META-ANALYSIS OF GENE EXPRESSION REVERSALS IN AGEING BRAIN

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

BY

HANDAN MELIKE DÖNERTAŞ

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

JUNE 2016

Aproval of the thesis:

META-ANALYSIS OF GENE EXPRESSION REVERSALS IN AGEING BRAIN

submitted by **HANDAN MELIKE DÖNERTAŞ** 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:

Prof. Dr. Mesut Muyan
Biology Dept., METU _____

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

Assoc. Prof. Dr. Tolga Can
Comp. Eng. Dept., METU _____

Assist. Prof. Dr. Emre Akbaş
Comp. Eng. Dept., METU _____

Assist. Prof. Dr. Can Alkan
Comp. Eng. Dept., İhsan Doğramacı Bilkent Uni. _____

Date: 06.06.2016

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	: HANDAN MELIKE DONERTAS
Signature	:

ABSTRACT

META-ANALYSIS OF GENE EXPRESSION REVERSALS IN AGEING BRAIN

Dönertaş, Handan Melike
M.S., Department of Biology
Supervisor : Assoc. Prof. Dr. Mehmet Somel

June 2016, 108 pages

Brain ageing is characterised by disruptive changes in cognitive abilities, histology, and anatomy. The underlying molecular nature of brain ageing, on the other hand, is little understood, partly due to the stochastic and heterogeneous nature of ageing process. In this study, using published microarray studies spanning 22 brain regions with 1,015 samples, gene expression changes in ageing are analysed in comparison to those in postnatal development. A previous observation that mRNA abundance of a large number of genes in the ageing prefrontal cortex reverses toward pre-adolescent levels, is shown to be a widespread phenomenon across different brain regions. Furthermore, functional analysis reveals that gene expression reversals are consistently associated with decline in neuronal / synaptic gene expression across all studied brain regions, and thus may be linked to ageing-related phenotypes such as decline in cognitive functions. Regulatory analysis show that the genes increasing in expression in development and decreasing in expression in ageing are associated with several trans-regulators, whereas there is no consistent association with any potential trans-regulator for the genes decreasing in expression in development and increasing in expression in ageing.

Overall, the results show that meta-analysis is crucial for ageing studies due to the stochastic nature of ageing and that studying gene expression change in ageing in the context of changes in development is a promising approach to discover the molecular mechanisms of ageing.

Keywords: ageing, gene expression, microarray, brain

ÖZ

GEN ANLATIM GERI DÖNÜŞLERININ YAŞLANAN BEYINDE META-ANALIZI

Dönertaş, Handan Melike
Yüksek Lisans, Biyoloji Bölümü
Tez Yöneticisi : Doç. Dr. Mehmet Somel

Haziran 2016, 108 sayfa

Beyin yaşlanması, bilişsel kabiliyetler, histoloji ve anatomide yıkıcı değişimlerle karakterize edilir. Buna karşın, beyin yaşlanması altında yatan moleküler yapı, kısmen yaşlanma sürecinin stokastik ve heterojen yapısından dolayı, az anlaşılmıştır. Bu çalışmada, 22 beyin bölgesini kapsayan 1,015 örnek içeren yayınlanmış mikrodizin çalışmaları kullanılarak, yaşlanmada yaşanan gen anlatım değişimleri, doğum sonrası gelişime kıyasla analiz edilmiştir. Yaşlanan prefrontal kortekste çok sayıda gene ait mRNA miktarının ergenlik öncesi seviyeye döndüğüne dair gözlem, bir çok beyin bölgesinde yaygın olarak gözlemlenmiştir. Buna ek olarak, fonksiyonel analiz, gen anlatım geri dönüşlerinin tutarlı olarak çalışılan tüm beyin bölgelerinde nöronal / sinaptik gen anlatımında düşüşle ilişkili olduğunu, dolayısıyla da bilişsel kabiliyetlerde düşüş gibi yaşlanma ile ilişkili fenotiplerle ilişkili olabileceğini göstermiştir. Regülasyon analizi, gelişimde anlatımı artan, yaşlanmada ise anlatımı düşen genlerin bir takım trans-regülatörlerle ilişkili olduğunu göstermiştir; buna karşın gelişimde azalıp, yaşlanmada artan anlatıma sahip genlerle ilişkili herhangi bir potansiyel trans-regülator bulunamamıştır.

Genel olarak, sonuçlar yaşlanmanın stokastik doğasından dolayı meta-analizin yaşlanma çalışmaları için kritik olduğunu ve yaşlanmada yaşanan gen anlatım değişimlerinin gelişimde yaşanan değişimler bağlamında incelenmesinin yaşlanmanın moleküler mekanizmalarının keşfi için umut verici olduğunu göstermektedir.

Anahtar Kelimeler: yaşlanma, gen anlatımı, mikrodizin, beyin

To all graduate researchers who work day and night
but are only recognised as "students" with no stipend and no insurance

ACKNOWLEDGEMENTS

This work would have not been possible without help, support and encouragement of many people. First and foremost, I would like to thank my supervisor Mehmet Somel for giving me chance to do research in his group. I am grateful for all the time he spent in the meetings, reading my work and for sharing the knowledge and motivation whenever I needed. The lessons I learnt from his insight, guidance and patience is not only important for this work but will help me for all my research life.

I also want to thank Mehmet for giving me opportunity to be a member of a great team. I want to thank all my lab mates for their endless support. They were all there for me in the times of crisis. We are a lot more people than I can name in here; however, I must acknowledge the ageing group Hamit İzgi, Gözde Turan and Poorya Parvizi as well as Gülşah Kılınç. They showed continuous willingness to discuss all the academic questions I had. Of course, my special thanks to Hamit for the fruitful discussions and keeping me focused whenever I needed. Moreover, without the critical reading of Mehmet, Gülşah, Hamit, Poorya and Gözde, this thesis would be a lot less interesting and have a lot more spelling and grammar mistakes. Thesis writing does not only concern the contents. One of the real challenges is formatting the text. I appreciate the LaTeXtemplate that Ozan Kıratlı has kindly shared with me.

I am indebted to many people for non-academic reasons as well. I want to thank Hamit, Betül Taşkoparan, Buse İşbilir, and Mustafa Gırak. Without their continuous support and friendship, I would not be able to continue in this path.

I also thank Scientific and Technical Research Council of Turkey (TUBITAK) for financially supporting me since my undergraduate degree.

Lastly, my mother and grandmother deserves a special recognition for their constant love and support. I am grateful that they listen to me when I start to talk about my work and that they even share my enthusiasm.

CONTENTS

ABSTRACT v

ÖZ vii

ACKNOWLEDGEMENTS x

TABLE OF CONTENTS xii

LIST OF TABLES xv

LIST OF FIGURES xvii

LIST OF ABBREVIATIONS xix

CHAPTERS

1 INTRODUCTION 1

1.1 Ageing 1

1.1.1 Ageing Studies 2

1.1.2 Theories of Ageing 4

1.2 Structural Changes in Human Brain with Age 8

1.2.1 Human Brain Organisation 9

1.2.2 Embryonic Period 11

1.2.3 Fetal Period 12

1.2.4 Postnatal Period. 13

1.2.5 Ageing 14

1.3 Transcriptome Change in Brain with Age 14

1.4 Gene Expression Reversal 20

1.5 Research Objectives 22

2 MATERIAL AND METHOD 25

2.1 Datasets 25

2.2 Gene Expression Data Preprocessing 28
2.2.1 RMA correction 29
2.2.2 Summarization of probe-sets to genes 29
2.2.3 Log2 Transformation 31
2.2.4 Quantile Normalization. 31
2.3 Sample Quality Check. 33
2.4 Cell Type-Specific Expression Analysis 35
2.5 Age Test 36
2.5.1 Multiple Testing Correction 37
2.6 Comparison of Age-Related Gene Expression Change across Datasets 37
2.6.1 Correlation across Datasets 37
2.6.2 Hierarchical Clustering of Datasets 38
2.7 Gene Reversal Index 38
2.8 Analysis of Gene Function and Regulation. 38
2.8.1 GO Biological Process Functional Analysis. . 40
2.8.2 miRNA Regulation Analysis. 41
2.8.3 Transcription Factor Regulation Analysis. . . 41
2.8.4 miRNA and TF Network Construction. . . . 41
2.9 Permutation Test 42
3 RESULTS 43
3.1 Cell Type-Specific Expression 43
3.2 Age-Related Gene Expression Change across Datasets . . 45
3.3 Strength of Age-Related Gene Expression Change in Development and in Ageing 48
3.4 Direction of Age-Related Gene Expression Change . . . 49
3.5 Gene Expression Reversals 51
3.6 Functional Analysis of Gene Expression Reversals . . . 53
3.6.1 GO Biological Process Enrichment Analysis . 53
3.7 Regulation of Gene Expression Reversals 60
3.7.1 miRNA Enrichment Analysis 61
3.7.2 Transcription Factor Enrichment Analysis . . 63
4 DISCUSSION 67
4.1 Limitations and Possible Improvements 74

5 CONCLUSION 77
REFERENCES 79

APPENDIX 96
A LIST OF GO BIOLOGICAL PROCESS CATEGORIES WITH CONSISTENT OVER-REPRESENTATION OF DOWN-UP GENES . . 97
B LIST OF GO BIOLOGICAL PROCESS CATEGORIES WITH CONSISTENT OVER-REPRESENTATION OF MONOTONICALLY DECREASING GENES 99
C LIST OF GO BIOLOGICAL PROCESS CATEGORIES WITH CONSISTENT OVER-REPRESENTATION OF UP-DOWN GENES . . 101
D LIST OF GO BIOLOGICAL PROCESS CATEGORIES WITH CONSISTENT OVER-REPRESENTATION OF MONOTONICALLY INCREASING GENES 105
E LIST OF GO BIOLOGICAL PROCESS CATEGORIES WITH CONSISTENT OVER-REPRESENTATION OF UP-DOWN AND MONOTONICALLY DECREASING GENES 107

LIST OF TABLES

TABLES

Table 2.1 Dataset summary. The “number of samples” column represents the total number of samples in the study, whereas the numbers in parentheses show the number of individuals. 27

Table 2.2 Number of probe-sets and final number of genes for each platform after summarization as described in the text. ENSG represents Ensembl Gene IDs. 31

Table 2.3 Contingency Table for Up-Down pattern 39

Table 2.4 Contingency Table for Down-Up pattern 39

Table A.1 List of GO BP Categories having $OR > 1$ for Down-Up vs. Monotonic Decrease Comparison in all datasets. "↑" in the column names means increase in gene expression. 97

Table B.1 List of GO BP Categories having $OR < 1$ for Down-Up vs. Monotonic Decrease Comparison in all datasets. "↑" in the column names means increase in gene expression. 99

Table C.1 List of GO BP Categories having $OR > 1$ for Up-Down vs. Monotonic Increase Comparison in all datasets. "↑" in the column names means increase in gene expression. 101

Table D.1 List of GO BP Categories having $OR < 1$ for Up-Down vs. Monotonic Increase Comparison in all datasets. "↑" in the column names means increase in gene expression. 105

Table E.1 List of GO BP Categories having $OR > 1$ for Up-Down vs. Monotonic Increase Comparison and $OR < 1$ for Down-Up vs. Monotonic Decrease Comparison in all datasets. "↑" in the column names means increase in gene expression. 107

LIST OF FIGURES

FIGURES

Figure 1.1 Comparison of the leading causes of death in the world, 2000 and 2012 (World Health Organization, 2012). *COPD*: Chronic obstructive pulmonary disease 2

Figure 1.2 Major mRNA change patterns with age (Somel et al., 2010) . . 21

Figure 2.1 Distribution of ages across datasets 26

Figure 2.2 Summary of probe-set to gene summarization. The upper panel corresponds to a probe-set associated with multiple genes. This probe-set is removed from data. The lower panel shows a gene associated with multiple probe-sets. Expression value of this gene is represented with the mean expression value of the probe-sets. 30

Figure 2.3 Example boxplot for the Somel2010 dataset before and after quantile normalization. Each boxplot shows the distribution of a sample. 32

Figure 2.4 Example histogram for the Somel2010 dataset after each step. . 33

Figure 2.5 Example PCA for the Kang2011 dataset, Hippocampus brain region. Each number on plot shows the age of an individual. The x-axis corresponds to the first principle component and the y-axis corresponds to the second principle component. Percentages in parentheses represent the variance explained by each component. 34

Figure 3.1 Cell Type-Specific Expression Profile Change with Age 44

Figure 3.2 Correlation plot for age-related gene expression changes between all brain regions in development and ageing. The size and the colour of the squares change with the magnitude of the correlation coefficient between different regions as the large blue squares show strong positive correlation, and the large red squares show strong negative correlation. Row and column labels show the type of samples - green: development and red: ageing. Brain regions are ordered by hierarchical clustering of correlation coefficients between datasets. 46
Figure 3.3 Hierarchical clustering of brain regions in development and ageing based on age-related gene expression change 47
Figure 3.4 Magnitude of age-related gene expression change in development vs. ageing. Each boxplot shows the distribution of $|\rho_{dev}| - |\rho_{age}|$ for genes in a given dataset. Red line shows 0 which means no difference. Datasets are ordered according to their median $|\rho_{dev}| - |\rho_{age}|$ differences. 48
Figure 3.5 Direction of age-related gene expression change across all brain regions 50
Figure 3.6 Proportion of reversal genes 51
Figure 3.7 GO Biological Process enrichment result for down-up genes, summarized by REVIGO 54
Figure 3.8 GO Biological Process enrichment result for monotonically decreasing genes, summarized by REVIGO 56
Figure 3.9 GO Biological Process enrichment result for up-down genes, summarized by REVIGO 59
Figure 3.10 GO Biological Process enrichment result for monotonically increasing genes, summarized by REVIGO 60
Figure 3.11 Network representation of up-down gene regulation by miRNAs 61
Figure 3.12 Network representation of up-down gene regulation by TFs . . 64

LIST OF ABBREVIATIONS

CNS	Central Nerveous System
RMA	Robust Multi-Array Analysis
STC	Superior Temporal Cortex
EC	Entorhinal Cortex
HIP	Hippocampus
PCG	Postcentral Gyrus
SFG	Superior Frontal Gyrus
FC	Frontal Cortex
APFC	Anterior Prefrontal Cortex
PFC	Prefrontal Cortex
A1C	Primary Auditory Cortex
AMY	Amygdala
CBC	Cerebellar Cortex
DFC	Dorsolateral Prefrontal Cortex
IPC	Posterior Inferior Parietal Cortex
ITC	Inferior Temporal Cortex
M1C	Primary Motor Cortex
MD	Mediodorsal Nucleus of the Thalamus
MFC	Medial Prefrontal Cortex
OFC	Orbital Prefrontal Cortex
S1C	Primary Somatosensory Cortex
STC	Superior Temporal Cortex
STR	Striatum
V1C	Primary Visual Cortex
VFC	Ventrolateral Prefrontal Cortex

CHAPTER 1

INTRODUCTION

1.1 Ageing

Ageing is associated with time-dependent disruptive changes in a variety of complex biological processes. These changes are characterised by functional decline and loss of physiological integrity in the biological systems leading to reduced homoeostasis, vulnerability to many pathologies, and ultimately death.

In today's world, mostly as a result of decrease in mortality rate at young ages and infectious diseases the life expectancy of people in most of the countries is around 80 years of age. For the last 60 years, there has been a demographic shift towards higher ages in all countries (Department of Economic and Social Affairs Population Division, 2000). For example, in Turkey the percentage of people who are 60 or more was 10-19% in 2015 and the projections suggest this value will be 25-29% by the end of 2050 (World Health Organization, 2015). One may think that these numbers indicate there is no need to study ageing, as the populations are "successfully ageing" naturally. Unfortunately, this may not be the case. The numbers mostly reflect the decrease in mortality at younger ages as a result of socio-economic developments in the recent years, rather than indicating old people living longer. We may observe an aged population, but the critical question is "Do we live as healthy as we do in young ages or do we experience a poor health period during the added years?". And even if we have old people living longer, as suggested especially for high-income countries (Crimmins & Beltrán-Sánchez, 2011), the quality of life during the added years is still an important aspect to consider. Age-

ing is seen as the major risk factor for many pathologies including cardiovascular diseases, diabetes mellitus, hypertension, cancer, and neurodegenerative disorders (Niccoli & Partridge, 2012). Change in the leading causes of death over the last years indicate that ageing-related disorders are becoming more and more important as the rate of death due to infectious diseases decreases and impact of age-related diseases increases (**Figure 1.1**).

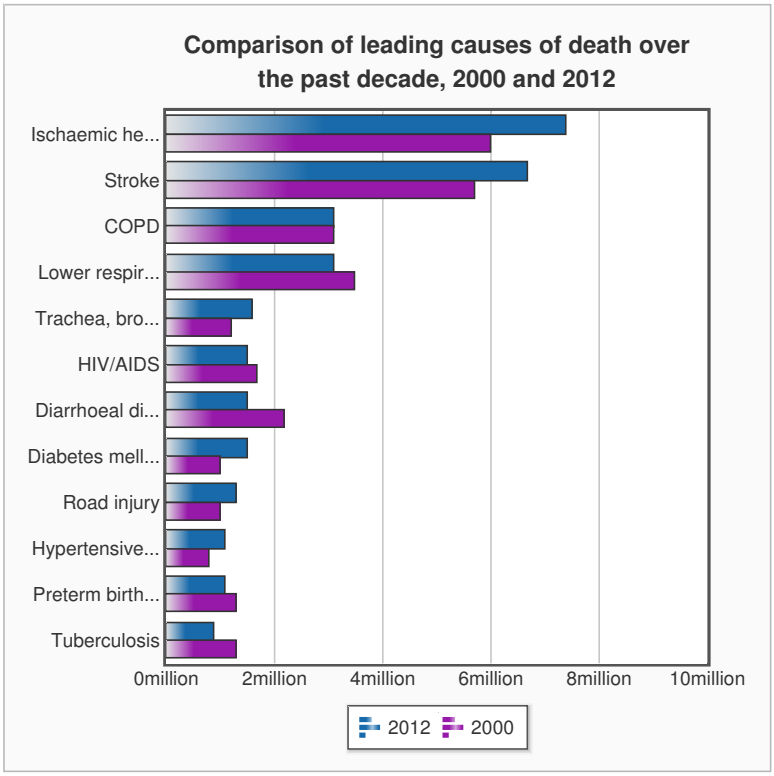

Figure 1.1: Comparison of the leading causes of death in the world, 2000 and 2012 (World Health Organization, 2012). *COPD*: Chronic obstructive pulmonary disease

1.1.1 Ageing Studies

As a result of the before-mentioned increase in the median population age and the corresponding increase in the prevalence of age-related diseases, we encounter more and more medical, economic and social challenges. For this reason, understanding ageing has been a particular interest for the research community especially within the last decade, thanks to the improvement in high-throughput technologies enabling simultaneous analysis of many factors together. As ageing is a complex

state with many contributors, it is crucial to collect information regarding different components. In contrast, the early days of ageing research lacked such technologies and studies were concentrated on understanding the modulation of lifespan through single gene mutations. The first genetic element modulating lifespan was discovered in *C. elegans* (Friedman & Johnson, 1988; Kenyon, Chang, Gensch, Rudner & Tabtiang, 1993) and following studies confirmed the role of homologous pathways on long-lived mutants of *D. melanogaster* and dwarf mice (G. M. Martin, 2011). This was exciting because it showed that organisms actually carried significant potential to extend their lifespan, and also provided models to study healthy ageing.

One of the earliest discoveries in ageing research, dating to the 1930's, was the observation that rodents live longer when subjected to dietary restriction (Bishop & Guarente, 2007). In subsequent studies, this effect was at least partly replicated (in the form of increased healthspan) in a diverse range of organisms, including yeast, nematodes, flies, and primates. Meanwhile, a number of chemicals have also been suggested to extend median or maximum lifespan, some physiologically mimicking dietary restriction (Fontana, Partridge & Longo, 2010).

Another early finding relevant to ageing was the discovery of replicative senescence. In the 1960's, a group of researchers demonstrated that non-transformed cells have limited division capacity (Hayflick, 1965; Hayflick & Moorhead, 1961). Following work showed that the number of divisions decrease as the age of the donor increases (Campisi & Robert, 2014). This could explain why tissues fail to regenerate with advanced age. Another important finding in this area was that cellular senescence works in accordance with tumour suppressor mechanisms (Campisi, 2001). However, apart from arresting the cell cycle, another common feature of senescent cells is senescence-associated secretory phenotype (SASP), which is known to include secretion of pro-inflammatory cytokines, growth factors, and proteases. Being referred as "the dark side of cellular senescence", SASP, can actually trigger development of neoplasm through secretion of mitogens by accumulated post-replicative senescent cells. Recent work on the epigenetic landscape of senescent cells also supports this idea, reporting senescence-related methylation changes that resemble those observed in cancer (Cruickshanks & Adams, 2011; Cruickshanks et al., 2013). The senescent phenotype has also been associated with other

age-related pathologies such as atherosclerosis, hypertension, skin thinning, degeneration, osteoarthritis, COPD, insulin resistance and many others (Campisi & Robert, 2014).

Both the discovery of single gene mutations and environmental manipulations that can alter lifespan, and the nature of replicative senescence have had serious contributions to our understanding of ageing today. However, until the advent of high-throughput genome and transcriptome analysis technologies, ageing was studied in reductionist manner. Studies were limited in addressing the complex nature of this multifaceted phenomenon. The revolution in ageing studies, as in many research areas in molecular biology, came with the application of global gene expression analysis. In the first such study, the authors analysed $\sim$6000 genes in neocortex and cerebellum in young and ageing mice, and found increased inflammatory response, oxidative stress and reduced neurotrophic support at old age (Lee, Weindruch & Prolla, 2000). Intriguingly, in the same study, caloric restriction was found to curb the expression of inflammatory and stress response genes observed at old age, implying that these expression changes may have relevance to the ageing phenotype. In the following years, many studies followed, investigating the ageing phenotype at the genomic (Faggioli, Wang, Vijg & Montagna, 2012; Forsberg et al., 2012; Moskalev et al., 2012), epigenomic (Horvath et al., 2012; T. Yuan et al., 2015), transcriptomic levels (de Magalhães, Finch & Janssens, 2010; Erraji-Benchekroun et al., 2005; Lu et al., 2004; Zahn et al., 2006), as well as at the systems level (McAuley et al., 2009; Xue et al., 2007). I will describe these later in this section.

1.1.2 Theories of Ageing

The phenomenon of ageing has been of great interest to the research community and more than 300 theories to explain "why and how ageing occurs?" has been postulated throughout the late 19^{th} and the 20^{th} centuries (Medvedev, 1990). However, there is no single theory that can fully explain the ageing process. Instead, we can think of ageing as a combination of diverse biological phenomena postulated by different theories. We can further classify the theories as describing the ultimate and proximal reasons / causes of ageing, which describe evolutionary and mechanistic (molecular and physiological) processes, respectively.

Evolutionary Theories

The basic evolutionary perspective on ageing suggests that it is a result of decline in selective force with age. The first clearly articulated evolutionary theory of ageing is the "**mutation accumulation**" theory (Medawar, 1952). The theory states that even if a *de novo* germline mutation has a negative effect on a survival/reproduction gene, such a mutation cannot be opposed by natural selection if this is a late-acting gene. As a result, such mutations can drift to fixation and accumulate in the genome over generations.

Another theory concerns pleiotropic genes that show opposing effects in early life and later ages (Williams, 1957). Thus, if a gene has beneficial effect during early years when natural selection is strong, but has detrimental effect in later life when selection is weak, it will still reach fixation. The idea is today known as "**antagonistic pleiotropy**" theory.

A similar idea, but described in terms of physiological trade-offs instead of purely genetic mechanisms, has been coined the "**disposable soma**" theory. The theory suggests that an organism needs to divide the metabolic resources between maintenance and reproductive functions (Kirkwood, 1977). Thus, since the organism has limited resources, and because allocating resources to reproduction is evolutionarily preferred, it will tend to allocate an insufficient amount of resource to maintenance. This will cause somatic damage to accumulate, leading to ageing.

Molecular Theories

At the molecular level, there are several characteristics associated with ageing. Each of these is associated with the normal ageing process, but none is capable of explaining ageing on their own. Rather, they act together to form what we call as the ageing phenotype, but their individual contributions and the interconnections among them is still an open research area. These characteristics have been recently reviewed and gathered under nine "hallmarks of ageing" (López-Otín, Blasco, Partridge, Serrano & Kroemer, 2013). These hallmarks are further divided into three groups:

1. **Primary Hallmarks :** These are the primary causes of the cellular damage.

   (a) **Genomic Instability:** Somatic mutation accumulation, especially mitochondrial mutation accumulation, have long been suspected as a mechanisms of ageing. Accumulation of genetic damage can happen due to extrinsic effects such as mutagens, or due to intrinsic mechanisms such as spontaneous replication errors. These can result in point mutations, translocations, chromosomal abnormalities, transposon integrations *etc.*. All these alterations in the genome can affect essential genes and transcription. However, somatic damage is not limited to nuclear DNA. Mitochondrial DNA and nuclear architecture are also possible targets, which result in defects in energy metabolism and in the transcription of nuclear genes, respectively. Contribution of genetic instability to ageing has also been shown to be important since many progeria syndromes are associated with deficits in repair mechanisms and nuclear structures(Gregg et al., 2012; Hoeijmakers, 2009).

   (b) **Telomere Attrition:** After the discovery of replicative senescence, telomere shortening was first suggested, and later shown to be a causal factor (Olovnikov, 1996). Most of the mammalian somatic cells do not have telomerase expression, which causes time-dependent and cumulative loss of telomere sequences. Also, pathological telomere deficiency is associated with decreased lifespan, and induced telomerase activation can reverse some premature and normal ageing phenotypes in mice (Bernardes de Jesus et al., 2012; Jaskelioff et al., 2011).

   (c) **Epigenetic Alterations:** Age-related epigenetic alterations involve global DNA hypomethylation, region-specific DNA hypermethylation, increased or decreased histone acetylation, methylation on specific histone proteins. Many such changes can have strong effects on genomic integrity and ageing. For example, a 2011 study identified an epigenetic perturbation (H4K16 hypoacetylation) linked to progeria, and the authors could further reverse the progeroid syndrome by re-establishing normal H4K16 acetylation (Krishnan et al., 2011). However, to what degree observed genome-wide epigenetic changes represent spontaneous cellular damage, how much they directly contribute to ageing, and how much they may represent programmed responses to other age-related

cellular processes (see below), still remains unknown.

(d) **Loss of Proteostasis:** Proteostasis or protein homeostasis is the total of processes evolved to preserve a stable and functional proteome. Proteostasis involves chaperons and protein modifiers that ensure efficient protein folding and function, as well as the proteasome or lysosome for protein degradation. Accumulation of non-functional and misfolded polypeptides is associated with many age-related pathologies such as Alzheimer's and Parkinson's, especially in non-mitotic tissues. Accordingly, improvement of proteostasis via genetic alterations has been shown in a number of experiments to delay mammalian ageing (Zhang & Cuervo, 2008).

2. **Antagonistic Hallmarks:** These are the responses to compensate the effect of damage induced by the first category. In excess and with time, these can become pathological.

   (a) **Deregulated Nutrient Sensing:** It is suggested that ageing is accelerated by anabolic signalling, which decreases lifespan (Fontana et al., 2010). Moreover, restricted nutrition or drugs mimicking this state, such as rapamycin, have been shown to extend longevity in various model organisms (Harrison et al., 2009).

   (b) **Mitochondrial Dysfunction:** Ageing is associated with decrease in the efficiency of energy metabolism, resulting in low levels of ATP production. The change might be due to accumulation of reactive oxygen species, mtDNA mutations, and involve the process called mitohormesis (mitochondrial mild stress signals that invigorate the cell against further stress) (Yun & Finkel, 2014).

   (c) **Cellular Senescence:** It is the most complex contributor of ageing. Because senescent cells accumulate in tissue with age, it was initially suggested as an ageing mechanism. However, senescence at the same time prevents damaged cells from proliferation. So, cellular senescence can be seen as a beneficial system that decreases the possibility of oncogenic transformations. Conversely, considering the SASP phenomenon, which was mentioned before (see **Section 1.1.1**), cellular senescence's relation to ageing becomes more complicated. Accordingly, limited

up-regulation of senescence-inducing tumour suppressors are known to increase lifespan (Matheu et al., 2007), while removal of senescent cells from a progeria mouse model also delays age-related pathologies (Baker et al., 2011). These observations underscore the complex relationship between senescence and ageing.

