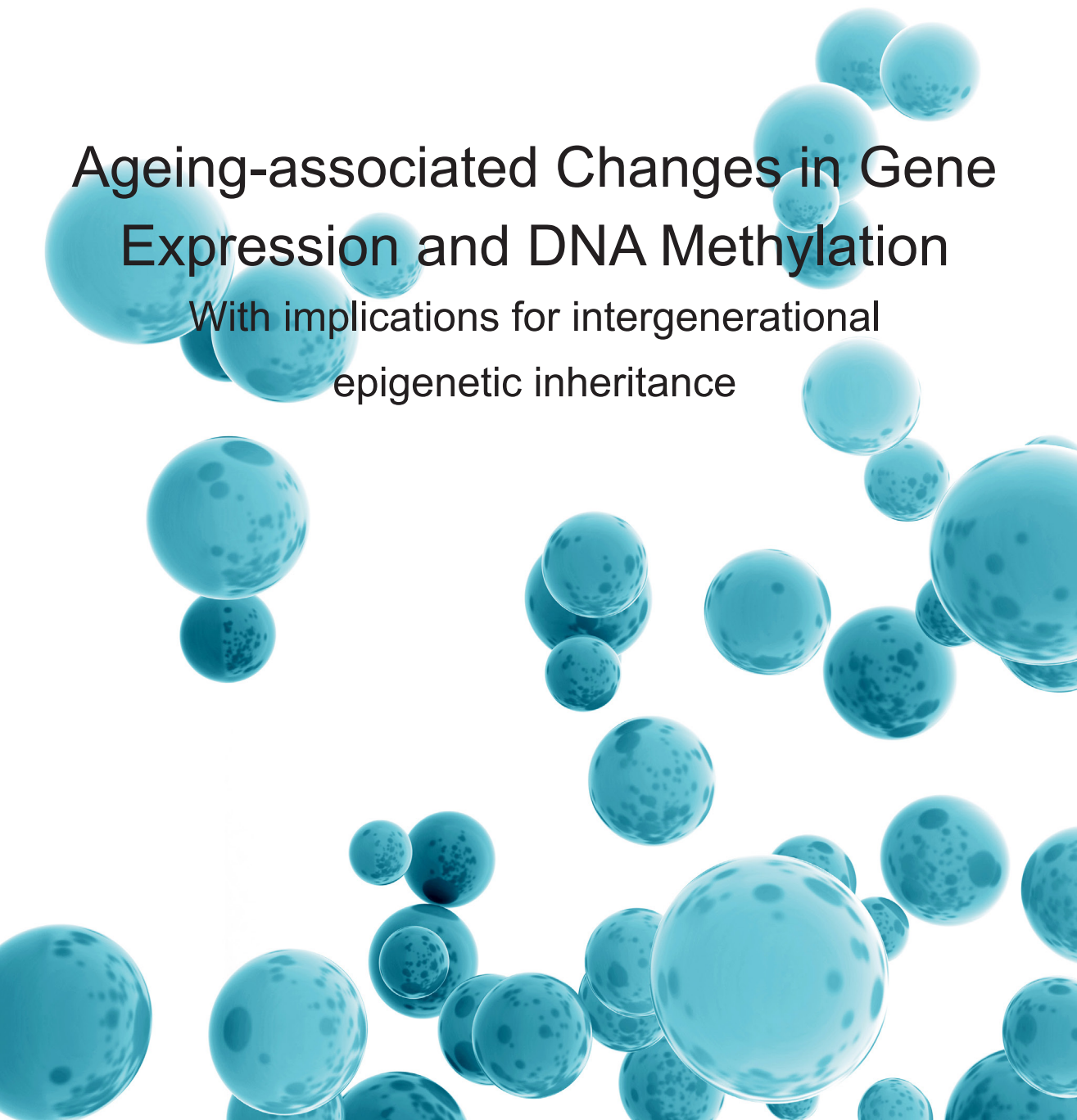
Acta Universitatis Tamperensis 2150

SAARA MARTTILA

Ageing-associated Changes in Gene Expression and DNA Methylation

With implications for intergenerational epigenetic inheritance



SAARA MARTTILA

Ageing-associated Changes in Gene Expression and DNA Methylation

With implications for intergenerational epigenetic inheritance

ACADEMIC DISSERTATION
To be presented, with the permission of
the Board of the School of Medicine of the University of Tampere,
for public discussion in the small auditorium of building M,
Pirkanmaa Hospital District, Teiskontie 35, Tampere,
on 20 May 2016, at 12 o'clock.

UNIVERSITY OF TAMPERE

SAARA MARTTILA

Ageing-associated Changes in Gene Expression and DNA Methylation

With implications for intergenerational epigenetic inheritance

Acta Universitatis Tamperensis 2150
Tampere University Press
Tampere 2016

UNIVERSITY
OF TAMPERE

ACADEMIC DISSERTATION
University of Tampere, School of Medicine
Tampere Graduate Program in Biomedicine and Biotechnology
Finland

Supervised by
Professor Mikko Hurme
University of Tampere
Finland

Reviewed by
Docent Antero Salminen
University of Eastern Finland
Finland
Docent Cilla Söderhäll
Karolinska Institutet
Sweden

The originality of this thesis has been checked using the Turnitin OriginalityCheck service in accordance with the quality management system of the University of Tampere.

Cover design by
Mikko Reinikka

Distributor:
verkkokauppa@juvenesprint.fi
https://verkkokauppa.juvenes.fi

Acta Universitatis Tamperensis 2150
ISBN 978-952-03-0072-2 (print)
ISSN-L 1455-1616
ISSN 1455-1616

Acta Electronica Universitatis Tamperensis 1649
ISBN 978-952-03-0073-9 (pdf)
ISSN 1456-954X
http://tampub.uta.fi

Suomen Yliopistopaino Oy – Juvenes Print
Tampere 2016

Muorille

Contents

List of Original Communications 8

Abbreviations 9

Abstract 12

Tiivistelmä 14

1 Introduction 17

2 Review of the Literature 19

2.1 Ageing 19

2.1.1 Immunosenescence and inflamm-aging 20

2.1.2 Manifestation of ageing at the cellular and molecular level 22

2.1.3 Ageing theories 25

2.2 Lifespan and longevity 27

2.2.1 Sexual dimorphism of lifespan and longevity 27

2.2.2 Heritability of lifespan and longevity 30

2.2.3 Genetics of longevity 31

2.2.4 Extension of lifespan 32

2.3 Epigenetics and transcriptomics 34

2.3.1 DNA methylation 34

2.3.2 Regulatory functions of DNA methylation 38
2.3.3 Ageing-associated changes in DNA methylation 40
2.3.4 Other epigenetic mechanisms and ageing-associated changes 43
2.3.5 Ageing-associated gene expression changes 45
2.4 Transgenerational and intergenerational inheritance 46
2.4.1 Transgenerational epigenetic inheritance in model organisms 47
2.4.2 Trans- and intergenerational inheritance in humans 49
2.4.3 Epigenetic reprogramming during development 50
3 Aims of the Study 52
4 Materials and Methods 53
4.1 Study subjects 53
4.1.1 Vitality 90+ study (I, II & IV) 53
4.1.2 Young Finns Study (III) 54
4.2 Methods 54
4.2.1 Sample collection 54
4.2.2 Extraction of DNA and RNA 55
4.2.3 Determination of cell type proportions 55
4.2.4 Gene expression analysis (I & II) 56
4.2.5 DNA methylation analysis (II, III & IV) 58
4.2.6 Availability of microarray data 60
4.2.7 Pathway analyses 60

4.2.8 Ethics 61

5 Results 62

5.1 Ageing-associated gene expression changes (I) 62

5.1.1 Differentially expressed individual genes 62

5.1.2 Pathways associated with differentially expressed genes 64

5.2 Ageing-associated changes in DNA methylation (II & III) 66

5.2.1 CpG sites differentially methylated with advancing age 66

5.2.2 Location of a-CpGs 68

5.2.3 Functions of a-CpGs 70

5.2.4 Ageing-associated DNA methylation changes and sex 76

5.2.5 Ageing-associated DNA methylation changes and differences in cell type proportions 77

5.3 Association between ageing-associated DNA methylation changes and gene expression (I & II) 79

5.4 Manifestation of parental lifespan in the DNA methylation profile of the progeny (IV) 81

6 Discussion 86

6.1 Ageing-associated gene expression changes (I) 86

6.2 Ageing-associated DNA methylation changes (II & III) 88

6.3 The relationship between ageing-associated DNA methylation changes and gene expression (I & II) 93

6.4 Effects of parental lifespan on the DNA methylation profile of the progeny (IV) 94

6.5 Role of developmental pathways in ageing and longevity (II, III & IV) 96

6.6 Limitations of the study 97

6.6.1 Study population 97

6.6.2 Gene expression analysis 97

6.6.3 DNA methylation analysis 98

7 Summary and Conclusions 100

8 Acknowledgements 102

9 References 104

10 Original Communications 128

List of Original Communications

This dissertation is based on the following original communications, referred to in the text with their roman numerals (I-IV). In addition, this thesis contains unpublished data.

I Marttila S, Jylhävä J, Nevalainen T, Nykter M, Jylhä M, Hervonen A, Tserel L, Peterson P, Hurme M. Transcriptional analysis reveals gender-specific changes in the aging of the human immune system. PLoS One. 2013 Jun 11;8(6)e66229.

II Marttila S, Kananen L, Häyrynen S, Jylhävä J, Nevalainen T, Hervonen A, Jylhä M, Nykter M, Hurme M. Ageing-associated changes in the human DNA methylome: genomic locations and effects on gene expression. BMC Genomics. 2015 Mar 14;16:179.

III Kananen L, Marttila S, Nevalainen T, Jylhävä J, Mononen N, Kähönen M, Raitakari OT, Lehtimäki T, Hurme M. Aging-associated DNA methylation changes in middle-aged individuals: the Young Finns Study. BMC Genomics. 2016 Feb 9;17:103.

IV Marttila S, Kananen L, Jylhävä J, Nevalainen T, Hervonen A, Jylhä M, Hurme M. Length of paternal lifespan is manifested in the DNA methylome of their nonagenarian progeny. Oncotarget. 2015 Oct 13;6(31):30557-67.

Abbreviations

5caC	5-carboxylcytosine
5fC	5-formylcytosine
5hmC	5-hydroxymethylcytosine
5mC	5-methylcytosine
a-CpG	ageing-associated CpG site
AID	activation-induced cytidine deaminase
APOBEC	apolipoprotein B mRNA editing enzyme, catalytic polypeptide
APOE	apolipoprotein E
BER	base excision repair
BH	Benjamini-Hochberg
bp	base pair
BPA	bisphenol-A
CCR7	chemokine (C-C motif) receptor 7
cDNA	complementary DNA
CGI	CpG island
CpG	cytosine-phosphate-guanine, sequence of CG in a nucleotide chain
CR	calorie restriction
CRP	C-reactive protein
CV	coefficient of variation
DNMT1	DNA (cytosine-5-)-methyltransferase 1
DNMT3A	DNA (cytosine-5-)-methyltransferase 3 alpha
DNMT3B	DNA (cytosine-5-)-methyltransferase 3 beta
DNMT3L	DNA (cytosine-5-)-methyltransferase 3-like
DR	dietary restriction
EGFP	enhanced green fluorescent protein
ELOVL2	fatty acid elongase 2
ERV	endogenous retrovirus
FC	fold change

FDA	Food and Drug Administration
FDR	false discovery rate
GEO	Gene Expression Omnibus
GO	gene ontology
GOrilla	Gene Ontology enRIchment anaLysis and visuaLizAtion tool
GWAS	genome wide association study
HAT	histone acetyl transferase
HDAC	histone deacetylase
HP1α	heterochromatin protein 1α
HPA-axis	hypothalamic–pituitary–adrenal-axis
IFN-γ	interferon γ
IGF-1	insulin like growth factor 1
IGF2R	insulin-like growth factor 2 receptor
IIS	insulin and IGF-1 signalling
IL-6	interleukin 6
IPA	QIAGEN's Ingenuity® pathway analysis
LEF1	lymphoid enhancer-binding factor 1
LINE	long interspersed element
lncRNA	long non-coding RNA
LRRN3	leucine rich repeat neuronal 3
m^6A	N^6-methyl-adenosine
miRNA	micro-RNA
MMSE	Mini Mental State Examination
mRNA	messenger RNA
mTOR	mechanistic target of rapamycin
mTORC1	mTOR complex 1
mTORC2	mTOR complex 2
ncRNA	non-coding RNA
NGS	next generation sequencing
NK cells	natural killer cells
non-CGI	non-CpG island
PBMC	peripheral blood mononuclear cell
PC	principal component
PCA	principal component analysis

PGC	primordial germ cell
PUFA	polyunsaturated fatty acid
qPCR	quantitative polymerase chain reaction
ROS	reactive oxygen species
rsn	robust spline normalization
RQ	relative quantification
SAH	S-adenosyl-L-homocysteine
SAM	S-adenosyl-L-methionine
SIRT	sirtuin
SNP	single nucleotide polymorphism
TDG	thymine DNA glycosylase
TET	ten-eleven translocation
TET1	tet methylcytosine dioxygenase 1
TET2	tet methylcytosine dioxygenase 2
TET3	tet methylcytosine dioxygenase 3
TF	transcription factor
TNF-α	tumor necrosis factor α
TSS	transcription start site
TSS200	genomic region 200 bp upstream of TSS
TSS1500	genomic region 1500 bp upstream of TSS
UTR	untranslated region
V90+	Vitality 90+ study
vst	variance stabilizing transformation
YFS	Cardiovascular Risk in Young Finns Study

Abstract

Ageing can be defined as the decreasing functionality and compromised homeostasis of cells and tissues, leading to decreased functionality, increased morbidity and, eventually, to death. The detrimental phenotypic changes associated with ageing are believed to be due to changes at the molecular level, including changes in gene expression and in epigenetic mechanisms, such as DNA methylation. The expression of protein coding genes as well as non-coding RNAs (ncRNAs) is affected, and both global hypomethylation and promoter-specific hypermethylation are known to occur.

The aims of this study were to identify ageing-associated gene expression changes in nonagenarians (I), to identify ageing-associated DNA methylation changes in nonagenarians and to analyse how these changes are associated with the level of gene expression (II), to identify ageing-associated DNA methylation changes in middle-aged individuals (III) and to investigate whether parental lifespan manifests itself in the DNA methylation profile of progeny (IV).

The studies were conducted in two populations, namely, the Vitality 90+ (I, II & IV) and the Young Finns Study (III). Using commercial array techniques, we analysed gene expression levels using peripheral blood mononuclear cells (I) and the DNA methylation level from the same cells (II & IV) or from whole blood (III). Gene expression was analysed with Illumina HumanHT-12 v4 BeadChip, and DNA methylation was analysed with Illumina Infinium HumanMethylation450 BeadChip. The data were primarily analysed with the R programming language as well as SPSS and Chipster software. In addition, bioinformatic tools were used to identify enriched GO terms and canonical pathways.

Our results indicate that the ageing-associated changes in gene expression differ between males and females. Genes where the level of expression was associated with age were associated primarily with immune system functions (I). CpG sites differentially methylated with age in our study (II) were unequally distributed across the genome, with hypermethylation being enriched in CpG-islands (CGIs) and regions adjacent to transcription start sites (TSSs). The identified ageing-associated hyper- and hypomethylation differ also in terms of

associated genes; hypermethylated genes were associated with development and morphogenesis as well as DNA binding and transcription, whereas hypomethylated genes did not cluster to any specific process. That hyper- and hypomethylation differ in terms of location and associated genes during ageing implies that the causes and consequences of these processes also differ. The results that were obtained from nonagenarians (II) and from middle-aged population (III) were highly similar. The association between identified ageing-associated DNA methylation changes and gene expression in nonagenarians was poor (I & II). In the last study, we identified DNA methylation sites where the methylation level is associated with paternal lifespan even at the age of 90. These sites were primarily located outside of CGIs, and the genes harbouring these sites were associated with cell signalling as well as development and morphogenesis (IV).

Our results further confirm the role of immune system changes and sexual dimorphism in the ageing process. The results of DNA methylation analysis support previously reported findings and underline the complex nature of ageing-associated epigenetic changes. These results are also the first to show that the entire parental lifespan affects the DNA methylation profile of the progeny. As ageing-associated hyper- and hypomethylation show distinct features, we propose that ageing-associated hypermethylation is due to programmed changes, whereas ageing-associated hypomethylation appears to be due to environmental and stochastic effects. Methylation changes in genes associated with developmental processes were identified to be associated with both ageing and paternal lifespan, supporting the hyperfunction theory of ageing.

Tiivistelmä

Vanheneminen on solujen ja kudosten rakenteen ja toiminnan heikkenemistä, joka johtaa toimintakyvyn vajauksiin, sairastuvuuden lisääntymiseen ja lopulta kuolemaan. Vanhenemiseen liittyvien haitallisten muutosten taustalla ajatellaan olevan muutokset solu- ja molekyylitasolla, muun muassa geenien ilmenemisessä ja sitä säätelevissä epigeneettisissä mekanismeissa, kuten DNA:n metylaatiossa. Sekä proteiineja koodaavien geenien että muiden RNA-molekyylien ilmenemisessä tapahtuu muutoksia. DNA metylaation tiedetään vähenevän globaalisti (hypometylaatio) ja toisaalta lisääntyvän tietyillä promoottorialueilla (hypermetylaatio).

Tämän tutkimuksen tarkoitus oli tunnistaa ikääntymiseen liittyviä geeni-ilmenemisen muutoksia 90-vuotiailla (I), tunnistaa ikääntymiseen liittyviä DNA:n metylaation muutoksia 90-vuotiailla ja selvittää kuinka ne ovat yhteydessä geeni-ilmenemisen tasoon (II), tunnistaa ikääntymiseen liittyviä DNA:n metylaation muutoksia keski-ikäisessä väestössä (III) ja kartoittaa vanhempien eliniän mahdolliset vaikutukset jälkeläisten DNA:n metylaatioprofiiliin (IV).

Tutkimukset toteutettiin kahdessa aineistossa: Tervaskannot 90+ (I, II & IV) ja Lasten ja Nuorten Sepelvaltimotaudin riskitekijät (III). Näytteistä määritettiin geeni-ilmenemisen taso veren mononukleaarisista valkosoluista (I), DNA:n metylaatioaste samoista soluista (II & IV) tai DNA:n metylaatioaste kokoverestä (III) kaupallisilla sirutekniikoilla. Geeni-ilmeneminen määritettiin Illuminan HumanHT12 v4 BeadChip:llä ja DNA:n metylaatio Illuminan Infinium HumanMethylation450 BeadChip:llä. Tulokset analysoitiin pääasiassa R-ohjelmointikielellä sekä SPSS- ja Chipster-ohjelmistoilla. Lisäksi käytettiin GO-termejä ja erilaisia signalointireittejä tunnistavia bioinformaattisia työkaluja.

Tuloksemme osoittavat, että vanhenemiseen liittyvät geeni-ilmenemisen muutokset poikkeavat toisistaan naisten ja miesten välillä. Nämä muutokset liittyivät pääasiassa immuunijärjestelmän toimintaan (I). Tunnistamamme ikääntymiseen liittyvät DNA:n metylaation muutokset (II) eivät jakaudu tasaisesti ympäri genomia, vaan iän myötä hypermetyloituneet kohdat keskittyvät CpG-saariin ja lähelle transkription aloituskohtia. Lisäksi geenit, joissa iän myötä

tapahtuu hyper- tai hypometylaatiota, poikkeavat toisistaan. Hypermetyloituneet geenit liittyvät DNA:han sitoutumiseen ja transkription aloitukseen sekä kasvuun ja kehitykseen, kun taas iän myötä hypometyloituvat geenit eivät muodosta yhtenäistä ryhmää. Voidaan olettaa että hyper- ja hypometylaatioon johtavat tapahtumat ovat erilaisia ja siten myös näiden prosessien seuraukset poikkeavat toisistaan. Huomionarvoista oli, että vanhenemismuutokset 90-vuotialla ja keski-ikäisillä olivat suurelta osin samansuuntaisia ja samoihin prosesseihin keskittyneitä (II & III). Tuloksemme osoittavat myös, että vanhenemiseen liittyvien DNA:n metylaation muutosten yhteys geeni-ilmenemisen tasoon on heikko (I & II). Neljännessä osatyössä tunnistettiin joukko metylaatiokohtia, joissa metylaatioaste on yhteydessä isän eliniän pituuteen vielä 90-vuotiaanakin. Nämä metylaatiokohdat sijaitsivat pääasiassa CpG-saarien ulkopuolella. Geenit, joiden alueella nämä muutokset sijaitsivat, liittyivät solusignalointiin sekä kasvuun ja kehitykseen (IV).

