Annual Review of Gerontology and Geriatrics
doi: 10.1891/0198-8794.33.233

Chapter 11: Gene Expression Biomarkers and Longevity

David Melzer *[1], Luke C. Pilling[1], Alexander D. Fellows[2], Alice C. Holly[2], Lorna W. Harries[2], and Luigi Ferrucci[3]

[1]Epidemiology and Public Health Group, Medical School, University of Exeter, Exeter EX1 2LU, UK
[2]Institute of Biomedical and Clinical Sciences, Medical School, University of Exeter, Exeter EX1 2LU, UK
[3]National Institute on Aging, Baltimore, Maryland, USA

1 Introduction

Chronological age, a count of how many orbits of the sun an individual has made as a passenger of planet earth, is a useful but limited proxy of aging processes. Some individuals die of age related diseases in their sixties, while others live to double that age. As a result, a great deal of effort has been put into identifying biomarkers that reflect the underlying biological changes involved in aging. These markers would provide insights into what processes were involved, provide measures of how much biological aging had occurred and provide an outcome measure for monitoring the effects of interventions to slow ageing processes.

Our DNA sequence is the fixed reference template from which all our proteins are produced. With the sequencing of the human genome we now have an accurate reference library of gene sequences. The recent development of a new generation of high throughput array technology makes it relatively inexpensive to simultaneously measure a large number of base sequences in DNA (or RNA, the molecule of gene expression). In the last decade, array technologies have supported great progress in identifying common DNA sequence differences (SNPs) that confer risks for age related diseases, and similar approaches are being used to identify variants associated with exceptional longevity [1]. A striking feature of the findings is that the majority of common disease-associated variants are located not in the protein coding sequences of genes, but in regions of the genome that do not produce proteins. This indicates that they may be involved in the regulation of nearby genes, or in the processing of their messages.

While DNA holds the static reference sequences for life, an elaborate regulatory system influences whether and in what abundance gene transcripts and proteins are produced. The relative abundance of each transcript is a good guide to the demand for each protein product in cells (see section 2 below). Thus, by examining gene expression patterns or signatures associated with aging or age related traits we can peer into the underlying production processes at a fundamental level. This approach has already proved successful in clinical applications, for example using gene signatures to classify cancer subtypes [2]. In aging research, recent work conducted in the InCHIANTI cohort has identified gene-expression signatures in peripheral leucocytes linked to several aging phenotypes, including low muscle strength, cognitive impairment, and chronological age itself. In the sections that follow we provide a brief introduction to the underlying processes involved in gene expression, and summarize key work in laboratory models of aging. We then provide an overview of recent work in humans, thus far mostly from studies of circulating white cells.

2 Introducing gene expression

Since the early 1900s a huge worldwide research effort has lead to the discovery and widespread use of genetic science (see the NIH website [3] for a comprehensive review of the history of the subject, and a more detailed description of the transfer of genetic information). The human genome contains the information needed to create every protein used by cells. The information in the DNA is transcribed into an intermediate molecule known as the messenger RNA (mRNA), which is then translated into the sequence of aminoacids (proteins) which ultimately determine the structural and functional characteristics of cells, tissues and organisms (see figure 1 for a summary of the process). RNA is both an intermediate to proteins and a regulatory molecule; therefore the transcriptome (the RNA

*Address correspondence to Prof. David Melzer, Epidemiology and Public Health Group, Medical School, University of Exeter, Exeter EX1 2LU, UK. E-mail: D.Melzer@exeter.ac.uk

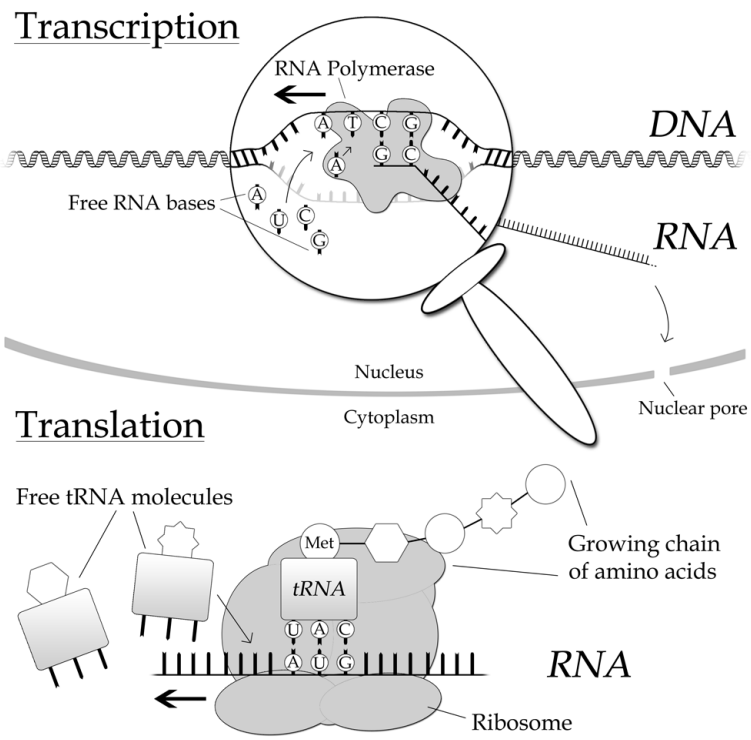

Figure 1: **Representation of the transcription and translation processes from DNA to RNA to Protein** — DNA makes RNA makes Protein. This is the central dogma of molecular biology, and describes the transfer of information from DNA (made of four bases; Adenine, Guanine, Cytosine and Thymine) to RNA to Protein (made of up to 20 different amino acids). Machinery known as RNA polymerase carries out transcription, where a single strand of RNA is created that is complementary to the DNA (i.e. the sequence is the same, but inverted although in RNA thymine (T) is replaced by uracil (U)). Not all RNA molecules are messenger RNA (mRNA) molecules: RNA can have regulatory functions (e.g. micro RNAs), and or can be functional themselves, for example in translation transfer RNA (tRNA) molecules have an amino acid bound to one end (the individual components of proteins) and at the other bind to a specific sequence of RNA (a codon again, this is complementary to this original sequence) for instance in the figure a tRNA carrying methionine (Met) can bind to the sequence of RNA, and the ribosome (also in part made of RNA) attaches the amino acids together to form a protein.

production of a particular cell, or sample of cells, at a given time) is of particular interest in determining the underlying molecular mechanisms behind specific traits and phenotypes. Genes are also regulated at the post-transcriptional level, by non-coding RNAs or by post-translational modifications to the encoded proteins.