3. **Integrative Hallmarks:** These are the culprits of ageing resulting from the above two processes and the ultimate causes of decline in functionality.

   (a) **Stem Cell Exhaustion:** Decreased proliferative capacity of tissues with age is one of the most prominent characteristics of ageing. Although low proliferation of stem / progenitor cells limits removal of damaged cells, increased proliferation of stem cells is also detrimental, as it causes stem cell exhaustion. This phenomenon is associated with more than one other characteristics of ageing, including loss of proteostasis and cellular senescence.

   (b) **Altered Intercellular Communication:** Multicellular organisms also demonstrate changes in intercellular communication including hormonal and non-hormonal communication. One of the most prominent consequences is "inflammaging" which can be the result of accumulation of pro-inflammatory tissue damage, defective immune system, SASP-related cytokine secretion *etc.*. (Salminen, Kaarniranta & Kauppinen, 2012).

The authors only described two integrative hallmarks; one regarding stem cell pool and the other regarding the intercellular communication. Even if it does not satisfy the criteria to be a hallmark, another important consequence of the primary causes corresponds to intracellular changes. Cellular damage accumulation, mainly as a result of genome instability and loss of proteostasis, can result in cellular dysfunction. This aspect is especially important for the non-mitotic tissues and can ultimately contribute to the decline in functionality.

1.2 Structural Changes in Human Brain with Age

In this section, organisation of the human brain, structural and functional changes in human brain with age, and summary of the previous brain ageing transcriptome studies will be given.

1.2.1 Human Brain Organisation

In this section, before describing ageing-related phenotypes observed in the human brain, I will first describe the organisation of the human brain, based on *"Elements of Molecular Neurobiology"* (Smith, 2002). There are two types of cells in the nervous systems: **neurons** and **glia**. Information process in the nervous systems is possible thanks to the neuronal networks, which are specific organization of afferent, efferent and inter-neurons so that input retrieval, output generation and integration can occur. In vertebrates, **brain** and **spinal cord** are responsible for information processing, storage, and retrieval. These are called the **central nervous system (CNS)**. Sensory and efferent neurons and their supporting cells which are distributed across body are called **peripheral nervous system (PNS)**.

Neurons

Neurons can transmit messages from, to or within the CNS. One of the most distinctive features of neurons is that almost all of the neurons lose their potential to proliferate after they differentiate. Different neuron types have been defined such as pyramidal cells, stellate cells *etc.*; however, all of these share a common architecture. Most of the neurons consist mainly of 4 regions: cell body, dendrites, axon(s), and axon terminals. Nucleus and most of the organelles are within the **cell body**. The projections branching out from the cell body are called **dendrites** and these are responsible for retrieval of information from other cells. Most of the neurons have one projection, which is much longer than the rest and is called the **axon**. Axons are responsible for carrying action potential -the physiological message- from the cell body towards the target cell. At the end of the axon a group of fine nerve endings with a swelling are present. These swellings are called **axon terminals** and are responsible for intercellular communication via **synapses**. Synapse is the structure responsible for inter-neuronal transmission. There are two types of synapses: electrical and chemical. Most of the synapses in vertebrates are chemical synapses. Chemical transmission is mediated by **neurotransmitters**, which are a group of chemical molecules. Action potential arriving at the end of the axon causes neurotransmitters to be released from axon terminal. These molecules then bind to the receptors on the target cells and thus transmit the message.

Glia

Glial cells were previously thought to act as a glue holding the nervous system together without any functional role. However, we now know that they have many other functions. The number of glia in the human brain is much more than neurons. There are several types of glia, each with different functions.

1. **Oligodendrocytes**, in the white matter, wrap around axons and produce a protein called myelin. As a result, tight multi-layered myelin sheaths are formed. Myelination increases the axonal conduction dramatically. In the grey matter, they have role in the metabolic interactions among neurons.

2. **Astrocytes** contribute to the blood-brain barrier, which is the structure responsible for protecting brain from toxic materials in the blood. Also, because of the presence of strong filaments in the cytoplasm, astrocytes confer tensile strength to nervous tissue. Astrocytes can secrete neurotransmitters, which act as growth factors in developing brain. Furthermore, by the secretion of cholesterol and apolipoprotein E into intercellular space, astrocytes play role in synapse formation and maintenance.

3. **Microglia** originate from bone marrow stem cells and function in the brain immune defence and ionic environment maintenance.

Neuron Organisation in Brain

There are two types of substances in human brain:

1. **White matter** which includes many nerve fibres. It appears white due to the presence of myelin sheats surrounding the neurons.

2. **Grey matter** which includes dendrites and cell body of the neurons and many glial cells. Since these components are not myelinated, it appears grey.

Unique Features of the Human Brain

When compared to the closest relatives, the human brain bears differences in phenotypic, anatomical and histological features. Phenotypic differences include increased cognitive abilities such as social intelligence, cooperative behaviour, and language (Somel, Rohlfs & Liu, 2014). Increased brain volume / body ratio is the

most prominent anatomical change, which is suggested to be a reflection of increased number of neurons. Regarding the histological changes, we observe a high proportion of glial cells (Rowitch & Kriegstein, 2010), increased intercellular connectivity and increased astrocyte diversity and complexity (Oberheim et al., 2009). Although these phenotypic, anatomical and histological differences are defined, the causal relationship among these differences is still not clear and needs more studies at the molecular level. Nevertheless, there are several suggested explanations linking these unique features to the human brain developmental process (Jiang & Nardelli, 2015).

1.2.2 Embryonic Period

Human brain development is a complex process necessary for centralized behaviour control, perception and cognitive abilities. This process involves the emergence of new cell types in the right place, right time and with correct numbers, as well as the establishment of connections between these neural and non-neural cell types. Brain development starts by the end of third week after gastrulation. After this stage, stem cell lines are produced, including the neuroectodermal stem cells. These stem cells can produce all cell types that forms the CNS and can be referred to as neural stem cells or neural progenitor cells. At 3^{rd}-4^{th} weeks of gestation the neural tube forms. By the end of the 8^{th} week of gestation, the 5 major subdivisions of the brain are formed (telencephalon, diencephalon, mesencephalon, metencephalon, myelencephalon), and these make up the overall structural organization of the CNS (Stiles & Jernigan, 2010). By the end of the embryonic period, general structure of neocortex sensorimotor regions, diencephalic and midbrain regions, spinal column and hindbrain segmental organization are established (Stiles & Jernigan, 2010). Although the general architecture is said to be formed, it is important to note that the basic structure and function of the brain is still open to changes as a result of input and experience, even during early postnatal development (Johnson, 2001).

1.2.3 Fetal Period

Most of the neurons are produced during mid-gestation. Until the 7^{th} week after conception, neural progenitors undergo symmetrical cell division, which increases the progenitor cell pool. After this stage, cell division becomes asymmetrical, which results in one neural progenitor and one neuron production (Jiang & Nardelli, 2015). The new progenitor continues to divide asymmetrically, whereas the post-mitotic neuron migrates to the developing cortex. Human cortical neurogenesis is almost complete at the day 108 after conception (Clancy, Darlington & Finlay, 2001). However, it is suggested that neurogenesis in the macaque neocortex continues even in the adulthood (Gould, 1999).

Neurons that migrate to different layers of cortex are different types of neurons, but all of them can be produced by the same type of neural progenitor cell (Desai & McConnell, 2000). However, the exact mechanism which controls the timing of production of different neuronal types is not well-defined (Molyneaux, Arlotta, Menezes & Macklis, 2007).

The next step is the formation of axons and dendrites, which are necessary for the establishment of neuronal networks. At the end of axons, a structure called "growth cone" is present and it is responsible for axon elongation and extension (Stiles & Jernigan, 2010). After reaching to the target cell, synaptic connections are formed.

Physiological events in the brain are not only restricted to progressive developmental processes. Instead, cell death and synaptic connection loss is also seen. Most of the neuronal cell death occurs during fetal period, whereas glia death and connection loss are mostly postnatal processes. There are two types of cell death: necrosis and apoptosis. The second one is triggered by cell intrinsic signals and is common in all neuronal cell types (Rakic & Zecevic, 2000). One of the established protection mechanisms against the apoptotic cascade is uptake of neurotrophic molecules, which are basically growth factors. Neurotrophic factors are produced by target neurons and if a neuron has more effective connections, the available neurotrophic factor amount is accordingly large, so that it can survive (Huang & Reichardt, 2001). In the brain, apoptotic cell death may serve for (1) effective and

functional neural network formation; (2) correcting errors in neuron production and migration; (3) elimination of neurons having transient function in development (Stiles & Jernigan, 2010). Apoptotic cell death is also suggested for the neural progenitor cells, but since these cells do not form synaptic connections, the mechanism to ensure their survival should be different from those of mature neurons.

1.2.4 Postnatal Period

Neurogenesis and the general brain region architecture is almost complete at birth. Postnatal neurogenesis is generally associated with hippocampal and neocortical areas and is suggested to be important for memory and learning (de Graaf-Peters & Hadders-Algra, 2006; Gould, 1999). Postnatal brain growth also involves further steps in neuronal maturation, and also proliferation of glial progenitors with important roles in functional organization (Bhardwaj et al., 2006). Synaptogenesis, axonogenesis, gliogenesis and myelination are the main processes during the postnatal brain growth (Huttenlocher & Dabholkar, 1997).

Postnatal proliferation in brain is mainly due to glial progenitors. These are present in the brain for the whole lifetime, which enables them to differentiate upon injury. They can proliferate as they migrate. Different glial cells have different destinations and functions. Once an oligodendrocyte progenitor cell (OPC) migrates and reaches its destination, it starts to differentiate, causing the production of myelin protein. As a result, tight multi-layered myelin sheaths are formed by oligodendrocytes. However, the functional result of these interactions is not limited to myelination, but involves maintenance of axonal integrity, neuronal survival and determination of neuronal size and axon diameter (Stiles & Jernigan, 2010). Astrocytes, the other principle macroglial cell type originating from the common multipotent progenitors, are also important in postnatal brain growth (Jiang & Nardelli, 2015).

Similar to fetal growth, postnatal period also involves regressive events such as glial apoptosis and synaptic pruning. During the initial myelination process, there is an excess amount of oligodendrocytes. However, receiving signals from axons, some oligodendrocytes survive, while others undergo apoptosis (McTigue & Tripathi, 2008).

Neuronal circuitry at birth is also rich in terms of synaptic connections and topography (Innocenti & Price, 2005). In contrast, towards adulthood the number of synaptic connections starts to decrease through a process called "synaptic elimination" (de Graaf-Peters & Hadders-Algra, 2006), and reaches almost the half of the number at birth. Competition for neurotrophic factors and afferent input are suggested as important contributors of modulation and stabilization of neuronal networks (Stiles & Jernigan, 2010).

1.2.5 Ageing

Postnatal changes in human brain is not limited to the age of 20. Instead, structural and histological changes continue until death (Peters, Sethares & Luebke, 2008). While some of these changes are the continuation of developmental processes, some might be direct outcomes of instability and damage. Between the ages of 20 and 50, myelination continues in some brain regions, whereas between the ages of 20 and 80, we observe a general decline in the total length of the myelinated fibres without a prominent change in the neuronal density (Sowell, Thompson & Toga, 2004). Starting from the 50's, brain weight starts to decline, which is a result of neuronal shrinkage. Reported differences during ageing also include change in grey and white matter volumes. Grey matter volume declines with age starting from the age of 7. The loss of grey matter volume between the ages of 7 and 20 is suggested to be associated with a combination of synaptic pruning and increased myelination, while changes after age of 20 until late adulthood is linked to neuronal shrinkage and continuation of myelination. Decrease in the late adulthood when the myelination is complete, can be the result of continuing degenerative effects (Sowell et al., 2004).. Unlike grey matter, white matter change is not linear in adulthood. It increases until 50's, consistent with the continuation of myelination, and declines afterwards. All these changes are accompanied by the phenotypic characteristics of human brain ageing; such as cognitive decline, memory loss, inclined occurrence of neurological disorders (Salthouse, 2009).

1.3 Transcriptome Change in Brain with Age

By definition, a phenotype is determined by the genotype and the expression profile -which can be influenced by environmental factors- of the genes within an organism. Several studies have showed that development and function of the brain

requires a well-orchestrated gene expression and changes in the regulation of this precise profile can result in neurodevelopmental disorders. Consistent with the physiological changes in the developing brain, genes important in neuronal differentiation, cell proliferation, and migration show high expression in the early stages of development, whereas in the later years of development, genes important for synaptic functions start to increase. In the mouse brain, it was shown that regional specialization of the transcriptome occurs in postnatal years, but the associated changes in the expression profile during thalamus and cortex specialization are different in human and mouse (Liscovitch & Chechik, 2013).

Although the physiological and phenotypic changes after the age of 20 years is well-studied, the molecular mechanisms behind these changes are not well-defined. During the last 15 years, there has been a significant number of studies focusing on the gene expression changes during human brain ageing, thanks to improvement in the high-throughput technologies.

The first global-scale transcriptome study in ageing brain was on mice and analysed $\sim$6000 genes in neocortex and cerebellum (Lee et al., 2000). The authors found gene expression signatures consistent with increased inflammatory response, oxidative stress and reduced neurotrophic support. The authors also showed that ageing mice brain shows changes in expression of lysosomal proteases, chaperons, and proteins associated with oxidative stress and inflammation. This result is in accordance with the changes observed in human neurodegenerative diseases. Furthermore, inflammatory and stress response genes that show age-related increase were shown to be repressed upon caloric restriction.

Another study on gene expression profile of ageing rat hippocampus also showed that inflammation, oxidative stress, protein processing, and decreased mitochondrial function are age-related pathways (Blalock et al., 2003). These authors also suggested a couple of novel pathways including decrease in early response signalling, emergence and maintenance of synapses, and increased myelin turnover, cholesterol synthesis, lipid and mono-amine metabolism, iron-related pathways, structural reorganization, and Ca^{2+} release pathways.

In 2004, another study was published investigating ageing-related changes in the human prefrontal cortex (Lu et al., 2004). Consistent with the previous studies, the authors showed decreased expression of genes related to synaptic plasticity, vesicular transport, and mitochondrial function, in contrast to increased expression of stress response and antioxidant genes. They argued that DNA repair genes show increased expression as a response to damage caused by high oxidative stress. The authors further claimed that cis-regulatory elements (DNA sequences controlling the expression of nearby genes) of the genes functional in learning, memory and neuronal survival are selectively vulnerable to DNA damage, and thus show decreased expression with age.

Another study on human prefrontal cortex showed that most of the genes do not show age-related change with age (Erraji-Benchekroun et al., 2005). Genes that increase in expression with age are mostly of glial origin, whereas decreasing ones are neuronal transcripts. The authors argued that most of the expression change occurs in early adult ages, and then show continuous expression throughout lifespan, which suggests detrimental age-related brain functions have the origins in these years. Another study showed that the rate of gene expression change is the highest before the age of 20 and it starts to decrease after maturation (Naumova et al., 2012).

In a study in 2007, the authors argue that different studies (Lu et al., 2004; Lund et al., 2002; Pletcher et al., 2002) find different genes as differentially expressed with age but if these are studied at the network level, these might be genes in the same modules (Xue et al., 2007). Integrating protein-protein interaction network, they showed that fruitfly and human brain ageing show similar pattern that cellular proliferation module shows decrease whereas differentiation module shows increase with age. The number of modules showing change with age is low and some of these changes are reversible.

Another work on rat hippocampal ageing suggested that defects in neuroenergetic pathways during ageing cause disruptions in neuronal and increased glial activity, which then result in cognitive dysfunction (Rowe et al., 2007).

Studying brain ageing focusing from an evolutionary perspective, another study showed that humans and rhesus macaques are different from mice in the repression of neuronal genes -especially the ones with synaptic function- during ageing (Loerch et al., 2008). They, however, could not find an overall neuronal or synaptic loss, despite gene expression down-regulation. The reason for the discrepancy between this result and the reported trend of synaptic loss in macaques during ageing remains unclear. In another study, the age-related changes in development is suggested to be more conserved than that in ageing (Somel et al., 2010).

Other studies have addressed the question of sex/gender differences in ageing. Analysing 4 different brain regions, a 2008 study showed that different brain regions are associated with different age-related changes, and also that expression profiles in males and females showed clear differences (Berchtold et al., 2008). Considering the different trajectory of neurodegenerative disorders in males and females, the authors suggest that normal brain ageing should also be studied on different sexes, separately. Later, the same data was tested for for differences in timing of ageing-related expression changes between sexes; perhaps unexpectedly, this study found accelerated ageing-related expression changes in females relative to males(Y. Yuan, Chen, Boyd-Kirkup, Khaitovich & Somel, 2012).

A number of studies have undertaken meta-analyses of ageing-associated expression changes across different tissues, including the brain. One such study identified common transcriptome patterns that accompany ageing in diverse tissues, including up-regulation of inflammatory, immune response, and lysosome genes, and down-regulation of collagen genes, energy metabolism, apoptosis, cell cycle and cellular senescence biomarkers (de Magalhães, Curado & Church, 2009).

A 2010 study, using gene, protein and miRNA expression profiles, showed that trans-regulation (regulation of expression by distant elements such as miRNAs and TFs) of gene expression is not only important in development but also has affect in ageing period. The authors further showed that many of the ageing-related changes are not unique to ageing but instead either continuation or reversal of the developmental process (Somel et al., 2010). That expression patterns observed in ageing can be continuation of developmental trends had been reported earlier (Errajji-Benchekroun et al., 2005). Gene expression reversals will be explained in detail

in the **Section 1.4**. The same phenomenon was observed in the same brain region -prefrontal cortex- by another group when they used much larger sample size (Colantuoni et al., 2011).

Another 2011 study investigated 16 distinct brain regions' transcriptomes across lifespan starting from the embryonic period. The authors found that 90% of analysed genes are differentially regulated across brain regions and/or time (Kang et al., 2011). Also, they showed that most of the age-dependent change occurs in prenatal years.

Brain ageing has also been investigated through use of second-generation sequencing-based RNA-sequencing (RNA-seq) technology in the recent years, and the results are in accordance with the microarray results. Similar functional groups, as those identified in microarray studies, are shown to be differentially expressed with age in RNA-seq studies (Naumova et al., 2012). One particular advantage of using RNA-seq is that it enables investigation of splicing patterns. Splicing changes with age have been shown to affect a third of genes expressed in the postnatal brain and 15% of these show differences between prefrontal cortex and cerebellum (Mazin et al., 2013).

Understanding ageing at the molecular scale also requires analyses of levels additional to mRNA expression, such as DNA methylation and chromatin marks. For example, there is a general observation indicating ageing-related global hypomethylation and promoter-specific hypermethylation during ageing (Calvanese, Lara, Kahn & Fraga, 2009). In addition, brain-specific analyses showed that nervous system development, neuron differentiation and neurogenesis genes show differential methylation with age (Horvath et al., 2012). The latter study also claimed that blood methylation patterns are in accordance with that of brain and blood can be used for screening of age-related effects.

Heterogeneity in Ageing

In the studies described earlier, genes important for synaptic functions, energy metabolism and inflammation have been reported to show significant mRNA expression changes during ageing. However, elucidating the regulation of these genes

and the interplay between them is not straightforward, because one of the prominent age-related gene expression changes is increased inter-individual heterogeneity (de Magalhães et al., 2009; Lu et al., 2004; Somel, Khaitovich, Bahn, Pääbo & Lachmann, 2006). Ageing, unlike development, is not thought of as a programmed event but a result of stochastic evolutionary and cellular dynamics, and accordingly, phenotypes have been observed to diverge among genetically identical individuals with age, in nematodes (Herndon et al., 2002) as well as in human twins (Fraga et al., 2005). There are several proposed evolutionary explanations on "why ageing cannot be a programmed process?" (Kirkwood, 2005), as well as molecular explanations. Analysing eight different expression datasets of rat and human, gene expression was found to become more variable among individuals with age. However, this result, the authors suggested, was not related with expression of germline mutations or difference in ageing rates among individuals. Instead, the heterogeneity is suggested to be a result of stochastic events such as accumulated damages and mutations in the somatic cells (Somel et al., 2006). In another study on mouse heart cells, age-related increase in heterogeneity due to stochastic effects was further tested and confirmed at the cellular level (Bahar et al., 2006). Another possible contributor to heterogeneity is environmental factors. Global hypomethylation and promoter-specific hypermethylation are suggested as age-related epigenetic changes (Calvanese et al., 2009). Showing the epigenetic changes in ageing is particularly important as it has a dynamic nature. Thus, environmental factors can be another contributor of the heterogeneous nature of the ageing through dynamic changes in the epigenome. Previous studies (Y. Yuan et al., 2012) also suggested gene expression changes in ageing can be influenced by psychological stress. All these imply that ageing is not only accompanied by molecular changes and functional impairments, but these changes and impairments appear to be driven by the accumulation of stochastic events, and influence different individuals to different extent.

Meta-analysis of Ageing Datasets

As described earlier, multiple experiments have been published that study ageing in different tissues and organisms, including the human brain. As the number of available datasets increase, meta-analysis of multiple datasets becomes possible. This is particularly important, considering the high inter-individual heterogeneity

observed in ageing, and the usually limited sample sizes used in ageing studies (due to difficulties in keeping old laboratory animals and sampling human tissue). The major advantages of meta-analyses are: 1) Because ageing is not a programmed process, integration of multiple datasets can reveal underlying common signatures in changes that are otherwise invisible; 2) Comparing profiles from different experiments can increase signal / noise ratio; 3) By integrating multiple experiments, confounding effects in individual datasets can be identified and removed.

Meta-analyses can help to confirm ageing-related patterns observed in primary publications, as well as identify novel patterns and test new hypotheses on ageing. Some of such studies have been mentioned before. For instance, one 27-dataset investigation confirmed age-related mRNA expression changes of energy metabolism, inflammatory and immune-related genes. The authors further argued that some of the age-related changes can be response to ageing rather than being the underlying mechanism (de Magalhães et al., 2009). Increased inter-individual heterogeneity was shown through meta-analyses, as described earlier. Along the same line, a re-analysis of AGEMAP database (Zahn et al., 2007) based on co-expression networks, showed decreased co-regulation among genes with age, also suggesting a loss of transcriptome coordination during ageing (Southworth, Owen & Kim, 2009). More generally, all these examples show the importance of meta-analysis and re-analysis of datasets using different approaches to improve our understanding of the underlying processes.

1.4 Gene Expression Reversal

Most of the studies on ageing have focused only on changes observed in adulthood. However, studying ageing with respect to the developmental processes can provide better understanding of what happens during ageing. Following this idea, there has been several studies profiling the transcriptome changes in brain throughout the lifespan (Colantuoni et al., 2011; Erraji-Benchekroun et al., 2005; Kang et al., 2011; Somel et al., 2010).

In one of these studies (Somel et al., 2010), it was suggested that ageing can be seen as either a continuation or reversal of the developmental process. Considering the age of 20 as a transition point from development to ageing, it was observed

that 64% of these genes show reversal pattern (clusters 2, 5, 6, 8 in **Figure 1.2**). Genes showing up-down pattern (clusters 2 and 5) were suggested to be important in energy metabolism, which is consistent with the suggestion that oxygen consumption increases until the age of 20, and then starts decreasing (Aanerud et al., 2012). Another cluster showing up-down pattern (cluster 8) was associated with neuronal and synaptic functions. This result is also in accordance with the studies suggesting synaptic loss and cognitive decline during ageing (Peters et al., 2008). Finally, genes showing down-up pattern (cluster 6) were found to be enriched in DNA damage response.

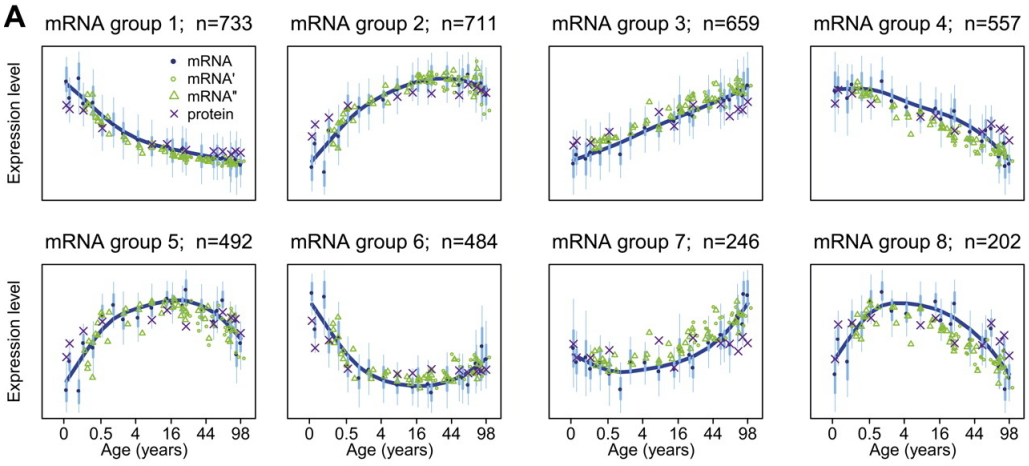

Figure 1.2: Major mRNA change patterns with age (Somel et al., 2010)

In another study, to relate the developmental functions with the reversal pattern, expression profiles of prenatal and infant samples were compared with the later postnatal years. 75% of the genes showing significant change with age were reported to show the reversal pattern. Those genes showing up-down pattern were enriched in neuronal development-related functions, whereas the down-up pattern was associated with cell division (Colantuoni et al., 2011). The results were in accordance with another study showing that genes enriched in differentiation and proliferation could show reversal patterns at the age of 85 (Xue et al., 2007).

If reversal is a common pattern, when does it initiate? So far, different studies reported different "turning points" or "break points" where the gene expression change direction changes. However the most common result indicates around 20 years of age (Colantuoni et al., 2011; Somel et al., 2010).

So far, two studies focusing on the same brain region (prefrontal cortex) suggested a high proportion of genes showing reversal patterns during ageing. However, the prevalence of such a pattern in other brain regions is an open question. Similarly, the regulatory processes behind these gene expression changes, and the mechanistic nature of these changes are still waiting to be uncovered. Several suggestions have been made, including potential miRNA and TFs as regulators and somatic mutations; however, a systematic investigation is necessary to comment on the prevalence, functional outcomes and regulation.

1.5 Research Objectives

In this study, the aim is to characterise the age-related gene expression changes in ageing in the context of development. The general hypothesis is that ageing is not a distinct phenomenon but a consequence of disruptive changes in gene regulation after adolescence.

Gene expression reversal, which is suggested for majority of the genes, is especially the focus of this study. One of the important changes during development is the activation of the differentiation pathways. As the organism approaches adulthood, activity of these pathways decreases. Considering that there are several ageing-related pathologies, including cancer and Alzheimer's disease, reported to render cells inclined to a de-differentiated state, a relevant question is whether gene-expression reversal causes cells to lose their cellular identity, which in turn may cause a predisposed state to the malignancies and/or dysfunction. With this motivation, I investigate the cellular pathways showing gene expression reversal over-representations.

Using published microarray gene expression data, I investigate the expression changes in ageing in the context of that in development, the possible functional associations of these changes, and their regulation.

The objectives are:

- To assess the nature of age-related gene expression changes in terms of heterogeneity and strength in development and ageing,
- To measure the prevalence of gene expression reversals among different brain regions,
- To characterise the functional pathways associated with the gene expression reversals,
- To elucidate possible drivers of the gene expression reversal by analysing trans-regulators of the reversal genes.

CHAPTER 2

MATERIAL AND METHOD

2.1 Datasets

In this study, gene expression data from 7 different published studies were analysed. All data are human microarray datasets. In total, gene expression profiles of 1,015 samples spanning the ages 0 to 106 years were analysed. The exact distribution of ages is given in **Figure 2.1**. The datasets have been generated using 4 different Affymetrix platforms. In total, 22 different brain regions were analysed. Among these, some are not mutually exclusive but are subregions of others (see **Chapter 4**). Summary of the datasets is presented in **Table 2.1**.