Tuloksemme vahvistavat, että ikääntymiseen liittyy muutoksia immuunijärjestelmässä ja että sekä immuunijärjestelmän toiminta että vanheneminen poikkeavat toisistaan sukupuolten välillä. DNA:n metylaatiota koskevat tuloksemme vahvistavat aiempia havaintoja. Huomionarvoista on vanhenemiseen liittyvien epigeneettisten muutosten monipuolisuus ja monimutkaisuus. Tuloksissamme myös osoitetaan ensimmäistä kertaa vanhemman koko elinkaareen pituuden merkitys jälkeläisen DNA:n metylaatioprofiiliin. Ikääntymiseen liittyvä hyper- ja hypometylaatio poikkeavat selvästi sekä toiminnaltaan että sijainniltaan. Esitämme, että ikääntymiseen liittyvä hypermetylaatio on säädellyn prosessin tulosta, kun taas hypometylaation taustalla vaikuttavat ympäristötekijät ja sattuma. Kasvua ja kehitystä säätelevissä geeneissä tunnistettiin sekä ikääntymiseen että isän elinikään liittyviä muutoksia DNA:n metylaatiossa. Nämä löydökset tukevat vanhenemisen hyperfunktio-teoriaa.

1 Introduction

Ageing has been defined in various ways, but in its essence, it is the declining function and integrity of cells, tissues and organs that leads to an increased risk of diseases and disabilities and eventually to death (Kirkwood, 2005; López-Otín et al., 2013; Moskalev et al., 2014; Rose et al., 2012). As the human lifespan has increased, and with it the proportion of old individuals, the detrimental effects of ageing concern an ever growing population. While ageing affects all individuals, the manifestations and speed of this process vary greatly. Understanding of the molecular mechanisms behind ageing-associated changes would help in the understanding of why some individuals are relatively healthy on their 90th birthday while others fail to see their 75th.

Epigenetic features, including DNA methylation, can be inherited through cell division and, in some cases, from parent to progeny, but such features also constantly change during an individual's lifespan. Epigenetic mechanisms have an effect on gene expression but do not alter the underlying DNA sequence. DNA methylation can control the expression of single genes but also silence large sections of chromatin. The methylated cytosine base is occasionally called the fifth base of DNA, underlining its important role in the regulation of gene expression (D'Aquila et al., 2013).

Ageing is known to be associated with changes in both gene expression and DNA methylation. Ageing affects the expression level of various protein coding genes as well as that of small non-coding RNAs (López-Otín et al., 2013). In addition, heterogeneity of gene expression has been shown to be increased with ageing (Bahar et al., 2006). Ageing is also characterised by global loss of methyl groups i.e. hypomethylation of the genome, but individual regions, primarily promoter sequences, are known to acquire methyl groups i.e. be hypermethylated with ageing (Zampieri et al., 2015).

Epigenetics is also strongly linked with inheritance of acquired traits. The concept of the inheritance of acquired traits was completely rejected after it was first proposed by Lamarck in the 19th century, but recent evidence suggests it is possible even in mammals (Anway et al., 2005; Martos et al., 2015). It is believed that DNA methylation, along with other epigenetic mechanisms, mediates the

inheritance of acquired traits (Grossniklaus et al., 2013; Heard & Martienssen, 2014).

We have studied ageing-associated gene expression and DNA methylation changes. In addition, the possibility of an effect of parental lifespan on the DNA methylation profile of the progeny was analysed. We were able to adjust for differences in cell type proportions in the DNA methylation analysis and to compare gene expression and DNA methylation data obtained from the same samples.

More specifically, in the first study, the ageing-associated gene expression changes between nonagenarians and young controls (aged 19 to 30 years) were analysed. In the second study, ageing-associated DNA methylation changes were analysed in the same population as in the first study. We also analysed the associations between the identified DNA methylation changes and gene expression. The aim of the third study was to characterise which ageing-associated DNA methylation changes can be identified also in a middle-aged population. In the final study, we sought to identify methylation sites in progeny that displayed an association with parental lifespan.

2 Review of the Literature

2.1 Ageing

Ageing is a process that occurs with passing time and is characterised by the declining function and integrity of cells, tissues and organs, leading to a diminished ability to respond to environmental and intrinsic challenges. This process manifests as an increased risk of diseases and disabilities and leads to death (Kirkwood, 2005; López-Otín et al., 2013; Moskalev et al., 2014; Rose et al., 2012). The pace and manifestation of the ageing process vary greatly between individuals, but it is nevertheless unavoidable in mammals, including humans, as well as in the great majority of other species (Kirkwood, 2005). Ageing is not a process of the old or the elderly; depending on the definition of ageing, it begins after maturation or at birth (or even at conception), and ageing-associated changes can be quantified beginning in early adulthood (Figure 1) (Belsky et al., 2015; Salthouse, 2009).

Figure 1. Female of Western European descent, aged (A) 17 years and (B) 85 years, showing typical facial features of ageing.

Ageing is associated with changes in all tissues and organ systems. These changes include a loss of muscle mass (sarcopenia), increased adiposity, decreased bone density (osteopenia), a loss of skin elasticity, neurodegeneration, declining cognitive functions and impaired immune system function, among other effects (Hunt et al., 2010). Frailty is a term that can be used to define individuals with several of these ageing-associated impairments. Frail individuals are at the limit of their physiological reserves and are at increased risk of death, institutionalisation and disability (Fried et al., 2001; Hubbard & Woodhouse, 2010).

Increased age is also a major risk factor for a majority of common diseases and disorders, such as cancer, diabetes mellitus, cardiovascular diseases and neurodegenerative conditions (Kolovou et al., 2014; Partridge 2010). With increasing age, mortality from infectious diseases also increases. Compared with young adults, mortality from pneumonia in old individuals is twice as high. From tuberculosis, mortality is ten-fold higher, and from appendicitis, this figure is nearly 20-fold (High, 2004).

2.1.1 Immunosenescence and inflamm-aging

Ageing is associated with declining function of the immune system, which contributes to the increased susceptibility to diseases observed in the elderly. The term immunosenescence can be used to describe ageing-associated changes in the immune system. The adaptive branch of the immune system is generally thought to be more affected by immunosenescence compared with innate immunity (Arnold et al., 2011; Pawelec et al., 2010). With age, the number of naïve T cells decreases due to declining function of haematopoietic stem cells (Wagner et al., 2008) and involution of the thymus (Aspinall & Andrew, 2000). There are also functional changes in naïve T cells, including increased production of IFN-γ (interferon γ), a pro-inflammatory cytokine (Pfister & Savino, 2008), shortened telomeres and a restricted T cell receptor repertoire (Pawelec et al., 2004). With decreasing numbers of naïve T cells, the proportions of memory and effector T cells increases. The increase in the number of memory cells is also due to the clonal expansions of memory T cells caused by persistent viral infections (Karrer et al., 2003).

Another typical feature of the aged immune system is the emergence of T cell populations that lack the costimulatory antigen CD28 (Arnold et al., 2011;

Pawelec et al., 2010), which produce pro-inflammatory cytokines, such as IFN-γ and TNF-α (tumor necrosis factor α) (Franceschi et al., 2000a). In both naïve and memory T cell populations, the observed changes are more profound in the CD8+ population, leading to changes in the CD4+/CD8+ T cell ratio.

The B cell compartment is also affected with ageing; the percentage and numbers of CD19-expressing B cells decrease with age (Ademokun et al., 2010; Paganelli et al., 1992). The B cells of elderly individuals also show a limited diversity, which has been shown to be associated with frail health (Gibson et al., 2009). The impairments in B cell function are also partially due to defects in T cell help (Ademokun et al., 2010; Yang et al., 1996).

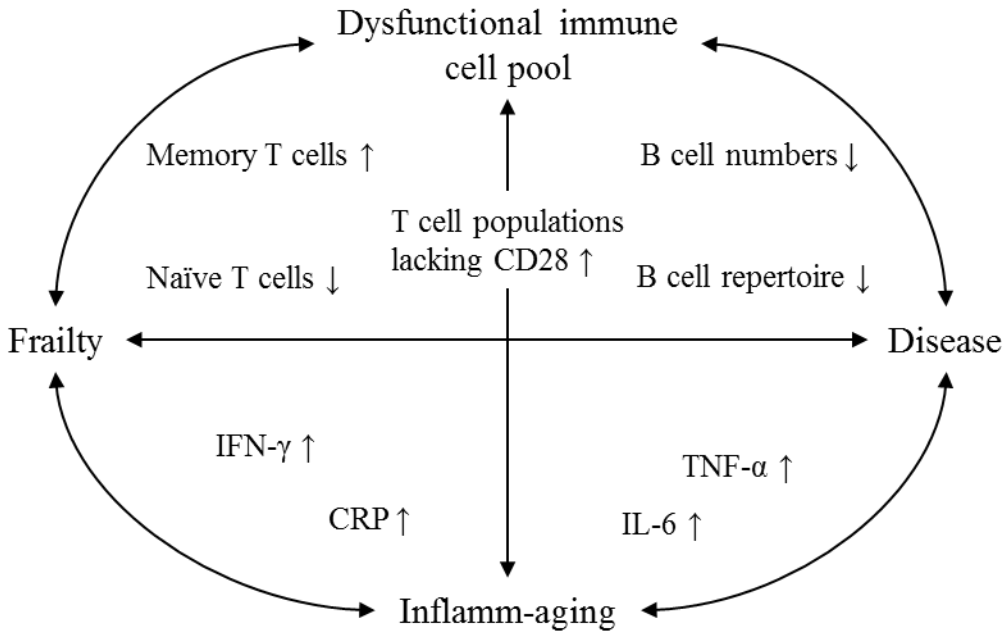


Figure 2. Ageing-associated phenotypes are all linked. Immune system impairment contributes to the increased susceptibility to diseases, diseases can contribute to inflamm-aging and inflamm-aging may be a cause of frailty, which in turn contributes to increased disease susceptibility (Franceschi & Campisi, 2014; Hubbard & Woodhouse, 2010; Hunt et al., 2010).

The low-grade, chronic, systemic inflammation associated with ageing is termed inflamm-aging. With age, the levels of various pro-inflammatory cytokines, such as IL-6 (interleukin 6), CRP (C-reactive protein), TNF-α and IFN-γ, are increased in the absence of infection (Franceschi & Campisi, 2014).

Inflamm-aging is associated with increased morbidity and mortality in the elderly (Franceschi et al., 2000a). The aetiology of inflamm-aging remains largely unknown, but several potential mechanisms have been suggested. As discussed before, immunosenescence leads to an increased number of pro-inflammatory cytokine-producing cells. In addition, the accumulation of damaged macromolecules and other self-debris, harmful products of commensal bacteria and the secretions of senescent cells may contribute to inflamm-aging (Biagi et al., 2011; Campisi & d'Adda di Fagagna, 2007; Franceschi et al., 2000b; Franceschi & Campisi, 2014). The interplay of immunosenescence, inflamm-aging and disease is summarised in Figure 2.

2.1.2 Manifestation of ageing at the cellular and molecular level

The phenotypes associated with ageing are believed to originate from functional impairments at the molecular and cellular level. Genomic instability, telomere attrition, epigenetic alterations, loss of proteostasis, deregulated nutrient sensing, mitochondrial dysfunction, cellular senescence, stem cell exhaustion and altered intercellular communication can be considered to be hallmarks of ageing, i.e., mechanisms that are commonly associated with ageing throughout species. Supporting evidence for their causal role in the ageing process also exists for these nine mechanisms (López-Otín et al., 2013).

The hallmarks of ageing can further be categorised into the following mechanisms: i) primary damage-causing mechanisms; ii) mechanisms that respond to this damage and initially try to mitigate it but they eventually enhance the damage; and iii) phenotypic consequences of the previously mentioned mechanisms. The primary damage-causing mechanisms include genomic instability, telomere attrition, epigenetic alterations and a loss of proteostasis (summarised in Figure 3).

Ageing-associated changes	Causal evidence
Genomic instability[1]	
-Abasic sites[2] ↑ -Translocations, insertions[3] ↑ -DNA oxidation[4] ↑ -DNA repair pathways[1] ↓	-Deficiencies in DNA repair cause accelerated ageing in mice[5] -Deficiencies in DNA repair in human progeroid syndromes[6] -Overexpression of mitotic checkpoint components extends healthy lifespan in mice[7]
Telomere attrition[8]	
-Telomere length in somatic cells ↓ 	-Telomerase deficiency associated with premature development of diseases[9] -Genetically modified mice show lifespan correlated with the length of telomeres[8] -In humans, short telomeres are associated with increased mortality risk[10]
Epigenetic changes[8]	
-Global DNA methylation level[11] ↓ -Histone proteins ↓ -Histone modifications ↑↓	-Sirtuin 6 (SIRT6) deficiency accelerates ageing[12], SIRT6 overexpression extends lifespan in mice[13] -Loss-of-function mutations of heterochromatin protein 1α (HP1α) reduces lifespan in flies, overexpression extends longevity[14] -Affecting histone acetyltransferases and deacetylases can improve the phenotype of progeroid mice and prevent ageing-associated memory impairment[15,16]
Loss of proteostasis[17]	
-Chaperones ↓ -Degradation via proteasome ↓ -Autophagy ↓ 	-Transgenic worms and flies overexpressing chaperones are long-lived[18] -Mutant mice deficient in chaperone components exhibit accelerated ageing[19] -Activation of autophagy is common in long-lived mutant worms[17] -Expression of autophagy genes extends lifespan in flies[17] -Fibroblasts from healthy human centenarians have more active proteasomes[20]

Figure 3. Four primary damage-causing hallmarks of ageing. ([1]Moskalev et al., 2013; [2]Atamna et al., 2000; [3]Ramsey et al., 1995; [4]Mecocci et al., 1999; [5]de Magalhães et al., 2009a; [6]Freitas & de Magalhães, 2011; [7]Baker et al., 2013; [8]López-Otín et al., 2013; [9]Armanios & Blackburn, 2012; [10]Boonekamp et al., 2013; [11]Talens et al., 2012; [12]Mostolavsky et al., 2006; [13]Kanfi et al., 2012; [14]Larson et al., 2012; [15]Peleg et al., 2010; [16]Krishnan et al., 2011; [17]Koga et al., 2011;[18]Walker & Lithgow, 2003; [19]Min et al., 2008; [20]Chondrogianni et al., 2003).

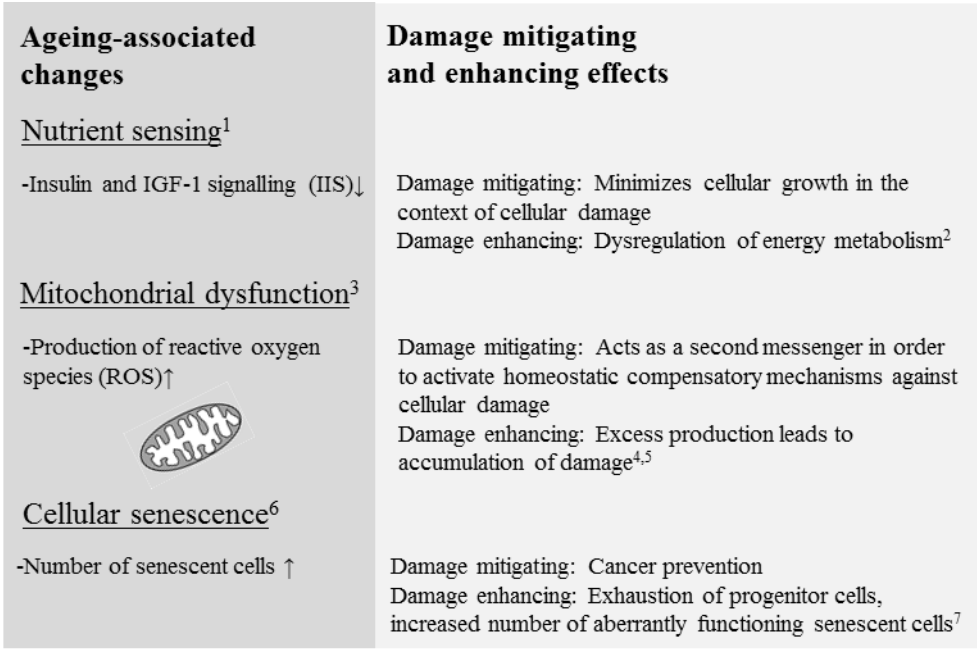


Figure 4. Three damage responsive hallmarks of ageing. IGF-1 = insulin like growth factor 1. ([1]Barzilai et al., 2012; [2]Garinis et al., 2008; [3]López-Otín et al., 2013; [4]Hekimi et al., 2011; [5]Sena & Chandel, 2012; [6]Wang et al., 2009; [7]Campisi & Robert, 2014).

Mechanisms responding to the damage caused by the four primary ageing mechanisms include deregulation of nutrient sensing, mitochondrial dysfunction and cellular senescence. Initially, these three mechanisms mitigate damage, but once a certain threshold is exceeded, they become deleterious themselves (summarised in Figure 4) (López-Otín et al., 2013). An example of this effect is the production of reactive oxygen species (ROS). According to new evidence, ROS act as second messengers to activate homeostatic compensatory mechanisms against cellular damage but in excessive quantities lead to increased cellular damage (Hekimi et al., 2011; Sena & Chandel, 2012).

Stem cell exhaustion and altered intercellular communication are consequences of the above-mentioned ageing-associated molecular mechanisms. Furthermore, many ageing-associated phenotypes are clear consequences of these effects on stem cells and cellular communication. For example, in the context of immune system, stem cell exhaustion leads to immunosenescence and altered intracellular communication leads to inflamm-aging (López-Otín et al., 2013).

2.1.3 Ageing theories

How the evidence regarding causes and consequences of ageing is interpreted depends largely on the theoretical framework through which the ageing process itself is viewed. Ageing theories address the question "what is the proximal cause of ageing?". The current theories have their roots in the understanding of evolution, as "nothing in biology makes sense except in the light of evolution" (Dobzhansky, 1973).

The first formal ageing theory was proposed by August Weismann in late 19th century. Since then, various ageing theories, both overlapping and completely contradictory to each other, have been proposed (Blagosklonny, 2013a; Harman, 1965; Jin, 2010; Zimniack, 2012). The proposed theories and their principles can be classified as adaptive/non-adaptive or as programmed/stochastic. It is important to note that the two categorizations are not interchangeable. Adaptive/non-adaptive refers to the evolutionary cause of ageing, i.e., whether or not ageing is positively selected for during evolution. Programmed/stochastic refers to the mechanism of ageing in the individual, i.e., whether ageing-associated changes are due to an organised process or due to random mistakes. While adaptive ageing by necessity is also programmed, a non-adaptive mechanism may be programmed or stochastic in nature.

An adaptive, programmed theory of ageing proposes that there is a positive evolutionary selection for the termination of life i.e. that mechanisms leading to the death of an individual are positively selected for in evolution (Kirkwood & Melov, 2011). The concept of adaptive programmed ageing has been widely discredited in contemporary literature. Adaptive programmed ageing was proposed to serve the good of the species at the expense of the individual. It has been proposed to prevent overpopulation, to accelerate evolution by speeding the cycling of generations and benefitting young animals by eliminating the old and less valuable individuals. However, the current view is that evolution functions primarily at the level of the individual, not the species. Additionally, for ageing to be positively selected for, a significant number of individuals in the wild should die of old age and senescence, which does not occur. If a programme is causing ageing, mutations that inactivates such a programme would offer a selective advantage for the individual and the ageing programme would disappear from the population. At the very least, immortal mutants should exist, but no know mutation abolishes ageing completely, ageing can merely be slowed by known mutations that affect lifespan. In addition, the notion that the purpose of ageing

and senescence is to eliminate the old and less valuable individuals is circular reasoning, as the old are less valuable precisely because of ageing (Blagosklonny, 2013a; Kirkwood & Melov, 2011). Defence for adaptive programmed ageing has recently been presented but it is based on examples that are difficult to generalise (Skulachev, 2011).

The most widely accepted theory of ageing states that it is the non-adaptive, stochastic accumulation of somatic damage due to a limited investment of resources on maintenance and repair. In wild populations, energy is a limited resource and its use must be optimised between reproduction and individual survival. According to the disposable soma theory, there is only a limited evolutionary pressure to select for maintenance and repair mechanisms. If 90% of individuals in a wild population die within the first year, mechanisms that ensure somatic integrity for five years are not selected for. It is enough to ensure the integrity of the soma for approximately one year, and the remaining energy resources can be invested in reproduction. However, when external causes of death are reduced, ageing of the individual is observed. The accumulation of somatic errors is not restricted to one type of damage but includes the accumulation of mutations, aggregated protein products and increased ROS, among others. It is also not specified what causes this damage, which can be due to both different environmental effects as well as metabolism by-products (Kirkwood, 2008; Kirkwood & Melov, 2011; Zimniak, 2012).