Transcription is a responsive process (many factors regulate transcription and translation in response to specific intra and extra-cellular signals), and thus the amount of RNA produced varies over time and between cell types and tissues. In addition to the gene and RNA transcript sequences that will determine the final protein sequence (so called exons) there are also intervening sections (the introns) that are removed by a process known as mRNA splicing. While it was once assumed that each gene produced only one protein, it is now clear that up to 90% of our genes can produce different versions of their protein through varying the number of exons included in the protein, a process called alternative splicing. Alteration in the functional properties of the protein can be introduced by varying which exons are included in the transcript, giving rise to different isoforms of the same gene. Many RNA regulatory factors govern this process, and variations to the DNA sequence can affect the binding of these factors (which can be thousands of base pairs from the gene itself) and alter when, where and for how long a particular transcript is produced.

The amount of mRNA produced for a protein is not necessarily directly related to the amount of protein produced or present, as other regulatory processes are involved. The amount of mRNA is broadly indicative of

the demand for a particular protein (more mRNA for gene A indicates that more protein A is required). RNA studies can also provide unique insight into the activity of non-protein-coding RNAs, such as small RNAs (microRNAs) or long no-coding RNAs (lncRNAs), which often have regulatory functions and influence both normal biology and disease states by causing degradation of target transcripts, or blocking the translation of those targets (miRNAs), or by influencing the activity of target genes by altering chromatin conformation, promoter methylation status or mRNA splicing patterns (lncRNAs).

3 Expression studies of aging in model organisms

Until recently, the majority of work examining gene expression in relation to aging and longevity has used animal models. Animal models offer advantages in that experimental interventions are possible and all tissues of interest are accessible. The main species used in aging research have been the nematode worm C.elegans, the Drosophila genus of fruitfly, and rodents, all chosen for being relatively short-lived, to expedite the research process. There are a limited number of studies on primates and higher mammals. Across these models, gene expression studies have been prominent in identifying the major pathways underlying aging processes, including energy sensing, telomere function and TOR signaling (controlling cell division, senescence and autophagy), mitochondrial function and inflammation.

3.1 Gene expression in C. elegans

Much of the initial work on age-related changes in gene expression was carried out in C. elegans nematode worms with mutations which affect lifespan. Genes involved in the insulin signaling pathway such as DAF-2 SKN-1 and DAF-16 have previously been implicated in aging in C. elegans [4]. Over-expression of the SIR-2 deacetylases have also been reported to increase lifespan in worms by interacting with DAF-16/FOX0 [4], but has also been implicated in maintaining chromatin structure at telomeres [5]. Genes involved in stress response have also been shown to be altered in aging in worms [6]. Genes such as DAF-16 and other genes involved in insulin signaling or energy sensing have also been identified in these microarray studies [7].

The TOR signaling pathway is also altered in aging in this species; the FOXA group of transcription factors, involved in autophagy, has been shown to be differentially expressed in aging worms [8].

Mutations in genes such as ISP-1 and CLK-1, involved in mitochondrial function, were also seen to result in altered lifespan [9]. The advent of whole transcriptome profiling has indentified over 1200 genes associated with aging in C. elegans. These studies have highlighted the role of GATA transcription factors such as ELT-3, ELT-5, and ELT-6 in aging in C. elegans [10].

3.2 Gene expression in Drosophila species

Microarray studies in Drosophila melanogaster have also identified several groups of genes associated with aging. As in C. elegans, genes involved in energy sensing and stress response were also found to be up-regulated in both aged and calorifically restricted flies [11] [12], whereas genes involved in mitochondrial function, reproduction, JUN kinase signaling and muscle contraction were found to be down-regulated in aged flies [13, 14]. Genes such as Attacin A, Attacin B, Attacin C, and Attacin D, which are associated with defense against bacterial pathogens have been found to show increases in gene expression with age [11].

3.3 Gene expression in rodents

Several whole transcriptome expression profiling studies have determined gene expression patterns in multiple tissues in the aging mouse [15-18]. Many of these studies have been curated into the AGEMAP (Atlas of Gene Expression in Mouse Aging Project) gene expression database. This work showed that gene expression patterns in mouse demonstrate tissue-specific differences in the transcripts that demonstrate expression differences during the aging process as they do in C. elegans. Zahn et al. discovered that several different patterns of aging occurred in mice; one common to neuronal tissues, one found in vascular tissues and a different pattern in steroid responsive tissues [16]. This tissue-specific pattern of age-related expression changes is also noted in other studies, with different patterns of transcripts being up- or down- regulated in different tissues, although some pathways were common to more than one tissue (for example inflammatory transcripts, histocompatibility transcripts and those involved in the complement cascade were up-regulated in both heart and brain tissues)[18].

Down-regulated genes showed more tissue-specific pattern of gene expression, with genes involved in glycolysis, energy metabolism and mitochondrial function being down-regulated in response to aging in muscle, whereas genes involved in fatty acid oxidation and beta-oxidation demonstrated decreases in expression in aging rodent heart [18].

3.4 Relevance to human aging

The animal models have given us a very good starting point for the analysis of gene expression changes in human aging. However, these models are usually genetically homogeneous, short-lived compared to man, and lack the complex interplay between genetics and environment apparent in the human population, as they are generally maintained in laboratory conditions. This raises questions about the transferability of results between the model species and man. Evidence is emerging that although useful, animal models may have critical limitations, and some of the findings from animal models do not directly translate to humans. Although evidence is building that alterations to specific pathways (such as insulin signaling, mTOR signaling, mitochondrial function) may be conserved throughout the animal kingdom [19], most age-related expression changes appear to be specific to a given species [20]. For example, while experimental models that down-regulate the IGF-1/Insulin signaling in invertebrates and in mice are generally associated with longer life span, insulin resistance in humans is associated with severe pathology and drastically reduced life expectancy. Furthermore, the lack of overall correlation of regulation of age-related traits even in mammalian species suggests that aging processes may differ between even closely related species [16]; this is exemplified by the observation that the pattern of expression changes noted in aging human brain cortex bears almost no resemblance to those observed in chimpanzee cortex, an organism with which we share 97% of our DNA [21]. These findings highlight the need for human studies to be able to fully unravel the processes underpinning the aging process in man.

4 Candidate gene expression studies and human longevity: p16

There have been few molecular markers which show promise as biomarkers of human aging, but perhaps the most promising is the cell cycle / cell senescence associated gene p16 (also called p16ink4a, or cyclin-dependent kinase inhibitor 2A (CDKN2A), located on chromosome 9: locus 9p21). Gene specific expression studies show good correlations with age in rodents and humans, and laboratory models indicate that this gene plays a causal role in aging. To add to this, a recent study removing p16 expressing senescent cells had a positive effect on age related pathology.