Three of the datasets have samples representing the whole lifespan (Kang et al., 2011; Somel et al., 2010, 2011), whereas the other four only consist of samples spanning adulthood (Barnes et al., 2011; Berchtold et al., 2008; Lu et al., 2004; Maycox et al., 2009). In human societies, including hunter-gatherer groups, 20 years of age roughly correspond to the age at first reproduction (R. Walker et al., 2006); earlier analyses had also shown that 20 years of age correspond to a global turning-point in brain gene expression trajectories (Colantuoni et al., 2011; Somel et al., 2010). Notably, processes such as myelination are known to continue well into early adulthood (see **Chapter 1.2**). Thus, in this study, 20 years of age was used to separate postnatal development from the ageing period. Accordingly, datasets with samples spanning both postnatal development and ageing were divided

into two datasets: (1) development (<20 years of age) and (2) ageing (>20 years of age). These datasets have comparable number of samples in development and ageing periods.

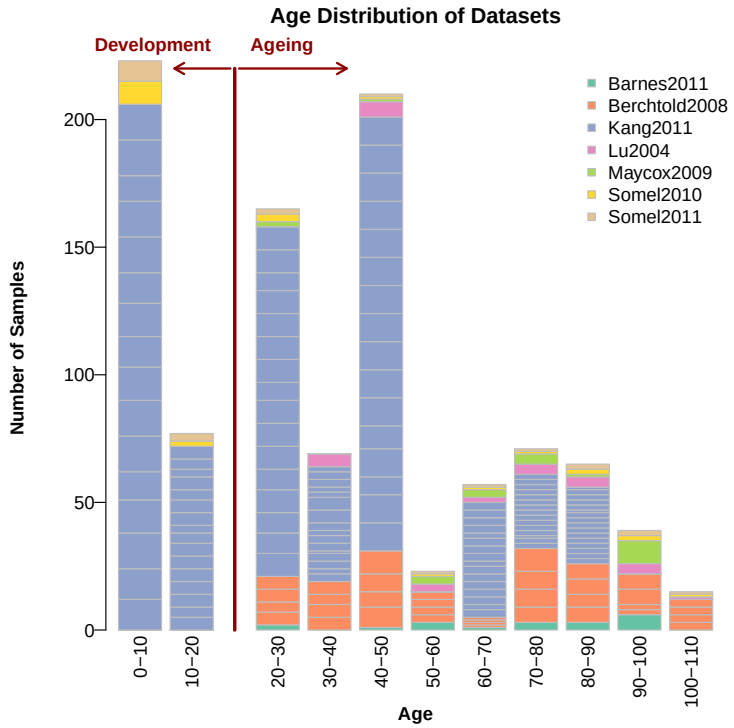

Figure 2.1: Distribution of ages across datasets

It is important to note that some of the samples in the Kang2011 dataset correspond to samples from the same individual's left and right hemispheres. These samples are treated as biological replicates rather than technical replicates for the following reasons:

1. A number of previous studies state that brain ageing might not be the same in right and left hemispheres (Dolcos, Rice & Cabeza, 2002; T. Sun et al., 2005).

2. Comparison of age-related gene expression change between the right and left hemisphere of the same individuals did not show strong correlation (Hamit İzgi, *unpublished*).

3. Although reasons are not specified, the original study also treats them as biological replicates (Kang et al., 2011).

4. There is no hemisphere information for the other datasets which are used to make comparisons. This also prevents division of left and right hemispheres into two separate datasets.

Table 2.1: Dataset summary. The "number of samples" column represents the total number of samples in the study, whereas the numbers in parentheses show the number of individuals.

GEO Acc	**Dataset Name**	**Brain Regions**	**# Samples**	**Platform**
GSE21935	Barnes2011	STC	19	HG-U133_Plus_2
GSE11882	Berchtold2008	EC, HIP, PCG, SFG	173 (57)	HG-U133_Plus_2
GSE1572	Lu2004	FC	30	HG_U95Av2
GSE17612	Maycox2009	APFC	23	HG-U133_Plus_2
GSE18069	Somel2010	PFC	23	HuGene-1_0-st
GSE22569	Somel2011	CBC	22	HuGene-1_0-st
GSE25219	Kang2011	A1C,AMY, CBC, DFC, HIP, IPC, ITC, M1C, MD, MFC, OFC, S1C, STC, STR, V1C, VFC	741 (33)	HuEx-1_0-st

2.2 Gene Expression Data Preprocessing

All of the microarray datasets in this study have been generated using Affymetrix microarray platforms. Microarray is an effective technology to estimate steady-state expression levels of transcripts. This estimation is performed through detection of light intensity, which is proportional to the abundance of fluorescent-labelled RNA molecules, and which, in turn, represents mRNA expression from a transcript. The pixel intensities are collected by the Affymetrix scanner and stored in .DAT file format. This information is used to calculate the light intensity levels for each probe. There are different Affymetrix programs that can perform this step. The converted data is stored in .CEL file format. As a result, each sample has an associated .CEL file containing a light intensity measurement for each probe on the chip. All .CEL files contain information about the light intensity, standard deviation of the intensity, and the number of pixels used for calculation. The exact content of the files depend on the version of .CEL format.

I started the analysis by downloading the .CEL files for each dataset, using the corresponding accession number from GEO database (Edgar, 2002).

As previously stated, some of the datasets have samples from more than one brain region. Since our interest is how expression changes with age within the region, each of the brain regions was processed separately. As a result, there are: (1) One ageing dataset for the Barnes2011; (2) Four ageing datasets for the Berchtold2008; (3) One ageing dataset for the Lu2004; (4) One ageing dataset for the Maycox2009; (5) One development and one ageing datasets for the Somel2010; (6) One development and one ageing datasets for the Somel2011; and (7) 16 development and 16 ageing datasets for the Kang2011.

The analysis was done in the R programming environment. The preprocessing involves mainly four steps: summarization of probe level data into probe-set level using RMA convolution model, summarization of probe-sets to genes, log2 transformation and quantile normalization.

2.2.1 RMA correction

RMA (log scale robust multi-array analysis) convolution method is an algorithm that is used for background normalization and probe to probe-set summarization. The information stored in .CEL files is at the probe level. In custom Affymetrix platforms, each transcript is represented by 11-20 probes. Summarization of probe level intensity values is necessary, so that each transcript receives one intensity value to represent its expression level. The probes corresponding to one transcript are grouped into "probe-sets" defined by Affymetrix.

Apart from probe to probe-set summarization, the RMA convolution method involves removal of technical artefacts so that the neighbouring measurements corresponding to different probes on the array do not affect each other (Irizarry, 2003).

For the RMA correction of "HG-U133_Plus_2" and "HG_U95Av2" platforms, "affy" (Gautier, Cope, Bolstad & Irizarry, 2004) and for "HuGene-1_0-st" and "HuEx-1_0-st", "oligo" R libraries were used (Carvalho & Irizarry, 2010).

There are two pairs of technical replicates in the Somel2010 dataset. As the correlation between replicates was high, they were represented by the mean value of the probe-set level expression values in the rest of the analysis.

2.2.2 Summarization of probe-sets to genes

In order to enable cross-platform analysis, expression levels should be defined as universal entities such as genes. Thus, probe-sets should be converted to corresponding gene IDs. However, the relation between probe-sets and genes is rarely one to one. There can be a probe-set associated with two genes, as well as two probe-sets associated with one gene (**Figure 2.2**).

If a probe-set corresponds to multiple genes, it was removed from the data so that the expression values were not duplicated based on a single observation. Genes having more than one probe-set annotation were represented by taking the mean expression level of the corresponding probe-sets.

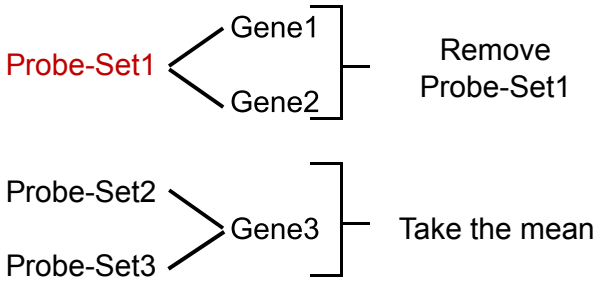

Figure 2.2: Summary of probe-set to gene summarization. The upper panel corresponds to a probe-set associated with multiple genes. This probe-set is removed from data. The lower panel shows a gene associated with multiple probe-sets. Expression value of this gene is represented with the mean expression value of the probe-sets.

An alternative approach would be to identify the probe-set, for each gene, with the maximum (or minimum) expression value across individuals in a dataset. That probe-set could then be used as representative for the gene. There are two specific reasons for using the mean expression value across probe-sets, instead of choosing one probe-set with the minimum/maximum expression values:

1. The expression level of different isoforms of a gene is known to change with age (Mazin et al., 2013; Rodríguez et al., 2016). It is possible that one isoform increases with age and the other one decreases. If so, choosing the probe-set with the minimum/maximum expression level would be misinterpreted as a specific gene's activity increasing or decreasing with age. However, in reality one isoform is just replaced by another isoform. These isoforms may function in the same manner or not, and studying changes in isoform abundance can be important for understanding ageing process. However, this is outside the scope of this study. Also, isoform abundance patterns are more conveniently studied using RNA-sequencing instead of microarrays.

2. We use data generated using different platforms, meaning that the same isoforms may not be targeted across different platforms. Probe-set definitions are also different for each platform. Thus, taking the minimum/maximum expression value among probe-sets per gene would complicate comparison across datasets.

For each platform, the number of probe-sets and genes after this step is given in **Table 2.2**. For "HG-U133_Plus_2", "HG_U95Av2", and "HuGene-1_0-st" Ensembl version 84 (Yates et al., 2015) annotations were used, whereas for "HuEx-1_0-st" original GPL file deposited in GEO database was used.

Table 2.2: Number of probe-sets and final number of genes for each platform after summarization as described in the text. ENSG represents Ensembl Gene IDs.

Platform	**# PS**	**# PS with ENSG**	**# PS >1 ENSG**	**Final # PS**	**# Genes**
HG-U133_Plus_2	54675	42009	2970	39039	21323
HG_U95Av2	12625	11504	1024	10480	8258
HuGene-1_0-st	33297	28985	4154	24831	22714
HuEx-1_0-st	22011	17392	2852	14540	14356

2.2.3 Log2 Transformation

Log2 transformation is a standard procedure in gene expression data analysis in order to make the gene expression level variance across genes more comparable, and to make the distribution more symmetrical. Gene expression level distributions across the genome are usually highly right skewed, with many genes having low levels of expression and a few genes having very high levels (**Figure 2.4**). Moreover, the mean and the variance are positively correlated with each other. Log2 transformation enables better separation between lowly expressed genes and makes the variance similar across genes. Thus, although I did not use any statistical test assuming normality, I applied log2 transformation.

2.2.4 Quantile Normalization

Quantile normalization is a complete data method aimed at removing non-biological variation across samples. The procedure makes the distribution of samples the same (**Figure 2.3**), but preserves differential expression between samples (Bolstad, Irizarry, Astrand & Speed, 2003). Quantile normalization applied in transcriptomics assumes that only a minority of genes show expression level differences, while the

overall expression level distributions are comparable. Thus, any major differences between the centre or shape of the distributions are considered non-biological and removed. The assumption that only a minority of genes are differentially expressed may not hold when divergent samples (e.g. samples of different tissues or organisms) are compared. But the assumption is expected to hold when the samples originate from the same tissue (brain) and have been collected under similar conditions (postnatal ages and healthy state).

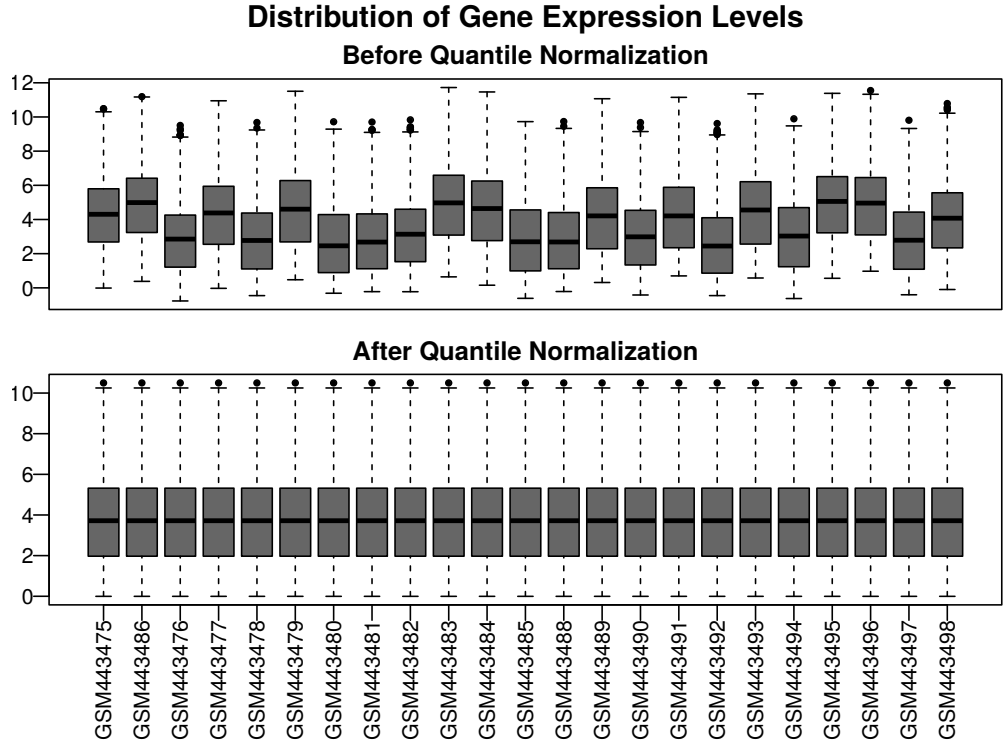

Figure 2.3: Example boxplot for the Somel2010 dataset before and after quantile normalization. Each boxplot shows the distribution of a sample.

The resulting distribution of probe-set/gene expression levels after each step is given in **Figure 2.4**.

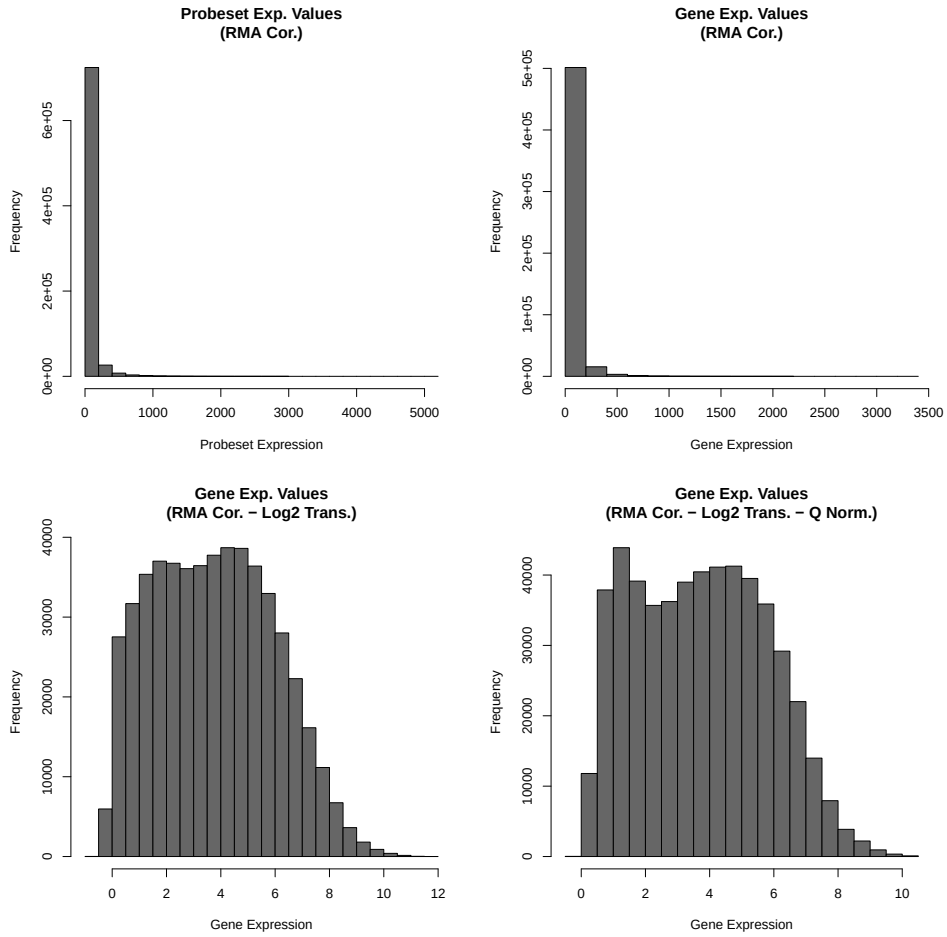

Figure 2.4: Example histogram for the Somel2010 dataset after each step.

2.3 Sample Quality Check

Principle components analysis (PCA) is a statistical approach that is widely used in gene expression studies to check variation among samples. It uses an orthogonal transformation approach to summarize correlated variables into uncorrelated ones, called principle components (PCs). Using this method, one can study transcriptome-wide clustering among samples in two or more dimensions.

Here I visually inspected PCA plots for each dataset, and depending on the distribution of the samples across the first and second components and a sample's

distance from the majority of other samples, I could pinpoint certain putative outlier samples. These are samples that show divergence from other samples, not due to age, but due to unknown variables (e.g. disease, diet, technical problems). Such outliers, influenced by unknown sources of variation, are expected to introduce noise in the data analysis aimed at identifying the effect of age on gene expression.

For the PCA calculations, R function "prcomp" was used with scaling option. Analysis was done on log2 transformed and quantile normalized probe-set expression levels. An example PCA is given in **Figure 2.5**.

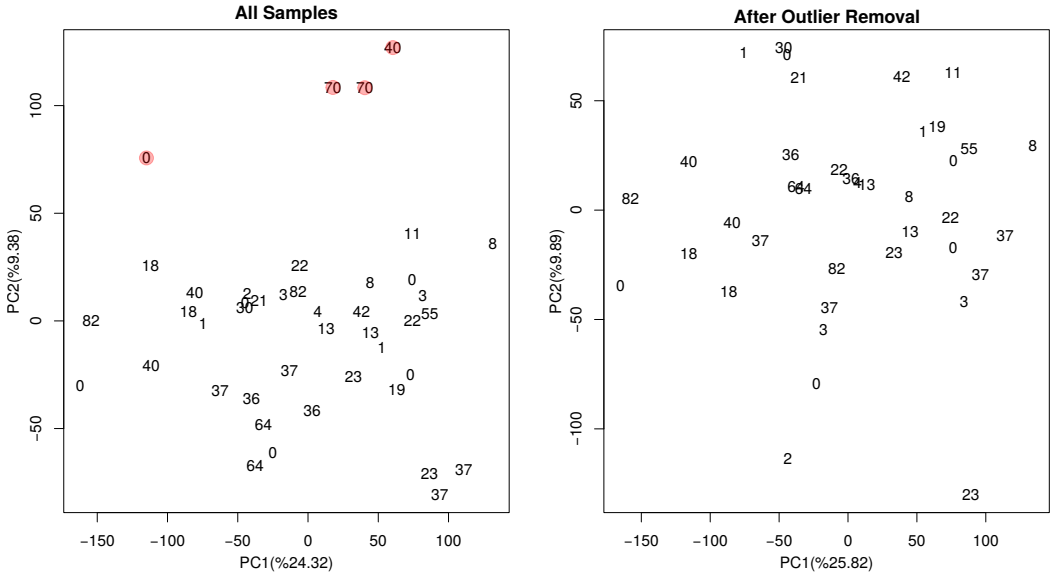

Figure 2.5: Example PCA for the Kang2011 dataset, Hippocampus brain region. Each number on plot shows the age of an individual. The x-axis corresponds to the first principle component and the y-axis corresponds to the second principle component. Percentages in parentheses represent the variance explained by each component.

Through visual inspection, the following putative outlier samples are discarded (numbers in the parenthesis represents the age of the sample):

1. Berchtold2008 HIP: *GSM300168 (98), GSM300169 (82), GSM300170 (84), GSM300171 (87), GSM300172 (82), GSM300185 (91), GSM300227 (91)*

2. Berchtold2008 PCG: *GSM300212 (74)*

3. Berchtold2008 SFG: *GSM300250 (85)*
4. Kang2011 A1C: *GSM705108 (3)*
5. Kang2011 CBC: *GSM705202 (42), GSM704438 (37)*
6. Kang2011 HIP: *GSM704567 (0), GSM704226 (70), GSM704227 (70), GSM704627 (40)*

Because outlier samples can influence the result of quantile-normalization, the corresponding datasets were preprocessed again after removal of these putative outlier samples.

2.4 Cell Type-Specific Expression Analysis

Cell type-specific expression analysis can be used to understand the relative contribution of different cell types to the transcriptome. Using a published mouse brain cell type-specific transcriptome dataset (Cahoy et al., 2008), I analysed the cell type-specific expression levels in the transcriptome of each sample in the datasets described in **Section 2.1**.

This mouse brain transcriptome dataset contains expression profiles for different cell types purified by using fluorescent-activated cell sorting (FACS). The expression values of 9,293 genes were measured for each cell type using microarrays (data provided by Mehmet Somel). First, for each gene, the mean expression level across the replicates of each main cell type was calculated: astrocytes (A), newly differentiated oligodendrocytes (OL), myelinating oligodendrocytes (M_OL), oligodendrocyte progenitor cells (OPC) and neurons (N). These mean expression levels for each cell type were then used to calculate that cell type's relative contribution to the transcriptome profile of each sample, according to the following linear regression model:

$$E = \alpha + \beta_A E_A + \beta_{OL} E_{OL} + \beta_{M\_OL} E_{M\_OL} + \beta_{OPC} E_{OPC} + \beta_N E_N + \epsilon \quad (2.1)$$

Where;

E : Mean Expression Level
α : Intercept
β : Regression Coefficients
ϵ : error

The regression coefficients were used to represent the relative contribution of the corresponding cell type in a given sample. For more information about the interpretation of the results see **Section 3.1**.

2.5 Age Test

To express the association between gene expression level and age, I used the non-parametric Spearman's rank correlation test. Since age is not expected to be normally distributed, and to remove the influence of possible outliers, non-parametric test was preferred. In this way, a correlation coefficient (ρ) was assigned to each gene. The sign of the correlation coefficient shows the direction of association. If the expression level tends to increase as age increases, the sign is positive. Conversely, if expression level decreases as the age increases, the sign is negative. The magnitude of correlation coefficient indicates the strength of the association. ρ can only change between -1 and 1. A value closer to 1 means there is strong positive monotonic association whereas a value closer to -1 means the association is monotonic and strong but in negative direction.

Spearman's rank correlation can only describe the monotonic association. However, postnatal gene expression trajectories in the brain can frequently attain non-linear forms, where the direction of gene expression change reverses between development and ageing (see **Section 1.4**). To account for this, I analysed development and ageing periods separately. To calculate the Spearman correlation, the "cor.test" function in R was used by specifying the method as "spearman". The function also calculates a p-value showing the significance of association.

2.5.1 Multiple Testing Correction

Analysis of high-throughput data generally requires application of a statistical test repeatedly, which causes the p-values to be unreliable (explained below). In the literature, this situation is referred to as "multiple testing problem".

The problem arises from the fact that p-values are themselves random variables and increasing the number of applied tests causes Type I error, which is the rejection of a true null hypothesis. That means, although a statistical test is applied, some of the reported results are not correct.

There are several approaches to control the error rates of a testing procedure. Using "p.adjust" function in R, I used the "Benjamini-Hochberg" approach, which keeps the false discovery rate (FDR) under a certain level α. I used FDR cutoff as $\alpha = 0.1$, because decreasing this value too much can increase Type II error rate, which is the failure to reject a false null hypothesis.

2.6 Comparison of Age-Related Gene Expression Change across Datasets

In order to compare the age-related gene expression changes in development and ageing, I compared datasets using correlation measure and hierarchical clustering.

2.6.1 Correlation across Datasets

Here, I used ρ-values calculated from the age test for each gene. To assess similarities in age-related gene expression change across datasets, I compared each pair of datasets with respect to these ρ-values, across expressed genes. For these pairwise comparisons, I used the Spearman's rank correlation test. The correlation across a pair of datasets was calculated by taking the pairwise complete set of genes. To visualize the correlations, "corrplot" R library was used (Wei, 2013).

2.6.2 Hierarchical Clustering of Datasets

To cluster the datasets according to age-related gene expression change, I used an agglomerative hierarchical clustering method. To decide on the hierarchy (combination of the clusters), I used the Euclidean distance measure and complete-linkage criteria, which finds similar clusters. This is implemented using the "hclust" function in R.

2.7 Gene Reversal Index

For the quantification and comparison of gene expression reversals between development and ageing (see Introduction) across datasets, I calculated a "reversal index" for each gene. The calculation is as follows:

$$RI = -\rho_{dev}\rho_{age} \tag{2.2}$$

Where;

RI : Reversal Index
ρ_{dev} : Gene expression - age Spearman's rank corr. coef. in development
ρ_{age} : Gene expression - age Spearman's rank corr. coef. in ageing

According to this calculation, if a gene shows reversal, it has a positive; and if it shows a monotonic change with age throughout the lifespan, it has a negative reversal index. The genes that have positive RI are referred as "reversal genes" throughout the text.

2.8 Analysis of Gene Function and Regulation

In order to understand if the gene expression reversals are associated with any specific biological function or regulatory element, GO Biological Process categories (GO BP), micro RNAs (miRNAs), and transcription factors (TFs) were analysed. The analysis of all the categories was done using the same procedure. For simplicity, I will refer each GO BP, miRNA, and TF as a "category".

I tested functional/regulatory enrichment of an ageing-pattern (reversal or no reversal), while keeping the developmental pattern fixed. For example, I compared genes increasing with age in development and decreasing with age in ageing ("up-down" genes) with the ones monotonically increasing throughout the lifespan. Likewise, I compared genes decreasing with age in development and increasing with age in ageing ("down-up" genes) with the ones monotonically decreasing throughout the lifespan. Thus, only the ageing pattern was tested. In this way, I aimed to prevent the developmental expression changes from dominating the enrichment test results. As a result, for each category, 2 odds ratios (OR) were calculated; one for up-down pattern and the other for down-up pattern.

Table 2.3: Contingency Table for Up-Down pattern

	Up-Down	**Monotonic Increase**
Category X	A	B
Other Categories	C	D

Table 2.4: Contingency Table for Down-Up pattern

	Down-Up	**Monotonic Decrease**
Category X	X	Y
Other Categories	Z	W

$$OR_{ud} = \frac{A/B}{C/D} \qquad\qquad OR_{du} = \frac{X/Y}{Z/W} \tag{2.3}$$

$OR_{ud} > 1$ shows that up-down reversal genes are over-represented in Category X; whereas $OR_{ud} < 1$ shows under-representation. The same applies to down-up as well.

There are specific tests for identification of statistically significant over- and under-representation. However, in this study, I was interested in categories behaving the same across datasets, to overcome challenges raised by heterogeneity of ageing. For that reason, my goal was to identify categories that have consistent over-representation of reversal genes, independent of statistical significance of enrich-

ment in any of one dataset. The measure of consistency I choose was having $OR > 1$ or $OR < 1$ in all of the datasets. The statistical significance for consistent behaviour, in turn, was tested using a permutation scheme (see **Section 2.9**).

2.8.1 GO Biological Process Functional Analysis

The Gene Ontology (GO) database (Ashburner et al., 2000) was used for functional analysis. GO is a hierarchical functional annotation database that has three categories: biological process, molecular function and cellular component. I used the biological process category, which groups genes based on biological activities such as developmental process, macromolecule metabolic process, localization. GO database and the associated gene annotation accession were performed using "GO.db" (Carlson, n.d.-a), "AnnotationDbi" (Pages, Carlson, Falcon & Li, n.d.) and "org.Hs.eg.db" (Carlson, n.d.-b) libraries in R environment. Only categories having more than 10 genes in each dataset were analysed. In total, 5,371 GO Biological Process categories with a total of 15,273 genes were analysed.