The hyperfunction theory or pseudo-programmed theory of ageing states that while ageing is non-adaptive, it is programmed. This is the most recently proposed theory of ageing, and it has also been strongly criticised (Zimniak, 2012). The main principle of the hyperfunction theory is that ageing is the aimless continuation of developmental programmes that fail to be terminated. The continuation of developmental programmes leads to hyperfunction of tissues, which in turns leads to tissue damage that causes the observed ageing effects. This is in contrast to the disposable soma theory, where the damage that causes ageing effects is on the cellular and molecular level. The hyperfunction theory should not be confused with the adaptive programmed theory of ageing; while development is selected for in the evolution, and thus programmed, ageing is not selected but is still programmed. This is accordance with the principles of antagonistic pleiotropy, namely, that the force of natural selection diminishes with age, particularly after reproduction. Therefore, a feature that is beneficial (or even neutral) early in life but detrimental later is not selected against. In addition, the fidelity of developmental programmes is of great importance, and any change that

would be beneficial in later life would likely be deleterious in early development. Figuratively, ageing is the shadow of development and thus cannot be affected without affecting the actual developmental programme (Blagosklonny, 2012; Blagosklonny, 2013a; de Magalhães, 2012).

2.2 Lifespan and longevity

Longevity, i.e. a long lifespan, is, in addition to beauty, probably the most sought after phenotype in human populations. Although the proportion and number of nonagenarians and centenarians (individuals over 90 or 100 years of age, respectively) is increasing, a corresponding increase in the maximal attainable lifespan has not been observed. Jeanne Calment still holds the title of the person with the longest lifespan with 122 years, and she passed away in 1997. The maximal human lifespan appears to be approximately 110-120 years of age. There have been fewer than 2000 individuals to reach the age of 110, some 30 to reach 115 and approximately 10 to reach the age of 116 (http://www.grg.org/).

As life expectancy has increased and fertility decreased, both the absolute number and proportion of old people is increasing worldwide, including in Finland. By 2020, the number of over 65-year-olds will exceed the number of under 5-year-olds for the first time in recorded history. The fastest growing population group are the oldest old; the percentage of change from 2010 to 2050 for individuals over 85 years of age is estimated to be 351%. For individuals over 100 years of age, this value is 1004%, compared with 22% for individuals aged 0-64 years (WHO, 2011). In Finland, it is estimated that the proportion of over 65-year-olds will increase from 18% in 2010 to 26% by 2030 (SVT, 2015a). Due to these demographic changes, ageing and its detrimental effects concern a growing number of individuals for a longer period of time.

2.2.1 Sexual dimorphism of lifespan and longevity

Females in general have longer lifespan compared with males, and the sexual dimorphism of lifespan exists in the great majority of countries as well as across all time periods for which data exist. The difference in life expectancy between males and females differs depending on the actual length of the lifespan (Seifarth et al., 2012), for example in Finland in 2013 it was 6 years (life expectancy for

males 77.8 years and for females 83.8 years) (SVT, 2015b). Females are much more likely to reach the age of 100 years than males, and in Western countries, there are 5 to 7 centenarian females for each centenarian male (Candore et al., 2006). However, centenarian males and females differ in their health status, as males are more likely to achieve old age by escaping common age-related diseases, whereas females are more likely to reach extremely old age after surviving common morbidities, such as cancer or cardiovascular diseases (Evert et al., 2003). Biological, behavioural and sociological factors contribute to the sexual dimorphism of lifespan (Newman & Murabito 2013; Seifarth et al., 2012).

Genetic contributions to the female advantage in lifespan include the X chromosome and the maternal inheritance of mitochondria. Two X chromosomes may offer protection from unfavourable alleles. Specifically, such alleles are expressed in only half of the cells of the body due to random inactivation of the X chromosome via DNA methylation during development (Seifarth et al., 2012). Also, the mosaic organisation of females may be beneficial due to cooperative mechanisms between the two cell populations (Dobyns et al., 2004; Ørstavik 2009). Typically the ratio of maternal and paternal X is 50:50 in young females, but it may be skewed in certain conditions (Bolduc et al., 2008; Minks et al., 2008). Ageing is associated with increased skewing of the ratio between the inactivated X (i.e. maternal or paternal X) (Hatakeyama et al., 2004; Sandovici et al., 2004) but this skewing is delayed in the offspring of centenarians (Gentilini et al., 2012) indicating that retaining the mosaic cell population is beneficial for longevity. Interestingly, the longevity advantage in many species is for the homogametic sex, females in mammals (XX) and males in birds (ZZ) (Seifarth et al., 2012; Liker & Szekely, 2005).

In majority of animals, including humans, the mitochondria are strictly inherited from the mother via oocyte (Birky 2001; Song et al., 2014). It has been suggested that the mitochondrial genome has evolved to optimally function with the female nuclear genome as natural selection can act predominantly on the mitochondrial-nuclear genome interaction in females (Tower, 2006). In addition, evidence exists that aged female mitochondria function as those of younger males (Borrás et al., 2003).

Sex hormones have a role in the sexual dimorphism of lifespan and longevity via their effects on immune function. Oestrogens stimulate the production of anti-inflammatory cytokines and inhibit production of pro-inflammatory cytokines. Oestrogen is also considered to be an enhancer of humoral immunity, whereas androgens and progesterone act as immunosuppressants (Seifarth et al., 2012).

There is evidence that lower levels of androgens in males lead to a more robust immune system (Seifarth et al., 2012; Voltz et al., 2008). In general, females are considered to be more immunocompetent than males given that males are more susceptible to infectious diseases and cancer. However, females are more susceptible to autoimmune diseases (Markle & Fish, 2014; Nunn et al., 2009). A meta-analysis has shown a post-pubertal male bias in cutaneous leishmaniasis (incidence ratio 3.64), pulmonary tuberculosis (1.91), lepromatous leprosy (2.94) and meningococcal meningitis (1.39) (Guerra-Silveira & Abad-Franch, 2013). On the other hand the ratio of female to male patients is 2:1 in multiple sclerosis, 9:1 in Sjögren's syndrome, 2:1 in rheumatoid arthritis and 9:1 in systemic lupus erythematosus (Wang et al., 2015). The higher degree of female immunocompetence is not restricted to humans but is present in many vertebrates (Folstad & Karter, 1992) and, interestingly, in some insects, even though insects lack sex-specific hormones (Joop et al., 2006; Nunn et al., 2009).

Even though oestrogen has been suggested to have several beneficial effects, the true fluctuations of sex hormone levels through the entire lifespan complicate this interpretation. In addition, neither oestrogen supplementation in females or androgen supplementation (in female to male transsexuals) have been shown to affect mortality or morbidity (Gooren et al., 2008; Seifarth et al., 2012).

In addition to immune function, sex hormones regulate bone mineral density, play a role in combatting oxidative stress and affect the hormonal and cellular responses to stress (Seifarth et al., 2012). For example, females are twice as likely to experience fractures because of falls as compared to males (Seifarth et al., 2012), oestrogens may function as antioxidants (Behl et al., 1997; Ozacmak & Sayan, 2009) and males show a more pronounced HPA (hypothalamic–pituitary–adrenal) -axis stress response as compared to females (Dahl et al., 1992; Kirschbaum et al., 1992), all contributing to the dimorphism in lifespan and longevity.

There are differences in storage and metabolism of lipids between the sexes. As compared to males, females have larger adipose storages, which are located in hips, thighs and buttocks as opposed to abdominal region in males (Lemieux et al., 1993; Nielsen et al, 2004; Power & Schulkin, 2008). As compared to adipose tissue in hips, thighs and buttocks, visceral fat has a more detrimental secretory profile and it shows increased rate of lipolysis. These contribute to the differences in cardiovascular morbidity and mortality between males and females (Candore et al., 2006; Seifarth et al., 2012).

In addition to the biological factors, behavioural and environmental factors have a major contribution to sexual dimorphism of lifespan. Excess male mortality is partly due to work-related risks in industrial activity, car accidents, smoking and consumption of alcohol (Abbott, 2004; Candore et al., 2006). However, as females engage in more "typically male" behaviour, and there are changes in male smoking habits and employment patterns, the difference in life expectancy due to behavioural reasons is narrowing. For example, it has been predicted that the life expectancy of males born in the UK in 2000 who reach the age of 30 will be equal or even exceed that of females of the same birth cohort (Mayhew & Smith, 2014; Seifarth et al., 2012).

2.2.2 Heritability of lifespan and longevity

The heritability of age at death in adulthood has been estimated to be approximately 15-30%, depending on the study population. In twin studies, the heritability is estimated to be higher, 20 to 30%, whereas in population-based samples, it is estimated to be 15-25% (Brooks-Wilson, 2013; Murabito et al., 2012). The heritability of lifespan also varies by ethnicity. In African Americans, the heritability of lifespan was only 4%, whereas heritability was 29% in Caribbean Hispanics (Lee et al., 2004). The heritability of lifespan also increases with advancing age i.e. the heritability of a high age at death is higher than that of a low age at death. Before the age of 55-60, the heritability of lifespan is negligible, but increases thereafter (Hjelmborg et al., 2006; Willcox et al., 2006). Male and female siblings of US centenarians show a 17-fold and 8-fold increased likelihood to reach the age of 100, respectively (Perls et al., 2002). In another study, the heritability of living to 100 was estimated to be 33% in females and 48% in males (Sebastiani & Perls, 2012).

Healthy ageing is also heritable, and the offspring of long-lived parents show delayed onset of ageing-associated diseases (Atzmon et al., 2004; Terry et al., 2004). The offspring of long-lived parents show slower cognitive decline compared with progeny of non-long-lived parents (Dutta et al., 2014). In male twins, "wellness" (defined as achieving the age of 70 free of heart attack, coronary surgery, stroke, diabetes or prostate cancer) had a heritability exceeding 50% (Reed & Dick, 2003). It has also been shown parental survival past 65 years of age is associated with decreased all-cause mortality rate and a lower incidence of

cancer. Every attained decade of parental age further decreases all-cause mortality (Dutta et al., 2013).

Heritability of lifespan has been reported to be dependent on the sex of the parent and/or progeny, but the results are inconsistent. It has been reported that maternal longevity outweighs paternal longevity and vice versa, and it has also been reported that daughters benefit more from the longevity of parents (You et al., 2010). In a Chinese population, longevity was heritable between mothers and daughters and fathers and sons, but not between parents and progeny of the opposite sex (You et al., 2010). Another recent study reported no effects of sex on the heritability of longevity (Dutta et al., 2013).

2.2.3 Genetics of longevity

Although the heritability of lifespan is 15-30% (Brooks-Wilson, 2013; Murabito et al., 2012), few genes or genomic loci have been associated with longevity in multiple studies or in studies performed with different methods. *APOE* (apolipoprotein E) is the gene most frequently found to be associated with longevity in both candidate gene studies and GWAS (Genome Wide Association Study). *APOE* has three common polymorphic alleles (ε2, ε3 and ε4), of which ε4 can be considered a risk allele and ε2 can be considered a protective allele for longevity. In addition, ε4 is associated with increased Alzheimer's disease risk, and both ε2 and ε4 are associated with cardiovascular disease risk. APOE is involved in cholesterol and lipid transport, inflammation and oxidative stress, and its effect on lifespan and disease risk is believed to be mediated via these mechanisms (Brooks-Wilson, 2013; Shadyab & LaCroix, 2015). It needs to be mentioned that *APOE* did not reach genome-wide significance in some large GWAS analyses (Newman et al., 2010; Walter et al., 2011). Other genes more inconsistently associated with longevity include *FOXO3* and *CETP* (Brooks-Wilson, 2013).

Longevity is a complex trait, and it has been suggested that it is dependent on the number of small-effect genetic variants and on the interactions of these. In various studies, different groups of SNPs (single nucleotide polymorphisms) and genes have been associated with longevity. The identified genes and SNP locations have been associated with insulin/IGF-1 signalling, telomere maintenance and ageing-associated diseases, such as Alzheimer's disease (Deelen et al., 2013; Sebastiani et al., 2012; Yashin et al., 2010). In addition, a "genetic

risk score" composed of >700 SNPs associated with common traits and diseases was designed and found to be significantly associated with time-to-death (Ganna et al., 2013).

The identified risk alleles do not compromise human longevity, as centenarians carry the same number of risk alleles for common ageing-associated diseases as an average member of the population. It has been suggested that the presence of protective alleles is more important than the absence of risk alleles. Also, the effects of risk alleles may be buffered by favourable alleles in other genes (Beekman et al., 2010; Brooks-Wilson, 2013).

The study of genetic effects on longevity is also complicated by problems in study design. Extreme longevity and, particularly, healthy ageing can be defined in various ways, complicating the interpretation of results. The selection of the control population is critical, as DNA samples from the ideal control group (individuals of the same birth cohort) are usually not available, and using a younger population leads to confounding factors related to environmental and lifestyle factors. Some genetic variants may offer a longevity advantage only in certain populations, further complicating the interpretation of results (Brooks-Wilson 2013; Jylhävä, 2014; Shadyab & LaCroix, 2015). The genetic studies on longevity fail to explain the majority of heritability of lifespan, implying that heritability of lifespan is mediated, at least partly, via other heritable features than DNA nucleotide sequence.

2.2.4 Extension of lifespan

At this point, no particular environmental trait, such as diet or socioeconomic status has been found to be essential or sufficient for achieving advanced age. However, high consumption of vegetables and low consumption of red meat is the most often reported lifestyle associated with longevity (Kolehmainen et al., 2015; Orlich et al., 2013; Willcox et al., 2014; Zbeida et al., 2014).

The most efficient intervention in modulating lifespan is calorie restriction (CR) or dietary restriction (DR). Notably, the terms are not necessarily interchangeable; CR refers to a reduction in energy availability without malnutrition, while DR can be defined as different types of interventions, such as intermittent fasting or controlling the proportions of macronutrients (Ingram & Roth, 2015; Kaeberlein, 2013). For simplicity, the term CR is used here throughout.

CR was demonstrated to extend the lifespan on rats in the 1930s (McCay et al., 1935), and CR has since been shown to extend the lifespan of *S. cerevisiae* (Anderson et al., 2003), *C. elegans* (Lee et al., 2006), *D. melanogaster* (Partridge et al., 2005), mice (Weindruch et al., 1986) and rhesus monkeys (Bodkin et al., 2003; Colman et al., 2014). CR can also delay the onset of ageing-associated diseases and disabilities in mice, rats and rhesus monkeys (Fontana & Partridge, 2015). There is also evidence of the benefits of CR in humans. A 2-year CR regimen decreased the level of cardiometabolic risk factors in non-obese individuals aged 21-51 years (Ravussin et al., 2015). However, not all studies have confirmed the lifespan-extending effects of CR, and it has been speculated that the observed effect may be due to, or at least amplified, by the use of inbred laboratory strain animals (Sohal & Forster, 2014). Another caveat of CR is the reported increased susceptibility and mortality to infections reported in mice (Goldberg et al., 2015; Kristan, 2007).

Multiple processes contribute to the lifespan and healthspan advantage produced by CR. CR is associated with reduced inflammation, a decrease in growth-promoting hormones, changes in the activity of nutrient-sensing pathways, enhanced glucose homeostasis, decreased adiposity, the preservation of stem cell function and enhanced genomic stability and protein homeostasis, including increased autophagy (Fontana & Partridge, 2015; Kaeberlein, 2013). The physiological effects of CR are widespread, however studies in humans on possible side effects as well as efficacy and correct timing of CR regimen are still lacking.

The molecular components mediating these effects include FOXO, AMPK and sirtuins (Fontana & Partridge, 2015), but mTOR (mechanistic target of rapamycin) appears to be the central mediator of the effects of CR (Johnson et al., 2013; Kapahi et al., 2010). mTOR is a serine/threonine protein kinase that functions in two complexes, mTOR complex 1 and 2 (mTORC1 and mTORC2). The function of mTORC1 is more thoroughly understood; it promotes mRNA and protein synthesis, lipid biosynthesis, represses autophagy and regulates glucose metabolism. mTORC1 functions downstream of CR, as it has been shown that CR does not extend lifespan in organisms where mTORC1 has been inactivated either pharmacologically or genetically. The inhibition of mTORC1 is also sufficient to extend the lifespan of both invertebrates and mice under non-CR conditions (Johnson et al., 2013).

The CR conditions needed to extend lifespan are generally thought to be difficult to maintain voluntarily in humans; thus, pharmacological inhibition of

mTORC1 is considered a possible method for lifespan extension in humans. Rapamycin or other rapalogues (everolimus, deforolimus) have been approved by the FDA (Food and Drug Administration) to treat different cancers and for transplant patients to inhibit host rejection (Johnson et al., 2013). However, there are studies demonstrating toxic effects of rapamycin, such as insulin resistance and glucose intolerance, as well as suppression of the immune system (Ingram & Roth, 2015). It has been proposed that these issues may be circumvented by correct dosing and timing of the rapalogue treatment (Blagosklonny, 2014).

2.3 Epigenetics and transcriptomics

Epigenetics can be broadly defined as "the sum of all those mechanisms necessary for the unfolding of the genetic programme for development" (Holliday, 2006). The term "epigenetics" was first used by C. H. Waddington in the 1940s (Waddington, 1942) to link developmental biology and genetics, which, at that point, were considered separate disciplines.

In modern terms, epigenetics refers to mechanisms that have effects on the gene expression of the cell that are not based on the nucleotide sequence of the DNA strand and which can be inherited by cell division and, in some cases, from parent to progeny. These mechanisms include DNA methylation, histone modification and other chromatin remodelling mechanisms. Different RNA species can also be classified as epigenetic features.

2.3.1 DNA methylation

DNA methylation refers to the covalent modification of the cytosine base in DNA, where a methyl group ($-CH_3$) is added to the aromatic ring. This 5-methylcytosine (5mC) can be termed the fifth base of DNA. Approximately 4% of cytosines in human DNA are methylated, but as 5mC is predominantly found in symmetric CpG dinucleotides, it is not evenly distributed across the genome. Of the 28 million CpG dinucleotides (CpG sites) in the human genome, 80% are methylated. As the methylation sites are symmetric, the 5mCs are present in both strands of the DNA; thus, the methylation pattern can be faithfully propagated through DNA replication (Figure 5) (Breiling & Lyko, 2015; D'Aquila et al., 2013).

5mC can spontaneously deaminate to thymine, resulting in the underrepresentation of CpGs in the human genome. The existing CpGs form CpG-islands (CGI), 1 kb stretches of DNA with higher that average CG-contents. The human genome contains approximately 24000-27000 CGIs. Typically, CpGs outside genes and in introns, in the CpG-poor regions of the genome, are heavily methylated; in contrast, CpGs in CGIs and overlapping transcription start sites (TSSs) are unmethylated (D'Aquila et al., 2013). CGIs overlap the TSSs of the majority of human genes and are typically associated with a transcriptionally permissive chromatin state, making CGI-promoters the most common promoter type in the human genome. CGIs colocalise with the promoters of all constitutively expressed genes and 40% of genes showing tissue specific expression. (Deaton & Bird, 2011; Illingworth & Bird, 2009).

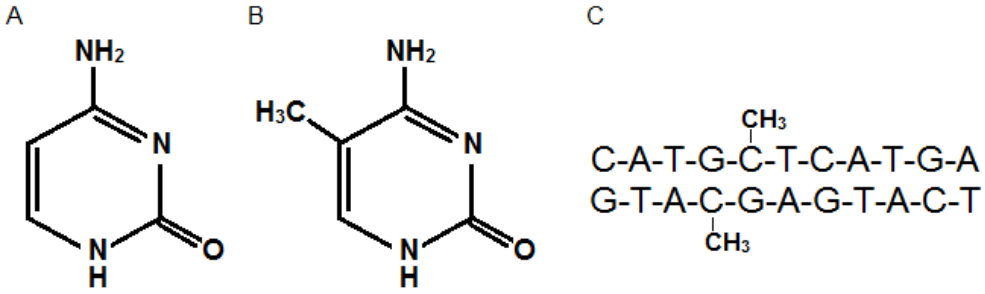


Figure 5. The structure of (A) cytosine (B) methylcytosine and (C) a methylated CpG site in the DNA chain.

DNA methylation is maintained through the DNA methyltransferases DNMT1 (DNA (cytosine-5-)-methyltransferase 1), DNMT3A (DNA (cytosine-5-)-methyltransferase 3 alpha), DNMT3B (DNA (cytosine-5-)-methyltransferase 3 beta) and DNMT3L (DNA (cytosine-5-)-methyltransferase 3-like). These enzymes transfer a methyl group from S-adenosyl-L-methionine (SAM) to deoxycytosine (Denis et al., 2011). DNMT1 is primarily a maintenance methyltransferase; during cell division, it is responsible for the maintenance of the DNA methylation landscape. DNMT3A and DNMT3B are *de novo* methyltransferases and are responsible for methylation during embryonic development. However, DNMT3A and DNMT3B are also needed for methylation maintenance (Jones, 2012; Jones & Liang, 2009). DNMT3L is not a catalytically active methyltransferase but is a regulatory protein essential for *de novo* methylation by DNMT3A (Jia et al., 2007). Each of the three catalytically active

methyltransferases are necessary for embryonic development (Li et al., 1992; Okano et al., 1999).

