promoter proximal region DNA binding
GO:0001227	transcriptional repressor activity, RNA polymerase II transcription regulatory region sequence-specific binding
GO:0001228	transcriptional activator activity, RNA polymerase II transcription regulatory region sequence-specific binding
GO:0003677	DNA binding
GO:0003682	chromatin binding
GO:0003700	transcription factor activity, sequence-specific DNA binding
GO:0003713	transcription coactivator activity
GO:0004842	ubiquitin-protein transferase activity
GO:0004871	signal transducer activity
GO:0004888	transmembrane signaling receptor activity
GO:0004930	G-protein coupled receptor activity
GO:0008134	transcription factor binding
GO:0017016	Ras GTPase binding
GO:0019787	ubiquitin-like protein transferase activity
GO:0031267	small GTPase binding
GO:0038023	signaling receptor activity
GO:0043565	sequence-specific DNA binding
GO:0043566	structure-specific DNA binding
GO:0044212	transcription regulatory region DNA binding
GO:0051020	GTPase binding
GO:0061630	ubiquitin protein ligase activity