Summarization of GO BP Analysis Results

As a result of GO BP functional analysis, a list of GO categories that show similar behaviour across datasets was obtained. A possible problem with using GO for functional analysis is that the outcome can be a long and redundant list that is hard to biologically interpret. The problem mainly arises from the fact that there is a hierarchical relationship between GO categories and many genes are found in more than one GO category. I used REVIGO (Supek, Bošnjak, Škunca & Šmuc, 2011) to summarize the GO categories of interest, using a clustering algorithm based on the semantic similarity measure. Semantic similarity in this context means the overlapping genes in categories. Using this pairwise measure, clustering was performed. REVIGO also calculates "uniqueness" and "dispensability" values for each category, which is used for the determination of cluster representatives. For the visualization of clusters, I used the "treemap" R package (Tennekes, 2015). Each square generated by "treemap" corresponds to a representative, where these are further joined into superclusters which are shown with distinct colours. The size of the rectangles represents the uniqueness of categories.

2.8.2 miRNA Regulation Analysis

The “biomaRt” (Durinck et al., 2005; Durinck, Spellman, Birney & Huber, 2009) library in R was used to access Ensembl (Yates et al., 2015) and TarBase (Vlachos et al., 2015) databases, to retrieve miRNA-target gene associations. Only miRNAs having more than 10 gene targets in each dataset were analysed. In total, 1,081 miRNA with 12,857 target genes were analysed.

2.8.3 Transcription Factor Regulation Analysis

For the transcription factor binding site (TFBS) determination:

1. +/- 2000 basepairs of transcription start site for each gene was extracted using Ensembl gene annotations

2. Within these sequences, transcription factor binding sites were predicted using TRANSFAC database and Match algorithm (Kel, 2003)

3. For each TFBS, phastCon scores were calculated using UCSC Genome Browser 17-way vertebrate Conserved Element table (Siepel et al., 2005)

4. Conserved TFBS were defined if 80% or more nucleotides had defined phastCon score and if the average score was 0.6 or more.

Only TFBS having at least 10 gene association were analysed. In total 202 TFBS with 16,217 associated genes were analysed (data courtesy of Xiling Liu).

2.8.4 miRNA and TF Network Construction

The "igraph" package in R was used for the network representation (Csardi & Nepusz, 2006). Interaction information was extracted from the same databases used for functional and regulatory analysis. Only miRNA and TFs having consistent over-representation of reversal genes were represented in the networks. Among their targets, only the ones showing the fixed developmental expression pattern in all datasets are present in the network. For example, if the network was constructed for miRNAs that have consistent over-representation of up-down genes, only

the genes increasing with age in all development datasets are present in the network. The colour of the target genes shows the proportion of datasets in which that gene shows reversal. Having a dark red colour means that the gene shows reversal in all datasets, whereas white means it does not show reversal in any.

2.9 Permutation Test

To test the significance of finding categories (GO BP categories, miRNAs, and TFs) behaving in the same manner across datasets, based on the procedure defined in **Section 2.8**, I used a permutation scheme. The aim of this test is to assess the probability of finding the same or higher number of categories having consistent over-representation as the observed number, and to estimate the false discovery rate.

For each dataset, the ages of the samples were randomly mixed using the R "sample" function, age test was applied, and the reversal index for each gene was calculated. This was only applied to the ageing parts (post 20 years) of the datasets; for the age-related gene expression change in development, the original values were used. The reason is that the aim of study is to analyse changes in ageing, in the context of development. Thus, the pattern in development was accepted as fixed and only the pattern in ageing was randomized. I can thus find the categories that have under- or over-representation of reversal genes independent of ageing, *i.e.* by chance.

After calculating 1000 such permuted reversal indices for each gene in each dataset, I followed the procedure described above by calculating the odds ratio for categories, as I did for the real data. Based on the permutation scheme, for each permutation, the number of categories showing $OR > 1$ or $OR < 1$ was recorded. The average (median) number of observations gives the false discovery rate and the number of cases having the same or higher number of observations divided by the total number of permutations gives the p-value showing the significance of the result.

CHAPTER 3

RESULTS

3.1 Cell Type-Specific Expression

The datasets used in this study includes whole tissue samples, *i.e.* not cell type-specific. Since brain is a heterogeneous tissue with different cell types, it is important to know if the samples have comparable proportions of cell types. To analyse the relative contribution of different cell types to the transcriptome of the samples, I applied a deconvolution approach using published cell type-specific gene expression data (Cahoy et al., 2008). **Figure 3.1** shows the proportion of transcriptome that can be explained by each cell type. In all datasets, the expression profile of the all samples mostly reflects the neuronal expression. Moreover, neuron-specific expression decreases with age in all datasets, but the expression profile continues to be dominated by neuron-specific expression even in old ages. Oligodendrocyte progenitor cell-specific expression is seen only at the beginning of the life and then disappears. Myelinating oligodendrocyte-specific expression increases with age. Similarly, astrocyte-specific expression, although is stable until late ages, starts to increase after 60s.

Having a cell type-specific expression profile that is comparable across datasets imply that technical issues (e.g. different proportions of grey vs. white matter sampled in different datasets), do not direct the expression profile.

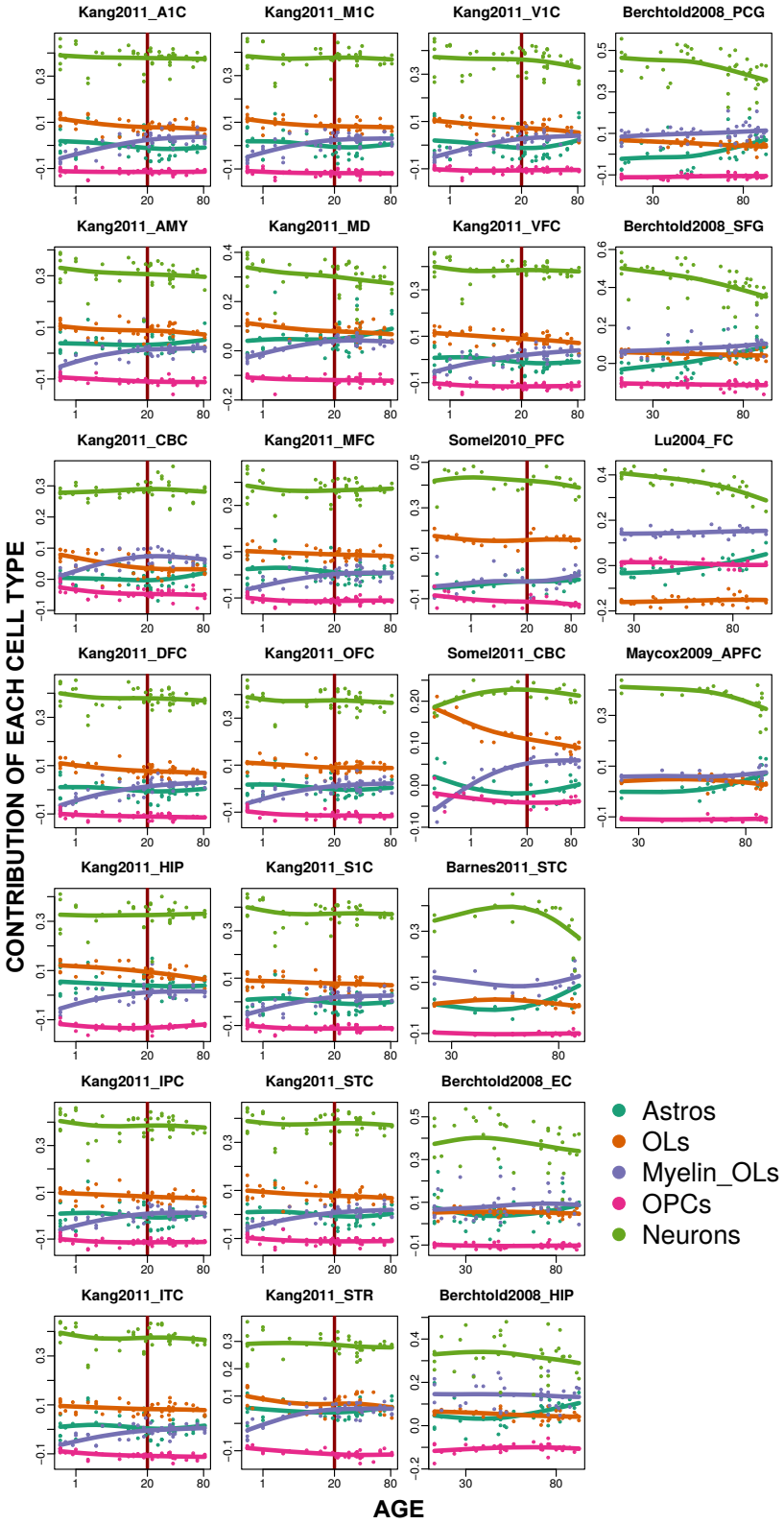

Figure 3.1: Cell Type-Specific Expression Profile Change with Age

It is important to note that, the only information used for this analysis is the gene expression levels. Although the change in coefficients of cell type profiles might be a consequence of the change in the cell type proportions, it may also reflect changes in gene expression *within* a cell type. For example, decrease in neuronal expression that we see in late ages might be an indicator of neuronal loss, as well as loss of expression regulation resulting in loss of cellular identity.

3.2 Age-Related Gene Expression Change across Datasets

As the next step, age-related gene expression changes were compared across datasets to provide a direct estimate of consistency of the changes among datasets. I used the ρ value calculated for each gene through the Age Test (see **Section 2.5**), to compare the gene expression changes. The ρ value gives information about the strength and the direction of age-related expression change. Since ρ is a measure of monotonic change and one may expect differences between development (<20 years of age) and ageing (>20 years of age), I investigated these periods separately. To compare the datasets, Spearman's rank correlation test was used (**Figure 3.2**).

One prominent result from this comparison is that correlation among datasets in the developmental period is stronger than that in ageing ($Median\rho_{dev} = 0.52$, $Median\rho_{age} = 0.30$, Mann Whitney U Test $p < 0.001$).

Furthermore, ageing datasets cluster in two distinct groups, not reflecting brain regions (an expected source of biological variation), but the laboratory of origin; *i.e.* the Kang2011 brain regions appear all similar to each other with respect to ageing-related expression changes, whereas the others resemble each other. Such a stark pattern is not observed for developmental expression changes, which tend to group according to source brain region (although data from only two laboratories were compared in this case).

Hierarchical clustering of brain regions based on ρ values representing age-related gene expression change (**Figure 3.3**) shows that gene expression change can differentiate ageing and development periods. Moreover, in development, similar brain regions tend to be clustered together (note the colour code).

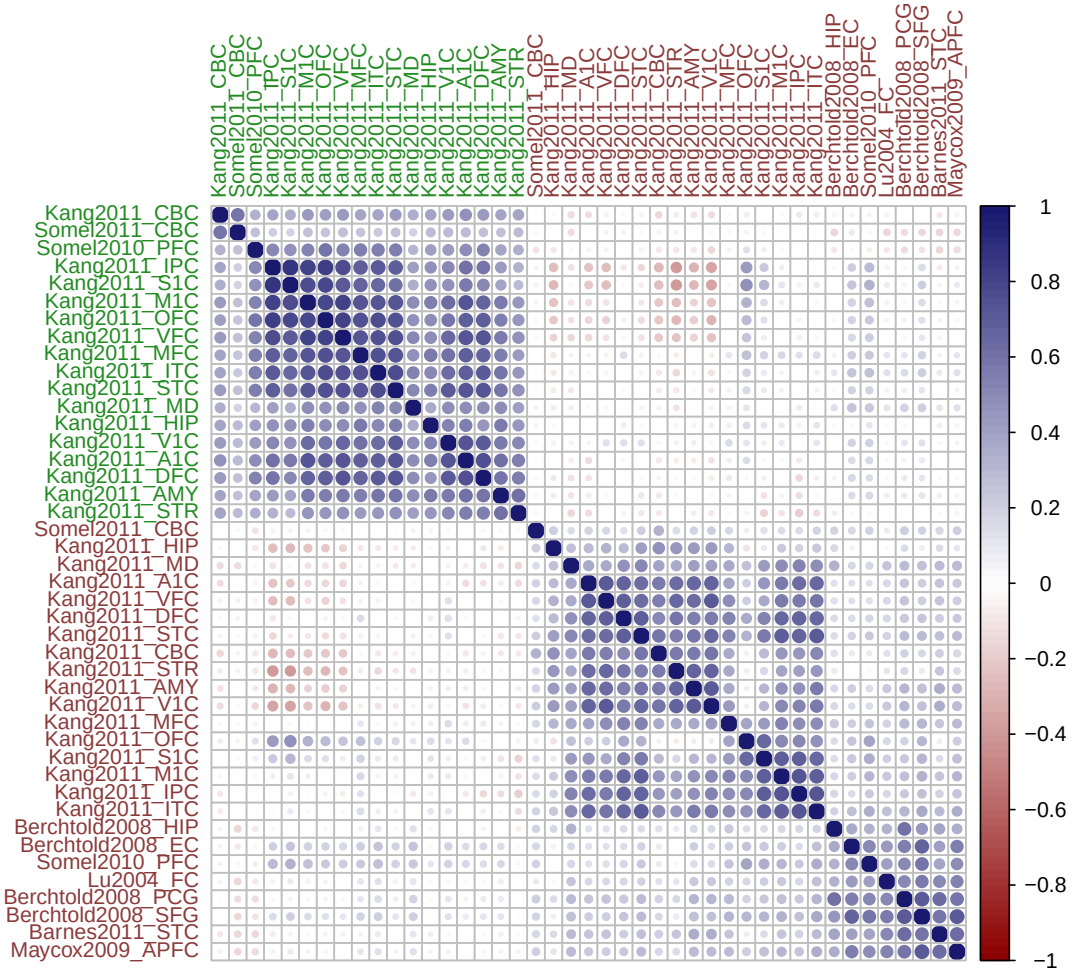

Figure 3.2: Correlation plot for age-related gene expression changes between all brain regions in development and ageing. The size and the colour of the squares change with the magnitude of the correlation coefficient between different regions as the large blue squares show strong positive correlation, and the large red squares show strong negative correlation. Row and column labels show the type of samples - green: development and red: ageing. Brain regions are ordered by hierarchical clustering of correlation coefficients between datasets.

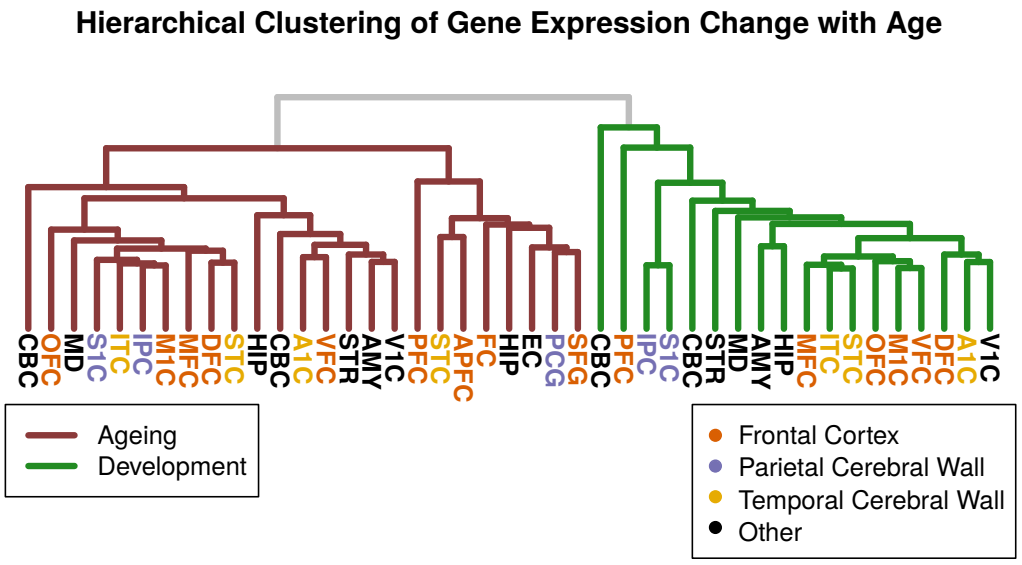

Figure 3.3: Hierarchical clustering of brain regions in development and ageing based on age-related gene expression change

The cell type-specific expression analysis (**Section 3.1**) and the hierarchical clustering of expression changes in development (**Figure 3.3**), both suggest that there is no prominent technical bias among datasets (such as some datasets behaving very differently from others with respect to age-related changes).

At the same time, the grouping of ageing datasets in the correlation plot based on laboratory of origin, and showing no clustering according to brain region might have multiple explanations. First, this might be a further indication that unlike development, ageing is not a programmed process. Specifically, age-related expression changes in ageing may tend to be more random than development, with more inter-individual heterogeneity. This would decrease similarity among datasets. Second, expression changes in ageing are also expected to be weaker in magnitude than those in development (see Introduction), so that technical noise becomes more prominent when measuring ageing-related expression changes, than developmental changes.

3.3 Strength of Age-Related Gene Expression Change in Development and in Ageing

Seeing that gene expression change in development is more homogeneous and reproducible, an interesting question is whether expression change in development is more pronounced than that in ageing. Using only the datasets which have samples representing both postnatal development and ageing (Kang et al., 2011; Somel et al., 2010, 2011), I investigated if the gene expression change in development, measured by the Spearman correlation coefficient, is stronger than in ageing (i.e. how much of the variance in expression is explained by age).

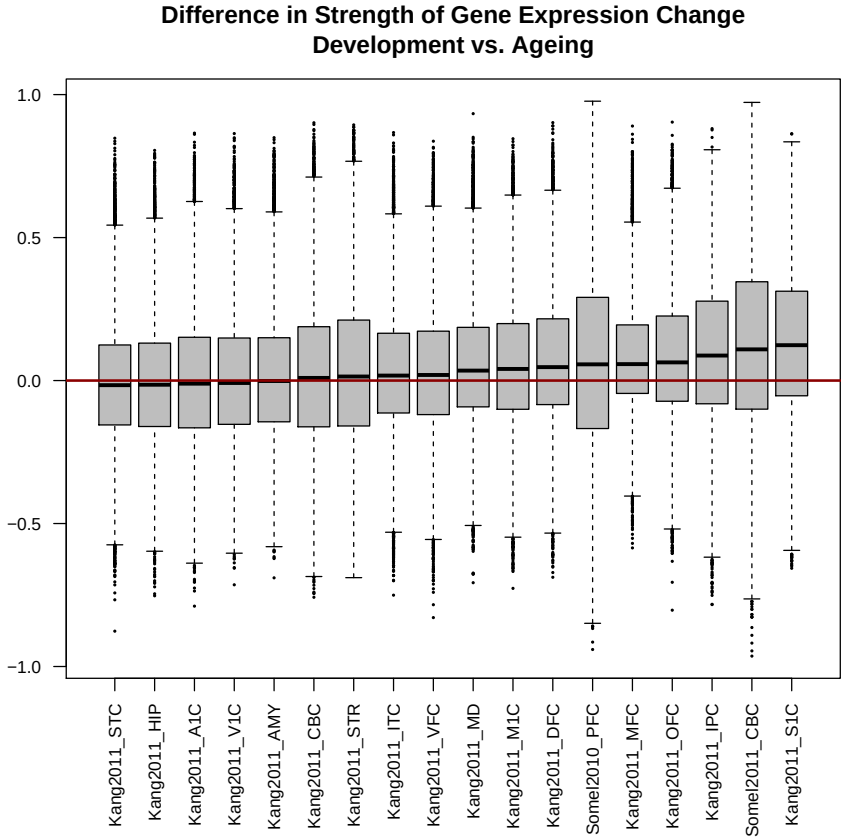

Figure 3.4: Magnitude of age-related gene expression change in development vs. ageing. Each boxplot shows the distribution of $|\rho_{dev}| - |\rho_{age}|$ for genes in a given dataset. Red line shows 0 which means no difference. Datasets are ordered according to their median $|\rho_{dev}| - |\rho_{age}|$ differences.

In order to get an overview of age-related gene expression change, I compared the the gene expression change strength (absolute value of Spearman's rank correlation coefficients) in development and ageing. **Figure 3.4** shows the distribution of the differences between expression change in development and ageing for each gene. Since the calculation was done as $|\rho_{dev}| - |\rho_{age}|$, a positive value shows that change in development is stronger than in ageing, whereas a negative value shows weaker change in development. For most of the datasets, the median difference is positive indicating that the gene expression change in development is more prominent than in ageing. Indeed, paired Wilcoxon signed rank test showed that 14 out of 18 datasets show more prominent changes in development (Benjamini-Hochberg FDR $q < 0.05$ for each dataset). Furthermore, assuming that the datasets are independent, observing this in 14 out of 18 datasets is not very likely to be due to chance (Binomial Test, $p = 0.03$).

3.4 Direction of Age-Related Gene Expression Change

Here, whether the gene expression change favours one specific direction (up or down) was tested in development and ageing datasets. For each dataset, the proportion of genes showing significant age-related change after multiple test correction (FDR q<0.1) was calculated (**Figure 3.5**). In development, all datasets show at least some degree of significant change. However, ageing is more heterogeneous, with a much lower degree of significant change. There are some datasets having only a few genes showing significant change and some having substantial number of genes showing significant change in ageing. By comparing the number of genes that show significant age-related change in either direction, I tested for general trends in development and ageing. I found that the direction of age-related gene expression change in development is generally in the positive direction (paired Wilcoxon signed rank test, $p = 0.054$), whereas in ageing expression change direction favours the negative direction (paired Wilcoxon signed rank test, $p = 0.001$).

Given the difference between the direction of gene expression change in development and ageing, a possible question to ask is whether these are linked to gene expression reversals previously mentioned in a limited number of ageing studies, and if yes, whether these bear any functional properties.

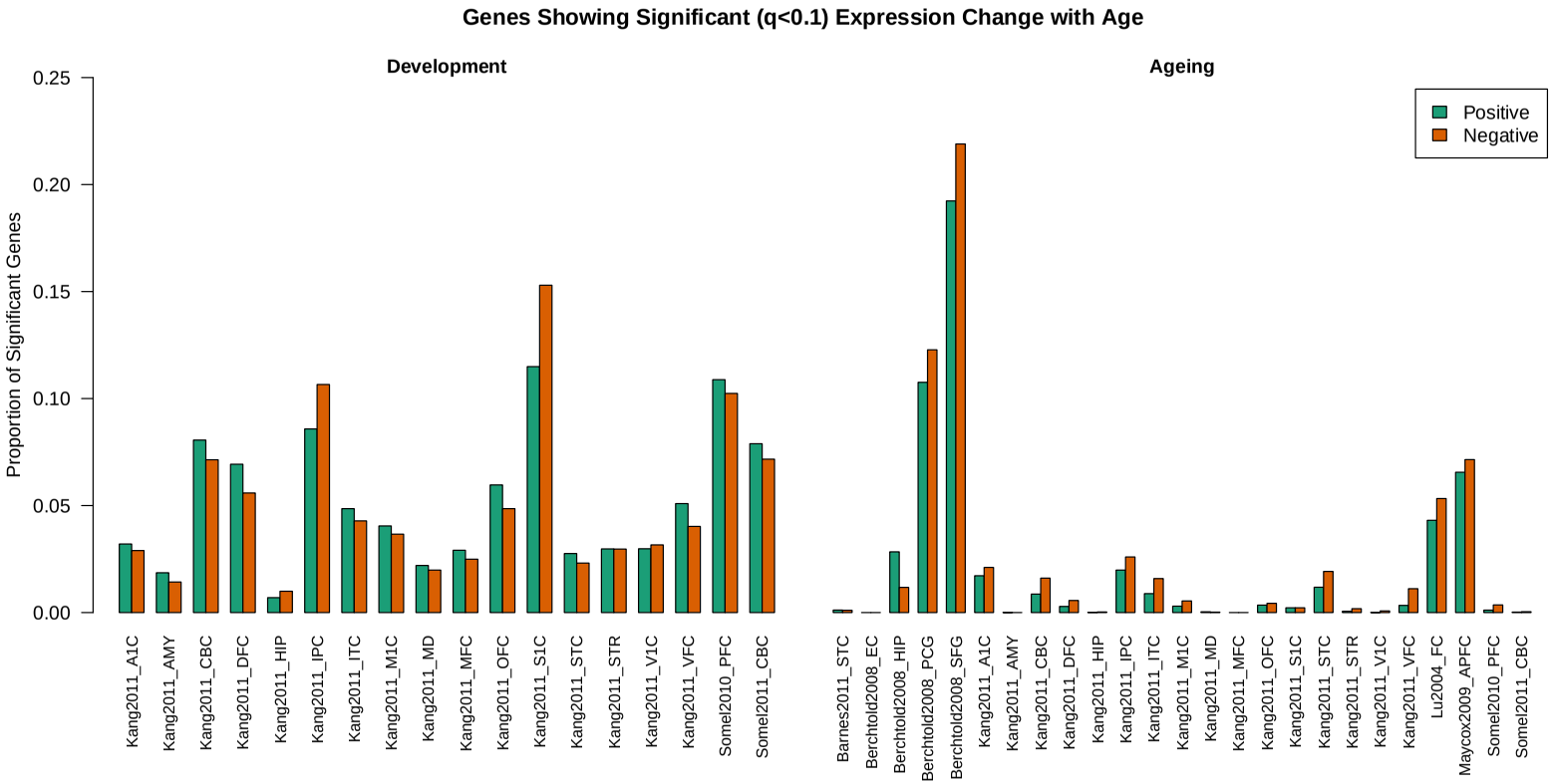

Figure 3.5: Direction of age-related gene expression change across all brain regions

3.5 Gene Expression Reversals

For the analysis of gene expression reversals, only three datasets (Kang et al., 2011; Somel et al., 2010, 2011), consisting in total of 17 brain regions were analysed, since these are the only datasets having samples representing both postnatal development and ageing periods.

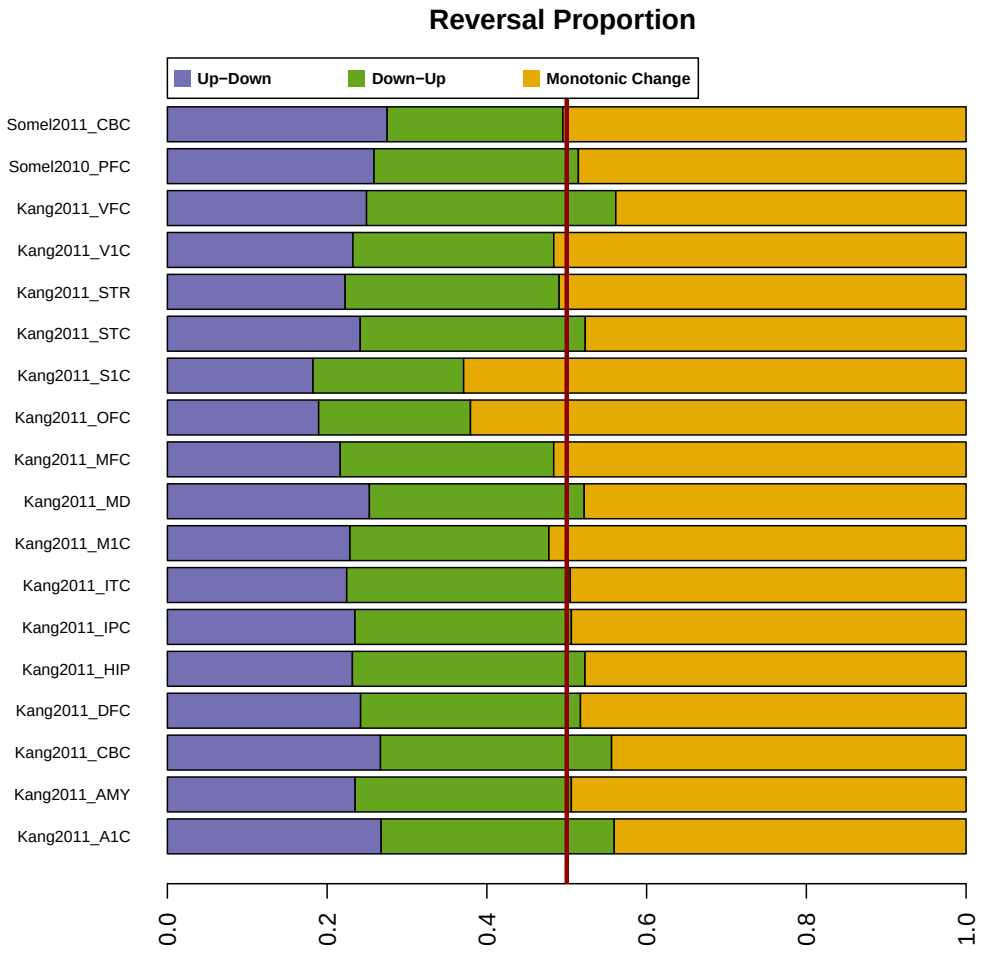

Figure 3.6: Proportion of reversal genes

Here the only criteria I used to consider a gene as a "reversal gene" is that it has a positive reversal index (see **Section 2.7**). The proportion of reversal genes varies between 37% (S1C - Kang2011) and 56% (VFC – Kang2011), with a median of 51%. Among all brain regions, S1C and OFC seem to have smaller reversal proportions (**Figure 3.6**).