While methylation of DNA is always an active process, demethylation can occur both actively and passively. Passive demethylation occurs when DNA methylation patterns are not properly maintained through cell divisions. Active demethylation of DNA involves modification of the methylated cytosine and additionally its replacement via base excision repair (BER) (Bhutani et al., 2011; Hill et al., 2014; Pastor et al., 2013). Ten-eleven translocation (TET) enzymes TET1 (tet methylcytosine dioxygenase 1), TET2 (tet methylcytosine dioxygenase 2) and TET3 (tet methylcytosine dioxygenase 3) catalyse the oxidation of 5mC to 5-hydroxymethylcytosine (5hmC) and further to 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC). 5hmC, 5fC and 5caC can be deaminated by various enzymes, including TDG (thymine DNA glycosylase), AID (activation-induced cytidine deaminase) and APOBEC (apolipoprotein B mRNA editing enzyme, catalytic polypeptide). The formed lesion is then repaired via BER (Hill et al., 2014; Pastor et al., 2013). The role of TDG in active demethylation of DNA is supported by various studies by multiple laboratories, but the role of AID and APOBEC remains controversial (Pastor et al., 2013). In addition, the maintenance methylase DNMT1 has a weaker affinity for hemi-5hmC than to hemi-5mC (Hashimoto et al., 2012; Valinluck & Sowers, 2007), thus the activity of TET enzymes also leads to a passive loss of DNA methylation during subsequent cell divisions. See Figure 6 for summary of methylation and demethylation.

Cytosines adjacent to adenine, thymine and guanine can also be methylated. This non-CpG methylation has been observed in human embryonic stem cells as well as adult tissues, such as skeletal muscle and brain. This type of methylation has been shown to regulate the expression of certain genes (*PGC1α* in skeletal muscle (Barres et al., 2009), *IFN-γ* in T cells (White et al., 2009)), but the mechanism remains to be elucidated. The level of non-CpG methylation seems to be influenced by the de novo methyltransferases, DNMT3A and DNMT3B (Pinney, 2014).

In addition to its role as an intermediate in the 5mC demethylation pathway, 5hmC is also in itself an epigenetic marker (Breiling & Lyko, 2015). Approximately 0.1% of cytosines are hydroxymethylated in mammalian tissues; however, in brain tissue, the frequency of 5hmC can be as high as 1% (Kriaucionis et al., 2009). A subset of 5hmCs are stable and present in mammalian promoters, in gene bodies of actively transcribed genes and at active enhancers. 5hmC has

been suggested to be an epigenetic marker specifically important for neuronal development (Breiling & Lyko, 2015).

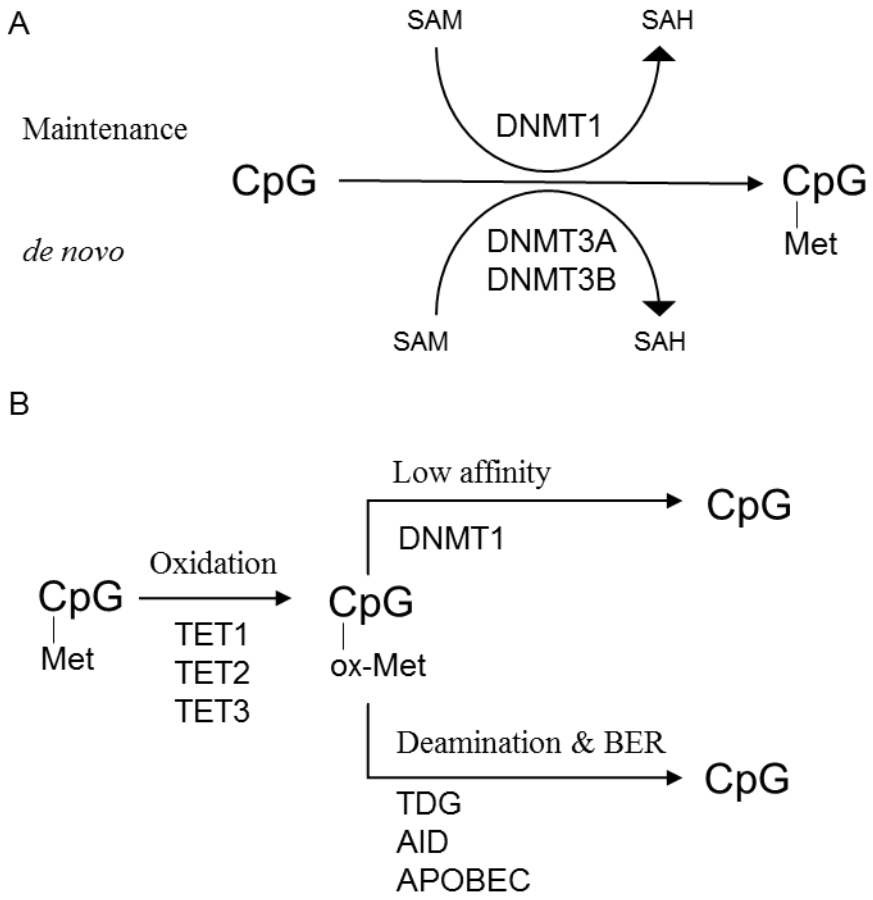


Figure 6. Addition of methyl group to CpG dinucleotides is performed by DNA methyltransferases, shown in (A). DNMT1 is responsible for maintaining DNA methylation patterns during cell division, whereas DNMT3A and DNMT3B are mainly responsible for *de novo* methylation. Active demethylation of CpG dinucleotides, shown in (B), begins with oxidation of the methyl group by TET enzymes, typically followed by deamination and base excision repair (BER). DNMT1 also has a low affinity for oxidised methyl groups, leading to passive demethylation during subsequent cell divisions. SAM = S-adenosyl-L-methionine, SAH = S-Adenosyl-L-homocysteine.

2.3.2 Regulatory functions of DNA methylation

DNA methylation was suggested to control the activity of genes in 1975 (Holliday & Pugh, 1975; Riggs 1975) and this was established in the 1980s. Unmethylated CGIs are associated with active transcription (Bird et al., 1985; Stein et al., 1983) and methylation of previously unmethylated CGIs leads to transcriptional inactivation (Pollack et al., 1980; Wigler et al., 1981). In addition to the regulation of individual genes, DNA methylation has an important role in the inactivation of repetitive elements and in the inactivation of X chromosome in cells containing more than one X chromosome. Most studies regarding the role of DNA methylation have focused on CGIs associated with TSSs; thus, this mechanism is understood the best (Jones, 2012).

The majority of CGIs are unmethylated in somatic cells, and an unmethylated status is associated with transcriptionally permissive chromatin. CGIs are relatively nucleosome deficient; for this reason, unmethylated CGI chromatin is accessible without ATP-dependent nucleosome displacement (Ramirez-Carrozzi et al., 2009). Nucleosomes that are present contain the histone variant H2A.Z and are marked with the trimethylation of histone proteins H3 and H4, markers of transcriptionally active chromatin (Kelly et al., 2010). The function of CGIs has been suggested to help distinguish promoter regions from transcriptionally irrelevant parts of the chromatin (Illingworth & Bird, 2009). This hypothesis is supported by findings that transcription factor (TF) binding sites are highly abundant throughout the genome, implying that promoters cannot be identified solely based on the presence of TF binding sites (Prestridge & Burks, 1993).

The methylation of CGIs represses transcription, but the initial mechanisms proposed, i.e., that methylation directly inhibits the binding of TFs to DNA (Riggs, 1975), is true only for a minority of genes. Transcription factors such as EF2 and CREB are regulated by this direct mechanism, in which methylation of the target sequence inhibits their binding and thus transcription (Campanero et al., 2000; Iguchi-Ariga & Schaffner, 1989). The other mechanism by which CGIs repress transcription is through the recruitment of proteins containing methyl-binding domains. These proteins recruit other chromatin-modifying proteins, which produce a repressive chromatin state. This mechanism is also mediated by certain zinc-finger proteins (Bogdanović & Veenstra, 2009).

It is still debated whether DNA is silenced because it is methylated or whether it is methylated because it is silenced. Early experiments in mice cells showed that silencing precedes methylation in the X chromosome (Lock et al., 1987), and

results in human cancer cells are consistent with this finding (Ohm et al., 2007; Schlesinger et al., 2007; Widschwendter et al., 2007). The results obtained with haematopoietic cells of mice, however, suggest that DNA methylation may have an instructive role in the silencing process (Challen et al., 2011). The majority of evidence suggests that silencing comes before DNA methylation but that DNA methylation is needed for the maintenance of silencing (Deaton & Bird, 2011; Jones, 2012).

In contrast to TSSs and promoters, gene body methylation is permissive for transcription, as has been shown for the active X chromosome (Hellman & Chess, 2007). Gene body CGI methylation is permissive for transcription, even though these CGIs are associated with the repressive histone methylation pattern (H3K9me3). The role of gene body methylation outside CGIs was initially thought to be a mechanism for the repression of repetitive DNA elements, i.e., to prevent the initiation of their transcription while allowing the host gene to be transcribed (Jones, 2012). It was later suggested that gene body methylation may have a role in the regulation of alternative splicing. Exons are more highly methylated compared with introns, and the change in the degree of methylation occurs at the intron-exon boundary (Laurent et al., 2010).

DNA methylation may also have a role in regulating the activity of other genetic elements, such as enhancers and insulators. Enhancers are typically CpG-poor, but it appears that enhancer methylation leads to decreased activity. In general, enhancer sequences are inconsistently methylated (Schmidl et al., 2009). It appears that methylation also has a negative effect on insulator function, but there are only a few studies on the role of DNA methylation on insulator elements (Bell & Felsenfeld, 2000; Jones, 2012)

Although DNA methylation is a regulator of gene expression, many genes show a poor correlation between CGI methylation and the expression level of the corresponding gene. The association between gene expression and CGI methylation is complicated by the existence of intragenic CGIs. These intragenic CGIs may represent an alternative TSS, such as in *PARP12* (Rauch et al., 2009). Intragenic CGIs may also be present at sites of antisense non-coding RNA (ncRNA) transcriptional initiation. When the sense transcript is negatively regulated by ncRNA, hypermethylation of the CGI leads to increased expression of the sense transcript, such as is observed in the *HOXD* cluster (Illingworth & Bird, 2009; Rinn et al., 2007).

2.3.3 Ageing-associated changes in DNA methylation

Ageing is associated with profound changes in the DNA methylation profile. A global decrease in DNA methylation is characteristic of ageing, as is promoter-specific hypermethylation of certain genes. Certain ageing-associated DNA methylation changes appear to be programmed, whereas others are caused by environmental and stochastic effects (Jones et al., 2015; Zampieri et al., 2015).

Global hypomethylation is the most profound change in the ageing DNA methylation profile. This finding was demonstrated in early studies, in which global methylcytosine/cytosine ratios were analysed by HPLC or colorimetric assays (Drinkwater et al., 1989; Wilson et al., 1987). More recent studies based on microarray technologies (Zampieri et al., 2015) and next generation sequencing (NGS) gave similar results (Heyn et al., 2012). The report by Heyn et al. (2012) covered more than 90% of CpG sites of the human genome, and showed that ageing-associated hypomethylation occurs in all genomic regions, including promoters, exons, introns and intragenic regions.

Global hypomethylation is also clearly evident in repetitive elements, even though hypomethylation does not occur to an equal degree for different types of repetitive sequences (Bollati et al., 2009; Jintaridth & Multirangura, 2010). It can be assumed that hypomethylation contributes to ageing-associated genomic instability (Vijg & Dollé, 2007). It has also been shown that ageing-associated hypomethylated regions colocalise with the binding sites of chromatin regulatory proteins and histone modifications associated with active chromatin, indicating that ageing-associated hypomethylation may contribute to global changes in chromatin structure (McClay et al., 2014).

Ageing-associated hypermethylation most typically occurs in CGI-associated gene promoters, as shown by candidate gene approaches and array-based methods (Zampieri et al., 2015). Compared to ageing-associated hypomethylation, hypermethylation is a relatively rare phenomenon, as only 13% of ageing-associated differentially methylated regions were shown to be hypermethylated based on NGS (Heyn et al., 2012).

The candidate gene approach showed that several tumour-suppressors or genes associated with differentiation and growth are hypermethylated with advancing age, including *p16INK4A* (So et al., 2006), *MYOD1* (Ahuja et al., 1998) and *IGF-2* (Issa et al., 1996). In array studies, hypermethylated genes have been shown to belong to pathways/categories relevant for ageing-related diseases and

phenotypes such as cancer and senescence (Bell et al., 2012; Hannum et al., 2013, McClay et al., 2014; Rakyan et al., 2010; Xu & Taylor, 2014).

Changes in DNA methylation have also been observed in various ageing-associated diseases and conditions, including neurodegenerative conditions (e.g., Alzheimer's disease), autoimmune disorders (e.g., rheumatoid arthritis) and cancer (Cribbs et al., 2015; Lardenoije et al., 2015; Paska & Hudler, 2015; Salminen et al., 2015; Zhang & Zhang, 2015). Particularly, cancer shows similar hypermethylation events as ageing. Many sites hypermethylated with ageing overlap gene promoters that have bivalent chromatin marks (i.e., marks of both transcriptionally active and repressive chromatin states, such as H3K4me3 and H3K27me3) in stem cells and are a target of the polycomb repressive complex 2 (Hannum et al., 2013; Heyn et al., 2012; Rakyan et al., 2012; Teschendorff et al., 2010; Xu & Taylor, 2014). These sites are hypermethylated also in cancer (Teschendorff et al., 2010). In addition, conditions associated with both ageing and cancer, such as obesity, inflammation and cigarette smoking, show similar hypermethylation events (Issa, 2011; Issa et al., 2001; Selamat et al., 2012; Suzuki et al., 2009; Xu et al., 2013). It has been proposed that epigenetic changes associated with ageing directly predispose individuals to ageing-associated diseases and conditions (Zampieri et al., 2015; Zane et al., 2014).

The identified ageing-associated DNA methylation changes include examples of both programmed, genetically determined changes and those caused by stochastic and environmental effects. Studies on monozygotic twins have shown that their epigenomes become increasingly discordant with advancing age (Fraga et al., 2005; Poulsen et al., 2007). Comparison of the methylomes of a newborn and a centenarian revealed that the methylome of the centenarian was less homogenously methylated as compared to that of the newborn (Heyn et al., 2012). In contrast, some methylation sites have been identified as ageing-associated in multiple studies. Within one study population, these sites show a strong correlation with chronological age (Table 1).

Given the complex role of DNA methylation in the regulation of gene expression, it is self-evident that the ageing-associated changes in DNA methylation may lead to various changes in gene expression and further to changes in molecular mechanisms and phenotypic features. As for example hypermethylation can lead to both upregulation and downregulation of a given transcript, depending on the location of the CpG site (Rauch et al., 2009; Rinn et al., 2007; Schmidl et al., 2009), the results from DNA methylation studies on ageing should not be interpreted without caution.

Table 1. Methylation sites repeatedly identified as having a strong association with chronological age. For each site, the Illumina Infinium probe number (ID) is given, along with the location of the CpG site in relation to CGIs and genes. For the genes harbouring these sites the abbreviation, full name, function and chromosome is indicated. TSS200= CpG is located in a region 200 bp upstream of TSS; TSS1500= CpG is located in a region 1500 bp upstream of TSS.

Gene	Name	Chr	Function	ID	Location (CGI)	Location (Gene)	Change with age	Reported as ageing-associated
ELOVL2	ELOVL fatty acid elongase 2	6	Elongation of polyunsaturated fatty acids (PUFAs)[1]	cg16867657	Island	TSS1500	↑	[7,8,9,10]
				cg24724428	Island	TSS1500	↑	[7,8,9,10]
				cg21572722	Island	TSS1500	↑	[7,8,9,10]
FHL2	four and a half LIM domains 2	2	A scaffolding protein, associated with integrins and involved in regulation of NF-κB and MAPK signalling[2]	cg22454769	Island	TSS200	↑	[7,8,9,10]
				cg24079702	Island	TSS200	↑	[7,8,9,10]
				cg06639320	Island	TSS200	↑	[7,8,9,10]
PENK	proenkephalin	8	A prohormone that is processed at multiple cleavage sites to generate various enkephalin peptides[3]	cg16419235	Island	TSS1500	↑	[8,9,10]
OTUD7A	OTU deubiquitinase 7A	15	Deubiquitinating enzyme, participates in NF-κB signalling[4,5]	cg04875128	Island	Body	↑	[7,8,9]
EDARADD	EDAR-associated death domain	1	A scaffold protein necessary to the normal formation of ectodermal appendages[6]	cg09809672	Shore	TSS1500	↓	[7,8,10,11,12]

[1]Jakobsson et al., 2006; [2]Verset et al., 2015; [3]Lu et al., 2012; [4]Mevissen et al., 2013; [5]Hu et al., 2013; [6]Sadier et al., 2014; [7]Heyn et al.,2012; [8]Hannum et al., 2012; [9]Garagnani et al., 2012; [10]Florath et al., 2014; [11]Teschendorff et al., 2010; [12]Xu & Taylor 2014.

Methylation sites where the methylation level is strongly associated with chronological age have also been used to construct methylation-based age-predictors. Three different predictors based on methylation sites in three genes (Bocklandt et al., 2011; Weidner et al., 2014) or on 71 methylation sites (Hannum et al., 2013) have been reported to produce methylation ages that deviate five or less years from the chronological age. However, in addition to being based on methylation samples from only one tissue, in two of these studies, the considered study population was rather small (n=~100) (Bocklandt et al., 2011; Weidner et al., 2014). Horvath (2013) constructed an age predictor based on over 8000 samples representing 51 healthy tissues and tissue types. In the test data, the correlation between methylation age and chronological age was shown to be 96% and the error 3.6 years. Increased methylome age predicted by the Horvath algorithm has been shown to be associated with decreased mental or physical fitness in elderly individuals (Marioni et al., 2015a) and higher mortality in individuals aged 69-79 years (Marioni et al., 2015b). Down syndrome patients also have increased methylome ages (Horvath et al., 2015), and in the initial report of the methylome age predictor, it was reported that cancer tissues exhibit increased methylome age compared with healthy tissue (Horvath, 2013).

2.3.4 Other epigenetic mechanisms and ageing-associated changes

In addition to DNA methylation, epigenetic mechanisms include the post-translational modification of histones, chromatin remodelling and both the expression and posttranscriptional modifications of non-coding RNAs (ncRNAs) (Ben-Avraham, 2015; D'Aquila et al., 2013).

Post-translational modification of histones and chromatin remodelling are tightly intertwined. The core component of chromatin is the nucleosome, which is approximately 147 bp of DNA wrapped around histone proteins (H2A, H2B, H3, H4). In addition, histone H1 links the nucleosomes together (Henikoff & Furuyama, 2012). The histone tails, and to some extent the histone core, can be acetylated, methylated, ubiquitylated or phosphorylated. These modifications lead to changes in nucleosome structure, contributing to the compartmentalization of the genome. In addition, modified histones may directly interact with certain protein complexes (Cosgrove & Wolberger, 2005).

Histone acetylation and methylation are the most known histone modifications. Histone H3 and H4 acetylation, typically in promoter regions, is

associated with transcriptional activation (Marmorstein & Roth, 2001). Histone acetylation is regulated by histone acetyl transferases (HATs) and histone deacetylases (HDACs), including the Sirtuin family (SIRT) of NAD+ -dependent deacetylases (Vaquero, 2009). Histone methylation is more complex, as it can lead to transcriptional activation or repression, depending on the histone and the number of methyl groups (mono-, di- or trimethylation) (Hublitz et al., 2009; Wu et al., 2007).

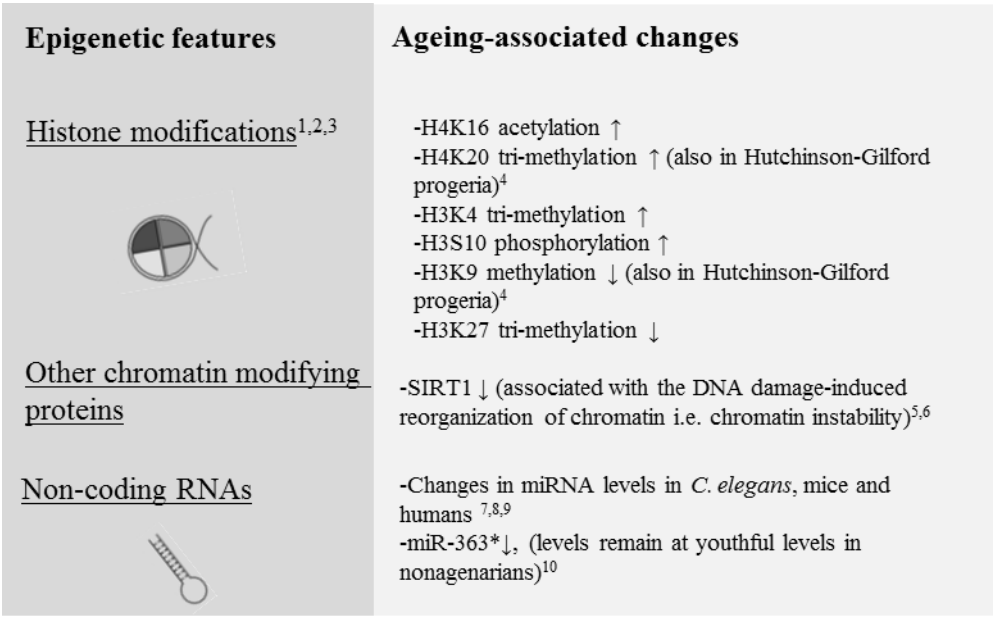


Figure 7. Examples of ageing-associated epigenetic changes other than DNA methylation. [1]Bártová et al., 2008; [2]Fraga & Esteller, 2007; [3]Han & Brunet, 2012; [4]McCord et al., 2009; [5] Oberdoerffer et al., 2008; [6]Sommer et al., 2006; [7]ElSharawy et al., 2012; [8]Inukai et al., 2012; [9]Pincus et al., 2011; [10]Gombar et al., 2012.