Two distinct isoforms are expressed from the CDKN2A, namely p16INK4a (p16) and p14ARF (ARF). Functionally the two proteins are similar in that they are both tumor suppressors and thus negatively regulate the cell cycle, however their mechanisms of action are distinct [22]. Isoform p16 binds to and inactivates CDK4, an important regulator of retinoblastoma (Rb, a potent tumor suppressor), thereby promoting the Rb pathway. In contrast, ARF promotes p53 activity (another potent tumor suppressing protein) by binding to and inhibiting the MDM2 protein (a negative regulator of p53). Their importance as tumor suppressors is emphasized by the fact that elements of the p16/ARF pathways are disrupted in nearly all human cancers [22] [23].

While the mechanism that cause cells to transition into senescence has not been fully elucidated, it is well know that multiple types of stimuli can induce this process, including telomere erosion, DNA damage too severe to be repaired, mitogenic signals, physical and chemical stress in cell cultures, oxidative stress, P53 induction and probably many others. Growth and proliferation arrest, expression of p16ink4a and the development of a pro-inflammatory secretory phenotype are all hallmarks of senescence. Although the physiological function of the senescence program remain unknown, it is generally accepted that activation of senescence is a protective strategy to avoid cancer, but that the accumulation of senescent cells contribute to the reduction in tissue function that is associated with aging (see Figure 2).

Increased p16 expression is a major player in preventing cell division and proliferation of damaged cells and cellular senescence[24]. Expression of p16 increases with age [25] in animal species and in man. The up-regulation of p16 levels during aging is due to a loss of binding of the Polycomb Repressive Complexes 1 and 2 (PRC1, PRC2) to the genomic region containing the CDKN2A and CDKN2B genes during cell senescence [26]. The PRC1 and 2 complexes direct the addition of methyl residues to specific parts of the histone scaffold that determines its conformation, and thus the accessibility of the gene promoters to transcription factors. The loss of PRC1 and 2 binding to CDKN2A and CDKN2B genes causes a change in histone modification with a resultant up-regulation of p14, p15 and p16 activity.

Recent evidence in mice shows that p16-associated senescent cells actually cause age-related tissue dysfunction, hypothesized to be due to the secretion of pro-inflammatory factors [27]. By targeting a cell killing drug to cells expressing p16, Baker et al showed delayed onset of age-related pathologies (in tissues such as adipose, skeletal muscle and the eye) in a mouse model of progeria [28]. Even if clearance of the p16-positive senescent cells occurred after the onset of age-related pathologies, disease progression was stalled by the re-

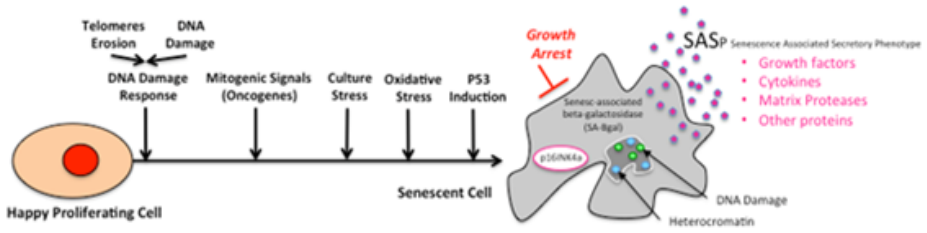

Figure 2: **Diagrammatic representation of the transition between healthy proliferating cells and cell senescence** — Senescent cells are those that have permanently left the cell-cycle and no longer proliferate, which is hypothesized to be an evolved tumor suppressing mechanism. Their phenotypes change as they secrete a milieu of factors known as the Senescence Associated Secretary Phenotype, known to be pro-inflammatory and pro-senescent (as well as, ironically, pro-cancer). The senescent phenotype is induced by a series of insults to the cell, mostly arising from the normal functioning and exposures of a proliferative cell.

moval of the senescent cells. This marks an exciting step forward in aging research; not only has a gene expression biomarker (p16) been causally implicated in the aging process, but Baker et al have shown that progression of some age-related disease can be postponed by removal of cells expressing p16 protein.

5 Analyzing genome wide expression data from humans

In recent years the gene expression field has changed from focusing on small sets of candidate genes to studying very large numbers of transcripts simultaneously. Current generations of gene expression chips are capable of determining the relative amounts of mRNA being produced for tens of thousands of transcripts across the genome, providing an increasingly comprehensive or genome wide view of gene expression. Below we discuss the challenges of analyzing such data and then review some of the emerging findings.

RNA is a fragile molecule and therefore on collection of samples from participants RNA needs to be stabilized as soon as possible. If this stabilization step does not happen quickly, a bias towards the survival of more robust longer lived transcripts can be introduced (some RNA molecules have half-lives of only a matter of minutes). Levels of some RNA transcripts can also increase during storage. In the authors studies, PAXgene collection tubes (from Qiagen) were used to preserve RNA levels of peripheral blood (leukocytes) as these were extracted from the respondents vein. The development of inexpensive and flexible blood collection systems such as PAXgene has accelerated the research in the field of aging, as other fields, allowing the collection of large cohorts suitable for epidemiological studies. Following extraction and stabilization, a microarray chip was used to generate a relative expression value for each measured transcript. The number of transcripts measured depends on the coverage of the specific microarray chip used these can be genome-wide, specific (only certain transcripts), or exonic (i.e. only protein-coding transcripts are measured).

Gene expression data generated using microarrays essentially reports the relative abundance of each probe, which correspond (or map) to one or more specific gene transcript. These resulting databases of study participant specific gene transcript abundances can then be statistically analyzed to identify the transcripts that are differentially expressed with a phenotype of interest. Until recently this was reserved for matched case/control studies using relatively small sample sizes, however due to cost reduction this technology is now available for larger samples of hundreds or even thousands of human participants in epidemiological studies of aging.

Methods for data extraction or curation vary depending on the technology used and the context of the experiment. Here we aim to provide an overview of the different methods for data analysis. Array expression data is commonly Z-transformed so that each value represents the difference in standard deviations from the mean probe value, which is useful for two reasons; as the original probe data has no units this gives the values some tangibility to the reader; it also allows direct comparison between genes that might in real terms be expressed at very different concentrations. Typically a quality control step is applied to the microarray data to remove any outlier samples (e.g. not enough probes differ from the background noise), and any probes that were not expressed. Non-expression can be due to several reasons including technical failures (e.g. RNA fragment did not bind to the probe correctly, or the RNA sample was degraded), or the RNA was not produced in high enough concentrations in the studied cell type.

Assessing statistical associations between gene expression levels and gross phenotypes, for instance a disease-state or a blood marker, and inferring biological

meaning from the results are two distinct challenges.

5.1 Determining expression associations

The mainstay of analysis is testing for associations between the expression level of each gene of interest (as measured by probe binding intensity) and the phenotype under assessment, using well established multivariate statistical methods, albeit repeated as many times as there are valid probes measured. Statistical associations between two variables can be estimated using multiple regression analysis (logistic, linear, etc as appropriate), where the size of the association (the coefficient) is derived, as well as indicators of statistical confidence (the confidence intervals and associated P-value) in the reported association. The advantages of this approach are that confounding factors can be adjusted for, and thus improve the robustness of the result. By individually assessing each probes association with a particular outcome it is possible to determine those transcripts most closely associated with a particular trait.