For the rest of the analysis, genes are classified according to their expression patterns in development and ageing, *i.e.* decreasing or increasing. According to this, all reversal genes are either “down-up” or “up-down” genes, where the first direction indicates the direction in development and the second one shows the direction in ageing. In general, all of the brain regions show comparable amount of down-up and up-down genes. The non-reversal genes are accordingly named as monotonically increasing or monotonically decreasing genes.

The number of common reversal genes across all 18 datasets is only 72 in 13,686 commonly expressed genes. This overlap is found to be statistically significant assuming independence of datasets (Permutation test $p<0.001$, $n_{expected}=0$), but is admittedly minute. One interpretation could be that expression reversal in the brain is not a biologically reproducible and relevant phenomenon. An alternative interpretation could be related to the highly stochastic nature of ageing, in accordance with previous discussion in the literature (Kirkwood, 2005) and as suggested by the earlier results of this study. Specifically, if the reasons of the gene expression reversals are partly stochastic somatic mutations, epimutations, or random environmental effects, if there is high inter-individual heterogeneity in expression changes (and on top of these, in the face of technical noise), not seeing exactly the same reversal pattern for the same genes across all 18 datasets may not be surprising.

Under this second interpretation, one may still, assume that reversal genes identified in each dataset can be functionally important for the ageing process, and investigation of their functional roles and regulation could provide a new perspective for brain ageing. If so, we may expect that not necessarily the same genes, but the same functional processes could show reversal patterns consistent across 18 datasets.

3.6 Functional Analysis of Gene Expression Reversals

In order to analyse the biological relevance of gene expression reversals, I performed functional analysis through the analysis of biological processes and trans-regulators (miRNAs and TFs).

3.6.1 GO Biological Process Enrichment Analysis

To study the biological functions that are associated with the reversal pattern, I used the Biological Process category of GO Database. I compared the functions of the reversal genes with the functions of the genes showing monotonic change with age, with the developmental pattern kept fixed. That means the up-down genes were compared with the genes showing monotonic increase and down-up genes were compared with the genes showing monotonic decrease. Thus, the results are not dominated by the functional properties of developmental expression change. Using this approach, I assigned to each functional process a measure of propensity for being enriched in reversal genes.

Instead of searching for particular functional processes within each dataset, the processes associated with reversal pattern consistently across all 18 datasets were investigated. In this way, technical noise and the noise due to stochastic nature of ageing was overcome. The results were tested using a permutation scheme (see **Section 2.9**).

Down-Up Pattern vs. Monotonic Decrease

There are 16 GO Biological Process categories having consistent over-representation of down-up genes when compared to the monotonically decreasing ones ($n_{expected}$= 0, permutation test $p < 0.001$). **Figure 3.7** shows the result summarized by REVIGO (see **Section 2.8.1**). The full list of GO groups is given in the **Appendix A**.

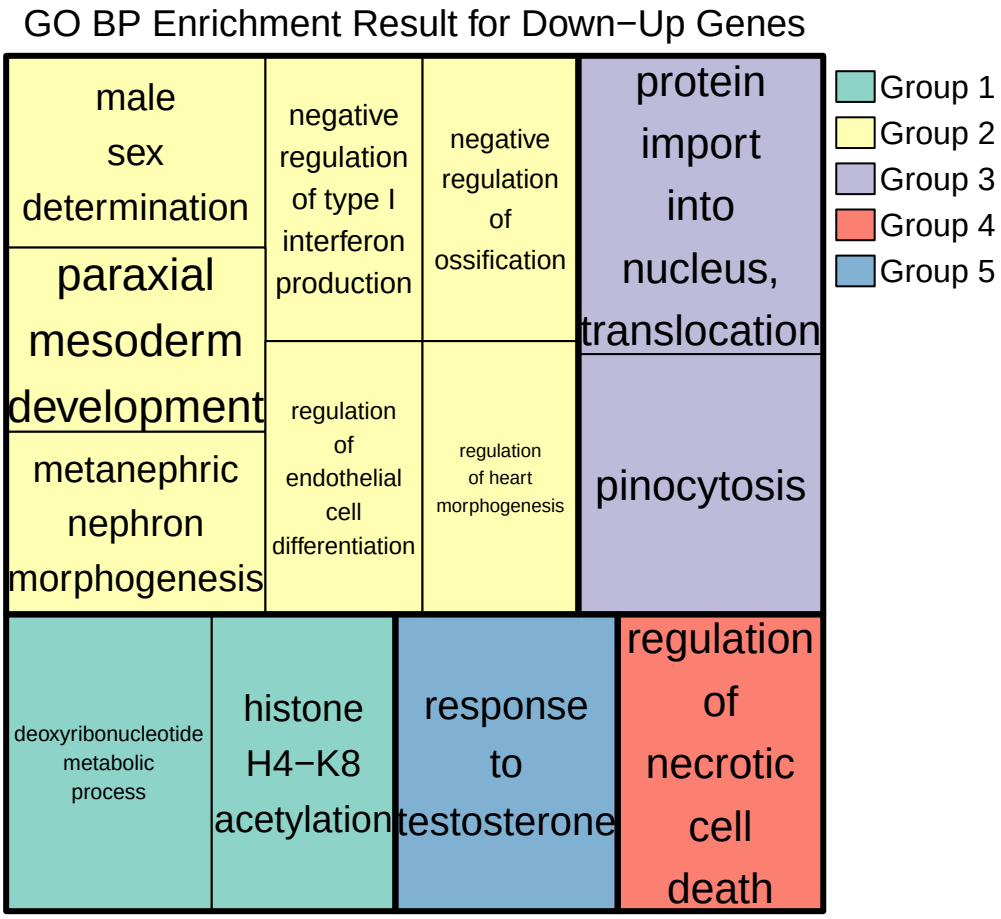

Figure 3.7: GO Biological Process enrichment result for down-up genes, summarized by REVIGO

REVIGO grouped all categories under 5 superclusters:

- **Group 1:** This category includes deoxyribonucleotide metabolism and histone acetylation. Expression change in deoxyribonucleotide metabolism genes can be a sign of change in DNA stability as well as change in cellular energetics, as dNTP metabolism-related genes are also in this category. Although the treemap (**Figure 3.7**) only shows one type of histone acetylation (H4K8), the full list of categories include H4K5 acetylation. The reason why only one of them is shown is that most of the genes are common between the

two categories. So, the result should not be interpreted as down-up pattern of H4K8 acetylation particularly, but histone acetylation in general. Histone acetylation is generally associated with open chromatin and increases in gene expression (Eberharter & Becker, 2002). It is possible that genes in other categories showing the down-up pattern are regulated by this epigenetic process.

- **Group 2:** Notably, in this group there are several development and differentiation-related categories. Some of these categories does not seem to be related to the brain cells. However, it is possible that the regulators of these differentiation pathways are the same. As a result of the presence of these common genes, these categories might be found to show consistent over-representation in the brain expression profile. Assuming that the up-regulation of diverse differentiation-related genes is related to loss of cellular identity, these results might indicate a possible link between gene expression reversal and loss of differentiation.

- **Group 3:** Pinocytosis -or "cell drinking"- also shows consistent over-representation of down-up genes. It may have functions in intercellular communication, growth factor and neurotransmitter uptake. Protein import into nucleus category contains the transporters which carry proteins into the nucleus. Regulation of this group is especially important for the signal transduction events which generally alter gene expression.

- **Group 4:** The only category in this group is the regulation of necrotic cell death. Necrotic cell death is characterised by increase in cell volume, organelle swelling, membrane rupture and loss of intracellular contents. Many of the genes under this category are cytokines, which may also have role in senescence associated secretory phenotype.

- **Group 5:** Response to testosterone category includes all genes that have role in change in state of cell upon testosterone stimulus. Among the genes in this category, there are protein modifiers including kinases and acetylases as well as transcription factors. These proteins are all involved in pathways induced by testosterone. Testosterone level is known to decrease in ageing (Rajfer, 2003), and the increase in expression of response genes might be related to complex feedback loops in signalling pathways.

While studying these GO groups having more than expected amount of reversal genes is interesting, GO groups that are depleted in reversal genes in all datasets can also provide valuable information. In total there are 52 such GO Biological Process categories listed in **Appendix B** ($n_{expected}$= 0, permutation test $p < 0.001$). The summary is given in **Figure 3.8**.

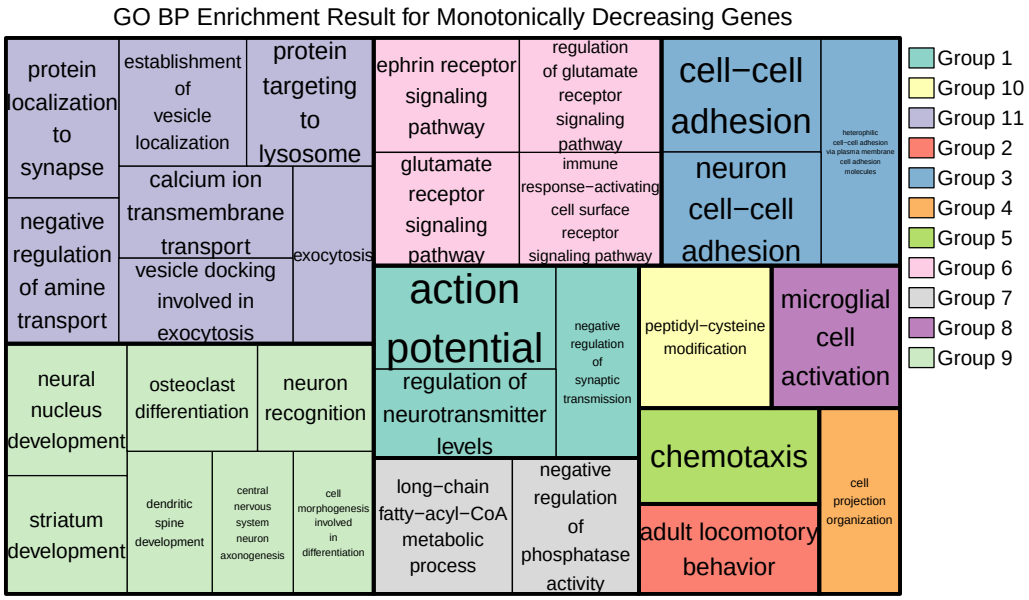

Figure 3.8: GO Biological Process enrichment result for monotonically decreasing genes, summarized by REVIGO

The summary gives 11 main clusters including nervous system development, intercellular communication, macromolecule modification, cellular signalling, neuronal functions, cell projection, transport, microglia cell activation categories. Since these show monotonic change throughout the lifespan and they decrease continuously, these are continuation of changes in early postnatal years. However, this does not necessarily mean that these are continuation of development. Instead it might be possible that for some of these groups, ageing-related disruptive changes has started in early years.

Up-Down Pattern vs. Monotonic Increase

Figure 3.9 shows the summarized list of functional categories which have consistent over-representation of up-down genes in comparison to monotonically increas-

ing genes. The full list contains 114 GO BP categories and is given in **Appendix C** ($n_{expected}$= 0, permutation test $p < 0.001$).

REVIGO grouped all 114 categories under 9 superclusters:

- **Group 1:** This group is mostly related to CNS functions and development. It includes startle response, which is related to sudden movement upon unexpected stimulus. The categories dendrite extension, neuron recognition, cellular component morphogenesis, and cell development are related to cellular differentiation and development, crucial for neuronal function. Similarly, cerebellar cortex development, which involves the cell differentiation, formation of a mature structure with layers of different neuronal cells, shows up-down pattern. Cerebellar cortex is important for cognition, memory and motor function, which are also enriched in up-down genes.

- **Group 2:** The only category in this group is locomotory behaviour, which includes motor functions such as walking in multicellular organisms.

- **Group 3:** This group includes cellular localization/transport functions. The up-down pattern in these functions, including protein, macromolecule and ionic transport, might be further indication of loss of cellular homoeostasis. Endocytic recycling, which is important for membrane composition, also shows up-down pattern.

- **Group 4:** Membrane docking, which is an ancestor term of synaptic vesicle docking, is an important category for cellular functions as well as synaptic transmission. It is also involved in intercellular communication.

- **Group 5:** Neuron cell-cell adhesion involves many proteins important for synapse formation and synaptic transmission.

- **Group 6:** This group is mainly composed of cellular signalling pathways and synaptic transmission-related functions. Since these are grouped together, the signalling pathways are also most probably induced by synaptic transmission. The fact that we see increase in development and decrease in ageing could be the reflection of decreased synaptic activity in ageing.

- **Group 7:** Phosphatidylinositol metabolism, which is important for membrane trafficking and lipid signalling, shows up-down pattern. Sphingolipids,

which have role in maintaining integrity of cellular components, show a similar trend.

- **Group 8:** This supercluster contains macromolecule modification-related functions including protein ubiquitination and dephosphorylation. One of the hallmarks of ageing is loss of proteostasis (see **Section 1.1.2**), and change in genes important for protein function can be important for this phenomenon. The same supercluster also contains maturation of small subunit rRNA, which might be an indication of decreased translation in ageing, reported earlier (Hands, Proud & Wyttenbach, 2009). Synapse organization, which is also found in this supercluster, is important for development, maturation and maintenance of synapses. Similarly, cell projection organization, which involves axonogenesis, also shows enrichment in up-down genes.

- **Group 9:** This group mainly involves response pathways including response to ions, glucagon, FGF stimulus and unfolded proteins. Ageing-related decrease in the activity of ionic response (lead and ammonium) and cocaine response might be linked to decrease in toxic defence. Similarly, unfolded protein response, which can again be an indicator of loss of proteostasis, is related with how cell compensates the accumulation of non-functional proteins, which are toxic to cell. Another group involved in FGF stimulus is important in cell proliferation.

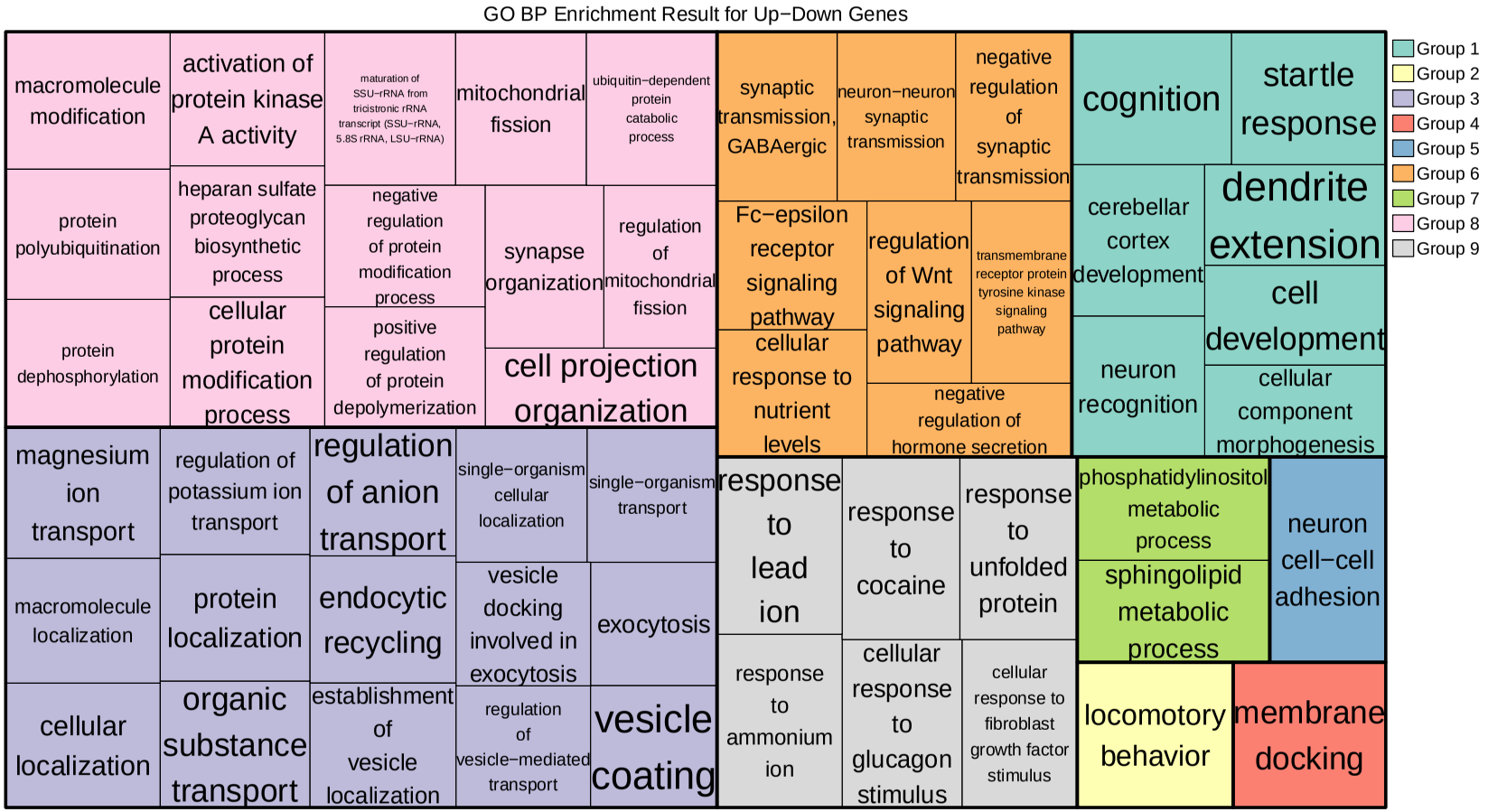

Figure 3.9: GO Biological Process enrichment result for up-down genes, summarized by REVIGO

Number of groups depleted in up-down genes and showing mostly monotonic increase is 35 ($n_{expected}$= = 0, permutation test $p < 0.001$). These groups are summarized by REVIGO in **Figure 3.10** and the full list is given in **Appendix D**.

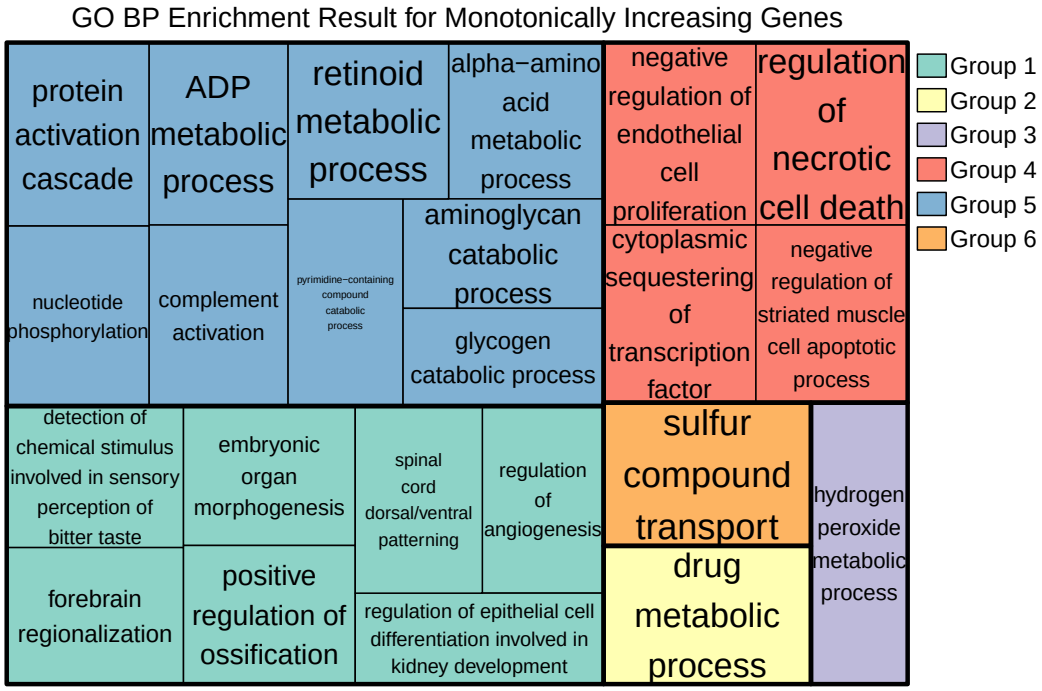

Figure 3.10: GO Biological Process enrichment result for monotonically increasing genes, summarized by REVIGO

The summary yielded 6 main clusters including metabolic processes -mostly catabolic functions, cell cycle and cell death, development, transport, drug metabolism and hydrogen peroxide metabolism.

3.7 Regulation of Gene Expression Reversals

The above analysis showed that gene expression reversals are consistently associated with several biological functions. This brings into question the regulation of these expression changes. Here I analysed trans-regulators of these genes, as the investigation of cis-regulation would require extra data such as DNA sequence and chromatin structure of the same samples, which is not available.

3.7.1 miRNA Enrichment Analysis

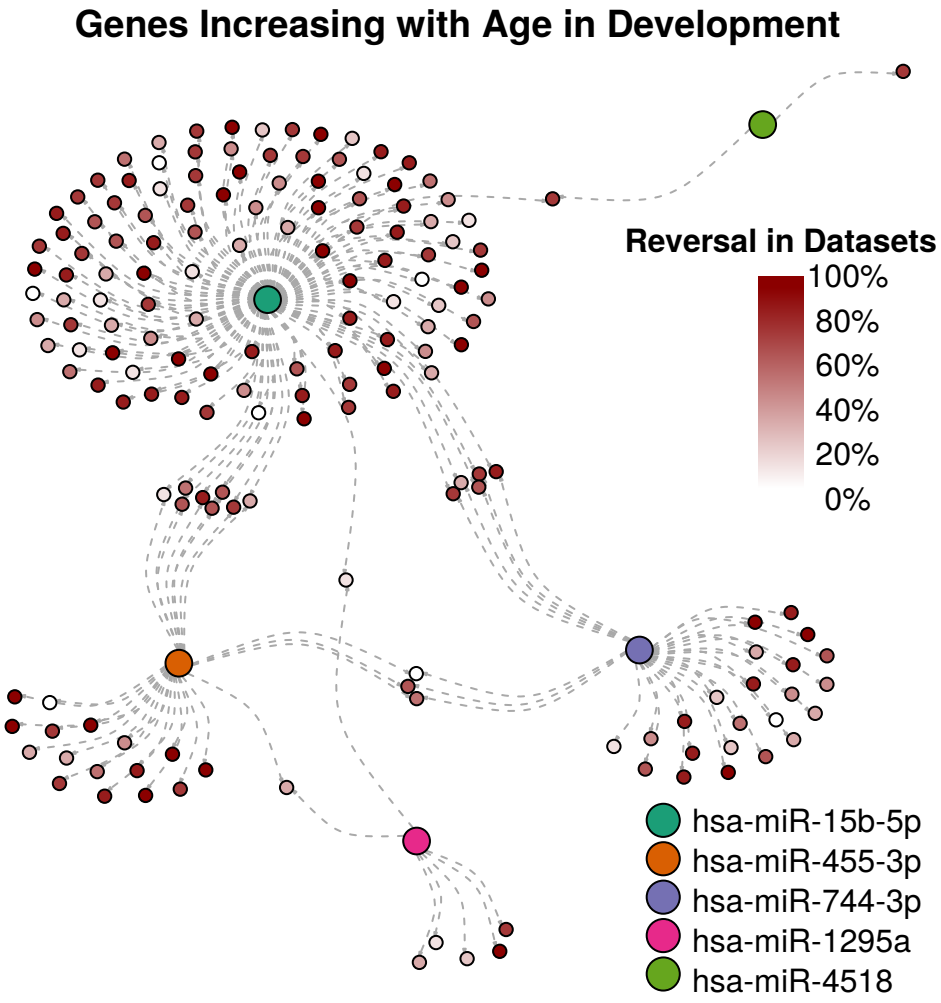

Figure 3.11: Network representation of up-down gene regulation by miRNAs

It is known that miRNAs have important role in development (Rajfer, 2003). However, whether there are miRNAs targeting reversal genes is an open question. In order to find this, for each miRNA, I asked whether its putative target genes have more than expected amount of reversal genes, compared to putative targets of all other miRNAs. As in the GO Enrichment Analysis, reversal genes are compared with the genes with monotonic expression change, keeping the developmental pattern fixed.

None of the miRNAs show consistent over-representation of down-up pattern. This suggests that this pattern is not mainly regulated by miRNA activity.

Up-down genes on the other hand, show consistent over-representation in 5 miRNA targets ($n_{expected}$= 0, permutation test $p < 0.001$). These are summarized in **Figure 3.11**. miRNAs are represented by larger nodes. Smaller nodes are the target genes. The network only shows the genes which show increase in all development datasets. Colour of the targets change between white and dark red, according to the number of datasets in which that gene shows reversal. The darker the colour, the more consistent the reversal pattern.

If we look at the miRNAs more closely:

1. **hsa-miR-15b-5p**: This miRNA has 1391 putative target genes overlapping with the analysed datasets. On average, 731 of them ($\sim$53%) increase with age in development and 416 of these ($\sim$57%) decrease with age in ageing. Circulating hsa-miR-15b-5p is suggested as biomarker for Alzheimer's disease (Kumar et al., 2013), bipolar disorder (R. M. Walker et al., 2015), and several cancer types, including glioma (Pang et al., 2015; G. Sun et al., 2014). Having been linked to glioma and Alzheimer's, hsa-miR-15b-5p might be important for the age-related changes in brain. This miRNA also has the largest number of targets among the other miRNAs having consistently over-represented reversal genes.

2. **hsa-miR-455-3p**: This miRNA has 270 putative target genes overlapping with the analysed datasets. On average 133 of them ($\sim$49%) increase with age in development and 76 of these ($\sim$57%) decrease with age in ageing. Hypoxia signalling, some cancer and differentiation pathways are associated with this miRNA (Lalevée, Lapaire & Bühler, 2014; Ujifuku et al., 2010). Some of the targets are in common with hsa-miR-15b-5p and most of the targets show reversal consistently.

3. **hsa-miR-744-3p**: This miRNA has 257 putative target genes overlapping with the analysed datasets. On average 133 of them ($\sim$52%) increase with age in development and 79 of these ($\sim$59%) decrease with age in ageing. This miRNA is suggested to inhibit DNA repair mechanisms and have role in prostate cancer(Hatano et al., 2015). Also, by controlling the post-translational

modifiers of TGF-β1, have role in inflammation, cancer and fibrosis (J. Martin et al., 2011). This miRNA has common targets with both hsa-miR-15b-5p and hsa-miR-455-3p.

4. **hsa-miR-1295a**: This miRNA has 72 putative target genes overlapping with the analysed datasets. On average 36 of them ($\sim$51%) increase with age in development and 22 of these ($\sim$61%) decrease with age in ageing. There is not much study on this miRNA, thus there is no specific function associated with it.

5. **hsa-miR-4518**: This miRNA has 24 putative target genes overlapping with the analysed datasets. On average 13 of them ($\sim$53%) increase with age in development and 9 of these ($\sim$72%) decrease with age in ageing. Similar to hsa-miR-1295-a there is not much study on this miRNA.

There are 3 consistently under-represented miRNAs for up-down genes, meaning these miRNAs are associated with monotonically increasing genes more than expected. The number of miRNAs showing consistent under-representation for down-up genes is 2. One of the miRNAs associated with the monotonically decreasing genes is hsa-miR-455-3p, which also shows consistent over-representation of up-down genes.

3.7.2 Transcription Factor Enrichment Analysis

Similar to miRNA analysis, the aim here is to find if any transcription factor is related with the reversal pattern. However, transcription factor-target gene interactions are generally not as well defined as for miRNAs. Hence, I used a transcription factor binding site (TFBS) prediction approach, and further find transcription factors associated with TFBSs from MSigDB (Subramanian et al., 2005).