The best-characterised ncRNA species is micro-RNAs (miRNAs). Other types of ncRNAs include Piwi-interacting RNAs, small interfering RNAs, long non-coding RNAs (lncRNAs), promoter-associated RNAs and enhancer RNAs. These species have been shown to participate in gene expression regulation in both pathological and non-pathological states (D'Aquila et al., 2013).

The field of epigenetics is further complicated by the posttranscriptional modifications of lncRNAs and mRNAs. Over 100 different posttranscriptional modifications have been identified, and the most common is N^6-methyl-adenosine (m^6A), which on average is present in more than 3 sites per mRNA

molecule. m^6A is suggested to affect mRNA splicing, transport and stability, among other processes (Liu & Pan, 2015).

Examples of ageing-associated changes in epigenetic mechanisms, other than DNA methylation, are summarised in Figure 7. Histone modifications and chromatin structure along with ncRNAs affect each other and regulate gene expression, and thus may contribute to ageing-associated gene-expression changes. Changes in epigenetic features also lead to phenotypic changes, as changes in histone modifications can have an effect on the length of lifespan in *D. melanogaster*, *C elegans* and mice (Han & Brunet, 2011; Kawahara et al., 2009) and also improve the ageing-associated memory impairment in progeroid mice (Krishnan et al., 2011; López-Otín et al., 2013; Peleg et al., 2010).

2.3.5 Ageing-associated gene expression changes

That ageing leads to gene expression changes is somewhat self-evident given that epigenetic mechanisms that regulate gene expression undergo major ageing-associated changes. The observed physiological changes associated with ageing also necessarily have their origins and/or consequences in the level of gene expression.

A meta-analysis of rodent and human age-associated gene expression studies performed in 2009 reported that the overexpression of inflammatory and immune response genes, as well as genes associated with lysosomes, were the most prominent ageing-associated gene expression changes (de Magalhães et al., 2009b). More recent studies in humans have also reported ageing-associated changes in the expression of immune system-associated genes (Bektas et al., 2014; Nakamura et al., 2012; Remondini et al., 2010). In addition, ageing-associated gene expression changes in humans have been observed to be associated with RNA processing and chromatin remodelling, but the results are dependent on the tissue type studied (Gheorghe et al., 2014; Glass et al., 2013; Kumar et al., 2013).

Age-prediction algorithms based on gene expression data have also been constructed. An age-prediction model based on the expression of only six genes has been reported by Harries et al. (2011). With an accuracy of 95%, the expression levels of *LRRN3*, *CD248*, *CCR6*, *GRAP*, *VAMP5* and *CD27* were able to distinguish between study subjects aged under 65 years from those aged over 75 years. In the study by Peters et al. (2015), an age predictor based on more than

10000 transcripts was constructed. On average, the predicted transcriptomic age deviated 7.8 years from chronological age. This study also showed that higher predicted age was associated with adverse phenotypes and thus reflected accelerated biological ageing.

Ageing has also been associated with increased cell-to-cell heterogeneity in gene expression. The expression levels of housekeeping genes as well as heart-specific genes were shown to vary more significantly from cell to cell in the cardiomyocytes of aged mice compared with young mice (Bahar et al., 2006).

2.4 Transgenerational and intergenerational inheritance

The inheritance of acquired traits was first suggested by Lamarck in the early 19th century. In the 20th century, this theory was completely rejected, but evidence published during the last two decades suggests that the inheritance of acquired traits is possible even in mammals (Anway et al., 2005; Crews et al., 2014; Franklin et al., 2010; Martos et al., 2015). This inheritance of acquired traits and environmental effects is believed to be mediated via epigenetic mechanisms, such as DNA methylation (Grossniklaus et al., 2013; Heard & Martienssen, 2014). Of note, the concept of inheritance of acquired traits is also an elegant example of the self-correcting nature of natural sciences, where evidence against old truths are not rejected and hid, but embraced.

A distinction between the terms of intergenerational and transgenerational inheritance should be made. In the case of a female being exposed during pregnancy, both the foetus (F_1) and its germline (future F_2) are also exposed to a given environmental factor. Thus, effects observed in these generations are intergenerational. Only when an effect is observed in the F_3 generation can it be called truly transgenerational. When a male is exposed, so is his germline (future F_1), and effects observed in F_1 are considered intergenerational. Effects in F_2 and subsequent generations can be considered transgenerational (Figure 8) (Heard & Martienssen, 2014; Szyf, 2015)

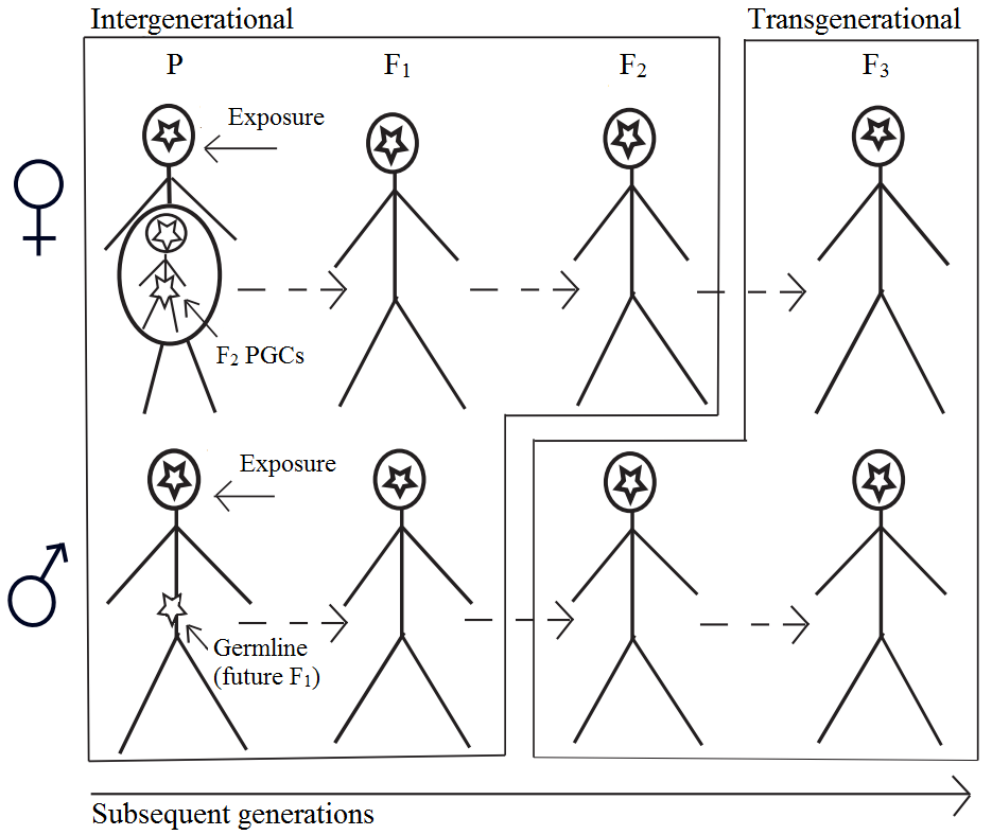


Figure 8. Difference between transgenerational and intergenerational inheritance. When a pregnant female is exposed, F_3 and subsequent generations can be considered to show transgenerational characteristics. If male is exposed, F_2 and subsequent generations show transgenerational effects. PGC=primordial germ cell.

2.4.1 Transgenerational epigenetic inheritance in model organisms

Transgenerational epigenetic inheritance is extensively reported in plants and non-mammalian model organisms. In plants, phenotypes such as coloration and ripening of fruit have been shown to be affected by transgenerational epigenetic inheritance. These effects are mediated by DNA methylation, chromatin remodelling and ncRNAs. In *C. elegans* and *D. melanogaster*, epigenetic silencing of certain genes has been shown to be inherited for several tens of generations. These organisms lack DNA methylation; therefore, the epigenetic effects are mediated via ncRNAs and chromatin remodelling (Heard & Martienssen, 2014).

Interestingly, transgenerational epigenetic inheritance of longevity was shown by Greer et al. (2011) in *C. elegans*. Alterations of the components of H3K4me3 methyltransferase complex in the parent resulted in a 20 to 30% increase in the lifespan of F_3 and F_4 progeny; the effect was lost in F_5 progeny. This transgenerational effect was proposed to be due to local changes in H3K4me3 in certain genes. Specifically, certain gene expression changes were observed in F_3 and F_4 progeny but lost in F_5 progeny.

Transgenerational epigenetic inheritance in mammals was first reported by Anway et al. (2005). Female rats were exposed to vinclozolin, an endocrine-disrupting, non-mutagenic chemical, during gestation, and decreased fertility was observed in male progeny in the F_1-F_4 generations. Changes in DNA methylation patterns were observed for F_2 and F_3 offspring. These results were replicated (Anway et al., 2006), and expression changes in genes associated with histone modification and DNA methylation were identified in the testis of F_1-F_3 offspring (Anway et al., 2008). However, studies showing no effect after vinclozolin exposure have also been reported (Inawaka et al., 2009; Schneider et al., 2008; Schneider et al., 2013).

In addition to vinclozolin, several other endocrine-disrupting chemicals, such as jet fuel JP-8, bisphenol-A (BPA) and DDT, have been shown to induce transgenerationally heritable epigenetic and phenotypic effects in rats. The observed phenotypes include abnormalities of the immune system, tumours and neurological and behavioural effects (Crews et al., 2014; Martos et al., 2015). Foetal alcohol exposure has also been shown to be associated with gene hypermethylation that is inherited paternally through three generations (Sarkar, 2016). In rats, also maternally transmitted effects have been recently reported (Manikkam et al., 2014; Skinner et al., 2013).

There is also some evidence of non-chemical induced phenotypes causing transgenerationally inherited phenotypes and epigenetic alterations. A depressive phenotype can be induced in mice by separating them from the mother during early life. It has been shown that this phenotype can be transmitted up to the F_3 generation. A study by Franklin et al. (2010) suggests that changes in DNA methylation mediate the transmission of this phenotype, as DNA methylation changes were shown in F_1 sperm and the F_2 brain. However, in a study by Gapp et al. (2014) it is suggested that behavioural phenotypes are mediated by ncRNA in sperm.

The evidence for transgenerational epigenetic inheritance is more robust in plants and non-mammalian model organisms than in mammals. This may in part

be due to more extensive study possibilities for lower organisms. However, it has been proposed that short generation times and acute environmental exposures have predisposed these species to favour epigenetic transgenerational inheritance over germline reprogramming. In immobile plants and invertebrates with short generation times, the developing individual is most likely exposed to the same environmental effects as the parent. Thus, an individual that is primed for the environment has an advantage over the unprimed individual (Grossniklaus et al., 2013; Heard & Martienssen, 2014).

2.4.2 Trans- and intergenerational inheritance in humans

Evidence for true transgenerational epigenetic inheritance is more difficult to obtain in humans compared with model organisms. While definitive proof is still lacking, there are implications of intergenerational epigenetic inheritance as well as of phenotypes being inherited in a transgenerational manner.

Early life experiences and maternal environmental exposures during pregnancy have been shown to be associated with epigenetic changes. Early life abuse has been associated with DNA methylation changes in middle-aged males (Suderman et al., 2014), and some of these changes are shared between rats and humans (Suderman et al., 2012). Stress induced by a natural disaster experienced by mothers during pregnancy has been shown to affect the DNA methylation profile of the progeny (Cao-Lei et al., 2014). The progeny of mothers who were exposed to the Dutch famine in 1944 during the first trimester of pregnancy were more obese in adulthood and showed hypermethylation of the insulin-like growth factor 2 receptor (*IGF2R*) gene (Heijmans et al., 2008; Painter et al., 2005). However, it should be noted that these phenomena do not represent inheritance as the individual showing the epigenetic feature or phenotype is also subjected to the environmental insult (see Figure 8). However, these are examples of the plasticity of the epigenome during development.

The lifestyle choices of fathers have been shown to affect the health of sons. Paternal smoking, if started before the age of 11, is associated with increased adiposity in sons aged 11-17 years. In addition, paternal exposure to betel quid is associated with an increased risk of an early manifestation of metabolic syndrome in the progeny (Grossniklaus et al., 2013; Pembrey et al., 2006).

The early-life food supply of paternal grandparents has been shown to be associated with variation in all-cause mortality in the grandchildren in a sex-

dependent manner. Specifically, the food supply of paternal grandfather had an effect on mortality of the grandson and the food supply of paternal grandmother had an effect on the mortality of the granddaughter (Pembrey et al., 2006; Pembrey, 2010). Sharp changes in the food supply of paternal grandmother before puberty have also been associated with excess risk for cardiovascular mortality for the granddaughter (Bygren et al., 2014). In the Dutch famine study, the adverse effects of foetal exposure to famine were inherited only paternally. The progeny of exposed fathers had a higher BMI compared with progeny of unexposed fathers, whereas no effect was identified for the progeny of exposed mothers (Veenendaal et al., 2013).

Sexual dimorphism is observed in the majority of results on the effects of parental lifestyle and nutritional factors on progeny phenotype. The results of these studies also imply that environmental effects that contribute to the health of progeny occur before puberty. It is also noteworthy that while in animal studies the simultaneous intergenerational inheritance of a phenotype and an epigenetic feature have been shown (Anway et al., 2005; Crews et al., 2014; Martos et al., 2015), corresponding evidence in human studies is still lacking. In studies on humans, each individual study has shown either the inheritance of a phenotype or that of an epigenetic feature, but not both.

2.4.3 Epigenetic reprogramming during development

The issue raised against transgenerational epigenetic inheritance is the genome-wide epigenetic reprogramming that occurs twice in the mammalian life cycle, both in the zygote and in developing primordial germ cells (PGCs). After fertilization, the paternal genome is demethylated actively and maternal genome passively. The genome is remethylated after implantation, but the cells destined to become PGCs go through another cycle of demethylation and remethylation, which is completed by birth in males and between birth and puberty in females (Martos et al., 2015; Sharma, 2015).

Recently it has been shown that the reprogramming is not complete, as it has been shown that certain genomic loci retain their DNA methylation patterns throughout development. Certain repetitive elements, such as imprinted genes and certain LINEs (long interspersed elements), are resistant to zygotic reprogramming, PGC reprogramming or both in mice (Seisenberger et al., 2012; Smith et al., 2012). In addition, non-imprinted genes and single-copy sequences

have been shown to escape reprogramming in mouse PGCs (Borgel et al., 2010; Guibet et al., 2012). A genome-wide DNA methylation profiling identified 4730 repetitive sequence-associated genomic loci and an additional 233 single-copy loci that escape reprogramming in mouse PGCs (Hackett et al., 2013).

Histone modifications have also been shown to be transmitted to the embryo through sperm in both mice and humans. The retained nucleosomes (nucleosomes not replaced with protamines) are enriched at genes involved in development of the embryo (Brykczynska et al., 2010; Hammoud et al., 2009). In humans it has been shown that canonical histone modifications (constitutive heterochromatin) are retained in sperm, transmitted to the oocyte and further propagated through embryonic cleavage divisions (van de Werken et al., 2014).

The other issue raised against transgenerational epigenetic inheritance is the transmission of information between soma and germline. In order for acquired traits to be inherited via a transgenerational epigenetic mechanism, information must be transferred from soma to germline. Recently it has been demonstrated that RNA expressed in soma can be transmitted to the germline, as xenografted mice injected with cells expressing enhanced green fluorescent protein (EGFP) were shown to harbour EGFP RNA both in circulating exosomes and in sperm heads (Cossetti et al., 2014). In *C. elegans*, neurons have been shown to transmit double-stranded RNA to germline, where it can initiate the transgenerational silencing of the corresponding gene (Devanapally et al., 2015).

3 Aims of the Study

The studies in this thesis were conducted to analyse ageing-associated changes in gene expression and DNA methylation and to consider how these two types of changes correlate with each other. In addition, the possibility of epigenetic inheritance of lifespan effects from parent to progeny was investigated.

Specifically, the main aims of the present study were as follows:

1. Characterise the ageing-associated gene expression changes between nonagenarians and young adults
2. Characterise the ageing-associated DNA methylation changes between nonagenarians and young adults and also in middle-aged individuals
3. Identify the associations between ageing-associated DNA methylation changes and gene expression in nonagenarians
4. Identify DNA methylation features that are associated with the length of parental lifespan in nonagenarians

4 Materials and Methods

4.1 Study subjects

4.1.1 Vitality 90+ study (I, II & IV)

Vitality 90+ study (V90+) is a prospective population-based study consisting of home-dwelling and institutionalised individuals aged 90 and over living in the city of Tampere (Goebeler et al., 2003). The individuals in studies I, II and IV were from the 2010 cohort (total n=166, 119 females, 47 males) and were all born in 1920. The study participants included in the study had not had any infections or receive any vaccinations in the 30 days prior to blood sample collection.

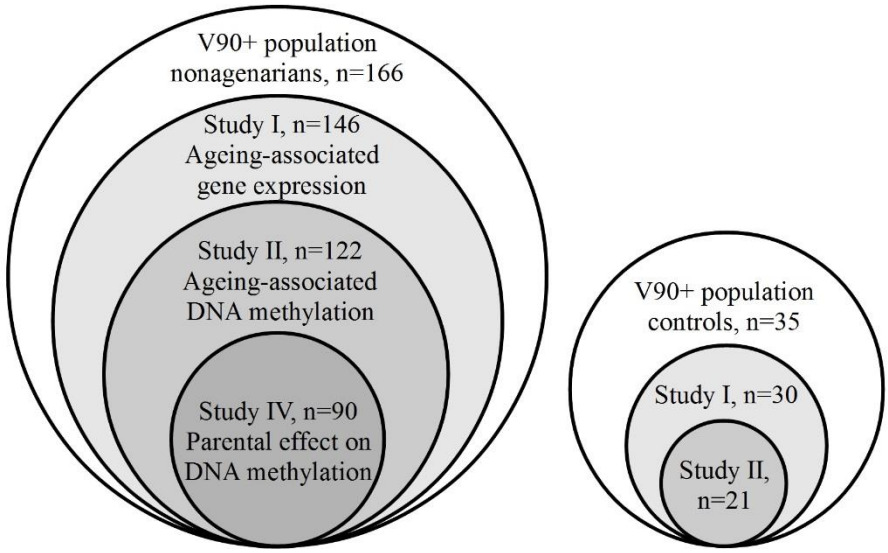


Figure 9. V90+ population was used in studies I, II and IV. Gene expression data was available for 146 nonagenarians and 30 young controls. Methylation data was available for 122 nonagenarians and 21 young controls, and from all of these individuals gene expression data was also available. Of the nonagenarians from whom methylation data was available, the maternal and paternal age at death was available (and exceeded 39 years) for 90 individuals. Sizes of circles are not in scale with number of individuals.

The control group in studies I and II (total n=35, 25 females, 10 males aged 19 to 30 years) consisted of healthy laboratory personnel who did not have any medically diagnosed chronic illnesses, were non-smokers and had not had any infections or received any vaccinations within the two weeks prior to blood sample collection. All of the participants were of Western European descent.

In studies I, II and IV a subpopulation of the total V90+ population has been used, due to availability of biological material and questionnaire data. Overlap of the study population is summarised in Figure 9.

4.1.2 Young Finns Study (III)

The Cardiovascular Risk in Young Finns Study (YFS) is a multi-centre follow-up study by five university hospitals in Finland (Raitakari et al., 2008). The participants were originally randomly selected from the national population register in 1980, when they were aged 3 to 18 years. From the follow-up cohort of 2011, a sub-population of 184 individuals was randomly selected and used in study III. In this sub-population, the individuals were aged 40 (n=50; 29 females, 21 males), 43 (n=44; 30 females, 14 males), 46 (n=55; 31 females, 24 males) and 49 (n=35; 21 females, 14 males) years. All of the participants were of Western European descent.

4.2 Methods

4.2.1 Sample collection

In the V90+ (studies I, II & IV) blood samples were collected into EDTA-containing tubes (3x9 ml) by a trained medical student during a home visit. All of the blood samples were drawn between 8 am and 12 am. The samples were directly subjected to leucocyte separation on a Ficoll-Paque density gradient (Ficoll-Paque™ Premium, cat. no. 17-5442-03, GE Healthcare Bio-Sciences AB, Uppsala, Sweden). The PBMC (peripheral blood mononuclear cell) layer was collected and was suspended in 1 ml of a freezing solution (5/8 FBS, 2/8 RPMI-160 medium, 1/8 DMSO) and stored in liquid nitrogen. In addition to blood sample collection, the examination of participants included an interview and

physical examination. Information on the age at death of parents and siblings and the age of living siblings was collected during the interview.