An important consideration when screening more than one transcript for statistical associations is choosing a suitable significance level or p-value, which will be taken to suggest significant association. The conventional p-value=0.05 threshold equates to a 5% probability that a single test result could have arisen by chance alone: probabilities lower than this are widely accepted as reasonable statistical evidence for a real association.

The probability of falsely reporting a positive association where none exists inevitably increases with the number of tests performed. False-positive results generated by testing for multiple associations can be controlled for using one of several techniques; these include the Bonferroni correction [29], which achieves good control of false positives, but tends to have low power to detect real differences (high numbers of false negatives). An alternative approach is the false discovery rate (FDR)[30], which yields better power at the cost of slightly higher chosen likelihood of false positives [31].

Many small-scale studies, such as one by Welle et al [32], find many hundreds of genes to be differentially expressed between old and young participants. Some of these associations may be true, but due to the small sample sizes involved (e.g. 7 young and 8 old in the cited study) many are likely to be false-positives. Online databases such as the Digital Ageing Atlas [33] collect the results of these studies for broader views of gene and transcriptome associations with age-related phenotypes.

In performing a 1-by-1 analysis of transcript expression associations with a phenotype certain assumptions are made. For instance, a linear regression model assumes that the outcome, or dependent, variable is normally distributed. To ensure that this assumption is not violated probe abundance values can be transformed into a normal distribution (for example log2 transformation) prior to analysis. This is completely separate to data normalization for known systematic biases (see Mark Reimers website for information about data normalization [34]).

An alternative analysis approach is to use machine learning algorithms to determine underlying relationships in a given set of data. These methods fall into two broad categories; supervised and unsupervised analyses. In both cases the machine learning algorithms aim to group similar samples together based on the data. In a supervised analysis the investigator may supply classification information, e.g. cancerous/non-cancerous cells [35], to train the algorithm with regard to that specific phenotype so that later, in a discovery set, the algorithm may correctly classify the samples based purely on the data. In an unsupervised analysis no information is given other than the data to be analyzed this is known as clustering and often reveals underlying structure within a dataset.

5.2 Biological inference

Once each transcript has been assessed for statistical associations with an outcome, interpretation of the results can begin. Techniques vary depending on the number of significant associations if very few genes are associated, then information can easily be found using databases such as GeneCards [36] or the National Center for Biotechnology web interface (http://www.ncbi.nlm.nih.gov/guide/). Alternatively if many associations are reported, it may be more practical to consider the broader biological processes that the genes represent.

A popular method of converting a list of genes into meaningful results is to use one of the many bioinformatics tools and resources available, to infer the higher-order biological pathways, functions or processes represented by the genes. In table 1 we have listed some of the freely available web tools.

Public repositories of information about higher-order biological processes, and how the individual genes and proteins fit into each function or pathway, are used to infer functions that are over-represented in a list of genes, above what you might expect by chance. These include Gene Ontology [37] and KEGG (the Kyoto Encyclopedia of Genes and Genomes) [38]. Biological processes are grouped in these systems so that every gene and pathway is classified into a hierarchical set of categories, with Biological Process at the top and with each child becoming gradually more specific. In figure 3 we have given an example of this data structure as displayed by WebGestalt (most tools can give an equivalent

Table 1: Examples of freely available bioinformatic tools for gene set analysis of enriched functional categories

Package	Details
DAVID	Database for Annotation, Visualization and Integrated Discovery. (http://david.abcc.ncifcrf.gov)
BiNGO	Biological Network Gene Ontology tool. (http://www.psb.ugent.be/cbd/papers/BiNGO)
PANTHER	Protein ANalysis THrough Evolutionary Relationships. (http://www.pantherdb.org)
WebGestalt	WEB-based GEne SeT AnaLysis Toolkit. (http://bioinfo.vanderbilt.edu/webgestalt)

network diagram), in the context of 6 immune-system related genes. Highlighted categories (appear in red) indicate which biological processes are overrepresented in the genes.

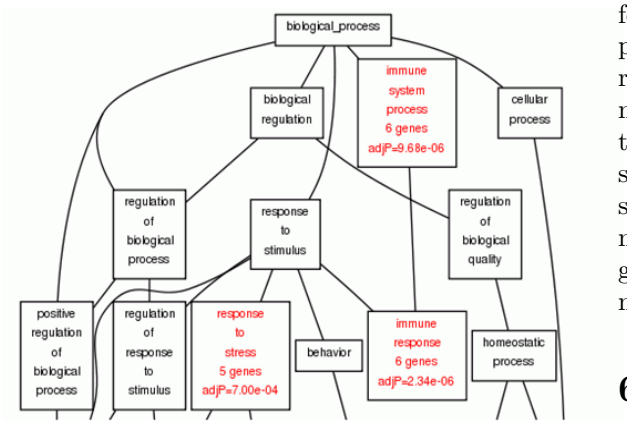

Figure 3: **Example diagram of gene transcript sets overrepresented in associations with a phenotype and summarized into biological processes (strong associations in red)** — Example of an enrichment analysis of Biological Processes (as defined by Gene Ontology) in a list of 6 immune system-related genes (CCR7, LTB, CCL5, CCR6, CD27, LAG3) using WebGestalt, a web-based bioinformatic tool used to infer the over-represented processes associated with a list of genes. The writing in red indicates an associated process (adjusted P-value also supplied) processes with black writing make up the rest of the tree structure. This image has been cropped.

5.2.1 Gene set enrichment analysis

Gene Set Enrichment Analysis (GSEA) [39] is another example of a free tool that makes use of public gene ontology databases to infer biological function, but it uses the expression data itself to generate an enrichment score (coefficient) for each process as a whole, and from there determines statistical significance. There are advantages to this type of approach, in that the direction and magnitude of the association for each gene and thus each pathway can also be taken into account, giving another level of information not available when just analyzing the lists. Disadvantages include no option to adjust the analysis for potential confounding factors, and a limited selection of algorithms.

Alternative programs to GSEA include PathVar [40], which, apart from performing GSEA-like analyses, also has the ability to extract overall pathway expression variance information (i.e. an individual phenotype for each functional pathway, with a value for each sample). This extracted data can then undergo adjusted regression analysis like any other however this usefulness is also limited in that it is not possible to determine the contribution of each gene to the pathway variance score. There are also commercial packages available, such as MetaCore [41] and Ingenuity [42], which offer much the same functionality as described above, but generally with much easier to use interfaces and many more custom/specific features.