Similar to miRNA analysis result, there is no TFBS that shows consistent over-representation of down-up pattern genes. However, 9 TFBS show consistent over-representation among up-down genes. These are summarized in **Figure 3.12**. Similar to the miRNA network, larger nodes represent the TFBS and smaller ones represent the target genes. The colour gets darker as the number of datasets showing reversal increases.

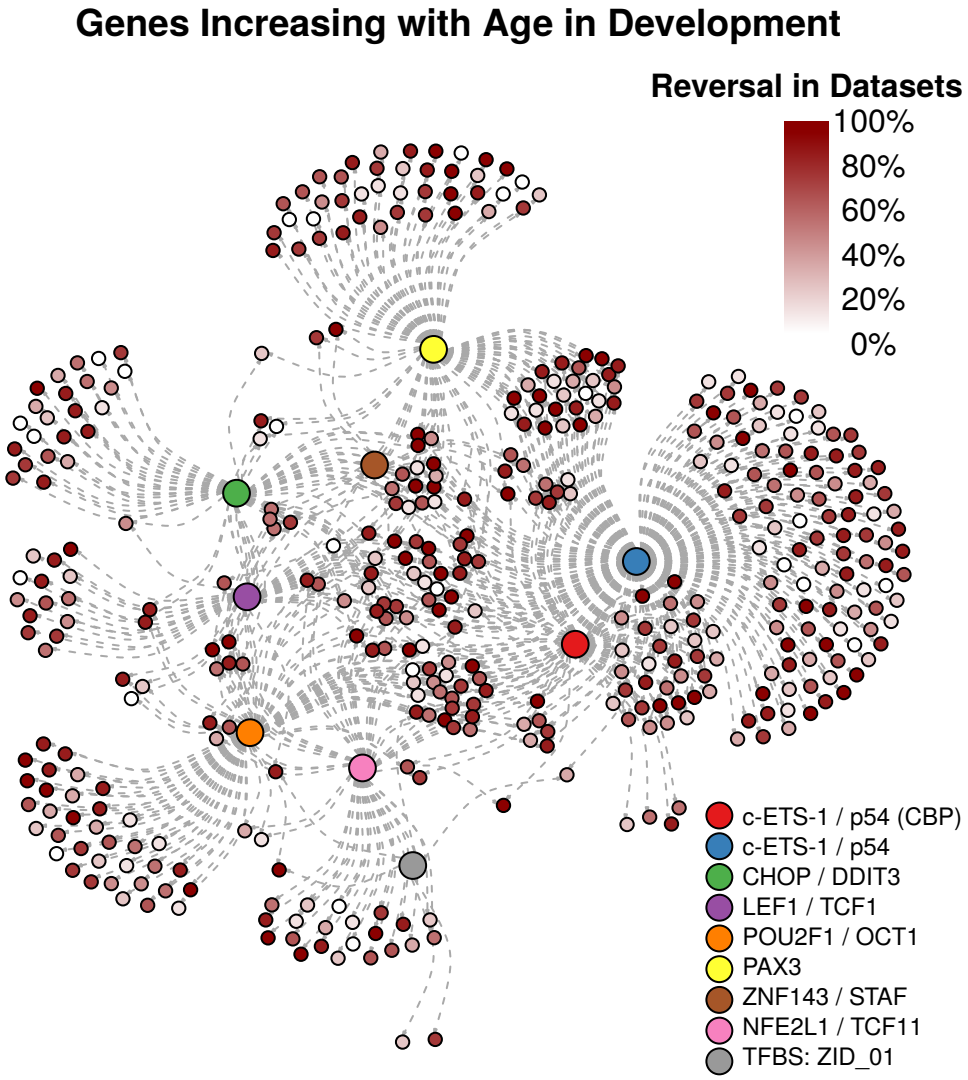

Figure 3.12: Network representation of up-down gene regulation by TFs

The list of transcription factors (TFs) associated with these 9 TFBSs are given below. Unless specified, the functional annotations for the TFsare retrieved from the UniProt Database (The UniProt Consortium, 2014).

1. **c-ETS-1 / p54**: This TF is associated with two enriched TFBS. One of them also requires core binding protein (CBP) to bind. The TFBS is found in the promoter regions of 1134 genes overlapping with the analysed datasets, and 445 of them (∼39%) increase with age in development and 237 of these (∼53%) decrease with age in ageing. The binding site that does not require CBP binding has 4028 target genes and 1620 of them (∼40%) increase with

age in development and 847 of these (∼52%) decrease with age in ageing. c-ETS-1 can act as both activator and repressor of gene expression. It controls cytokine and chemokine genes. This TF has been linked to stem cell development, cell senescence and tumorigenesis. Also, it was found that it controls migration and invasion in endothelial cells, which is related with angiogenesis (Yordy et al., 2004).

2. **CHOP / DDIT3**: The TFBS is found in the promoter regions of 1185 genes overlapping with the analysed datasets and 466 of them (∼39%) increase with age in development and 253 of these (∼54%) decrease with age in ageing. CHOP is activated upon ER stress and induces cell cycle arrest and apoptosis. Also, activation of several caspases by CHOP also has effect on inflammatory response.

3. **LEF1 / TCF1**:The TFBS is found in the promoter regions of 991 genes overlapping with the analysed datasets and 365 of them (∼37%) increase with age in development and 196 of these (∼54%) decrease with age in ageing. LEF1 is an activator. It is associated with GPCR signalling and cancer pathways. It is important in Wnt signalling pathway. There are alternatively spliced forms of this TF. One of these isoforms regulates e-cadherin expression and causes cell migration. Also, it causes cell proliferation in pancreatic tumours.

4. **POU2F1 / OCT1**:The TFBS is found in the promoter regions of 1830 genes overlapping with the analysed datasets and 697 of them (∼38%) increase with age in development and 372 of these (∼53%) decrease with age in ageing. It is the first identified TF in POU family. It is linked to primary amebic meningoencephalitis and herpes simplex. It regulates expression of several snRNAs, histone H2B, and immunoglobulins.

5. **PAX3**:The TFBS is found in the promoter regions of 1988 genes overlapping with the analysed datasets and 764 of them (∼38%) increase with age in development and 411 of these (∼54%) decrease with age in ageing. This TF is important in neural crest differentiation and chromatin organization. It is associated with cell proliferation, migration and apoptosis. It functions in neural development and myogenesis.

6. **ZNF143 / STAF**:The TFBS is found in the promoter regions of 211 genes overlapping with the analysed datasets and 87 of them (∼41%) increase with

age in development and 49 of these (∼57%) decrease with age in ageing. It is a transcriptional activator. It activates selenocysteine tRNAs. Also has role in snRNA activation.

7. **NFE2L1 / TCF11**:The TFBS is found in the promoter regions of 900 genes overlapping with the analysed datasets and 353 of them (∼39%) increase with age in development and 193 of these (∼55%) decrease with age in ageing. It is linked to globin gene expression in erythrocytes. It is suggested to regulate Alzheimer's disease genes (Acquaah-Mensah & Taylor, 2016). There are several articles suggesting that it affects proteasome function, oxidative stress and ER stress response, differentiation and inflammation (Kim, Han & Chan, 2016).

8. **TFBS: ZID_01**:The TFBS is found in the promoter regions of 53 genes overlapping with the analysed datasets and 21 of them (∼40%) increase with age in development and 14 of these (∼65%) decrease with age in ageing. There is no TF associated with this binding matrix.

There is only 1 consistently under-represented TFBS for up-down genes, meaning it contains more than expected amount monotonically increasing genes across all datasets. The number of TFBS consistently under-represented among down-up genes is 8. One of the TFBS associated with the monotonically decreasing genes is LEF1TCF1_Q4, which is bound by LEF1. This TF has also a consistent over-representation of up-down genes.

CHAPTER 4

DISCUSSION

In this study, I investigated the age-related gene expression changes in 1,015 samples across 22 different brain regions. In total, seven different datasets were used. By incorporating different datasets, I aimed to overcome the issues raised by the heterogeneous and stochastic nature of ageing.

Three of the analysed datasets (Kang et al., 2011; Somel et al., 2010, 2011) have samples representing the whole lifespan, and four of them (Barnes et al., 2011; Berchtold et al., 2008; Lu et al., 2004; Maycox et al., 2009) have samples only for the ageing period. In human societies, age of 20 is considered as the beginning of adulthood. Similarly, previous studies have found that age of 20 represents the turning point for expression of many genes in the brain (see **Section 2.1**). Thus, I used 20 years of age to separate datasets as development and ageing periods. However, as I will mention later (see **Section 4.1**), this aspect corresponds to one of the limitations of this study.

In total, the datasets consist of 22 different brain regions, 17 of which can be analysed throughout the whole lifespan. As part of the workflow of this study, I searched for consistent patterns across these brain regions. It is possible that the brain regions, which are functionally and histologically different, undergo different age-related changes. However, by searching for common patterns across brain regions, I focused only on the shared age-related changes, and ignored possible region-specific patterns. Moreover, requiring consistency across different brain re-

gions and datasets from different laboratories is useful for the study of ageing-related changes, as focusing on only one dataset might allow technical noise to dominate the identified patterns.

The data corresponds to whole tissue samples, which means it is not cell type-specific. A possible problem when one works with whole tissue samples is that the interpretation can be complicated by the heterogeneous structure of the tissue. Thus, as the first step, I analysed the relative contribution of different cell types in the brain to the transcriptome of the samples. The results showed that neuron-specific expression dominates the expression profile. Given that neurons represent only a small fraction of brain cells in number, this result might be surprising. However, if one considers the size of the neuron cell body in comparison to that of glia, it is natural that the expression profile mostly reflects the neuronal expression.

Moreover, I found that the neuron-specific expression decreases in old ages. The analysis also showed that the astrocyte- and myelinating oligodendrocyte-specific expression also increases in elderly. This analysis is important as it suggests that all datasets had similar cell type-specific expression profiles, which even showed similarity in ageing-related changes; thus, they appear comparable at least with respect to histology.

As the next step, I showed that the consistency in age-related expression change patterns among datasets is higher in development than that in ageing. Ageing datasets clustered into two separate groups, which do not reflect the similarity of brain regions but reflects the laboratory of origin. One of the most likely reasons is the presence of technical issues. However, the age-related changes in development datasets, which share the same experimental setup, show clustering somewhat more according to the brain regions, notice clustering of frontal cortex or parietal cortex regions in **Figure 3.3**.

Why do the ageing datasets cluster according to the laboratory of origin, but not according to the brain regions and why do they show lower correlation among each other relative to correlation among development datasets? Cell type-specific expression analysis suggests that the datasets are histologically comparable. Thus, it is plausible that there is no prominent technical bias among datasets. There can

be several other possible answers to the question. The first, the gene expression change in ageing might be more random than development, considering the increased heterogeneity among individuals at advanced age. The second, the gene expression change in development might be stronger than that in ageing. Both can result in technical noise to become prominent in the measured expression levels. In turn, ageing-related changes become harder to detect than it is for the developmental changes.

In order to test the idea that stronger age-related gene expression change accompanies the developmental period, I compared age-related gene expression change strengths in development and ageing for each dataset that includes samples representing the whole lifespan. The results showed that ageing is associated with weaker age-related expression changes.

Considering that the strength of the age-related gene expression change differs between development and ageing, an important question is whether the direction differs too. Analysis showed that gene expression tend to increase more in development, whereas in ageing, gene expression change is mostly in the negative direction (**Figure 3.5**). One possible explanation for the difference in the direction of gene expression change in development and ageing is that genes having opposing trends in development and ageing are different genes: While developmental period involves up-regulation of many genes, ageing involves down-regulation of others. The second explanation, on the other hand, is that individual genes having increasing expression in development start to decrease in expression in ageing, or vice versa. This explanation, raises the question that whether what we see as opposing expression change trends in development and ageing, is linked to the gene expression reversals reported previously in the literature (see **Section 1.4**).

In order to understand if the gene expression change trajectory of the individual genes shows opposing trends in development and ageing, I calculated the prevalence of this pattern across different brain regions, using a statistic I developed for this purpose. Results showed that, on average, half of the genes expressed in each brain region show expression reversal, thus gene expression reversal is not rare.

There can be two different explanations for this pattern. The first one is that in developmental period, one cell type is prominent and in the ageing period another cell type dominates the tissue. As a result, for a given gene, developmental period reflects the expression pattern of the first cell type, whereas the expression level of the second cell type becomes more prevalent in the ageing period. So, the reversal pattern actually reflects the change in cell type proportions.

The second explanation, on the other hand, suggests that the reversal pattern is due to the gene expression change within a cell type. The expression profile of a single gene shows opposing trends in development and in ageing. The cell type-specific expression analysis (see **Section 3.1**) can give a crude idea about whether the first explanation is relevant. The relative contribution of specific cell types to the transcriptome of a sample can be discovered by using this analysis. My results showed that, although there is a slight decrease in neuron-specific gene expression in late ages, the magnitude of that decrease most likely reflects the change in the expression profile of the sample rather than the change in cell type proportions (**Figure 3.1**). Another possible approach to analyse cell type proportions is; instead of using the whole expression profile, using certain "marker" genes for each cell-type (van de Leemput et al., 2014). In this way the result can be interpreted more easily. In future I may adopt such approach, which might confirm this results.

Having shown that gene expression reversals most probably reflect intracellular expression changes, rather than showing the change in cell types, it is tempting to ask which genes show expression reversal. Among 13,686 commonly expressed genes, only 72 show reversal in 18 datasets. Although this overlap is statistically significant, the number is quite low. It is possible to interpret this result as gene expression reversal not being a biologically reproducible phenomenon, at least in the ageing human brain. However, given the stochastic nature of ageing, not seeing the reversal pattern for exactly the same genes is not unexpected. Considering the heterogeneity in ageing and that stochastic somatic mutations, DNA damage, epimutations or environmental factors might have influence over gene expression reversals, it is even more logical not to expect the reversal pattern for the same genes. A relevant question then is, if the reversal genes are over-represented in any of the GO Biological Process categories. To answer this question, I determined the categories having consistent over-representation of reversal genes across

18 datasets. The results are promising as most of the ageing-related known phenotypes were found to be associated with the reversal pattern, even though the genes are not the same across all datasets. The categories include neuronal functions and macromolecule modification pathways. As synaptic loss, decrease in cognitive abilities and loss of proteostasis are known ageing-related phenotypes (see **Section 1**), gene expression reversals might have important role in brain ageing.

Another interesting observation is that some of the GO BP categories showing consistent representation of up-down genes are in common with the GO BP categories with over-representation of monotonic decrease. One possible explanation is that gene expression reversal for some of the genes in these categories start in very early ages, e.g. around birth. As already discussed in the literature (Colantuoni et al., 2011), some genes have a turning point at the age of 20, whereas, some have turning points just after birth. As we do not include any prenatal expression data in this study, we may not detect such an early turning point. But it is possible that development of these functions is completed at prenatal period, and after birth, these functions start to be suppressed, perhaps in a disruptive way. The list of the functions that are in common between up-down and monotonic decrease GO BP categories is given in **Appendix E**. All of these functions are related to nervous system functions such as neuronal development, morphogenesis, and synaptic transmission. Further analysis of these groups can enable better interpretation of these changes. Some of the possible improvements to the analysis are: 1) instead of taking 20 years of age as a global turning point, several different ages can be defined as turning points based on the expression profiles of the genes; 2) Gene expression patterns within these GO categories can be further analysed so that presence of multiple turning points can be detected and further functional analysis can be applied.

A further notable outcome of the functional analysis is that there are several functional groups that are development/differentiation-related and have consistent over-representation of down-up genes. The down pattern in development is expected for these functional pathways, as they are mostly active during the early postnatal years and starts to decrease. However, the reason why expression of these genes start to increase in ageing is not so easy to interpret. An intriguing question is whether cells undergo expression changes such that they reverse back to their un-

differentiated state. Considering that the neuronal and synaptic functions also show reversal, it is possible that cells tend to lose their differentiated state towards elderly. Some of the age-related diseases such as cancer (Hanahan & Weinberg, 2011) and Alzheimer's disease (Arendt et al., 2000; Vincent, Rosado & Davies, 1996) have previously been linked to de-differentiation. So, a relevant question would be whether ageing itself gives cells a predisposed state to the malignancies. The investigation of a direct link between gene expression reversals and loss of differentiation needs further analysis.

Gene expression reversals are shown here to be widespread and functionally relevant phenomenon in the human brain. Moreover, some of the biological functions are in accordance with the characteristics of brain ageing. This raises a critical question: what are the possible drivers of the reversal pattern? There are several possible explanations but the most reasonable two are: 1) accumulation of random somatic damage; 2) cellular response to the accumulation of damages such as mutations and reactive oxygen species (ROS). Regulatory analysis of the reversal genes can give an idea about whether this phenomenon is a regulated response or due to random damage.

Accumulation of random somatic damage mostly concerns the cis-regulatory elements of the genes. These elements are the DNA sequences controlling the expression of the proximal genes. Unfortunately, without additional data from different layers such as genomic sequence or epigenomic signatures, it is not possible to confirm the effect of cis-regulatory elements. A possible way to interpret the data could be analysis of mutational load of the cis-regulatory elements to understand if the ones associated with the reversal genes are more prone to mutations (Lu et al., 2004). However, tissue specificity of the reversal genes should also be considered in this analysis, as most of the genome data is not brain-specific. So, either one should first confirm that gene expression reversals are widespread among different tissues and use genome information, or the analysis should be done using genome sequence of the brain tissue.

In this study, I focused on the trans-regulators and asked if there are miRNA or TFs associated with reversal genes. At the end of this analysis, if one can find significant association, two possible explanations are possible: 1) Gene expres-

sion reversals are controlled by trans-regulators that are activated/repressed upon a cellular signal (e.g. damage accumulation, cellular stress, environmental signals); 2) Stochastic changes affect the expression/function of the trans-regulator which causes expression change in the target genes.

I found that none of the miRNAs/TFs are associated with the down-up genes consistently in all of the datasets. However, up-down genes were found to be associated with five miRNAs, three of which have been linked to differentiation, cell signalling, cancer and neurodegenerative diseases. Two of the miRNAs are not widely studied. Further study on these miRNAs can provide additional information about the ageing.

Similarly, the TFs associated with up-down genes are related to development, cell cycle, inflammation and cell death functions. Some are also associated with phenotypes such as cell senescence and cancer. The TFBS identified as related to up-down genes is not yet mapped to any particular TF, and studying that TFBS to elucidate the TF bears potential to understand ageing.

Overall, the regulatory analysis suggests that there are miRNAs and TFs showing consistent over-representation of reversal genes. However, considering that the reversal genes are not the same across datasets, an important question is whether these associated trans-regulators actually regulate the reversal genes. It is important to note that the analysis only shows an association but not direct functional relationship. A possible way to further check the association is analysing the expression patterns of the regulators. For example, if a miRNA that is associated with up-down genes, shows age-related expression decrease in ageing, the functional relationship might actually be missing. When considering the expression patterns of the regulators, another important aspect is the post-transcriptional/translational regulation of the regulators. For example, it is important to find out if there is age-related alternative splicing of the regulator. The mature form favoured by the cell can show differences with age. Analysis of these alternative forms can also provide understanding on how/if gene expression reversal is regulated.

Although some of the reversal patterns are associated with trans-regulators, there can be reversal genes that are not trans-regulated. These genes might have pro-

moters that are more vulnerable to damages, resulting in expression change in later ages. For example, if these genes are expressed at high levels throughout the lifetime, that means the chromatin structure of the promoter regions is different. The more active the gene, generally the more open the chromatin structure. That, in turn, causes DNA to be more prone to mutations (Schwer et al., 2016); conversely, transcription-coupled repair might also limit some types of mutation accumulation (Yadav, DeGregori & De, 2016). Another valuable information would be the chromosomal location of the reversal genes. If these genes are generally close to each other, or close to a mutational hotspot, this can also imply a possible reason why gene expression reversals are observed.

4.1 Limitations and Possible Improvements

1. In this study, I accepted 20 years of age as a global turning point for the gene expression change. However, the results imply that some genes might show an early reversal. A possible improvement would be grouping genes according to the turning points and then classifying them accordingly.

2. The Kang2011 dataset contains 16 brain regions, in total 741 samples from 33 individuals. Throughout the study, I treated different brain regions as they are independent from each other. However, since some of the samples for different regions are taken from the same individuals, to test the categories with consistent over-representation of reversal genes, I can use an improved version of permutation scheme that takes controls for individual identity.

3. For the cell type analysis, I used the whole transcriptome of each cell type. This in turn can give information on whether the cell type proportion changes in the tissue. However, the result can also be interpreted as a change in cell type-specific expression, thus loss of cellular identity. Without further single-cell transcriptome data generated, a possible improvement would be confirming the results using just the "marker" genes for each cell type.

4. After dividing datasets into brain regions and development/ageing periods, there are in total 43 datasets analysed. However, not all of them have comparable number of samples. This would affect the significance measure of the age-expression correlations (also for this reason, significance cut-offs were not used in this study except for in **Figure 3.5**). A possible improvement

would be repeating the analysis by taking random N samples from each dataset, N being the number of samples in the smallest dataset.

5. The regulatory analysis only gives an idea about the possible association between reversal genes and trans-regulators. However, for the correct interpretation, the expression profiles of the trans-regulators should also be considered.

6. Genomic features of the reversal genes are not analysed in this study. Investigation of features such as chromatin structure or chromosomal location of these genes can provide information about the possible causes of gene expression reversal pattern.

CHAPTER 5

CONCLUSION

Although it is still disputed whether ageing is itself a disease or not, it is widely accepted that ageing is the major risk factor for many diseases. One of the other major consequences of ageing is decline in cognitive abilities and memory. This does not only affect the individual's health but also affects the social life. For these reasons, brain ageing is a particular interest to this study.

To obtain insight into the underlying molecular nature of human brain ageing, I used published microarray data from 7 different studies, containing 1,015 samples from 22 brain regions. Age-related gene expression has been widely studied; however, this study provides a new approach by combining different studies together and studying ageing in the context of development. The major outcomes of this study are:

- Age-related gene expression change in ageing shows more variability among datasets, which can possibly be attributed to higher inter-individual variability; thus, expression change in ageing is more heterogeneous than that in development.

- Age-related gene expression change is stronger in development than that in ageing.

- In development, the proportion of genes showing age-related expression decrease is higher than that of genes with age-related expression increase. In

contrast, ageing is mostly accompanied by genes showing age-related gene expression decrease.

- Gene expression reversal is observed in all brain regions and have consistent functional associations. These functional associations include many ageing-related phenotypic changes, including decrease in neuronal activity, loss of proteostasis and decline in cognitive abilities.

- Apart from decrease in neuronal activity, functional processes related to cell differentiation / development also tend to reverse back to the early years of life. This may suggest that cells are inclined to lose their differentiated state / cellular identity.

- Up-down genes are associated with several miRNA and TFs, which suggests that at least some of the reversal genes might be regulated.

- Down-up genes do not have a consistent association with any miRNA and TFs.

Overall, the results showed that meta-analysis is a useful approach for transcriptome studies of ageing due to the stochastic nature of ageing, and that studying ageing-related gene expression changes in the context of development is a promising approach to discover the molecular mechanisms of ageing. Further work including analysis of reversals in the context of diseases and the regulatory mechanisms can provide better understanding into the process, and may also provide some potential targets to alleviate the consequences of ageing.

REFERENCES

Aanerud, J., Borghammer, P., Chakravarty, M. M., Vang, K., Rodell, A. B., Jónsdottir, K. Y., ... Gjedde, A. (2012, jul). Brain energy metabolism and blood flow differences in healthy aging. *Journal of cerebral blood flow and metabolism : official journal of the International Society of Cerebral Blood Flow and Metabolism*, *32*(7), 1177–87. Retrieved from http://jcb.sagepub.com/content/32/7/1177.full doi: 10.1038/jcbfm.2012.18

Acquaah-Mensah, G. K. & Taylor, R. C. (2016, apr). Brain in situ hybridization maps as a source for reverse-engineering transcriptional regulatory networks: Alzheimer's disease insights. *Gene*, *586*(1), 77–86. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/27050105 doi: 10.1016/j.gene.2016.03.045

Arendt, T., Holzer, M., Stöbe, A., Gärtner, U., Lüth, H. J., Brückner, M. K. & Ueberham, U. (2000). Activated mitogenic signaling induces a process of dedifferentiation in Alzheimer's disease that eventually results in cell death. *Annals of the New York Academy of Sciences*, *920*, 249–255. doi: 10.1111/j.1749-6632.2000.tb06931.x

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., ... Sherlock, G. (2000, may). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nature genetics*, *25*(1), 25–9. Retrieved from http://dx.doi.org/10.1038/75556 doi: 10.1038/75556

Bahar, R., Hartmann, C. H., Rodriguez, K. A., Denny, A. D., Busuttil, R. A., Dollé, M. E. T., ... Vijg, J. (2006, jun). Increased cell-to-cell variation in gene expression in ageing mouse heart. *Nature*, *441*(7096), 1011–4. Retrieved from http://dx.doi.org/10.1038/nature04844 doi: 10.1038/nature04844

Baker, D. J., Wijshake, T., Tchkonia, T., LeBrasseur, N. K., Childs, B. G., van de Sluis, B., ... van Deursen, J. M. (2011, nov). Clearance of p16Ink4a-positive

senescent cells delays ageing-associated disorders. *Nature*, *479*(7372), 232–6. Retrieved from http://dx.doi.org/10.1038/nature10600 doi: 10.1038/nature10600

Barnes, M. R., Huxley-Jones, J., Maycox, P. R., Lennon, M., Thornber, A., Kelly, F., ... De Belleroche, J. (2011, aug). Transcription and pathway analysis of the superior temporal cortex and anterior prefrontal cortex in schizophrenia. *Journal of Neuroscience Research*, *89*(8), 1218–1227. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/21538462 doi: 10.1002/jnr.22647

Berchtold, N. C., Cribbs, D. H., Coleman, P. D., Rogers, J., Head, E., Kim, R., ... Cotman, C. W. (2008, oct). Gene expression changes in the course of normal brain aging are sexually dimorphic. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(40), 15605–10. Retrieved from http://www.pnas.org/content/105/40/15605.long doi: 10.1073/pnas.0806883105

Bernardes de Jesus, B., Vera, E., Schneeberger, K., Tejera, A. M., Ayuso, E., Bosch, F. & Blasco, M. A. (2012, aug). Telomerase gene therapy in adult and old mice delays aging and increases longevity without increasing cancer. *EMBO molecular medicine*, *4*(8), 691–704. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3494070{&}tool=pmcentrez{&}rendertype=abstract doi: 10.1002/emmm.201200245

Bhardwaj, R. D., Curtis, M. A., Spalding, K. L., Buchholz, B. A., Fink, D., Björk-Eriksson, T., ... Frisén, J. (2006, aug). Neocortical neurogenesis in humans is restricted to development. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(33), 12564–8. Retrieved from http://www.pnas.org/content/103/33/12564.long doi: 10.1073/pnas.0605177103

Bishop, N. A. & Guarente, L. (2007, nov). Genetic links between diet and lifespan: shared mechanisms from yeast to humans. *Nature reviews. Genetics*, *8*(11), 835–44. Retrieved from http://dx.doi.org/10.1038/nrg2188 doi: 10.1038/nrg2188

Blalock, E. M., Chen, K.-C., Sharrow, K., Herman, J. P., Porter, N. M., Foster, T. C. & Landfield, P. W. (2003, may). Gene microarrays in hippocampal aging: statistical profiling identifies novel processes correlated with cognitive impairment. *The Journal of neuroscience : the official journal of the Society*

for Neuroscience, *23*(9), 3807–3819. Retrieved from http://www.jneurosci.org/content/23/9/3807.long doi: 23/9/3807[pii]

Bolstad, B., Irizarry, R., Astrand, M. & Speed, T. (2003, jan). A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics*, *19*(2), 185–193. Retrieved from http://bioinformatics.oxfordjournals.org/content/19/2/185.abstract doi: 10.1093/bioinformatics/19.2.185

Cahoy, J. D., Emery, B., Kaushal, A., Foo, L. C., Zamanian, J. L., Christopherson, K. S., ... Barres, B. A. (2008, jan). A transcriptome database for astrocytes, neurons, and oligodendrocytes: a new resource for understanding brain development and function. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *28*(1), 264–78. Retrieved from http://www.jneurosci.org/content/28/1/264.full doi: 10.1523/JNEUROSCI.4178-07.2008