In the YFS (study III) blood samples were drawn after an overnight fast. The sample collection of 2011 is described in more detail in Nuotio et al., 2014.

4.2.2 Extraction of DNA and RNA

For the V90+ (studies II & IV), DNA was extracted from PBMCs using the QIAamp DNA Mini kit (Qiagen, CA, USA). The DNA was eluted in 60 µl of AE elution buffer and stored at -20°C. For the YFS (study III) DNA was obtained from whole blood (EDTA) using a Wizard® Genomic DNA Purification Kit (Promega Corporation, Madison, WI, USA) according to the manufacturer's instructions.

For the V90+ (studies I & II), RNA was extracted from PBMCs with a miRNeasy mini kit (Qiagen, CA, USA) according to the manufacturer's protocol with on-column DNase digestion (AppliChem GmbH, Darmstadt, Germany). The concentration and quality of the RNA were assessed with the Agilent RNA 6000 Nano Kit on an Agilent 2100 Bioanalyzer (Agilent Technologies, CA, USA).

4.2.3 Determination of cell type proportions

In the V90+ (studies II & IV), the proportions of different lymphocyte populations were determined through FACS analysis (BD FACSCanto II), and the results were analysed with BD FACS Diva, version 6.1.3 (BD Biosciences, Franklin Lakes, NJ, USA). The antibodies employed in this analysis were FITC-CD14 (cat. no. 11-0149), PerCP-Cy5.5-CD3 (45-0037), APC-CD28 (17-0289) (eBioscience, San Diego, CA, USA), PE-Cy™7-CD4 (cat. no. 557852) and APC-Cy™7-CD8 (557834) (BD Biosciences).

In the YFS (study III) the cell proportions were determined from the methylation data (obtained with Illumina Infinium HumanMethylation450 BeadChip, see section 4.2.5) using an estimation algorithm implemented in the `estimateCellCounts` function of the `minfi` Bioconductor package (Jaffe & Irizarry, 2014) using R software (R $\geq$ 2.15.3). The algorithm utilises the selection of 600 control probes that represent specific signatures of CD8+ T cells and CD4+ T cells, monocytes, granulocytes, and natural killer (NK) and B cells.

The reference data used in the estimation is available in the `FlowSorted.Blood.450K` Bioconductor package (Jaffe & Irizarry, 2014).

4.2.4 Gene expression analysis (I & II)

The level of gene expression in PBMCs from the V90+ samples was determined using Illumina Human HT12 v4 Expression BeadChip. The gene expression data were available for 146 nonagenarian study subjects (103 females, 43 males) and for 30 young control subjects (19-30 years of age, median age 22.5 years, 21 females, 9 males).

4.2.4.1 Illumina Human HT-12 v4 Expression BeadChip

For gene expression array analysis, labelled cDNA (complementary DNA) was prepared from total RNA (Illumina TotalPrep RNA Amplification Kit, Ambion Inc., TX, USA). In total, 1,500 ng of labelled cRNA was hybridised to a HumanHT-12 v4 Expression BeadChip (Cat no. BD-103-0204, Illumina Inc., CA, USA) overnight, according to the Illumina protocol. This procedure was performed in the Core Facility at the Department of Biotechnology, University of Tartu. The chips were scanned using a Beadscan (Illumina Inc.).

4.2.4.2 Data preprocessing

In study I, the preprocessing, filtering and analysis of the gene expression data were performed with the Chipster v2.3 programme (Kallio et al., 2011; CSC, Espoo, Finland). In study II, the gene expression data were preprocessed as a `Lumibatch` object with the `lumi` pipeline using R software (Du et al., 2008). In both studies, the `Array_Address_ID` package was used as a probe identifier and background correction was performed with the `bgAdjust.affy` package. The data were transformed with the vst (variance stabilizing transformation) method and normalised with the rsn (robust spline normalization) method.

4.2.4.3 Statistical analyses

For study I, the non-expressed probes and probes whose expression did not change between study groups were filtered out based on the coefficient of variation (CV). The top 5% (2367 probes) with the highest CV were included in the analysis. The study groups were compared with an empirical Bayes two-group test in the `limma` package (Smyth, 2004) using the Benjamini-Hochberg false discovery rate (FDR) for multiple testing correction. The threshold for significance for p-values was set to 0.05. From these genes, we classified those with a linear fold change (FC) above 1.5 or below -1.5 as differentially expressed.

For study II, transcripts with transformed expression values greater than 7.5 in 20% of the samples were included in the analysis. The associations between gene expression and methylation levels (level of DNA methylation in CpG sites located within a given gene) were examined using bivariate correlation (Pearson) analyses; these analyses were performed separately for young and old individuals. In total, 2461 expression-CpG site pairs were tested. The nominal Benjamini-Hochberg-adjusted p-value was set to 0.05.

4.2.4.4 qPCR verification of the expression results

The gene expression results obtained with the array method in study I were validated with qPCR. In total, 300 ng of RNA was converted to cDNA using a High Capacity cDNA Reverse Transcription Kit (Part No. 4368814, Applied Biosystems, CA, USA). A pre-amplification step using TaqMan® PreAmp Master Mix (Part No. 4348266, Applied Biosystems) was performed, as the amount of cDNA was limited. Briefly, 15 ng of cDNA was amplified for 10 cycles according to the manufacturer's instructions using the same assays with which the actual qPCRs were performed (CD83 Hs01077168_g1, IL8 Hs00174103_m1, LRRN3 Hs00539582_s1, PLCG1 Hs01008225_m1 and GUSB Hs00939627_m1 as endogenous control).

The transcript levels were determined with the single gene assays described above using TaqMan® Gene Expression Master Mix (Part No. 4369016, Applied Biosystems). To determine whether the transcripts were differentially expressed between the nonagenarians and the young controls, the RQ (relative quantification) values were calculated with RQ Manager Software (Applied Biosystems).

4.2.5 DNA methylation analysis (II, III & IV)

DNA methylation analysis was performed from PBMCs (V90+, studies II & IV) or from whole blood (YFS, study III) with Illumina Infinium HumanMethylation450 BeadChip. Methylation data were available for 122 nonagenarians (89 females, 33 males) and 21 young controls (14 females, 7 males) from the V90+ and for all study subjects from the YFS sub-population (n=184, 111 females, 73 males).

4.2.5.1 Illumina Infinium HumanMethylation450 BeadChip

DNA methylation profiling was performed at the Institute for Molecular Medicine Finland (FIMM) Technology Centre of the University of Helsinki. Bisulphite conversion of 1 µg of DNA was performed using the EZ-96 DNA Methylation Kit (Zymo Research, Irvine, CA, USA) according to the manufacturer's instructions. A 4-µl aliquot of bisulphite-converted DNA was subjected to whole-genome amplification and then enzymatically fragmented and hybridised to the Infinium HumanMethylation450 BeadChip (Illumina, San Diego, CA, USA) according to the manufacturer's protocol. The BeadChips were scanned with the iScan reader (Illumina). The methylation analysis of V90+ was performed at two separate time points (with a 6 month interval), and this batch effect was adjusted for in the analysis.

4.2.5.2 Data preprocessing

The methylation data from the V90+ (studies II & IV) and YFS (study III) subjects were processed in a similar manner. The methylation data were preprocessed as a `methylumiset` object using R software with the `wateRmelon` array-specific package from Bioconductor (Pidsley et al., 2013). The annotation information was based on the GRCh37/hg19 genome assembly from February 2009. Prior to any processing, all unspecific or polymorphic sites were removed based on database information (Chen et al., 2013). Samples and target sites of a technically poor quality were filtered out by excluding sites with a bead count of <3 in 5% of the samples and sites for which 1% of the samples showed a detection p-value >0.05. Background correction and quantile normalisation via the `dasen` method were conducted individually for the two

applied chemistries (Infinium I and II) as well as for the intensities of methylation (m) and un-methylation (u). After `dasen` treatment, the u and m intensities were transformed to beta (β) and M values. β is the ratio of the methylated probe (m) intensities to the overall intensities (m + u + α), where α is the constant offset (100). Thus, β ranges linearly from 0 (non-methylated, 0%) to 1 (completely methylated, 100%). M values were derived from β values using the equation $M=\log2(\beta /(1- \beta))$. Next, the batch effect of the chemistries was adjusted using the `BMIQ` method, which is based on beta mixture models and the EM algorithm (Teschendorff et al., 2013). For V90+ methylation data (studies II & IV), the batch effect of two laboratory days (time interval of 6 months) was corrected using an algorithm based on Empirical Bayes methods and implemented in the R package `Combat` (Johnson et al., 2007).

4.2.5.3 Statistical analyses

The association of the methylation level at each individual CpG site and the phenotype in question (age group in study II, age in study III or parental age in study IV) was assessed with a generalised regression model, referred to as variable dispersion beta regression (Cribari-Neto & Zeileis 2010; Ferrari, 2004). In the analysis, the phenotype in question was employed as a predictor of the site-specific methylation outcome in the form of β-values (ranging from 0 to 1) in each equation, where the mean model with a linker function of logit was utilised. The nominal Benjamini-Hochberg adjusted p-value (studies III & IV) or Bonferroni-adjusted p-value (study II) was set to 0.05. Where appropriate, cell type proportions, sex and batch were adjusted for. The regression analyses were performed using R software and with algorithms implemented in the `betareg` package (Cribari-Neto & Zeileis 2010; Ferrari, 2004).

In study II, to detect CpG sites showing substantial differences in DNA methylation level between nonagenarians and young adults, the sites displaying the largest difference in the absolute value of the methylation level were included in the analysis ($-1 > \Delta M > 1$). The average levels of the two groups were further compared with the Wilcoxon rank-sum test, and the nominal Benjamini-Hochberg-adjusted p-value was set to 0.05.

In study IV, CpG sites showing substantial differences in methylation between longest-living and shortest-living fathers were extracted by calculating the difference in the median methylation values at each CpG site for the progeny of

the longest- and shortest-living fathers. Only sites with -0.01> Δβ >0.01 were included for further analysis.

4.2.6 Availability of microarray data

The gene expression data and methylation data are available at Gene Expression Omnibus database (GEO; http://www.ncbi.nlm.nih.gov/geo/). The series numbers for expression data (studies I & II), the V90+ (studies II & IV) methylation data and the YFS (study III) methylation data are GSE40366, GSE58888 and GSE69270, respectively.

4.2.7 Pathway analyses

4.2.7.1 QIAGEN's Ingenuity® pathway analysis

QIAGEN's Ingenuity® pathway analysis (IPA) was used to identify canonical pathways associated with differentially expressed genes (study I), genes containing methylation sites associated with age group (study II) and genes containing methylation sites associated with parental age (study IV). Benjamini-Hochberg multiple testing correction was employed to calculate the p-values for the pathways, which were considered significant at $p < 0.25$ (study I) or $p < 0.05$ (studies II & IV) and when the pathway contained a minimum of 3 genes.

4.2.7.2 GOrilla

The Gene Ontology enrichment analysis and visualization tool (GOrilla) (Eden et al., 2007; Eden et al., 2009) was used to identify enriched GO (gene ontology) terms. GO terms were searched based on two unranked lists (target and background), and all genes with at least one probe in the 450K array were used as the background list. The target genes were the same as used in the IPA analysis. A Bonferroni-corrected p-value (studies II & III) or Benjamini-Hochberg (BH)-corrected p-value (study IV) of <0.05 was used as the threshold for significance.

4.2.7.3 PScan

PScan (Zambelli et al., 2009) was used to predict if a group of identified genes containing methylation sites associated with age group or age (studies II & III) were regulated by a common TF. The analysis was performed with the default settings, i.e., using the Jaspar database and the -450 - +50 bp region around the TSS. A Bonferroni-corrected p-value of <0.05 was used as a threshold for significance.

4.2.8 Ethics

All of the studies reported here have been conducted according the guidelines of the Declaration of Helsinki. All of the study participants gave their written informed consent. For V90+ (studies I, II & IV), the study protocol was approved by the ethics committee of the city of Tampere (1592/403/1996; 765/13.03.01/2008, PSHP 7/2014, ETL R14002), and for the YFS (study III), the protocol was approved by the Ethical Review Committee of Turku University Hospital and by the local ethics committees of the participating University Hospitals.

5 Results

5.1 Ageing-associated gene expression changes (I)

To elucidate the gene expression changes associated with ageing in immune system cells, the gene expression profile of PBMCs from nonagenarians and young adults were analysed with Illumina array technology. As both immune system and ageing are known to display sexual dimorphism, the analysis was performed separately for both sexes.

5.1.1 Differentially expressed individual genes

Gene expression analysis revealed 339 transcripts that were differentially expressed between nonagenarian females and young females and 248 transcripts that were differentially expressed between nonagenarian males and young males (BH-corrected p-value <0.05, -1.5> FC >1.5). Of these genes, 180 were common to both sexes. The transcripts displaying the largest differences between nonagenarians and young controls are presented in Tables 2 and 3.

The results obtained with the microarray were verified with qPCR, and genes with both high and low FC were tested. In males, the tested genes and corresponding FCs (microarray/qPCR) were *CD83* (1.73/1.90), *IL8* (3.46/7.26) and *LRRN3* (-4.68/-5.65). In females, the tested genes were *CD83* (1.70/1.71), *IL8* (4.85/6.15), *LRRN3* (-5.64/-7.81) and *PLCG1* (-1.63/-1.98).

Differences in leukocyte proportions do not contribute significantly to the identified differences in gene expression levels according to principal component analysis (PCA). The principal components (PCs) explaining the most variation in the expression data did not correlate statistically significantly with the proportions of different leukocyte subtypes (data not shown).

Table 2. The most up-regulated transcripts in nonagenarians compared with young controls. P-values are Benjamini-Hochberg (BH)-corrected. There were two *IL8* transcripts among the top 10 hits in females. Rank= rank of the given transcript among all of the up-regulated transcripts in the given sex, FC=fold change.

	Females			Males		
Gene	FC	p-value	Rank	FC	p-value	Rank
IL8	4.85 (3.14)	$<10^{-6}$	1 (4)	3.46	$7.8*10^{-3}$	1
PTGS2	3.79	$<10^{-6}$	2	2.68	$3.4*10^{-3}$	7
NR4A2	3.17	$<10^{-6}$	3	3.42	$1.0*10^{-6}$	2
RHOB	2.95	$<10^{-6}$	5	2.07	$3.1*10^{-4}$	23
CDKN1A	2.91	$<10^{-6}$	6	2.87	$<10^{-6}$	5
IL1B	2.84	$<10^{-6}$	7	2.56	$5.2*10^{-3}$	9
RGS1	2.77	$<10^{-6}$	8	2.94	$1.5*10^{-5}$	4
EGR1	2.61	$<10^{-6}$	9	1.84	$3.3*10^{-2}$	45
CCL3L3	2.60	$<10^{-6}$	10	2.36	$9.4*10^{-3}$	13
HBEGF	2.56	$<10^{-6}$	11	2.62	$4.2*10^{-4}$	8
JUN	2.53	$<10^{-6}$	14	3.05	$5.9*10^{-3}$	3
OSM	2.53	$<10^{-6}$	13	2.72	$7.8*10^{-4}$	6
ADM	2.26	$<10^{-6}$	22	2.43	$3.4*10^{-5}$	10

Table 3. The most down-regulated transcripts in nonagenarians compared with young controls. P-values are Benjamini-Hochberg (BH)-corrected. There were two *LRRN3* transcripts in females and males and two *CD79B* transcripts in males among the top 10 hits. Rank= rank of the given transcript among all of the down-regulated transcripts in the given sex, FC=fold change.

	Females			Males		
Gene	FC	p-value	Rank	FC	p-value	Rank
LRRN3	-5.64 (-3.98)	$<10^{-6}$	1 (2)	-4.68 (-3.22)	$<10^{-6}$	1 (2)
CCR7	-3.61	$<10^{-6}$	3	-3.05	$<10^{-6}$	4
LOC652694	-2.88	$<10^{-6}$	4	-2.42	$1.7*10^{-4}$	10
IGJ	-2.85	$<10^{-6}$	5	-2.30	$5.4*10^{-3}$	17
CD27	-2.75	$<10^{-6}$	6	-2.30	$6.0*10^{-6}$	16
CD79A	-2.68	$<10^{-6}$	7	-3.07	$3.4*10^{-5}$	3
CD19	-2.65	$<10^{-6}$	8	-2.96	$5.0*10^{-6}$	5
IGLL1	-2.64	$<10^{-6}$	9	-1.86	$3.7*10^{-2}$	41
SGK223	-2.64	$<10^{-6}$	10	-2.37	$<10^{-6}$	14
FCRLA	-2.63	$<10^{-6}$	11	-2.87	$1.0*10^{-5}$	6
CD79B	-2.27	$<10^{-6}$	21	-2.71 (-2.53)	$<10^{-6}$	7 (9)
NELL2	-2.39	$<10^{-6}$	17	-2.71	$<10^{-6}$	8

5.1.2 Pathways associated with differentially expressed genes

Signalling pathways associated with differentially expressed genes were identified using QIAGEN's Ingenuity® pathway analysis (IPA, Ingenuity® Systems, www.ingenuity.com). We identified 48 pathways that were significantly affected in females and 29 pathways in males; of these pathways, 24 were common to both sexes (p-value <0.05, FDR <0.25, minimum of 3 genes per pathway differentially expressed). The 24 pathways common to both sexes were almost exclusively associated with different functions of the immune system, such as B cell and T cell functions, communication between different types of immune system cells and cytokine production. The top 5 pathways affected in both sexes are presented in Table 4.

Table 4. Top canonical pathways associated with genes that were differentially expressed between nonagenarians and young controls. P-values are derived from Fisher's exact test. The rank denotes the position of the given pathway in the sex-specific list.

	Females		Males	
Canonical pathway	p-value	Rank	p-value	Rank
B Cell Development	$4.0*10^{-9}$	1	$1.4*10^{-8}$	1
Dendritic Cell Maturation	$5.7*10^{-9}$	2	$3.2*10^{-4}$	4
T Helper Cell Differentiation	$1.6*10^{-6}$	3	$1.7*10^{-4}$	2
Communication between Innate and Adaptive Immune Cells	$7.1*10^{-6}$	4	$4.5*10^{-4}$	5
Role of NFAT in Regulation of the Immune Response	$7.2*10^{-6}$	5	$1.4*10^{-3}$	8
Role of JAK family kinases in IL-6-type Cytokine Signaling	$9.8*10^{-4}$	16	$3.1*10^{-4}$	3

Only five male-specific pathways were identified, and the top two were associated with oestrogen. All male-specific pathways associated with genes differentially expressed between nonagenarians and young controls are presented in Table 5. The pathways that were affected only in females included pro-inflammatory pathways and those associated with T-cell function. All female-specific pathways associated with genes that were differentially expressed between nonagenarians and young controls are presented in Table 6.

Table 5. Male-specific canonical pathways associated with genes differentially expressed between nonagenarians and young controls. P-values are derived from Fisher's exact test.

Canonical pathways	p-value
Estrogen-mediated S-phase Entry	$3.9*10^{-3}$
PDGF Signaling	$1.9*10^{-2}$
CD27 Signaling in Lymphocytes	$3.0*10^{-2}$
PPAR Signaling	$3.3*10^{-2}$
Role of Pattern Recognition Receptors in Recognition of Bacteria and Viruses	$3.3*10^{-2}$

Table 6. Female-specific canonical pathways associated with genes differentially expressed between nonagenarians and young controls. P-values are derived from Fisher's exact test.

Canonical pathways	p-value
Prostanoid Biosynthesis	$4.6*10^{-4}$
CTLA4 Signaling in Cytotoxic T Lymphocytes	$8.7*10^{-4}$
CCR5 Signaling in Macrophages	$8.7*10^{-4}$
IL-15 Production	$1.3*10^{-3}$
IL-10 Signaling	$1.4*10^{-3}$
p38 MAPK Signaling	$4.7*10^{-3}$
P2Y Purigenic Receptor Signaling Pathway	$5.6*10^{-3}$
iNOS Signaling	$6.8*10^{-3}$
Cytotoxic T Lymphocyte-mediated Apoptosis of Target Cells	$7.4*10^{-3}$
Differential Regulation of Cytokine Production in Intestinal Epithelial Cells by IL-17A and IL-17F	$7.9*10^{-3}$
iCOS-iCOSL Signaling in T Helper Cells	$8.7*10^{-3}$
IL-4 Signaling	$9.1*10^{-3}$
Nur77 Signaling in T Lymphocytes	$9.3*10^{-3}$
PKCθ Signaling in T Lymphocytes	$1.3*10^{-2}$
TNFR2 Signaling	$1.4*10^{-2}$
Calcium-induced T Lymphocyte Apoptosis	$1.5*10^{-2}$
G Protein Signaling Mediated by Tubby	$1.8*10^{-2}$
Glucocorticoid Receptor Signaling	$1.8*10^{-2}$
Inhibition of Angiogenesis by TSP1	$2.1*10^{-2}$
Role of JAK1 and JAK3 in γc Cytokine Signaling	$2.3*10^{-2}$
ERK5 Signaling	$2.6*10^{-2}$
Antigen Presentation Pathway	$2.7*10^{-2}$
Production of Nitric Oxide and Reactive Oxygen Species in Macrophages	$4.5*10^{-2}$
Phospholipase C Signaling	$4.7*10^{-2}$

5.2 Ageing-associated changes in DNA methylation (II & III)

The association of ageing and DNA methylation changes was assessed in two populations, between nonagenarians and young controls (V90+, study II) with PBMCs and in middle-aged individuals (YFS, study III) with whole blood using Illumina Infinium HumanMethylation450 BeadChip. In the analyses, the differences in proportions of different blood cell subtypes between individuals were adjusted for.