6 Findings from human aging studies

Until recently, studies of gene expression in aging were restricted to small scale case/control experiments (<100 participants): results obtained from these can be hard to interpret due to difference in experimental design and the high chance of producing false positive results. Recently, larger scale studies have appeared with documentation of methods, adjustment for potential confounders and statistical significance methods suitable for multiple statistical testing.

6.1 The InCHIANTI study of aging

Over the past few years the author group has used InCHIANTI [43], a population study of aging conducted in Italy, to assess gene transcript associations in circulating leukocytes with aging and age-related traits. In this study peripheral blood samples from 698 participants aged 30 to 104 years were stabilized for RNA extraction at the point of collection, so the resulting expression abundance reflects the in-vivo expression in leucocytes. Immune cells circulating in the blood are one of the few relatively accessible sources of RNA from humans, and have been shown to interact with a wide range of tissues, and the age related changes in white cells are an important aspect of aging. Using the Illumina HT12 microarray chip 48,803 probes were scanned, of which 16,571 were expressed in at least 5%

of the sample. These were carried forwards into our analyses.

Using some of the analysis methods described in section 3 we have estimated transcript associations with age [44]. 2% (295) of these transcripts were robustly associated with age, much less than originally expected. This suggests that molecular aging is in a specific subset of RNAs, as opposed to random/global deregulation. Less than 1% of biological processes were associated with age in an analysis of pathways, which prominently included the processing of mRNAs. This provided an interesting hypothesis that disruption to alternative splicing is characteristic of aging cells, and may provide a mechanism for the buildup of abhorrent or mis-folded proteins within the cells. Evidence in support of this hypothesis included up-regulated of transcripts such as the ubiquitin and proteosomal pathway, which identify and dispose of aberrant or misfolded proteins. This could be of particular interest with regards to age-related diseases such as Alzheimers and Parkinsons, in which buildups of misfolded proteins are characteristic of the progression of these diseases.

Genes most prominent in the analysis of individual transcripts included LRRN3, LEF1 and CD248, all of which are involved in immune cell function, and all of which were found to be down-regulated with age. This may represent tissue-specific changes in expression rather than more general markers of aging, but could be critical given that immune-system changes are characteristic of, and possibly contributory to, age-related phenotypes such as increased susceptibility to infection and chronic low-grade inflammation [45]. Similar results have been reported from the San Antonio Family Heart Study (see section 6.2), and more recently the Leiden Longevity Study (see section 6.3).

In an exploratory analysis, we determined that expression of just 6 genes expression discriminated between old (>75 years) and young (<65 years) participants. The genes selected for inclusion in the test half of the dataset were LRRN3, CD248, CCR6, GRAP, VAMP5 and CD27 (see Figure 4 for associations with age) based on their (not over-lapping) contribution to the classification. In the replication data this gene set had exceptionally high discrimination power (area under the ROC curve >95%) Current work in progress includes examining whether these markers or alternatives are responsive to change over time and whether they predict health outcomes.

In an analysis of the biological pathways associated with age we found genes involved in regulation of mRNA splicing or mRNA processing (5 capping of the mRNA molecule and polyadenylation) to be strongly significant (P <0.001). The remaining pathways featured genes involved in chromatin remodeling, which influences the transcriptional activity of genes. All these probes were negatively associated with age, indicating that loss of specialization and dedifferentiation may be a key characteristic of molecular aging.

Figure 4: **Spline point regression association between relative transcript abundance in leukocytes and age for 6 discriminant genes: data from the InCHIANTI study** — We can classify subjects as old (>75) or young (<65) correctly over 90% of the time using the expression levels of 6 genes. This plot shows the individual relationship of each gene with age in the InCHIANTI cohort (sample size = 695) using penalized cubic regression spline models, applying the methods as described in [48]. Only 1 of these genes (VAMP5) has a positive association with age; its function is related to the trafficking of vesicles to the cell membrane. The y-axis represents the expression levels at the age relative to the expression at the mean age (72 years) therefore, whilst the genes may vary in absolute expression, the relative expression and thus their relationships with age can here be compared. Confidence intervals are not included for graph-clarity, however the p-values for the spline terms were LRRN3 ($P<2x10^{-16}$), CD248 ($P<2x10^{-16}$), CCR6 ($P=9x10^{-11}$), GRAP ($P=1x10^{-12}$), VAMP5 ($P=4x10^{-10}$) and CD27 ($P=7x10^{-9}$), indicating very strong associations with age.

6.1.1 Muscle strength

Loss of muscle strength, or dynamopenia, is a core characteristic of aging, which is only partially explained by sarcopenia, the decline of lean body mass that occur with aging in virtually all men and women. Understanding the cause of dynamopenia is very important

because the decline in strength is a key cause of frailty and disability in the elderly. We aimed to determine any expression associations with muscle strength in InCHIANTI, hypothesizing that genes whose expression correlates with muscle strength would highlight the specific cellular mechanisms that are involved. We found only 1 transcript to be robustly associated with hand-grip strength in the cohort [46]; CCAAT/enhancer-binding-protein beta (CEBPB) is a transcription factor involved in macrophage function, which includes the mediation of repair to damaged muscle. We found that CEBPB expression increased proportionally with muscle strength across a wide age range (30years to 104years). Other genes involved in macrophage function were also moderately associated with strength, including TGFB3 (anti-inflammatory signaling).

Having studied the literature we noted that a mouse model lacking the ability to increase CEBPB expression had sarcopenia-like muscle loss and fibrosis [47]. This new link between CEBPB expression in blood leukocytes and muscle strength suggests that the relationship between tissue damage and macrophage-mediated tissue-repair is critical, and should be a target of further studies into age-related muscle decline.

6.1.2 Cognitive function

Inflammation is proposed to play a role in age-related cognitive decline, so we hypothesized that specific changes in gene expression in peripheral blood leukocytes may correlate with either recent decline in cognitive function, or independently with current cognitive status. Each participant in the study was assessed using the mini mental state examination (MMSE) of cognitive function, and we anticipated finding inflammatory-related pathways and mechanisms to be associated with the MMSE score. We only found a single transcript for chemokine receptor 2 (CCR2) to be robustly associated with MMSE score [48]. CCR2 was independently associated with change in cognitive score over 9 years in an analysis of inflammation-specific transcripts. Expression of CCR2 is also associated with the ApoE risk allele, one of the major risk factors for Alzheimers, indicating that CCR2 could have a more direct role in cognitive function in later life.

CCR2 has been previously associated with -amyloid deposition (a hallmark trait of Alzheimers) in mouse models [49]. It has been hypothesized that normal, CCR2 producing, monocytes (an immune cell subtype which includes macrophages) may contribute to the clearance of abnormal protein deposits. As the participants in the study only had their cognitive function assessed, we can only suppose that the link is also true in humans until further evidence has been gathered, but it is useful to note that several mouse models seem to have particular relevance to humans. Of singular interest is a study performed in mice whereby functional memory was restored by transplantation of monocytes with functional CCR2 [50]. The implications here for future studies are immense.