Calvanese, V., Lara, E., Kahn, A. & Fraga, M. F. (2009, oct). The role of epigenetics in aging and age-related diseases. *Ageing research reviews*, *8*(4), 268–76. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/19716530 doi: 10.1016/j.arr.2009.03.004

Campisi, J. (2001, nov). Cellular senescence as a tumor-suppressor mechanism. *Trends in Cell Biology*, *11*(11), S27–S31. Retrieved from http://www.sciencedirect.com/science/article/pii/S0962892401021511 doi: 10.1016/S0962-8924(01)02151-1

Campisi, J. & Robert, L. (2014, jan). Cell senescence: role in aging and age-related diseases. *Interdisciplinary topics in gerontology*, *39*, 45–61. Retrieved from http://www.karger.com/Article/FullText/358899 doi: 10.1159/000358899

Carlson, M. (n.d.-a). Go.db: A set of annotation maps describing the entire gene ontology [Computer software manual]. (R package version 3.2.2)

Carlson, M. (n.d.-b). org.hs.eg.db: Genome wide annotation for human [Computer software manual]. (R package version 3.2.3)

Carvalho, B. S. & Irizarry, R. A. (2010, oct). A framework for oligonucleotide microarray preprocessing. *Bioinformatics (Oxford, England)*, *26*(19), 2363–7. Retrieved from http://bioinformatics.oxfordjournals.org/content/26/19/2363.long doi: 10.1093/bioinformatics/btq431

Clancy, B., Darlington, R. & Finlay, B. (2001, jul). Translating developmental time across mammalian species. *Neuroscience*, *105*(1), 7–17. Retrieved from

http://www.sciencedirect.com/science/article/pii/S0306452201001713 doi: 10.1016/S0306-4522(01)00171-3

Colantuoni, C., Lipska, B. K., Ye, T., Hyde, T. M., Tao, R., Leek, J. T., ... Kleinman, J. E. (2011, oct). Temporal dynamics and genetic control of transcription in the human prefrontal cortex. *Nature*, *478*(7370), 519–23. Retrieved from http://dx.doi.org/10.1038/nature10524 doi: 10.1038/nature10524

Crimmins, E. M. & Beltrán-Sánchez, H. (2011, jan). Mortality and morbidity trends: is there compression of morbidity? *The journals of gerontology. Series B, Psychological sciences and social sciences*, *66*(1), 75–86. Retrieved from http://psychsocgerontology.oxfordjournals.org/content/66B/1/75.short doi: 10.1093/geronb/gbq088

Cruickshanks, H. a. & Adams, P. D. (2011). *Chromatin: A molecular interface between cancer and aging* (Vol. 21) (No. 1). doi: 10.1016/j.gde.2010.10.007

Cruickshanks, H. a., McBryan, T., Nelson, D. M., Vanderkraats, N. D., Shah, P. P., van Tuyn, J., ... Adams, P. D. (2013). Senescent cells harbour features of the cancer epigenome. *Nature cell biology*, *15*(12), 1495–506. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/24270890 doi: 10.1038/ncb2879

Csardi, G. & Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal*, *Complex Systems*, 1695. Retrieved from http://igraph.org

de Graaf-Peters, V. B. & Hadders-Algra, M. (2006, apr). Ontogeny of the human central nervous system: what is happening when? *Early human development*, *82*(4), 257–66. Retrieved from http://www.sciencedirect.com/science/article/pii/S0378378205002082http://www.earlyhumandevelopment.com/article/S0378378205002082/fulltext doi: 10.1016/j.earlhumdev.2005.10.013

de Magalhães, J. P., Curado, J. & Church, G. M. (2009, apr). Meta-analysis of age-related gene expression profiles identifies common signatures of aging. *Bioinformatics (Oxford, England)*, *25*(7), 875–81. Retrieved from http://bioinformatics.oxfordjournals.org/content/25/7/875.short doi: 10.1093/bioinformatics/btp073

de Magalhães, J. P., Finch, C. E. & Janssens, G. (2010, jul). Next-generation sequencing in aging research: emerging applications, problems, pitfalls and possible solutions. *Ageing research reviews*, *9*(3), 315–23. Retrieved from http://www.sciencedirect.com/science/article/pii/S1568163709000749 doi:

10.1016/j.arr.2009.10.006

Department of Economic and Social Affairs Population Division. (2000). Chapter 2: Magnitude and speed of population ageing. In *World population ageing 1950-2050* (pp. 11–13). Retrieved from http://www.un.org/esa/population/publications/worldageing19502050/

Desai, A. & McConnell, S. (2000, jul). Progressive restriction in fate potential by neural progenitors during cerebral cortical development. *Development*, *127*(13), 2863–2872. Retrieved from http://dev.biologists.org/content/127/13/2863.long

Dolcos, F., Rice, H. J. & Cabeza, R. (2002, nov). Hemispheric asymmetry and aging: right hemisphere decline or asymmetry reduction. *Neuroscience & Biobehavioral Reviews*, *26*(7), 819–825. Retrieved from http://www.sciencedirect.com/science/article/pii/S0149763402000684 doi: 10.1016/S0149-7634(02)00068-4

Durinck, S., Moreau, Y., Kasprzyk, A., Davis, S., De Moor, B., Brazma, A. & Huber, W. (2005, aug). BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis. *Bioinformatics (Oxford, England)*, *21*(16), 3439–40. Retrieved from http://bioinformatics.oxfordjournals.org/content/21/16/3439.long doi: 10.1093/bioinformatics/bti525

Durinck, S., Spellman, P. T., Birney, E. & Huber, W. (2009, jan). Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nature protocols*, *4*(8), 1184–91. Retrieved from http://dx.doi.org/10.1038/nprot.2009.97 doi: 10.1038/nprot.2009.97

Eberharter, A. & Becker, P. B. (2002, mar). Histone acetylation: a switch between repressive and permissive chromatin. Second in review series on chromatin dynamics. *EMBO reports*, *3*(3), 224–9. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1084017{&}tool=pmcentrez{&}rendertype=abstract doi: 10.1093/embo-reports/kvf053

Edgar, R. (2002, jan). Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Research*, *30*(1), 207–210. Retrieved from http://nar.oxfordjournals.org/content/30/1/207.full?ijkey=oxMPOWseARs7o{&}keytype=ref{&}siteid=nar doi: 10.1093/nar/30.1.207

Erraji-Benchekroun, L., Underwood, M. D., Arango, V., Galfalvy, H., Pavlidis, P., Smyrniotopoulos, P., ... Sibille, E. (2005, mar). Molecular aging in human prefrontal cortex is selective and continuous throughout adult life. *Biological psychiatry*, *57*(5), 549–58. Retrieved from http://www.biologicalpsychiatryjournal.com/article/S0006322304011187/fulltext doi: 10.1016/j.biopsych.2004.10.034

Faggioli, F., Wang, T., Vijg, J. & Montagna, C. (2012, dec). Chromosome-specific accumulation of aneuploidy in the aging mouse brain. *Human molecular genetics*, *21*(24), 5246–53. Retrieved from http://hmg.oxfordjournals.org/content/21/24/5246.long doi: 10.1093/hmg/dds375

Fontana, L., Partridge, L. & Longo, V. D. (2010, apr). Extending healthy life span–from yeast to humans. *Science (New York, N.Y.)*, *328*(5976), 321–6. Retrieved from http://science.sciencemag.org/content/328/5976/321.abstract doi: 10.1126/science.1172539

Forsberg, L. A., Rasi, C., Razzaghian, H. R., Pakalapati, G., Waite, L., Thilbeault, K. S., ... Dumanski, J. P. (2012, feb). Age-related somatic structural changes in the nuclear genome of human blood cells. *American journal of human genetics*, *90*(2), 217–28. Retrieved from http://www.sciencedirect.com/science/article/pii/S0002929711005441 doi: 10.1016/j.ajhg.2011.12.009

Fraga, M. F., Ballestar, E., Paz, M. F., Ropero, S., Setien, F., Ballestar, M. L., ... Esteller, M. (2005, jul). Epigenetic differences arise during the lifetime of monozygotic twins. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(30), 10604–9. Retrieved from http://www.pnas.org/content/102/30/10604.abstract doi: 10.1073/pnas.0500398102

Friedman, D. B. & Johnson, T. E. (1988, jan). A Mutation in the age-1 Gene in Caenorhabditis elegans Lengthens Life and Reduces Hermaphrodite Fertility. *Genetics*, *118*(1), 75–86. Retrieved from http://www.genetics.org/content/118/1/75.long

Gautier, L., Cope, L., Bolstad, B. M. & Irizarry, R. A. (2004, feb). affy–analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics (Oxford, England)*, *20*(3), 307–15. Retrieved from http://bioinformatics.oxfordjournals.org/content/20/3/307 doi: 10.1093/bioinformatics/btg405

Gould, E. (1999, oct). Neurogenesis in the Neocortex of Adult Primates. *Science*, *286*(5439), 548–552. Retrieved from http://science.sciencemag.org/content/286/5439/548.abstract doi: 10.1126/science.286.5439.548

Gregg, S. Q., Gutiérrez, V., Robinson, A. R., Woodell, T., Nakao, A., Ross, M. A., ... Niedernhofer, L. J. (2012, feb). A mouse model of accelerated liver aging caused by a defect in DNA repair. *Hepatology (Baltimore, Md.)*, *55*(2), 609–21. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3250572{&}tool=pmcentrez{&}rendertype=abstract doi: 10.1002/hep.24713

Hanahan, D. & Weinberg, R. A. (2011). *Hallmarks of cancer: The next generation* (Vol. 144) (No. 5).

Hands, S. L., Proud, C. G. & Wyttenbach, A. (2009, jul). mTOR's role in ageing: protein synthesis or autophagy? *Aging*, *1*(7), 586–97. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2806042{&}tool=pmcentrez{&}rendertype=abstract

Harrison, D. E., Strong, R., Sharp, Z. D., Nelson, J. F., Astle, C. M., Flurkey, K., ... Miller, R. A. (2009, jul). Rapamycin fed late in life extends lifespan in genetically heterogeneous mice. *Nature*, *460*(7253), 392–5. Retrieved from http://dx.doi.org/10.1038/nature08221 doi: 10.1038/nature08221

Hatano, K., Kumar, B., Zhang, Y., Coulter, J. B., Hedayati, M., Mears, B., ... Lupold, S. E. (2015, apr). A functional screen identifies miRNAs that inhibit DNA repair and sensitize prostate cancer cells to ionizing radiation. *Nucleic acids research*, *43*(8), 4075–86. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4417178{&}tool=pmcentrez{&}rendertype=abstract doi: 10.1093/nar/gkv273

Hayflick, L. (1965, mar). The limited in vitro lifetime of human diploid cell strains. *Experimental Cell Research*, *37*(3), 614–636. Retrieved from http://www.sciencedirect.com/science/article/pii/0014482765902119 doi: 10.1016/0014-4827(65)90211-9

Hayflick, L. & Moorhead, P. (1961, dec). The serial cultivation of human diploid cell strains. *Experimental Cell Research*, *25*(3), 585–621. Retrieved from http://www.sciencedirect.com/science/article/pii/0014482761901926 doi: 10.1016/0014-4827(61)90192-6

Herndon, L. A., Schmeissner, P. J., Dudaronek, J. M., Brown, P. A., Listner, K. M., Sakano, Y., ... Driscoll, M. (2002, oct). Stochastic and genetic factors influence tissue-specific decline in ageing C. elegans. *Nature*, *419*(6909), 808–14. Retrieved from http://dx.doi.org/10.1038/nature01135

doi: 10.1038/nature01135

Hoeijmakers, J. H. J. (2009, oct). DNA damage, aging, and cancer. *The New England journal of medicine*, *361*(15), 1475–85. Retrieved from http://www.nejm.org/doi/full/10.1056/NEJMra0804615 doi: 10.1056/NEJMra0804615

Horvath, S., Zhang, Y., Langfelder, P., Kahn, R. S., Boks, M. P. M., van Eijk, K., ... Ophoff, R. A. (2012, jan). Aging effects on DNA methylation modules in human brain and blood tissue. *Genome biology*, *13*(10), R97. Retrieved from http://genomebiology.biomedcentral.com/articles/10.1186/gb-2012-13-10-r97 doi: 10.1186/gb-2012-13-10-r97

Huang, E. J. & Reichardt, L. F. (2001, jan). Neurotrophins: roles in neuronal development and function. *Annual review of neuroscience*, *24*, 677–736. Retrieved from http://www.annualreviews.org/doi/abs/10.1146/annurev.neuro.24.1.677?url{_}ver=Z39.88-2003{&}rfr{_}dat=cr{_}pub{%}3Dpubmed{&}rfr{_}id=ori{%}3Arid{%}3Acrossref.org{&}journalCode=neuro doi: 10.1146/annurev.neuro.24.1.677

Huttenlocher, P. R. & Dabholkar, A. S. (1997, oct). Regional differences in synaptogenesis in human cerebral cortex. *The Journal of comparative neurology*, *387*(2), 167–78. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9336221

Innocenti, G. M. & Price, D. J. (2005, dec). Exuberance in the development of cortical networks. *Nature reviews. Neuroscience*, *6*(12), 955–65. Retrieved from http://dx.doi.org/10.1038/nrn1790 doi: 10.1038/nrn1790

Irizarry, R. A. (2003, feb). Summaries of Affymetrix GeneChip probe level data. *Nucleic Acids Research*, *31*(4), 15e–15. Retrieved from http://nar.oxfordjournals.org/content/31/4/e15.long doi: 10.1093/nar/gng015

Jaskelioff, M., Muller, F. L., Paik, J.-H., Thomas, E., Jiang, S., Adams, A. C., ... Depinho, R. A. (2011, jan). Telomerase reactivation reverses tissue degeneration in aged telomerase-deficient mice. *Nature*, *469*(7328), 102–6. Retrieved from http://dx.doi.org/10.1038/nature09603 doi: 10.1038/nature09603

Jiang, X. & Nardelli, J. (2015, jul). Cellular and molecular introduction to brain development. *Neurobiology of Disease*. Retrieved from http://www.sciencedirect.com/science/article/pii/S0969996115300097 doi: 10.1016/j.nbd.2015.07.007

Johnson, M. H. (2001, jul). Functional brain development in humans. *Nature*

reviews. Neuroscience, *2*(7), 475–83. Retrieved from http://dx.doi.org/10.1038/35081509 doi: 10.1038/35081509

Kang, H. J., Kawasawa, Y. I., Cheng, F., Zhu, Y., Xu, X., Li, M., ... Sestan, N. (2011, oct). Spatio-temporal transcriptome of the human brain. *Nature*, *478*(7370), 483–9. Retrieved from http://dx.doi.org/10.1038/nature10523 doi: 10.1038/nature10523

Kel, A. (2003, jul). MATCHTM: a tool for searching transcription factor binding sites in DNA sequences. *Nucleic Acids Research*, *31*(13), 3576–3579. Retrieved from http://nar.oxfordjournals.org/content/31/13/3576.long doi: 10.1093/nar/gkg585

Kenyon, C., Chang, J., Gensch, E., Rudner, A. & Tabtiang, R. (1993, dec). A C. elegans mutant that lives twice as long as wild type. *Nature*, *366*(6454), 461–4. Retrieved from http://dx.doi.org/10.1038/366461a0 doi: 10.1038/366461a0

Kim, H. M., Han, J. W. & Chan, J. Y. (2016, mar). Nuclear Factor Erythroid-2 Like 1 (NFE2L1): Structure, function and regulation. *Gene*, *584*(1), 17–25. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/26947393 doi: 10.1016/j.gene.2016.03.002

Kirkwood, T. B. L. (1977, nov). Evolution of ageing. *Nature*, *270*(5635), 301–304. Retrieved from http://dx.doi.org/10.1038/270301a0 doi: 10.1038/270301a0

Kirkwood, T. B. L. (2005). Understanding the odd science of aging. *Cell*, *120*(4), 437–447. doi: 10.1016/j.cell.2005.01.027

Krishnan, V., Chow, M. Z. Y., Wang, Z., Zhang, L., Liu, B., Liu, X. & Zhou, Z. (2011, jul). Histone H4 lysine 16 hypoacetylation is associated with defective DNA repair and premature senescence in Zmpste24-deficient mice. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(30), 12325–30. Retrieved from http://www.pnas.org/content/108/30/12325.long doi: 10.1073/pnas.1102789108

Kumar, P., Dezso, Z., MacKenzie, C., Oestreicher, J., Agoulnik, S., Byrne, M., ... Oda, Y. (2013, jan). Circulating miRNA biomarkers for Alzheimer's disease. *PloS one*, *8*(7), e69807. Retrieved from http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0069807 doi: 10.1371/journal.pone.0069807

Lalevée, S., Lapaire, O. & Bühler, M. (2014, jan). miR455 is linked to hyp-

oxia signaling and is deregulated in preeclampsia. *Cell death & disease*, *5*, e1408. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4540200{&}tool=pmcentrez{&}rendertype=abstract doi: 10.1038/cddis.2014.368

Lee, C. K., Weindruch, R. & Prolla, T. A. (2000, jul). Gene-expression profile of the ageing brain in mice. *Nature genetics*, *25*(3), 294–7. Retrieved from http://dx.doi.org/10.1038/77046 doi: 10.1038/77046

Liscovitch, N. & Chechik, G. (2013, jan). Specialization of gene expression during mouse brain development. *PLoS computational biology*, *9*(9), e1003185. Retrieved from http://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1003185 doi: 10.1371/journal.pcbi.1003185

Loerch, P. M., Lu, T., Dakin, K. A., Vann, J. M., Isaacs, A., Geula, C., ... Yankner, B. A. (2008, jan). Evolution of the aging brain transcriptome and synaptic regulation. *PloS one*, *3*(10), e3329. Retrieved from http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0003329 doi: 10.1371/journal.pone.0003329

López-Otín, C., Blasco, M. A., Partridge, L., Serrano, M. & Kroemer, G. (2013, jun). The hallmarks of aging. *Cell*, *153*(6), 1194–217. Retrieved from http://www.cell.com/article/S0092867413006454/fulltext 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), 883–91. Retrieved from http://dx.doi.org/10.1038/nature02661 doi: 10.1038/nature02661

Lund, J., Tedesco, P., Duke, K., Wang, J., Kim, S. K. & Johnson, T. E. (2002, sep). Transcriptional Profile of Aging in C. elegans. *Current Biology*, *12*(18), 1566–1573. Retrieved from http://www.cell.com/article/S0960982202011466/fulltext doi: 10.1016/S0960-9822(02)01146-6

Martin, G. M. (2011, nov). The biology of aging: 1985-2010 and beyond. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology*, *25*(11), 3756–62. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4050331{&}tool=pmcentrez{&}rendertype=abstract doi: 10.1096/fj.11-1102.ufm

Martin, J., Jenkins, R. H., Bennagi, R., Krupa, A., Phillips, A. O., Bowen, T.

& Fraser, D. J. (2011, jan). Post-transcriptional regulation of Transforming Growth Factor Beta-1 by microRNA-744. *PloS one*, *6*(10), e25044. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3186795{&}tool=pmcentrez{&}rendertype=abstract doi: 10.1371/journal.pone.0025044

Matheu, A., Maraver, A., Klatt, P., Flores, I., Garcia-Cao, I., Borras, C., ... Serrano, M. (2007, jul). Delayed ageing through damage protection by the Arf/p53 pathway. *Nature*, *448*(7151), 375–9. Retrieved from http://dx.doi.org/10.1038/nature05949 doi: 10.1038/nature05949

Maycox, P. R., Kelly, F., Taylor, A., Bates, S., Reid, J., Logendra, R., ... de Belleroche, J. (2009, dec). Analysis of gene expression in two large schizophrenia cohorts identifies multiple changes associated with nerve terminal function. *Molecular psychiatry*, *14*(12), 1083–94. Retrieved from http://dx.doi.org/10.1038/mp.2009.18 doi: 10.1038/mp.2009.18

Mazin, P., Xiong, J., Liu, X., Yan, Z., Zhang, X., Li, M., ... Khaitovich, P. (2013, jan). Widespread splicing changes in human brain development and aging. *Molecular systems biology*, *9*(1), 633. Retrieved from http://msb.embopress.org/content/9/1/633.abstract doi: 10.1038/msb.2012.67

McAuley, M. T., Kenny, R. A., Kirkwood, T. B. L., Wilkinson, D. J., Jones, J. J. L. & Miller, V. M. (2009, jan). A mathematical model of aging-related and cortisol induced hippocampal dysfunction. *BMC neuroscience*, *10*(1), 26. Retrieved from http://bmcneurosci.biomedcentral.com/articles/10.1186/1471-2202-10-26 doi: 10.1186/1471-2202-10-26

McTigue, D. M. & Tripathi, R. B. (2008, oct). The life, death, and replacement of oligodendrocytes in the adult CNS. *Journal of neurochemistry*, *107*(1), 1–19. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/18643793 doi: 10.1111/j.1471-4159.2008.05570.x

Medawar, P. B. (1952). *An Unsolved Problem of Biology*. doi: 10.1016/S0140-6736(00)99799-X

Medvedev, Z. A. (1990, aug). AN ATTEMPT AT A RATIONAL CLASSIFICATION OF THEORIES OF AGEING. *Biological Reviews*, *65*(3), 375–398. Retrieved from http://doi.wiley.com/10.1111/j.1469-185X.1990.tb01428.x doi: 10.1111/j.1469-185X.1990.tb01428.x

Molyneaux, B. J., Arlotta, P., Menezes, J. R. L. & Macklis, J. D. (2007, jun). Neuronal subtype specification in the cerebral cortex. *Nature reviews. Neur-*

oscience, *8*(6), 427–37. Retrieved from http://dx.doi.org/10.1038/nrn2151 doi: 10.1038/nrn2151

Moskalev, A. A., Shaposhnikov, M. V., Plyusnina, E. N., Zhavoronkov, A., Budovsky, A., Yanai, H. & Fraifeld, V. E. (2012, mar). The role of DNA damage and repair in aging through the prism of Koch-like criteria. *Ageing research reviews*, *12*(2), 661–84. Retrieved from http://www.sciencedirect.com/science/article/pii/S1568163712000086 doi: 10.1016/j.arr.2012.02.001

Naumova, O. Y., Palejev, D., Vlasova, N. V., Lee, M., Rychkov, S. Y., Babich, O. N., ... Grigorenko, E. L. (2012, nov). Age-related changes of gene expression in the neocortex: preliminary data on RNA-Seq of the transcriptome in three functionally distinct cortical areas. *Development and psychopathology*, *24*(4), 1427–42. Retrieved from http://journals.cambridge.org/abstract{_}S0954579412000818 doi: 10.1017/S0954579412000818

Niccoli, T. & Partridge, L. (2012, sep). Ageing as a Risk Factor for Disease. *Current Biology*, *22*(17), R741–R752. Retrieved from http://www.sciencedirect.com/science/article/pii/S0960982212008159 doi: 10.1016/j.cub.2012.07.024

Oberheim, N. A., Takano, T., Han, X., He, W., Lin, J. H. C., Wang, F., ... Nedergaard, M. (2009, mar). Uniquely hominid features of adult human astrocytes. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *29*(10), 3276–87. Retrieved from http://www.jneurosci.org/content/29/10/3276.long doi: 10.1523/JNEUROSCI.4707-08.2009

Olovnikov, A. M. (1996, jul). Telomeres, telomerase, and aging: Origin of the theory. *Experimental Gerontology*, *31*(4), 443–448. Retrieved from http://www.sciencedirect.com/science/article/pii/0531556596000058 doi: 10.1016/0531-5565(96)00005-8

Pages, H., Carlson, M., Falcon, S. & Li, N. (n.d.). Annotationdbi: Annotation database interface [Computer software manual]. (R package version 1.32.3)

Pang, C., Guan, Y., Zhao, K., Chen, L., Bao, Y., Cui, R., ... Wang, Y. (2015, jan). Up-regulation of microRNA-15b correlates with unfavorable prognosis and malignant progression of human glioma. *International journal of clinical and experimental pathology*, *8*(5), 4943–52. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4503059{&}tool=pmcentrez{&}rendertype=abstract

Peters, A., Sethares, C. & Luebke, J. I. (2008, apr). Synapses are lost during aging in the primate prefrontal cortex. *Neuroscience*, *152*(4), 970–81. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2441531{&}tool=pmcentrez{&}rendertype=abstract doi: 10.1016/j.neuroscience.2007.07.014

Pletcher, S. D., Macdonald, S. J., Marguerie, R., Certa, U., Stearns, S. C., Goldstein, D. B. & Partridge, L. (2002, apr). Genome-Wide Transcript Profiles in Aging and Calorically Restricted Drosophila melanogaster. *Current Biology*, *12*(9), 712–723. Retrieved from http://www.sciencedirect.com/science/article/pii/S0960982202008084 doi: 10.1016/S0960-9822(02)00808-4

Rajfer, J. (2003). *Decreased Testosterone in the Aging Male* (Vol. 5) (No. Suppl 1). MedReviews, LLC. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1502317/

Rakic, S. & Zecevic, N. (2000, aug). Programmed cell death in the developing human telencephalon. *European Journal of Neuroscience*, *12*(8), 2721–2734. Retrieved from http://doi.wiley.com/10.1046/j.1460-9568.2000.00153.x doi: 10.1046/j.1460-9568.2000.00153.x

Rodríguez, S. A., Grochová, D., McKenna, T., Borate, B., Trivedi, N. S., Erdos, M. R. & Eriksson, M. (2016, apr). Global genome splicing analysis reveals an increased number of alternatively spliced genes with aging. *Aging cell*, *15*(2), 267–78. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4783335{&}tool=pmcentrez{&}rendertype=abstract doi: 10.1111/acel.12433

Rowe, W. B., Blalock, E. M., Chen, K.-C., Kadish, I., Wang, D., Barrett, J. E., ... Landfield, P. W. (2007, mar). Hippocampal expression analyses reveal selective association of immediate-early, neuroenergetic, and myelinogenic pathways with cognitive impairment in aged rats. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, *27*(12), 3098–110. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/17376971 doi: 10.1523/JNEUROSCI.4163-06.2007

Rowitch, D. H. & Kriegstein, A. R. (2010, nov). Developmental genetics of vertebrate glial-cell specification. *Nature*, *468*(7321), 214–22. Retrieved from http://dx.doi.org/10.1038/nature09611 doi: 10.1038/nature09611

Salminen, A., Kaarniranta, K. & Kauppinen, A. (2012). Inflammaging: Disturbed

interplay between autophagy and inflammasomes. *Aging*, *4*(3), 166–175. doi: 100444[pii]

Salthouse, T. A. (2009, apr). When does age-related cognitive decline begin? *Neurobiology of aging*, *30*(4), 507–14. Retrieved from http://www.neurobiologyofaging.org/article/S0197458009000219/fulltext doi: 10.1016/j.neurobiolaging.2008.09.023

Schwer, B., Wei, P.-C., Chang, A. N., Kao, J., Du, Z., Meyers, R. M. & Alt, F. W. (2016, feb). Transcription-associated processes cause DNA double-strand breaks and translocations in neural stem/progenitor cells. *Proceedings of the National Academy of Sciences*, *113*(8), 201525564. Retrieved from http://www.pnas.org/content/113/8/2258.abstract doi: 10.1073/pnas.1525564113

Siepel, A., Bejerano, G., Pedersen, J. S., Hinrichs, A. S., Hou, M., Rosenbloom, K., ... Haussler, D. (2005, aug). Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome research*, *15*(8), 1034–50. Retrieved from http://genome.cshlp.org/content/15/8/1034.abstract doi: 10.1101/gr.3715005

Smith, C. U. M. (2002). *Elements of Molecular Neurobiology* (Third Edit ed.).