5.2.1 CpG sites differentially methylated with advancing age

In study II, ageing-associated CpG sites were identified using two different statistical methods. A regression model that was adjusted for blood cell type proportions (i.e., the ratio of CD4+ and CD8+ cells and the proportions of CD4+CD28-, CD8+CD28- and CD14+ cells), sex and batch resulted in the identification of 45507 CpG sites for which the methylation level was significantly associated with age group (Bonferroni-corrected p-value <0.05). The Wilcoxon rank-sum test identified 10083 CpG sites for which the methylation level differed between nonagenarians and young controls (ΔM-value >1, BH-corrected p-value <0.05). Of these, 8540 were identified with both methods, and these sites were then labelled as ageing-associated CpG sites (a-CpGs). Of these 8540 a-CpGs, 46% (3925 CpG sites) were hypermethylated in nonagenarians compared with controls. The top hits based on the regression model are presented in Table 7.

In the YFS data, only the regression model was used to identify a-CpGs. The regression model was adjusted for estimated blood cell type proportions (CD4+ and CD8+ T cells, monocytes, granulocytes, NK and B cells) and a sex*age interaction term. We identified 1202 CpG sites in which the methylation level was significantly associated with age (BH-corrected p-value <0.05). Of these a-CpGs, 48% (580 CpG sites) were hypermethylated with advancing age. The most significant hits are presented in Table 8.

Table 7. CpG sites that are most significantly associated with age group in the variable dispersion beta regression model in the V90+ study. The betareg estimate and betareg p-value are derived from the regression model. Δβ is the difference in methylation levels between nonagenarians and young controls.

ProbeID	Gene	betareg estimate	betareg p-value	Δβ	Wilcoxon p-value
cg16867657	*ELOVL2*	1.023	$6.38*10^{-66}$	0.243	$1.53*10^{-10}$
cg16762684	*MBP*	-1.486	$4.74*10^{-64}$	-0.168	$1.53*10^{-10}$
cg11344352	*ERCC1*	-1.202	$9.15*10^{-63}$	-0.153	$1.53*10^{-10}$
cg17110586	na	0.895	$1.46*10^{-59}$	0.200	$1.53*10^{-10}$
cg04875128	*OTUD7A*	1.514	$7.20*10^{-58}$	0.279	$1.53*10^{-10}$
cg08262002	*LDB2*	-0.710	$2.72*10^{-55}$	-0.197	$1.53*10^{-10}$
cg18618815	*COL1A1*	-0.941	$1.78*10^{-52}$	-0.225	$1.53*10^{-10}$
cg00748589	na	0.864	$1.36*10^{-51}$	0.179	$1.53*10^{-10}$
cg15416179	*MAP2K3*	-1.131	$2.38*10^{-51}$	-0.187	$1.53*10^{-10}$
cg12065799	*RRAGC*	-0.823	$8.15*10^{-51}$	-0.088	$1.53*10^{-10}$
cg23479922	*MARCH11*	0.940	$4.07*10^{-49}$	0.263	$1.53*10^{-10}$
cg07544187	*CILP2*	1.541	$2.35*10^{-48}$	0.252	$1.53*10^{-10}$
cg09038267	*C10orf26*	1.227	$1.48*10^{-47}$	0.150	$1.53*10^{-10}$
cg13033938	*IP6K1*	-0.699	$7.54*10^{-47}$	-0.061	$1.53*10^{-10}$
cg19283806	*CCDC102B*	-1.253	$9.82*10^{-47}$	-0.267	$1.53*10^{-10}$
cg07547549	*SLC12A5*	0.900	$5.02*10^{-46}$	0.245	$1.53*10^{-10}$
cg01949403	*APOL3*	0.807	$7.53*10^{-46}$	0.111	$1.53*10^{-10}$
cg01243823	*NOD2*	-1.280	$7.90*10^{-46}$	-0.232	$1.53*10^{-10}$
cg22242842	na	-0.952	$1.99*10^{-44}$	-0.206	$1.53*10^{-10}$
cg06007201	*FAM38A*	-0.932	$5.65*10^{-44}$	-0.156	$1.53*10^{-10}$

Of the 8540 CpG sites identified in the V90+ data and the 1202 CpG sites identified in the YFS data, 301 were identified in both studies as a-CpGs. However, when considering data sets obtained with comparable methods (only the regression model), 987 of the 1202 a-CpGs identified in the YFS are a-CpGs in the V90+ data.

Table 8. CpG sites that are most significantly associated with age in the variable dispersion beta regression model in YFS. The betareg estimate and betareg p-value are derived from the regression model. P-values are BH-corrected.

ProbeID	Gene	betareg estimate	p-value
cg16867657	*ELOVL2*	0.022	$<10^{-9}$
cg01528542	na	-0.028	$1.41*10^{-8}$
cg24724428	*ELOVL2*	0.021	$4.80*10^{-7}$
cg21572722	*ELOVL2*	0.013	$3.46*10^{-6}$
cg06639320	*FHL2*	0.018	$3.46*10^{-6}$
cg00059225	*GLRA1*	0.013	$5.13*10^{-6}$
cg08097417	*KLF14*	0.020	$1.87*10^{-5}$
cg22454769	*FHL2*	0.021	$5.03*10^{-5}$
cg17110586	na	0.014	$5.03*10^{-5}$
cg02650266	na	0.018	$6.07*10^{-5}$
cg07553761	*TRIM59*	0.016	$6.12*10^{-5}$
cg24024661	*HMHA1*	-0.020	$8.42*10^{-6}$
cg02705918	na	-0.020	$8.42*10^{-6}$
cg01588592	*ETV3L*	0.011	$1.14*10^{-4}$
cg10424974	na	-0.012	$1.14*10^{-4}$
cg07547549	*SLC12A5*	0.014	$1.28*10^{-4}$
cg04103761	na	0.016	$1.46*10^{-4}$
cg07197831	*DNAJC5G*	-0.018	$2.26*10^{-4}$

5.2.2 Location of a-CpGs

The a-CpGs identified in the V90+ data (study II) were distributed unevenly across the genome. There was an excess of a-CpGs on chromosomes 2, 3, 4, 5 and 8, whereas chromosomes 16, 17, 19 and 22 contained fewer than expected a-CpGs (hypergeometric test $p < 0.05$). The ratio between hypermethylated and hypomethylated a-CpGs was similar to the whole data set in the majority of these chromosomes, but chromosomes 18 and 19 contained an excess of hypermethylated a-CpGs. In these chromosomes, 72% and 75% of a-CpGs, respectively, were hypermethylated compared with 46% of hypermethylated a-CpGs in the whole data set. The a-CpGs identified in the YFS data did not show statistically significant enrichment in any chromosome (hypergeometric test $p > 0.05$), but the distribution of a-CpGs was similar to those in the V90+ data. The

excess of hypermethylated a-CpGs on chromosome 18 was also identified in the YFS data set.

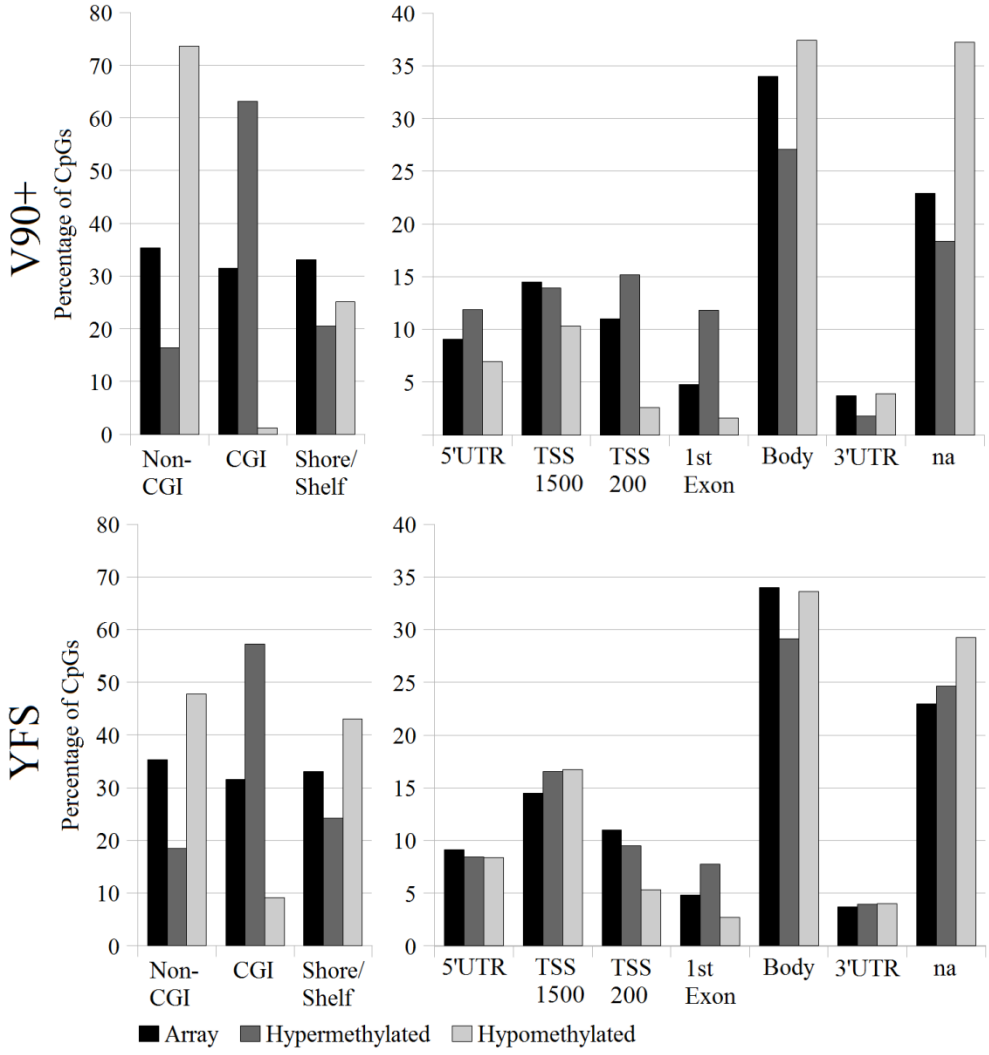


Figure 10. Distribution of hyper- and hypomethylated a-CpGs in V90+ and YFS data. UTR=untranslated region; TSS200= CpG is located in a region 200 bp upstream of TSS; TSS1500= CpG is located in a region 1500 bp upstream of TSS, na=CpG is not located in the region of annotated genes.

In both datasets, there were more than expected hypermethylated a-CpGs located in CGIs, whereas hypomethylated a-CpGs were most abundant in non-CGIs (non-CpG-islands) (hypergeometric test $p < 0.05$).

When the location of a-CpGs was investigated in regard to genes, an excess of hypermethylated a-CpGs in the V90+ data was identified in areas near the TSS and in the 1st exon. Hypomethylated a-CpGs were concentrated in areas outside of genes (hypergeometric test $p < 0.05$). For the YFS data, no statistically significant enrichment in regard to gene regions was observed. However, the trend was similar to that observed for the V90+ data. See Figure 10 for the distribution of a-CpGs in the V90+ and YFS data.

5.2.3 Functions of a-CpGs

To identify the pathways and processes that are affected by ageing-associated methylation changes, GOrilla (Eden et al., 2007; Eden et al., 2009), IPA (www.ingenuity.com) and PScan (Zambelli et al., 2009) were used. The pathway analyses were performed with genes containing the identified a-CpGs. In the V90+ data, the 3925 hypermethylated a-CpGs were located in 1832 different genes, and 4615 hypomethylated a-CpGs were located in 2057 different genes. In the YFS data, the 580 hypermethylated a-CpGs were located in 372 different genes, and the 622 hypomethylated a-CpGs were located in 417 different genes.

The genes containing hypermethylated a-CpGs identified in the V90+ data were enriched to 36 GO function terms and 265 GO process terms. In contrast, the genes containing hypomethylated a-CpGs identified in the V90+ data were enriched to only 27 GO function terms and to 53 GO process terms. Of these, 11 GO function terms and 41 GO process terms were common to hyper- and hypomethylated a-CpGs-containing genes. The hypermethylation-associated GO function terms were associated with sequence-specific DNA binding and transcription factor binding (Table 9). With respect to the hypomethylation-specific GO function terms, no common denominator was identified. The hypermethylation-specific GO process terms were associated with common processes, namely development and morphogenesis and metabolic processes (Tables 10 and 11). As with the GO function terms, no common denominator was identified for hypomethylation-specific GO process terms.

Table 9. Hypermethylation-specific GO function terms associated with DNA binding and transcription in the V90+ data. Also shown are the p-values from YFS for those GO terms that were identified in the YFS data. In total there were 36 significant GO function terms in V90+ data and 8 in YFS data Rank denotes the placement of the given GO term in the list of all significant GO function terms within one dataset.

		V90+		YFS	
GO Term	Description	p-value	Rank (out of 36)	p-value	Rank (out of 8)
GO:0043565	sequence-specific DNA binding	$1.18*10^{-32}$	1	$9.98*10^{-11}$	1
GO:0001071	nucleic acid binding transcription factor activity	$9.38*10^{-31}$	2	$1.90*10^{-7}$	3
GO:0003700	sequence-specific DNA binding transcription factor activity	$2.22*10^{-30}$	3	$1.90*10^{-7}$	4
GO:0003677	DNA binding	$6.48*10^{-16}$	4	$2.37*10^{-7}$	5
GO:0000981	sequence-specific DNA binding RNA polymerase II transcription factor activity	$2.60*10^{-15}$	5	$4.76*10^{-8}$	2
GO:0000976	transcription regulatory region sequence-specific DNA binding	$4.80*10^{-13}$	7		
GO:0044212	transcription regulatory region DNA binding	$7.92*10^{-12}$	9		
GO:0000975	regulatory region DNA binding	$2.22*10^{-11}$	10		
GO:0001067	regulatory region nucleic acid binding	$2.22*10^{-11}$	11		

The results obtained from YFS data mirror those of the V90+ data. The genes containing hypermethylated a-CpGs in the YFS data were enriched to 8 GO function terms and to 73 GO process terms, whereas genes containing hypomethylated a-CpGs were not enriched to any GO function or process terms. As in the V90+ data, the hypermethylation-associated GO function terms were associated with sequence-specific DNA binding and the GO process terms were associated with development and morphogenesis (See Tables 9, 10 and 11).

Table 10. Hypermethylation-specific GO process terms associated with development and morphogenesis in the V90+ data. Also shown are the p-values from YFS for those GO terms that were identified in the YFS data. In total there were 265 significant GO process terms in V90+ data and 73 in YFS data. Rank denotes the placement of the given GO term in the list of all significant GO process terms within one dataset.

		V90+		YFS	
GO Term	Description	p-value	Rank (out of 265)	p-value	Rank (out of 73)
GO:0048598	embryonic morphogenesis	$1.25*10^{-22}$	17	$8.36*10^{-8}$	28
GO:0048729	tissue morphogenesis	$2.99*10^{-19}$	22		
GO:0002009	morphogenesis of an epithelium	$6.94*10^{-18}$	26		
GO:0001763	morphogenesis of a branching structure	$1.84*10^{-17}$	29		
GO:0048754	branching morphogenesis of an epithelial tube	$1.26*10^{-15}$	49		
GO:0048562	embryonic organ morphogenesis	$6.12*10^{-14}$	67	$1.37*10^{-6}$	61
GO:0009887	organ morphogenesis	$1.13*10^{-13}$	71	$2.88*10^{-8}$	19
GO:0035107	appendage morphogenesis	$2.17*10^{-12}$	92	$2.79*10^{-8}$	17
GO:0035108	limb morphogenesis	$2.17*10^{-12}$	93	$2.79*10^{-8}$	18
GO:0030326	embryonic limb morphogenesis	$7.07*10^{-12}$	98	$2.71*10^{-7}$	34
GO:0035113	embryonic appendage morphogenesis	$7.07*10^{-12}$	99	$2.71*10^{-7}$	35
GO:0048704	embryonic skeletal system morphogenesis	$2.25*10^{-11}$	106		
GO:0048705	skeletal system morphogenesis	$1.04*10^{-10}$	113		
GO:0048732	gland development	$2.37*10^{-10}$	124		

Table 11. Hypermethylation-specific GO process terms associated with nucleotide metabolism, RNA metabolism and transcription in the V90+ data. Also shown are the p-values from YFS for those GO terms that were identified in the YFS data. In total there were 265 significant GO process terms in V90+ data and 73 in YFS data The rank denotes the placement of the given GO term in the list of all significant GO process terms within one dataset.

		V90+		YFS	
GO Term	Description	p-value	Rank (out of 265)	p-value	Rank (out of 73)
GO:0045935	positive regulation of nucleobase-containing compound metabolic process	$7.07*10^{-17}$	36		
GO:0051173	positive regulation of nitrogen compound metabolic process	$1.06*10^{-16}$	37		
GO:0031328	positive regulation of cellular biosynthetic process	$2.74*10^{-16}$	39	$4.69*10^{-7}$	41
GO:0009891	positive regulation of biosynthetic process	$3.61*10^{-16}$	41	$3.68*10^{-7}$	39
GO:0045893	positive regulation of transcription, DNA-templated	$5.05*10^{-16}$	45	$1.83*10^{-6}$	66
GO:0019219	regulation of nucleobase-containing compound metabolic process	$7.72*10^{-16}$	46	$1.24*10^{-6}$	56
GO:0010628	positive regulation of gene expression	$1.11*10^{-15}$	48	$1.33*10^{-8}$	12
GO:0006357	regulation of transcription from RNA polymerase II promoter	$5.05*10^{-15}$	53	$5.96*10^{-8}$	24
GO:0031326	regulation of cellular biosynthetic process	$1.07*10^{-14}$	57	$1.26*10^{-6}$	58
GO:0051171	regulation of nitrogen compound metabolic process	$1.28*10^{-14}$	58	$8.33*10^{-7}$	49
GO:0009889	regulation of biosynthetic process	$1.45*10^{-14}$	59	$1.24*10^{-6}$	55
GO:0051254	positive regulation of RNA metabolic process	$1.80*10^{-14}$	61	$6.97*10^{-7}$	47
GO:0006355	regulation of transcription, DNA-templated	$1.98*10^{-14}$	62	$4.88*10^{-7}$	42
GO:1902680	positive regulation of RNA biosynthetic process	$2.06*10^{-14}$	63	$1.52*10^{-6}$	63
GO:0045944	positive regulation of transcription from RNA polymerase II promoter	$2.94*10^{-14}$	65		
GO:0010557	positive regulation of macromolecule biosynthetic process	$6.72*10^{-14}$	68	$2.39*10^{-6}$	69
GO:0031323	regulation of cellular metabolic process	$1.09*10^{-13}$	70	$2.17*10^{-7}$	32
GO:2001141	regulation of RNA biosynthetic process	$1.37*10^{-13}$	72	$2.72*10^{-7}$	36

Table 11, continued from previous page.