6.2 The Leiden Longevity Study

It is well established that longevity is a moderately heritable trait, with the genetic component estimated to be 25% [51]. Determining the specific mechanisms and biological pathways by which this advantage is conferred is of upmost importance in understanding the aging process. The Leiden Longevity Study group has previously reported on the genetic enrichment for exceptional survival in long-lived families in their Dutch cohort [52], and in the current study they used nonagenarian spouse pairs, their middle-aged offspring and the partners of the offspring to determine gene transcripts associated with aging, and to determine which are associated with successful aging [53].

The study was split into 2 distinct sections; exploration and replication. The exploratory stage found that over 1,800 genes (almost 3,000 separate transcripts) were differentially expressed between long-lived (old) individuals and middle-aged controls. These included LRRN3 (3-fold decreased expression in nonagenarians) and interferon alpha-inducible protein 27 (IFI27; 2-fold increased expression in nonagenarians), which are involved in immune function and apoptosis (programmed cell death) respectively. Of the age-associated genes, 244 (360 transcripts) were differentially expressed at middle age between the children of long-lived parents and the controls (their partners). This list does not include LRRN3 or IFI27, the genes with the greatest decrease and increase in expression with age, respectively. The zinc finger protein 331 gene (ZNF331) was most differentially expressed between the offspring and controls (1.6-fold decrease every ten years). Upon exploration of the biological processes (as defined by Gene Ontology, GO) most prominent within the longevity-associated genes the group found that the Rho protein signal transduction pathway (GO:0007266) was significantly over-represented. This is a broad system involved in cell growth, cytoskeletal remodeling and lipid metabolism, and it is regulated by the mammalian Target Of Rapamycin (mTOR) pathway that has previously shown to regulate lifespan [54].

Replication of these findings was performed on separate, unrelated participants also from the Leiden Longevity Study. RT-qPCR was used to measure the expression of 25 specific genes in 79 nonagenarians, 332 of their offspring and 312 population controls (spouses of the offspring). Three genes were chosen based on their age or longevity associations (LRRN3, IFI27, and

ZNF331), the other 22 for prior literature relating to cellular aging. Over 80% of the genes replicated the associations observed in the microarray data, including LRRN3 and IFI27. After accounting for multiple testing, only 2 of the tested genes showed offspring vs controls associations; anti-silencing function 1 homolog A (ASF1A) and interleukin 7 receptor (IL7R) were both found to be significantly decreased in the long-lived families in comparison to the controls. ZNF331 was not found to be differentially expressed between the offspring of nonagenarians and their spouses. Functionally ASF1A is involved in histone modulation and chromatin remodelling one of several types of epigenetic gene expression regulation and IL7R is a receptor required for the development and maintenance of the immune system.

6.3 The San Antonio Family Heart Study

Originally initiated as a study of the genetics of atherosclerosis in Mexican Americans, the San Antonio Family Heart Study (SAFHS) study of gene expression in lymphocytes from 1240 individuals (aged 15-94) with quantification of 18,519 genes (using microarray technology) [55]. Unlike the other studies described in this section, the SAFHS assessed separated lymphocyte gene expression rather than whole blood, which captures the mean expression across the leukocyte cell sub-types (of which lymphocytes are one).

In 2008 Hong et al [55] published an analysis gene expression in human brain and lymphocyte senescence. We discuss the results specific to the brain in the next section, but the results gained from lymphocytes included associations with age. LRRN3 featured prominently as the gene with the strongest association with age. The analysis also reported associations with broader biological functions including mitochondrial function (negatively correlated with age) and signal transduction/cell communication (positively correlated with age).

In an interesting analysis into the variations in genomics architecture between the age-associated genes the group found that intron/exon sequence length ratios were larger amongst the genes that were negatively associated with age, which replicated in the brain tissue analysis (next section). They also looked into pre-mRNA length, coding sequence length, number of exons, total intron length and 5 and 3 untranslated region (UTR) lengths. This may be of relevance to the InCHIANTI study finding that the most deregulated biological pathways with age were RNA processing pathways [44].

6.4 In other tissues

Whole transcriptome quantification has been employed in only a few analyses of other humans tissues in the context of aging, which include analyses of skeletal [56] muscle, the brain [57] and the kidney [58].

6.4.1 Brain

Hong et al (in the same study described in section 6.3) also analyzed transcriptome changes with age in brain tissue [55], in a sample of 191 port-mortem brains with age-at-death ranging from 65 to 100 years. Fifty four of the 14,078 transcripts assessed were associated with age (after accounting for multiple testing). Functionally these represent a down-regulation of genes involved in synaptic function, and an increase in expression of genes involved in transcription and oxidative phosphorylation. These changes were not observed in the lymphocytes sample, so the authors conclude that tissue specific changes to mitochondrial function with age are likely. That genes involved in translational processes were not found to be differentially expressed implies that the increase in transcription is not in protein-coding genes, but in non-coding RNAs (ncRNAs) these tend to be regulatory in function, although some are structural (e.g. ribosomal RNAs or involved in protein biosynthesis e.g. transfer RNAs).

In a recent publication, Kang et al analyze differential transcript expression in different regions of the brain, at different ages [57]. The focus is not on aging, but rather development; 42% of the sample was pre-natal (ranging from 5.7 weeks post conception to 38 weeks), 23% were children, 30% were young/middle-aged adults, and the remaining samples were older (>60 years). Using clustering techniques the authors show that expression of genes across the age-range (in all regions) varied most between the youngest brain (5.7 weeks post-conception) and the young (<12 years). Comparatively, the young to old adults had very similar expression of genes however the different groups still distinctly cluster together, implying that in a study designed with enough power to detect early/mid/late-life changes in gene expression would find them.

6.4.2 Muscle

Sarcopenia is the age-related loss of skeletal muscle mass. Several studies into transcriptome changes in muscle fibers with age have appeared over the years, one of the most recent studied the effects of exercise on skeletal muscle gene expression. This study found that the response to exercise (at the gene expression level) was more pronounced in younger participants (24-25 years, vs 78-84 years), and also that basal level gene

expression differed between the groups. 661 genes associated with resistance exercise training regardless of the age of the participants, yet changes were seen in >1000 genes in the young participants, only half of which were also seen in the older participants. The cellular pathways most affected were (as expected) growth-related. The expression of 49 genes were found to be robustly associated with age in skeletal muscle (in both studies of their 2-study design). These included CDKN2B i.e. p16 (see above), CDKN1A and CCNG2 all tumor suppressing genes, and all up-regulated in the older participants. This corresponds well with age-related loss of muscle mass, because muscle fibers will be less able to enter the cell-cycle, and thus less able to replace any damaged/removed cells.