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 Research*, *20*(9), 1207–1218. Retrieved from http://genome.cshlp.org/content/20/9/1207.long doi: 10.1101/gr.106849.110

Somel, M., Khaitovich, P., Bahn, S., Pääbo, S. & Lachmann, M. (2006, may). Gene expression becomes heterogeneous with age. *Current biology : CB*, *16*(10), R359–60. Retrieved from http://www.cell.com/article/S0960982206014850/fulltext doi: 10.1016/j.cub.2006.04.024

Somel, M., Liu, X., Tang, L., Yan, Z., Hu, H., Guo, S., ... Khaitovich, P. (2011, dec). MicroRNA-driven developmental remodeling in the brain distinguishes humans from other primates. *PLoS biology*, *9*(12), e1001214. Retrieved from http://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.1001214 doi: 10.1371/journal.pbio.1001214

Somel, M., Rohlfs, R. & Liu, X. (2014, dec). Transcriptomic insights into human brain evolution: acceleration, neutrality, heterochrony. *Current opinion in genetics & development*, *29*, 110–9. Retrieved from http://www.sciencedirect.com/science/article/pii/S0959437X14000963 doi: 10.1016/

j.gde.2014.09.001

Southworth, L. K., Owen, A. B. & Kim, S. K. (2009, dec). Aging mice show a decreasing correlation of gene expression within genetic modules. *PLoS genetics*, *5*(12), e1000776. Retrieved from http://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.1000776 doi: 10.1371/journal.pgen.1000776

Sowell, E. R., Thompson, P. M. & Toga, A. W. (2004, aug). Mapping changes in the human cortex throughout the span of life. *The Neuroscientist : a review journal bringing neurobiology, neurology and psychiatry*, *10*(4), 372–92. Retrieved from http://nro.sagepub.com/content/10/4/372.abstract doi: 10.1177/1073858404263960

Stiles, J. & Jernigan, T. L. (2010, dec). The basics of brain development. *Neuropsychology review*, *20*(4), 327–48. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2989000{&}tool=pmcentrez{&}rendertype=abstract doi: 10.1007/s11065-010-9148-4

Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., ... Mesirov, J. P. (2005, oct). 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*, *102*(43), 15545–50. Retrieved from http://www.pnas.org/content/102/43/15545.abstract doi: 10.1073/pnas.0506580102

Sun, G., Shi, L., Yan, S., Wan, Z., Jiang, N., Fu, L., ... Guo, J. (2014, jan). MiR-15b targets cyclin D1 to regulate proliferation and apoptosis in glioma cells. *BioMed research international*, *2014*, 687826. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4065679{&}tool=pmcentrez{&}rendertype=abstract doi: 10.1155/2014/687826

Sun, T., Patoine, C., Abu-Khalil, A., Visvader, J., Sum, E., Cherry, T. J., ... Walsh, C. A. (2005, jun). Early asymmetry of gene transcription in embryonic human left and right cerebral cortex. *Science (New York, N.Y.)*, *308*(5729), 1794–8. Retrieved from http://science.sciencemag.org/content/308/5729/1794.abstract doi: 10.1126/science.1110324

Supek, F., Bošnjak, M., Škunca, N. & Šmuc, T. (2011, jan). REVIGO summarizes and visualizes long lists of gene ontology terms. *PloS one*, *6*(7), e21800. Retrieved from http://journals.plos.org/plosone/article?id=10.1371/

journal.pone.0021800 doi: 10.1371/journal.pone.0021800

Tennekes, M. (2015). treemap: Treemap visualization [Computer software manual]. Retrieved from http://CRAN.R-project.org/package=treemap (R package version 2.4)

The UniProt Consortium. (2014, oct). UniProt: a hub for protein information. *Nucleic Acids Research*, *43*(D1), D204–212. Retrieved from http://nar.oxfordjournals.org/content/43/D1/D204 doi: 10.1093/nar/gku989

Ujifuku, K., Mitsutake, N., Takakura, S., Matsuse, M., Saenko, V., Suzuki, K., ... Yamashita, S. (2010, oct). miR-195, miR-455-3p and miR-10a(*) are implicated in acquired temozolomide resistance in glioblastoma multiforme cells. *Cancer letters*, *296*(2), 241–8. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/20444541 doi: 10.1016/j.canlet.2010.04.013

van de Leemput, J., Boles, N. C., Kiehl, T. R., Corneo, B., Lederman, P., Menon, V., ... Fasano, C. A. (2014, jul). CORTECON: a temporal transcriptome analysis of in vitro human cerebral cortex development from human embryonic stem cells. *Neuron*, *83*(1), 51–68. Retrieved from http://www.sciencedirect.com/science/article/pii/S0896627314003997 doi: 10.1016/j.neuron.2014.05.013

Vincent, I., Rosado, M. & Davies, P. (1996, feb). Mitotic mechanisms in Alzheimer's disease? *The Journal of Cell Biology*, *132*(3), 413–425. Retrieved from http://jcb.rupress.org/content/132/3/413.abstract doi: 10.1083/jcb.132.3.413

Vlachos, I. S., Paraskevopoulou, M. D., Karagkouni, D., Georgakilas, G., Vergoulis, T., Kanellos, I., ... Hatzigeorgiou, A. G. (2015, jan). DIANA-TarBase v7.0: indexing more than half a million experimentally supported miRNA:mRNA interactions. *Nucleic acids research*, *43*(Database issue), D153–9. Retrieved from http://nar.oxfordjournals.org/content/43/D1/D153.long doi: 10.1093/nar/gku1215

Walker, R., Gurven, M., Hill, K., Migliano, A., Chagnon, N., De Souza, R., ... Yamauchi, T. (2006, jan). Growth rates and life histories in twenty-two small-scale societies. *American Journal of Human Biology*, *18*(3), 295–311. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/16634027 doi: 10.1002/ajhb.20510

Walker, R. M., Rybka, J., Anderson, S. M., Torrance, H. S., Boxall, R., Sussmann, J. E., ... Evans, K. L. (2015, mar). Prelimin-

ary investigation of miRNA expression in individuals at high familial risk of bipolar disorder. *Journal of psychiatric research*, *62*, 48–55. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4379383{&}tool=pmcentrez{&}rendertype=abstract doi: 10.1016/j.jpsychires.2015.01.006

Wei, T. (2013). corrplot: Visualization of a correlation matrix [Computer software manual]. Retrieved from http://CRAN.R-project.org/package=corrplot (R package version 0.73)

Williams, G. C. (1957). Pleiotropy, Natural Selection, and the Evolution of Senescence. *Evolution*, *11*(4), 398–411.

World Health Organization. (2012). *Comparison of leading causes of death over the past decade , 2000 and 2012.* Retrieved from http://www.who.int/mediacentre/factsheets/fs310/en/

World Health Organization. (2015). *World report on ageing and health*. World Health Organization. Retrieved from www.who.int/ doi: 10.1007/s13398-014-0173-7.2

Xue, H., Xian, B., Dong, D., Xia, K., Zhu, S., Zhang, Z., ... Han, J.-D. J. (2007, jan). A modular network model of aging. *Molecular systems biology*, *3*(1), 147. Retrieved from http://msb.embopress.org/content/3/1/147.abstract doi: 10.1038/msb4100189

Yadav, V. K., DeGregori, J. & De, S. (2016, feb). The landscape of somatic mutations in protein coding genes in apparently benign human tissues carries signatures of relaxed purifying selection. *Nucleic acids research*, *44*(5), 2075–2084. Retrieved from http://nar.oxfordjournals.org/content/44/5/2075.short?rss=1http://nar.oxfordjournals.org/content/44/5/2075.long doi: 10.1093/nar/gkw086

Yates, A., Akanni, W., Amode, M. R., Barrell, D., Billis, K., Carvalho-Silva, D., ... Flicek, P. (2015, dec). Ensembl 2016. *Nucleic acids research*, *44*(D1), D710–716. Retrieved from http://nar.oxfordjournals.org/content/44/D1/D710 doi: 10.1093/nar/gkv1157

Yordy, J. S., Li, R., Sementchenko, V. I., Pei, H., Muise-Helmericks, R. C. & Watson, D. K. (2004, aug). SP100 expression modulates ETS1 transcriptional activity and inhibits cell invasion. *Oncogene*, *23*(39), 6654–65. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/15247905 doi: 10.1038/sj.onc.1207891

Yuan, T., Jiao, Y., de Jong, S., Ophoff, R. A., Beck, S. & Teschendorff, A. E. (2015, feb). An integrative multi-scale analysis of the dynamic DNA methylation landscape in aging. *PLoS genetics*, *11*(2), e1004996. Retrieved from http://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.1004996 doi: 10.1371/journal.pgen.1004996

Yuan, Y., Chen, Y.-P. P., Boyd-Kirkup, J., Khaitovich, P. & Somel, M. (2012, oct). Accelerated aging-related transcriptome changes in the female prefrontal cortex. *Aging cell*, *11*(5), 894–901. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3470704{&}tool=pmcentrez{&}rendertype=abstract doi: 10.1111/j.1474-9726.2012.00859.x

Yun, J. & Finkel, T. (2014, may). Mitohormesis. *Cell metabolism*, *19*(5), 757–66. Retrieved from http://www.cell.com/article/S1550413114000175/fulltext doi: 10.1016/j.cmet.2014.01.011

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 genetics*, *3*(11), e201. Retrieved from http://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.0030201 doi: 10.1371/journal.pgen.0030201

Zahn, J. M., Sonu, R., Vogel, H., Crane, E., Mazan-Mamczarz, K., Rabkin, R., ... Kim, S. K. (2006, jul). Transcriptional Profiling of Aging in Human Muscle Reveals a Common Aging Signature. *PLoS Genetics*, *2*(7), e115. Retrieved from http://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.0020115 doi: 10.1371/journal.pgen.0020115

Zhang, C. & Cuervo, A. M. (2008, sep). Restoration of chaperone-mediated autophagy in aging liver improves cellular maintenance and hepatic function. *Nature medicine*, *14*(9), 959–65. Retrieved from http://dx.doi.org/10.1038/nm.1851 doi: 10.1038/nm.1851

APPENDIX A

LIST OF GO BIOLOGICAL PROCESS CATEGORIES WITH CONSISTENT OVER-REPRESENTATION OF DOWN-UP GENES

Table A.1: List of GO BP Categories having $OR > 1$ for Down-Up vs. Monotonic Decrease Comparison in all datasets. "↑" in the column names means increase in gene expression.

GO ID	GO Term	↑ in Dev	↑ in Age
GO:0000060	protein import into nucleus, transl	0.38 %	0.63 %
GO:0006907	pinocytosis	0.57 %	0.64 %
GO:0009262	deoxyribonucleotide metabolic proce	0.51 %	0.6 %
GO:0010939	regulation of necrotic cell death	0.44 %	0.71 %
GO:0030238	male sex determination	0.53 %	0.6 %
GO:0030279	negative regulation of ossification	0.41 %	0.54 %
GO:0032480	negative regulation of type I inter	0.58 %	0.56 %
GO:0033574	response to testosterone	0.44 %	0.6 %
GO:0043981	histone H4-K5 acetylation	0.61 %	0.55 %
GO:0043982	histone H4-K8 acetylation	0.61 %	0.55 %
GO:0045601	regulation of endothelial cell diff	0.27 %	0.68 %
GO:0048339	paraxial mesoderm development	0.26 %	0.63 %
GO:0072273	metanephric nephron morphogenesis	0.35 %	0.64 %
GO:2000826	regulation of heart morphogenesis	0.33 %	0.64 %

APPENDIX B

LIST OF GO BIOLOGICAL PROCESS CATEGORIES WITH CONSISTENT OVER-REPRESENTATION OF MONOTONICALLY DECREASING GENES

Table B.1: List of GO BP Categories having $OR < 1$ for Down-Up vs. Monotonic Decrease Comparison in all datasets. "↑" in the column names means increase in gene expression.

GO ID	GO Term	↑ in Dev	↑ in Age
GO:0000904	cell morphogenesis involved in diff	0.42 %	0.45 %
GO:0001505	regulation of neurotransmitter leve	0.5 %	0.43 %
GO:0001508	action potential	0.49 %	0.39 %
GO:0001774	microglial cell activation	0.49 %	0.4 %
GO:0002429	immune response-activating cell sur	0.51 %	0.45 %
GO:0006622	protein targeting to lysosome	0.67 %	0.41 %
GO:0006887	exocytosis	0.52 %	0.43 %
GO:0006904	vesicle docking involved in exocyto	0.65 %	0.32 %
GO:0006935	chemotaxis	0.43 %	0.46 %
GO:0007156	homophilic cell adhesion via plasma	0.4 %	0.35 %
GO:0007157	heterophilic cell-cell adhesion via	0.31 %	0.35 %
GO:0007158	neuron cell-cell adhesion	0.29 %	0.14 %
GO:0007215	glutamate receptor signaling pathwa	0.36 %	0.26 %
GO:0007269	neurotransmitter secretion	0.5 %	0.4 %
GO:0007409	axonogenesis	0.42 %	0.42 %
GO:0007411	axon guidance	0.41 %	0.42 %
GO:0007413	axonal fasciculation	0.29 %	0.18 %
GO:0008038	neuron recognition	0.27 %	0.19 %

GO:0008344	adult locomotory behavior	0.55 %	0.39 %
GO:0010923	negative regulation of phosphatase	0.53 %	0.38 %
GO:0017157	regulation of exocytosis	0.56 %	0.36 %
GO:0017158	regulation of calcium ion-dependent	0.5 %	0.31 %
GO:0018198	peptidyl-cysteine modification	0.51 %	0.46 %
GO:0021756	striatum development	0.53 %	0.28 %
GO:0021952	central nervous system projection n	0.28 %	0.29 %
GO:0021955	central nervous system neuron axono	0.37 %	0.37 %
GO:0022029	telencephalon cell migration	0.36 %	0.34 %
GO:0030030	cell projection organization	0.45 %	0.45 %
GO:0030316	osteoclast differentiation	0.49 %	0.46 %
GO:0031175	neuron projection development	0.44 %	0.42 %
GO:0035235	ionotropic glutamate receptor signa	0.37 %	0.23 %
GO:0035305	negative regulation of dephosphoryl	0.56 %	0.38 %
GO:0035308	negative regulation of protein deph	0.55 %	0.37 %
GO:0035336	long-chain fatty-acyl-CoA metabolic	0.39 %	0.38 %
GO:0035418	protein localization to synapse	0.27 %	0.24 %
GO:0042330	taxis	0.43 %	0.46 %
GO:0048013	ephrin receptor signaling pathway	0.39 %	0.38 %
GO:0048666	neuron development	0.44 %	0.43 %
GO:0048667	cell morphogenesis involved in neur	0.42 %	0.42 %
GO:0048812	neuron projection morphogenesis	0.43 %	0.42 %
GO:0048857	neural nucleus development	0.49 %	0.43 %
GO:0050805	negative regulation of synaptic tra	0.47 %	0.29 %
GO:0051648	vesicle localization	0.55 %	0.37 %
GO:0051650	establishment of vesicle localizati	0.55 %	0.37 %
GO:0051953	negative regulation of amine transp	0.41 %	0.33 %
GO:0060996	dendritic spine development	0.36 %	0.3 %
GO:0061462	protein localization to lysosome	0.68 %	0.4 %
GO:0061564	axon development	0.42 %	0.42 %
GO:0070588	calcium ion transmembrane transport	0.51 %	0.42 %
GO:0097485	neuron projection guidance	0.41 %	0.42 %
GO:0098609	cell-cell adhesion	0.44 %	0.46 %
GO:0098742	cell-cell adhesion via plasma-membr	0.37 %	0.38 %
GO:1900449	regulation of glutamate receptor si	0.37 %	0.29 %

APPENDIX C

LIST OF GO BIOLOGICAL PROCESS CATEGORIES WITH CONSISTENT OVER-REPRESENTATION OF UP-DOWN GENES

Table C.1: List of GO BP Categories having $OR > 1$ for Up-Down vs. Monotonic Increase Comparison in all datasets. "↑" in the column names means increase in gene expression.

GO ID	GO Term	↑ in Dev	↑ in Age
GO:0000209	protein polyubiquitination	0.51 %	0.44 %
GO:0000266	mitochondrial fission	0.67 %	0.32 %
GO:0000462	maturation of SSU-rRNA from tricist	0.59 %	0.32 %
GO:0001964	startle response	0.44 %	0.26 %
GO:0006464	cellular protein modification proce	0.49 %	0.48 %
GO:0006470	protein dephosphorylation	0.51 %	0.46 %
GO:0006511	ubiquitin-dependent protein catabol	0.54 %	0.46 %
GO:0006665	sphingolipid metabolic process	0.55 %	0.45 %
GO:0006672	ceramide metabolic process	0.53 %	0.46 %
GO:0006688	glycosphingolipid biosynthetic proc	0.47 %	0.39 %
GO:0006810	transport	0.5 %	0.48 %
GO:0006887	exocytosis	0.52 %	0.43 %
GO:0006900	membrane budding	0.51 %	0.41 %
GO:0006901	vesicle coating	0.55 %	0.4 %
GO:0006904	vesicle docking involved in exocyto	0.65 %	0.32 %
GO:0006986	response to unfolded protein	0.52 %	0.47 %
GO:0007006	mitochondrial membrane organization	0.53 %	0.45 %
GO:0007158	neuron cell-cell adhesion	0.29 %	0.14 %

GO:0007169	transmembrane receptor protein tyro	0.46 %	0.47 %
GO:0007173	epidermal growth factor receptor si	0.49 %	0.45 %
GO:0007270	neuron-neuron synaptic transmission	0.48 %	0.34 %
GO:0007409	axonogenesis	0.42 %	0.42 %
GO:0007411	axon guidance	0.41 %	0.42 %
GO:0007613	memory	0.47 %	0.37 %
GO:0007616	long-term memory	0.51 %	0.23 %
GO:0007626	locomotory behavior	0.49 %	0.39 %
GO:0008038	neuron recognition	0.27 %	0.19 %
GO:0008104	protein localization	0.51 %	0.48 %
GO:0008286	insulin receptor signaling pathway	0.51 %	0.47 %
GO:0008543	fibroblast growth factor receptor s	0.48 %	0.45 %
GO:0010288	response to lead ion	0.48 %	0.36 %
GO:0015012	heparan sulfate proteoglycan biosyn	0.43 %	0.28 %
GO:0015693	magnesium ion transport	0.71 %	0.36 %
GO:0016241	regulation of macroautophagy	0.58 %	0.43 %
GO:0016242	negative regulation of macroautopha	0.63 %	0.29 %
GO:0017157	regulation of exocytosis	0.56 %	0.36 %
GO:0019941	modification-dependent protein cata	0.54 %	0.46 %
GO:0021695	cerebellar cortex development	0.47 %	0.46 %
GO:0022008	neurogenesis	0.43 %	0.45 %
GO:0022406	membrane docking	0.59 %	0.35 %
GO:0030030	cell projection organization	0.45 %	0.45 %
GO:0030111	regulation of Wnt signaling pathway	0.42 %	0.49 %
GO:0030148	sphingolipid biosynthetic process	0.57 %	0.42 %
GO:0030201	heparan sulfate proteoglycan metabo	0.37 %	0.32 %
GO:0030534	adult behavior	0.48 %	0.38 %
GO:0031175	neuron projection development	0.44 %	0.42 %
GO:0031400	negative regulation of protein modi	0.49 %	0.48 %
GO:0031669	cellular response to nutrient level	0.54 %	0.48 %
GO:0032456	endocytic recycling	0.55 %	0.32 %
GO:0032989	cellular component morphogenesis	0.45 %	0.46 %
GO:0032990	cell part morphogenesis	0.45 %	0.45 %
GO:0033036	macromolecule localization	0.51 %	0.49 %
GO:0034199	activation of protein kinase A acti	0.51 %	0.3 %
GO:0034613	cellular protein localization	0.5 %	0.48 %
GO:0035249	synaptic transmission, glutamatergi	0.4 %	0.35 %
GO:0035964	COPI-coated vesicle budding	0.62 %	0.27 %
GO:0036211	protein modification process	0.49 %	0.48 %

GO:0038093	Fc receptor signaling pathway	0.51 %	0.45 %
GO:0038095	Fc-epsilon receptor signaling pathw	0.5 %	0.45 %
GO:0038127	ERBB signaling pathway	0.48 %	0.45 %
GO:0038179	neurotrophin signaling pathway	0.47 %	0.45 %
GO:0042220	response to cocaine	0.44 %	0.35 %
GO:0043161	proteasome-mediated ubiquitin-depen	0.53 %	0.46 %
GO:0043266	regulation of potassium ion transpo	0.56 %	0.4 %
GO:0043412	macromolecule modification	0.49 %	0.49 %
GO:0043632	modification-dependent macromolecul	0.54 %	0.46 %
GO:0044070	regulation of anion transport	0.5 %	0.41 %
GO:0044344	cellular response to fibroblast gro	0.48 %	0.45 %
GO:0044765	single-organism transport	0.5 %	0.48 %
GO:0045921	positive regulation of exocytosis	0.59 %	0.37 %
GO:0046467	membrane lipid biosynthetic process	0.57 %	0.43 %
GO:0046488	phosphatidylinositol metabolic proc	0.6 %	0.42 %
GO:0046513	ceramide biosynthetic process	0.52 %	0.42 %
GO:0046888	negative regulation of hormone secr	0.48 %	0.44 %
GO:0046907	intracellular transport	0.52 %	0.48 %
GO:0048011	neurotrophin TRK receptor signaling	0.48 %	0.45 %
GO:0048194	Golgi vesicle budding	0.55 %	0.31 %
GO:0048200	Golgi transport vesicle coating	0.62 %	0.27 %
GO:0048205	COPI coating of Golgi vesicle	0.62 %	0.27 %
GO:0048278	vesicle docking	0.62 %	0.34 %
GO:0048468	cell development	0.44 %	0.46 %
GO:0048666	neuron development	0.44 %	0.43 %
GO:0048667	cell morphogenesis involved in neur	0.42 %	0.42 %
GO:0048699	generation of neurons	0.43 %	0.45 %
GO:0048812	neuron projection morphogenesis	0.43 %	0.42 %
GO:0048858	cell projection morphogenesis	0.45 %	0.45 %
GO:0050805	negative regulation of synaptic tra	0.47 %	0.29 %
GO:0050808	synapse organization	0.42 %	0.36 %
GO:0050890	cognition	0.45 %	0.38 %
GO:0051234	establishment of localization	0.5 %	0.48 %
GO:0051641	cellular localization	0.5 %	0.48 %
GO:0051648	vesicle localization	0.55 %	0.37 %
GO:0051649	establishment of localization in ce	0.5 %	0.48 %
GO:0051650	establishment of vesicle localizati	0.55 %	0.37 %
GO:0051932	synaptic transmission, GABAergic	0.62 %	0.3 %
GO:0060359	response to ammonium ion	0.46 %	0.37 %

GO:0060627	regulation of vesicle-mediated tran	0.52 %	0.44 %
GO:0061564	axon development	0.42 %	0.42 %
GO:0070727	cellular macromolecule localization	0.5 %	0.48 %
GO:0071377	cellular response to glucagon stimu	0.4 %	0.35 %
GO:0071702	organic substance transport	0.51 %	0.49 %
GO:0071774	response to fibroblast growth facto	0.48 %	0.45 %
GO:0072657	protein localization to membrane	0.46 %	0.48 %
GO:0090140	regulation of mitochondrial fission	0.79 %	0.28 %
GO:0097484	dendrite extension	0.59 %	0.22 %
GO:0097485	neuron projection guidance	0.41 %	0.42 %
GO:1901379	regulation of potassium ion transme	0.55 %	0.41 %
GO:1901881	positive regulation of protein depo	0.33 %	0.41 %
GO:1902578	single-organism localization	0.5 %	0.48 %
GO:1902580	single-organism cellular localizati	0.49 %	0.48 %
GO:1902591	single-organism membrane budding	0.53 %	0.39 %
GO:1903859	regulation of dendrite extension	0.57 %	0.2 %
GO:1903861	positive regulation of dendrite ext	0.57 %	0.2 %
GO:2000785	regulation of autophagosome assembl	0.59 %	0.35 %

APPENDIX D

LIST OF GO BIOLOGICAL PROCESS CATEGORIES WITH CONSISTENT OVER-REPRESENTATION OF MONOTONICALLY INCREASING GENES

Table D.1: List of GO BP Categories having $OR < 1$ for Up-Down vs. Monotonic Increase Comparison in all datasets. "↑" in the column names means increase in gene expression.

GO ID	GO Term	↑ in Dev	↑ in Age
GO:0000272	polysaccharide catabolic process	0.46 %	0.66 %
GO:0001523	retinoid metabolic process	0.41 %	0.55 %
GO:0001580	detection of chemical stimulus invo	0.48 %	0.7 %
GO:0001937	negative regulation of endothelial	0.37 %	0.56 %
GO:0005980	glycogen catabolic process	0.45 %	0.65 %
GO:0006026	aminoglycan catabolic process	0.32 %	0.6 %
GO:0006027	glycosaminoglycan catabolic process	0.31 %	0.59 %
GO:0006721	terpenoid metabolic process	0.42 %	0.54 %
GO:0006801	superoxide metabolic process	0.52 %	0.63 %
GO:0006956	complement activation	0.45 %	0.69 %
GO:0009251	glucan catabolic process	0.45 %	0.65 %
GO:0010664	negative regulation of striated mus	0.53 %	0.63 %
GO:0010939	regulation of necrotic cell death	0.44 %	0.71 %
GO:0017144	drug metabolic process	0.49 %	0.66 %
GO:0021511	spinal cord patterning	0.39 %	0.69 %
GO:0021513	spinal cord dorsal/ventral patterni	0.41 %	0.7 %
GO:0021871	forebrain regionalization	0.27 %	0.61 %
GO:0042743	hydrogen peroxide metabolic process	0.46 %	0.63 %

GO:0042994	cytoplasmic sequestering of transcr	0.48 %	0.65 %
GO:0044247	cellular polysaccharide catabolic p	0.47 %	0.66 %
GO:0044724	single-organism carbohydrate catabo	0.58 %	0.59 %
GO:0045669	positive regulation of osteoblast d	0.32 %	0.56 %
GO:0045765	regulation of angiogenesis	0.44 %	0.57 %
GO:0045778	positive regulation of ossification	0.35 %	0.56 %
GO:0046031	ADP metabolic process	0.65 %	0.58 %
GO:0046939	nucleotide phosphorylation	0.61 %	0.58 %
GO:0048562	embryonic organ morphogenesis	0.38 %	0.57 %
GO:0050912	detection of chemical stimulus invo	0.47 %	0.68 %
GO:0072348	sulfur compound transport	0.41 %	0.63 %
GO:0072376	protein activation cascade	0.45 %	0.65 %
GO:0072529	pyrimidine-containing compound cata	0.49 %	0.68 %
GO:1901342	regulation of vasculature developme	0.45 %	0.57 %
GO:1901605	alpha-amino acid metabolic process	0.55 %	0.57 %
GO:1904037	positive regulation of epithelial c	0.41 %	0.64 %
GO:2000696	regulation of epithelial cell diffe	0.28 %	0.6 %

APPENDIX E

LIST OF GO BIOLOGICAL PROCESS CATEGORIES WITH CONSISTENT OVER-REPRESENTATION OF UP-DOWN AND MONOTONICALLY DECREASING GENES

Table E.1: List of GO BP Categories having $OR > 1$ for Up-Down vs. Monotonic Increase Comparison and $OR < 1$ for Down-Up vs. Monotonic Decrease Comparison in all datasets. "↑" in the column names means increase in gene expression.

GO ID	GO Term	↑ in Dev	↑ in Age
GO:0006887	exocytosis	0.52 %	0.43 %
GO:0006904	vesicle docking involved in exocyto	0.65 %	0.32 %
GO:0007158	neuron cell-cell adhesion	0.29 %	0.14 %
GO:0007409	axonogenesis	0.42 %	0.42 %
GO:0007411	axon guidance	0.41 %	0.42 %
GO:0008038	neuron recognition	0.27 %	0.19 %
GO:0017157	regulation of exocytosis	0.56 %	0.36 %
GO:0030030	cell projection organization	0.45 %	0.45 %
GO:0031175	neuron projection development	0.44 %	0.42 %
GO:0048666	neuron development	0.44 %	0.43 %
GO:0048667	cell morphogenesis involved in neur	0.42 %	0.42 %
GO:0048812	neuron projection morphogenesis	0.43 %	0.42 %
GO:0050805	negative regulation of synaptic tra	0.47 %	0.29 %
GO:0051648	vesicle localization	0.55 %	0.37 %
GO:0051650	establishment of vesicle localizati	0.55 %	0.37 %
GO:0061564	axon development	0.42 %	0.42 %
GO:0097485	neuron projection guidance	0.41 %	0.42 %