6.4.3 Kidney

Rodwell et al found almost 1000 genes in kidney tissue showed altered expression with advancing age [58]. They note an increase in expression of some immune-related genes, suggesting a possible link with age-related inflammation. Using an additive kidney score (appearance of the glomeruli, the tubules and the arterioles) to assess kidney function, the authors determined that many of the genes (n=447) also associated with the relative health/performance of the kidney in the older participants, offering a possible biological aging biomarker.

7 Future Prospects

As set out above, gene expression studies have become central to much laboratory work and have already produced clinically useful signatures, especially for cancers. In human aging, a new generation of array technology is supporting genome wide studies, with very promising results. We should note however that there are important limitations to this technology, including:

7.0.4 Tissue / cell specific expression

Gene expression varies across tissues, cell types and even across regions of the same organ. Access to most tissues from population representative human volunteers will always be severely limited, and care will be needed to avoid implying that changes seen in accessible tissues are necessarily mirrored in the rest of the body.

7.0.5 Attributing causation

Gene expression is influenced by potential confounders such as smoking status, diet, and geographical location [59-61]. Associations found with phenotypes can be the result of confounding and are not necessarily causal. Gene expression patterns can reflect responses to pathology rather than drivers.

7.0.6 Interpretation of signatures is somewhat subjective

Gene transcripts identified as associated with phenotypes of interest are frequently of unknown function or have large numbers of functions within multiple pathways. Investigators tend to focus on the subset of striking findings that can be interpreted as conforming to a known pathway or animal model. Caution is needed to regard such interpretations as unproven hypotheses, to ensure that these are robustly tested before being accepted. Thus, much work is needed to confirm that similarities with specific mouse models do in fact accurately reflect what is happening in the human.

Human volunteer or population-based studies are useful, but have limited utility for defining mechanisms. This is because such studies are often limited to observational (rather than experimental) statistical associations, and human tissue samples often contain a heterogeneous mix of many different cell types, each of which may behave differently. In-vitro studies provide a widely studied and controlled environment to research the mechanisms of aging (although generalisability to the in-vivo situation is not guaranteed). Many fundamental discoveries in the biology of human aging, including telomere shortening as a biomarker of aging, have been identified by in-vitro models [62]. In-vitro senescence models provide an important and accessible tool to investigate the pathways of aging and further research into the mechanism of senescence will advance our understanding of the aging process. Given the challenges ahead, it is likely that breakthroughs will come from integrating in-vitro, animal model and human study approaches.

Despite the challenges we are likely to see many exciting gene expression findings in aging, given the large study resources appearing and the development of consortia to replicate and pool data (e.g. the CHARGE consortium). We are also likely to see studies exploring the regulatory mechanisms controlling gene expression, including methylation, histone modification and miRNA/lncRNA control, as these may provide routes to interventions. Given the progress thus far, it seems likely that the goal of identifying robust biomarkers of aging processes might now be within reach.

8 References

1. Pilling, L.C., et al., Genomics and successful aging: grounds for renewed optimism? J Gerontol A Biol Sci Med Sci, 2012. 67(5): p. 511-9.

2. Harris, T.J. and F. McCormick, The molecular pathology of cancer. Nat Rev Clin Oncol, 2010. 7(5): p. 251-65.

3. NIH. Deciphering the Genetic Code: M. Nirenberg. 2010 12th May 2012]; Available from: http://history.nih.gov/exhibits/nirenberg.

4. Kenyon, C., The plasticity of aging: insights from long-lived mutants. Cell, 2005. 120(4): p. 449-60.

5. Dang, W., et al., Histone H4 lysine 16 acetylation regulates cellular lifespan. Nature, 2009. 459(7248): p. 802-7.

6. Golden, T.R., A. Hubbard, and S. Melov, Microarray analysis of variation in individual aging C. elegans: approaches and challenges. Experimental gerontology, 2006. 41(10): p. 1040-5.

7. Murphy, C.T., et al., Genes that act downstream of DAF-16 to influence the lifespan of Caenorhabditis elegans. Nature, 2003. 424(6946): p. 277-83.

8. Hansen, M., et al., A role for autophagy in the extension of lifespan by dietary restriction in C. elegans. PLoS Genet, 2008. 4(2): p. e24.

9. Feng, J., F. Bussiere, and S. Hekimi, Mitochondrial electron transport is a key determinant of life span in Caenorhabditis elegans. Developmental cell, 2001. 1(5): p. 633-44.

10. Budovskaya, Y.V., et al., An elt-3/elt-5/elt-6 GATA transcription circuit guides aging in C. elegans. Cell, 2008. 134(2): p. 291-303.

11. Lai, C.Q., et al., Candidate genes affecting Drosophila life span identified by integrating microarray gene expression analysis and QTL mapping. Mechanisms of ageing and development, 2007. 128(3): p. 237-49.

12. Pletcher, S.D., et al., Genome-wide transcript profiles in aging and calorically restricted Drosophila melanogaster. Current biology : CB, 2002. 12(9): p. 712-23.

13. Park, S.K., Genomic approaches for the understanding of aging in model organisms. BMB reports, 2011. 44(5): p. 291-7.

14. Girardot, F., et al., Specific age-related signatures in Drosophila body parts transcriptome. BMC genomics, 2006. 7: p. 69.

15. Weindruch, R., et al., Gene expression profiling of aging using DNA microarrays. Mechanisms of ageing and development, 2002. 123(2-3): p. 177-93.

16. Zahn, J.M., et al., AGEMAP: a gene expression database for aging in mice. PLoS Genet, 2007. 3(11): p. e201.

17. Englander, E.W., Gene expression changes reveal patterns of aging in the rat digestive tract. Ageing research reviews, 2005. 4(4): p. 564-78.

18. Park, S.K., et al., Gene expression profiling of aging in multiple mouse strains: identification of aging biomarkers and impact of dietary antioxidants. Aging cell, 2009. 8(4): p. 484-95.

19. McCarroll, S.A., et al., Comparing genomic expression patterns across species identifies shared transcriptional profile in aging. Nature genetics, 2004. 36(2): p. 197-204.

20. Kim, S.K., Common aging pathways in worms, flies, mice and humans. The Journal of experimental biology, 2007. 210(Pt 9): p. 1607-12.

21. Fraser, H.B., et al., Aging and gene expression in the primate brain. PLoS biology, 2005. 3(9): p. e274.

22. Collins, C.J. and J.M. Sedivy, Involvement of the INK4a/Arf gene locus in senescence. Aging Cell, 2003. 2(3): p. 145-50.

23. Castellano, M., et al., CDKN2A/p16 is inactivated in most melanoma cell lines. Cancer Res, 1997. 57(21): p. 4868-75.

24. Sharpless, N.E., Ink4a/Arf links senescence and aging. Experimental Gerontology, 2004. 39(11-12 SPEC. ISS.): p. 1751-1759.

25. Liu, Y., et al., Expression of p16(INK4a) in peripheral blood T-cells is a biomarker of human aging. Aging Cell, 2009. 8(4): p. 439-48.

26. Agherbi, H., et al., Polycomb mediated epigenetic silencing and replication timing at the INK4a/ARF locus during senescence. PloS one, 2009. 4(5): p. e5622.

27. Nelson, G., et al., A senescent cell bystander effect: senescence-induced senescence. Aging cell, 2012: p. 1-5.

28. Baker, D.J., et al., Clearance of p16Ink4a-positive senescent cells delays ageing-associated disorders. Nature, 2011: p. 1-6.

29. Bonferroni, C.E., Teoria statistica delle classi e calcolo delle probabilita‘. Pubblicazioni del R. Istituto superiore di scienze economiche e commerciali di Firenze. Vol. 8. 1936: Firenze : Seeber.

30. Benjamini, Y. and Y. Hochberg, Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society. Series B (Methodological), 1995: p. 289-300.

31. Shaffer, J.P., Multiple Hypothesis-Testing. Annual Review of Psychology, 1995. 46: p. 561-584.

32. Welle, S., et al., Skeletal muscle gene expression profiles in 20-29 year old and 65-71 year old women. Exp Gerontol, 2004. 39(3): p. 369-77.

33. Craig, T., C. Smelick, and J.P. de Magalhaes. The Digital Aging Atlas: http://ageing-map.org. 2010 24th April 2012].

34. Reimer, M. An Opinionated Guide to Microarray Data Analysis: Normalizing Expression Arrays. [cited 2012 28th March]; Available from: http://www.people.vcu.edu/ mreimers/OGMDA/normalize_expression.html.

35. Alizadeh, A.A., et al., Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. Nature, 2000. 403(6769): p. 503-11.

36. Safran, M., et al., GeneCards Version 3: the human gene integrator - www.genecards.org. Database-the Journal of Biological Databases and Curation, 2010.

37. Ashburner, M., et al., Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nature genetics, 2000. 25: p. 25-9.

38. Kanehisa, L. KEGG: Kyoto Encyclopedia of Genes and Genomes. 2012 [cited 2012 26th March]; Available from: http://www.genome.jp/kegg/.

39. Subramanian, A., et al., Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. Proceedings of the National Academy of Sciences of the United States of America, 2005. 102(43): p. 15545-15550.

40. Glaab, E. PathVar - Microarray analysis of pathway expression variance 2012 [cited 2012 26th March]; Available from: http://pathvar.embl.de.

41. Reuters, T. Pathway Analysis, MetaCore. 2012 [cited 2012 26th March]; Available from: http://www.genego.com/metacore.php.

42. Ingenuity. Ingenuity Systems Pathway Analysis. 2012 [cited 2012 26th March]; Available from: http://www.ingenuity.com.

43. Ferrucci, L., et al., Subsystems contributing to the decline in ability to walk: bridging the gap between epidemiology and geriatric practice in the InCHIANTI study. Journal of the American Geriatrics Society, 2000. 48: p. 1618-25.

44. Harries, L.W., et al., Human aging is characterized by focused changes in gene expression and deregulation of alternative splicing. Aging cell, 2011: p. 1-13.

45. Chung, H.Y., et al., The molecular inflammatory process in aging. Antioxidants and redox signaling, 2006. 8: p. 572-81.

46. Harries, L.W., et al., CCAAT-Enhancer-Binding Protein-Beta (CEBPB) Expression In-Vivo is Associated with Muscle Strength. Aging cell, 2012. 11(2): p. 262-268.

47. Ruffell, D., et al., A CREB-C/EBPbeta cascade induces M2 macrophage-specific gene expression and promotes muscle injury repair. Proc Natl Acad Sci U S A, 2009. 106(41): p. 17475-80.

48. Harries, L.W., et al., Leukocyte CCR2 Expression is Associated with Mini-Mental-State-Examination (MMSE) Score in Older Adults. Revujenation Research, 2012. Accepted, In Press.

49. Philipson, O., et al., Animal models of amyloid-beta-related pathologies in Alzheimer's disease. FEBS Journal, 2010. 277: p. 1389-1409.

50. Naert, G. and S. Rivest, Hematopoietic CC-chemokine receptor 2-(CCR2) competent cells are protective for the cognitive impairments and amyloid pathology in a transgenic mouse model of Alzheimer's disease. Molecular medicine (Cambridge, Mass.), 2011.

51. Herskind, A.M., et al., The heritability of human longevity: a population-based study of 2872 Danish twin pairs born 1870-1900. HumGenet, 1996. 97(3): p. 319.

52. Schoenmaker, M., et al., Evidence of genetic enrichment for exceptional survival using a family approach: the Leiden Longevity Study. Eur J Hum Genet, 2006. 14(1): p. 79-84.

53. Passtoors, W.M., et al., Transcriptional profiling of human familial longevity indicates a role for ASF1A and IL7R. PLoS ONE, 2012. 7(1): p. e27759.

54. Sharp, Z.D. and R. Strong, The role of mTOR signaling in controlling mammalian life span: what

a fungicide teaches us about longevity. The journals of gerontology. Series A, Biological sciences and medical sciences, 2010. 65: p. 580-9.

55. Hong, M.-G., et al., Transcriptome-wide assessment of human brain and lymphocyte senescence. PloS one, 2008. 3: p. e3024.

56. Raue, U., et al., Transcriptome Signature of Resistance Exercise Adaptations: Mixed Muscle and Fiber Type Specific Profiles in Young and Old Adults. J Appl Physiol, 2012.

57. Kang, H.J., et al., Spatio-temporal transcriptome of the human brain. Nature, 2011. 478(7370): p. 483-9.

58. Rodwell, G.E., et al., A transcriptional profile of aging in the human kidney. PLoS Biol, 2004. 2(12): p. e427.

59. Lampe, J.W., et al., Signatures of environmental exposures using peripheral leukocyte gene expression: tobacco smoke. Cancer Epidemiol Biomarkers Prev, 2004. 13(3): p. 445-53.

60. Cameron-Smith, D., et al., A short-term, high-fat diet up-regulates lipid metabolism and gene expression in human skeletal muscle. Am J Clin Nutr, 2003. 77(2): p. 313-8.

61. Storey, J.D., et al., Gene-expression variation within and among human populations. American journal of human genetics, 2007. 80: p. 502-9.

62. Harley, C.B., A.B. Futcher, and C.W. Greider, Telomeres shorten during ageing of human fibroblasts. Nature, 1990. 345(6274): p. 458-460.