GO Term	Description	V90+		YFS	
		P-value	Rank (out of 265)	P-value	Rank (out of 73)
GO:0031325	positive regulation of cellular metabolic process	$2.22*10^{-13}$	75	$1.46*10^{-6}$	62
GO:0045934	negative regulation of nucleobase-containing compound metabolic process	$2.80*10^{-13}$	77		
GO:0031327	negative regulation of cellular biosynthetic process	$3.75*10^{-13}$	78		
GO:0009893	positive regulation of metabolic process	$3.84*10^{-13}$	79	$8.89*10^{-9}$	10
GO:0009890	negative regulation of biosynthetic process	$3.84*10^{-13}$	80		
GO:0000122	negative regulation of transcription from RNA polymerase II promoter	$5.25*10^{-13}$	82	$3.49*10^{-6}$	72
GO:0051252	regulation of RNA metabolic process	$9.69*10^{-13}$	84	$3.41*10^{-7}$	38
GO:0051172	negative regulation of nitrogen compound metabolic process	$1.07*10^{-12}$	85		
GO:2000112	regulation of cellular macromolecule biosynthetic process	$1.14*10^{-12}$	86		
GO:0080090	regulation of primary metabolic process	$1.39*10^{-12}$	87		
GO:0010629	negative regulation of gene expression	$2.02*10^{-12}$	91		
GO:0010556	regulation of macromolecule biosynthetic process	$3.09*10^{-12}$	94		
GO:0045892	negative regulation of transcription, DNA-templated	$4.17*10^{-12}$	95		
GO:1902679	negative regulation of RNA biosynthetic process	$4.83*10^{-12}$	96		
GO:0019222	regulation of metabolic process	$1.56*10^{-11}$	102	$7.15*10^{-9}$	9
GO:0051253	negative regulation of RNA metabolic process	$1.94*10^{-11}$	104		
GO:0010468	regulation of gene expression	$1.35*10^{-10}$	117	$4.95*10^{-8}$	22
GO:0010558	negative regulation of macromolecule biosynthetic process	$1.39*10^{-10}$	119		
GO:0010604	positive regulation of macromolecule metabolic process	$1.49*10^{-10}$	123		
GO:2000113	negative regulation of cellular macromolecule biosynthetic process	$3.62*10^{-10}$	129		

In the V90+ data, IPA analysis identified 19 canonical pathways associated with hypermethylated a-CpGs-containing genes and 3 canonical pathways that were associated with hypomethylated-a-CpGs containing genes (BH-corrected p-value <0.05). Of these, one (*Axonal Guidance Signalling*) was common to both sets of genes. As these pathways were grouped to pathway categories, *Organismal growth and development*, *Cellular growth* and *Proliferation and development* were found to be the most enriched categories. In Figure 11, all of the affected pathway categories are shown. In the YFS data no statistically significant canonical pathways associated with a-CpGs-containing genes were identified.

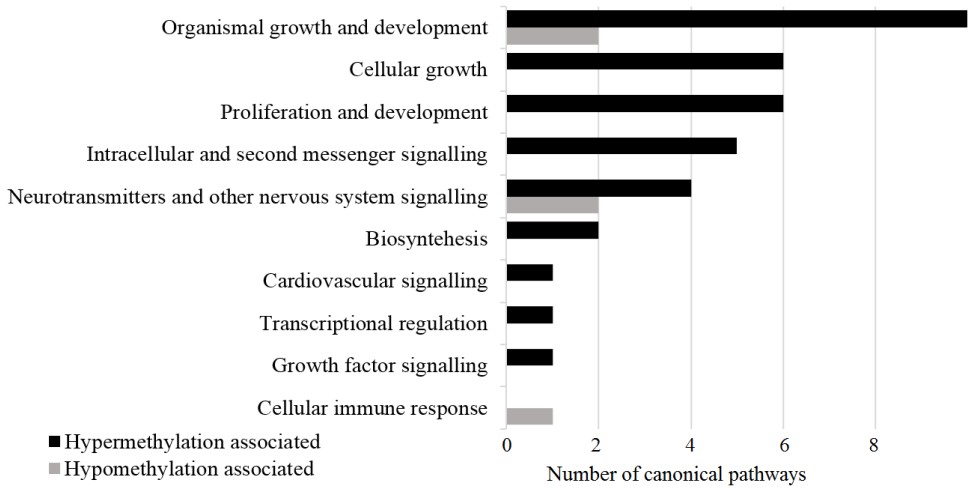


Figure 11. Canonical pathway categories associated with genes containing hyper- and hypomethylated a-CpGs in the V90+ data. One canonical pathway can belong to several pathway categories.

PScan (Zambelli et al., 2009) can be used to predict which TFs are common regulators for a group of genes. For genes containing hypermethylated a-CpGs in the V90+ data, we identified 24 common TFs, half of which were zinc-coordinating. For genes containing hypomethylated a-CpGs, only one common TF was identified. Genes containing hypermethylated a-CpGs in the YFS data were predicted to be regulated by 11 common TFs, 6 of which were zinc-coordinating. In the YFS data, no common TFs for hypomethylated a-CpGs containing genes were identified. Of the TFs predicted regulate the expression of hypermethylated a-CpGs-containing genes in the V90+ and YFS data, 4 were common to both data sets (SP1, EGR1, TFAP2A and E2F1).

Taken together, the number of hyper- and hypomethylated a-CpGs and genes containing these sites was comparable in both the V90+ and YFS datasets. In both datasets, the hypermethylated a-CPGs formed a more uniform and functional group, being enriched in more numerous GO terms and canonical pathways, and were predicted to have more common TFs. The results of both datasets are summarised in Table 12.

Table 12. Features of a-CpGs in the V90+ and YFS data. Presented are the location of enrichment in regard to CGIs and genes, the number of genes harbouring a-CpGs and the number of canonical pathways, GO terms and TFs for hyper- and hypomethylated a-CpGs in the V90+ and YFS data. The number of common a-CpGs, genes, GO terms and TFs between V90+ and YFS data are given in parentheses where applicable.

	V90+		YFS	
	Hypermethylated	Hypomethylated	Hypermethylated	Hypomethylated
n (a-CpGs)	3925	4615	580 (147)	622 (154)
CpG island location	CpG islands	Non-CGI	CpG islands	Non-CGI
Genomic location	TSS, 1st exon	Gene body, outside of genes	1st exon*	Outside of genes*
a-CpG associated genes	1832	2057	372 (187)	417 (176)
GO function terms	36	27	8 (7)	0
GO process terms	265	53	73 (63)	0
Canonical pathways	19	3	0	0
Transcription factors	24	1	11 (4)	0

*no statistically significant enrichment (hypergeometric test p >0.05)

5.2.4 Ageing-associated DNA methylation changes and sex

As we had identified sex-specific differences in ageing-associated gene expression changes (study I), we also wanted to investigate whether there are sex-specific differences in ageing-associated DNA methylation changes. However, in the V90+ data, we identified only 7 CpG sites where sex was a significant covariate, in addition to age group, in the regression model (Bonferroni corrected p-value <0.05). These sites were cg25990647 (*OPRK1*), cg18322569 (*BARHL2*), cg16355231 (*PEX10*), cg03078043 (*SYT12*), cg00552235, cg05316627 and cg01578875 (*ZNF827*). None of these genes were found to be differentially expressed between nonagenarians and young controls of either sex in study I.

In the YFS data, we identified more a-CpGs that were also associated with sex; in total, 79 CpG sites were located in 56 genes (Bonferroni-corrected p-value <0.05). Of these, only one was also found to be associated with sex in the V90+ data (cg16355231 in *PEX10*).

5.2.5 Ageing-associated DNA methylation changes and differences in cell type proportions

It has been shown (Jaffe & Irizarry, 2014) that differences in the proportions of blood cell types contribute significantly to observed DNA methylation differences. In our data sets, this result was verified using PCA. The first five PCs explained a significant proportion of variation in the methylation data: 2.0-20.5% in the V90+ data and 2.4-9.4% in the YFS data. These components were strongly associated with cell type proportions in both datasets (Tables 13 and 14).

Table 13. The correlation coefficient (Spearman) (r) between top five principal components (PCs) and proportions of different cell types in the V90+ data, p=p-value.

Principal component		PC1	PC2	PC3	PC4	PC5
Variance explained		20.5%	6.8%	3.5%	2.8%	2.0%
CD4+/CD8+ ratio	r	0.450	0.181	0.069	-0.122	0.127
	p	$3.96*10^{-8}$	0.035	0.423	0.159	0.141
CD14+	r	-0.295	0.210	0.617	-0.013	0.316
	p	$5.0*10^{-4}$	0.014	$1.20*10^{-15}$	0.885	$1.75*10^{-4}$
CD8+CD28-	r	-0.594	-0.029	0.366	0.010	0.151
	p	$2.52*10^{-14}$	0.737	$1.16*10^{-5}$	0.906	0.080
CD4+CD28-	r	-0.710	-0.081	0.443	-0.017	0.161
	p	$4.12*10^{-22}$	0.351	$6.66*10^{-8}$	0.841	0.061

Table 14. The correlation coefficient (Spearman) (r) between the top five principal components (PCs) and cell type proportions in the YFS data, p=p-value.

Principal component		PC1	PC2	PC3	PC4	PC5
Variance explained		9.4%	4.8%	3.1%	2.8%	2.4%
CD8 T cells	r	-0.684	0.109	0.175	0.047	-0.036
	p	$1.07*10^{-26}$	0.139	$1.76*10^{-2}$	0.526	0.613
CD4 T cells	r	-0.399	-0.237	0.241	-0.613	0.392
	p	$1.97*10^{-8}$	$1.22*10^{-3}$	$9.63*10^{-4}$	$2.33*10^{-20}$	$3.81*10^{-8}$
NK cells	r	-0.637	-0.048	-0.174	0.088	-0.087
	p	$2.51*10^{-22}$	0.521	$1.85*10^{-2}$	0.235	0.238
B cells	r	-0.270	0.197	0.108	-0.319	0.217
	p	$2.06*10^{-4}$	$7.32*10^{-3}$	0.145	$1.00*10^{-5}$	$3.08*10^{-3}$
Monocytes	r	0.307	-0.028	-0.186	0.075	-0.024
	p	$2.22*10^{-5}$	0.707	$1.13*10^{-2}$	0.310	0.749
Granylocytes	r	0.819	0.135	-0.128	0.431	-0.240
	p	$8.77*10^{-46}$	$6.74*10^{-2}$	$8.43*10^{-2}$	$9.76*10^{-10}$	$1.01*10^{-3}$

If the differences in cell type proportions are not adjusted for in the analysis, the number of a-CpGs can be overestimated. In the V90+ data, Wilcoxon's test identified 10083 CpG sites with a significant difference in the methylation level between nonagenarians and young controls. However, 1543 of these were not identified as significantly associated with age group in the regression model following adjustment for cell type proportions. This result indicates that the perceived difference in the methylation level was due to differences in cell type proportions and not age *per se*. The overestimation of numbers of a-CpGs is even more pronounced if the analysis is based solely on the regression model and not adjusted for cell type proportions. The regression model for the V90+ data identified 45507 CpG sites where age group was a significant covariate. When the analysis was performed without using cell type proportions as covariates, 94556 CpG sites were identified as ageing-associated, more than double the number of true a-CpGs.

5.3 Association between ageing-associated DNA methylation changes and gene expression (I & II)

DNA methylation is a regulator of gene expression. Thus, we sought to investigate if the ageing-associated gene expression differences identified in study I are associated with DNA methylation and if there are additional associations between the level of DNA methylation in a-CpGs and gene expression.

Table 15. Genes that were differentially expressed between nonagenarians and young controls of both sexes (study I) and for which the level of expression correlates with methylation level in nonagenarians. P-values are Benjamini-Hochberg (BH)-corrected. For genes that had several differentially expressed transcripts, fold changes (FC) for all of these are shown.

Gene	Probe	Correlation coefficient	p-value	FC females	FC males
ABLIM1	cg12649038	-0.26	0.021	-2.05/-2.32	-2.26/-2.00
	cg27290215	0.24	0.043		
ACSL1	cg27571769	-0.25	0.033	1.89	1.578
BACH2	cg03035849	-0.51	$1.77*10^{-7}$	-2.32	-2.26
BCL11A	cg09565597	0.24	0.037	-1.83/-1.60	-1.85/-1.58
	cg07469838	0.27	0.019		
CCR7	cg23663547	-0.42	$3.22*10^{-5}$	-3.61	-3.05
FAIM3	cg23088126	-0.45	$6.03*10^{-6}$	-2.19	-2.23
FAM117B	cg01745766	-0.24	0.037	-1.61/-1.60	-1.89/-1.78
FAM134B	cg15529432	-0.50	$3.46*10^{-7}$	1.56/1.50	1.51/1.62
	cg15529432	-0.42	$4.31*10^{-5}$		
LEF1	cg21600258	-0.25	0.031	-2.46/-2.41	-2.41/-2.34
LRRN3	cg09837977	-0.61	$<10^{-12}$	-5.64/-3.98	-4.68/-3.22
	cg19798735	-0.49	$5.07*10^{-7}$		
MAN1C1	cg10555744	-0.62	$<10^{-12}$	-1.87	-1.75
RORA	cg16964728	-0.24	0.042	1.93	1.70/2.08
STAP1	cg04398282	-0.42	$4.15*10^{-5}$	-1.63	-1.86
ZSCAN18	cg25784220	-0.37	$4.58*10^{-4}$	-1.97	-2.07

In study II, the analysis was performed by correlating the level of DNA methylation of a-CpGs to the expression of the gene in which the a-CpG was annotated. The gene expression data and the methylation data were obtained from

samples collected at the same time point. In the nonagenarians of the V90+ study, we identified 422 a-CpG-gene pairs with correlating levels of DNA methylation and gene expression (Pearson correlation, BH-corrected p-value <0.05). Of these 60% (255) showed an inverse correlation between DNA methylation level and gene expression.

One gene can contain several a-CpGs and several transcripts, and a given a-CpG can be located in a region of overlapping transcripts. These facts explain how 422 a-CpG-gene pairs consisted of 233 individual genes and 377 individual CpG sites. Of the 233 genes where the level of expression correlated with the level of DNA methylation, we identified 14 to be differentially expressed between nonagenarians and young controls in both sexes in study I (Table 15). An additional 14 were differentially expressed between nonagenarians and young controls in either sex.

Of the 422 correlating a-CpG-gene pairs, 31 showed a correlation coefficient above 0.5, including *LRRN3*-cg09837977. *LRRN3* was the most down-regulated gene between nonagenarians and young controls of both sexes in study I. A strong association between gene expression and DNA methylation was also observed for *ZNF154*, which is a zinc-finger gene located on chromosome 19. See Table 16 for the a-CpG-gene pairs with the strongest association.

Table 16. Gene-a-CpG-pairs where the level of expression is most strongly associated with DNA methylation (based on Pearson correlation coefficient) in nonagenarians of both sexes. P-values are Benjamini-Hochberg (BH)-corrected.

Gene	ID	Correlation coefficient	p-value
GFI1	cg04777348	-0.72	$<10^{-10}$
HOXC4	cg27138204	0.69	$<10^{-10}$
PYHIN1	cg19884600	-0.63	$<10^{-10}$
FCRL6	cg24833981	-0.63	$<10^{-10}$
MXRA7	cg00004089	-0.63	$<10^{-10}$
RUNX3	cg08544331	-0.63	$<10^{-10}$
PYHIN1	cg19884600	-0.62	$<10^{-10}$
MAN1C1	cg10555744	-0.62	$<10^{-10}$
FCRL6	cg19762800	-0.61	$<10^{-10}$
ZNF154	cg03234186	-0.61	$<10^{-10}$
LRRN3	cg09837977	-0.61	$<10^{-10}$
ZNF154	cg27049766	-0.61	$<10^{-10}$
ZNF154	cg08668790	-0.60	$<10^{-10}$
FAM134B	cg15529432	-0.59	$1.41*10^{-10}$

The genes displaying a correlation between expression and DNA methylation levels were enriched in 20 GO process terms (Bonferroni-corrected p-value <0.05), of which 30% (6) were associated with the immune system. In addition, many identified GO terms were associated with the reaction to environment. IPA revealed 15 affected canonical pathways (BH-corrected-p-value <0.05, Table 17), the majority of which were immune system associated or associated with cytoskeleton remodelling and endocytosis.

Table 17. Canonical pathways associated with genes where the level of expression is correlated with level of DNA methylation. P-values are Benjamini-Hochberg (BH)-corrected. The ratio is the number of identified genes in a given pathway divided by the number of total genes in the given pathway.

Ingenuity canonical pathways	p-value	Ratio
Integrin signalling	0.016	0.054
Actin cytoskeleton signalling	0.017	0.048
Tec kinase signalling	0.019	0.051
Agrin interactions at neuromuscular junction	0.019	0.090
Paxillin signalling	0.020	0.064
Reelin signalling in neurons	0.026	0.073
Phospholipase C signalling	0.030	0.041
Germ cell-Sertoli cell junction signalling	0.030	0.051
Crosstalk between dendritic cells and natural killer cells	0.030	0.066
Protein kinase A signalling	0.030	0.035
Antigen presentation pathway	0.030	0.100
Fcγ receptor-mediated phagocytosis in macrophages and monocytes	0.030	0.063
T helper cell differentiation	0.037	0.073
Ephrin receptor signalling	0.038	0.041
Caveolar-mediated endocytosis signalling	0.044	0.062

5.4 Manifestation of parental lifespan in the DNA methylation profile of the progeny (IV)

There are implications of the inheritance of acquired traits in mammals, and in some cases, the hereditary component may be epigenetic in nature. Therefore, we

investigated whether the length of parental lifespan is manifested in the DNA methylation profile of the progeny.

To minimise the bias caused by non-natural deaths at an early age, we included only individuals with both parents surviving >39 years. There were 90 such individuals in the V90+ study population (females n=66, males n=24).

The CpG sites where methylation level was associated with parental lifespan were identified with a regression model that was adjusted for cell type proportions, sex and batch. In addition, only sites with $\Delta\beta$ >1% between the progeny of longest-living and shortest living fathers or mothers were considered significant. We identified 659 CpG sites where the methylation level was associated with the length of paternal lifespan (BH-corrected p-value <0.05). No association with DNA methylation profile and the length of maternal lifespan was identified.

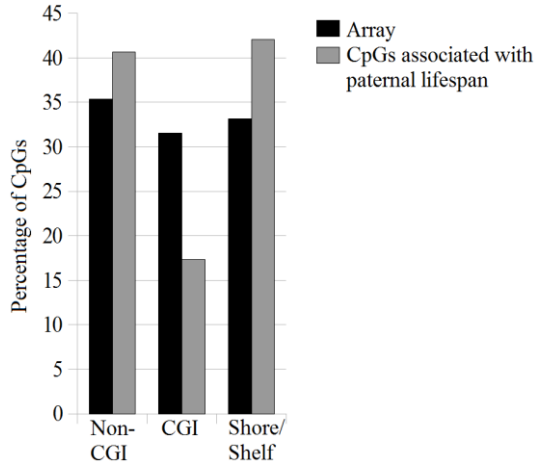


Figure 12. Location of CpG-sites associated with paternal lifespan. There were fewer than expected CpG sites in CGIs (hypergeometric test p <0.05).

Of the 659 CpG sites associated with paternal lifespan, the methylation level decreased with increasing paternal lifespan in 423 CpG sites (64%). The sites associated with the length of paternal lifespan were not enriched in any particular chromosome or gene location (hypergeometric test p >0.05), but there was fewer than expected CpG sites associated with paternal lifespan found in CpG islands (hypergeometric test p <0.05), see Figure 12.

The CpG sites where the difference in the methylation level between the progeny of longest-living and shortest-living fathers was largest are shown in Table 18. The largest difference was observed for cg19628988 in *CXXC5*. This

gene harboured 5 additional CpG sites where the level of expression was associated with paternal age at death.

Table 18. CpG sites with the largest difference in the methylation level between the progeny of the longest- and shortest-living fathers. P-values are Benjamini-Hochberg (BH)-corrected. Δβ denotes the difference in the methylation level between the progeny of the longest- and shortest-living fathers.

Gene	ID	p-value	Δβ
CXXC5	cg19628988	0.048	-0.082
NOTCH1	cg12076931	0.032	-0.080
KRT27	cg10747531	0.032	-0.077
na	cg11284147	0.047	-0.077
CXXC5	cg15165154	0.023	-0.072
MPZL1	cg04846203	0.035	-0.067
NOTCH4	cg06023661	0.038	-0.066
UEVLD	cg15846482	0.033	-0.065
SORT1	cg02175308	0.028	-0.065
DAP	cg14129473	0.032	-0.064
MORC2	cg23825480	0.047	0.055
RRAD	cg06410849	0.032	0.056
RESP18	cg19020434	0.032	0.057
ITPKB	cg23717186	0.037	0.059
na	cg00248242	0.041	0.059
CPA5	cg22664614	0.039	0.059
na	cg14828411	0.040	0.060
GULP1	cg16947583	0.034	0.062
EPM2AIP1	cg24607398	0.023	0.069
na	cg23644389	0.045	0.072

The 659 CpG sites associated with paternal age were located in 422 genes. These genes were enriched in 35 GO process terms (BH-corrected p-value <0.05) that were associated with development and morphogenesis as well as cell signalling (Table 19).

Table 19. GO process terms associated with genes where the methylation level is associated with paternal lifespan. P-values are Benjamini-Hochberg (BH)-corrected.

GO Term	Description	p-value
GO:0048523	negative regulation of cellular process	0.011
GO:0010646	regulation of cell communication	0.012
GO:0022603	regulation of anatomical structure morphogenesis	0.013
GO:0023051	regulation of signalling	0.014
GO:0040012	regulation of locomotion	0.016
GO:0044767	single-organism developmental process	0.016
GO:0009966	regulation of signal transduction	0.017
GO:0032502	developmental process	0.019
GO:0048519	negative regulation of biological process	0.022
GO:0030154	cell differentiation	0.022
GO:0009653	anatomical structure morphogenesis	0.024
GO:0051270	regulation of cellular component movement	0.024
GO:0050878	regulation of body fluid levels	0.024
GO:0050794	regulation of cellular process	0.025
GO:2000147	positive regulation of cell motility	0.025
GO:0044707	single-multicellular organism process	0.025
GO:0031325	positive regulation of cellular metabolic process	0.025
GO:0040017	positive regulation of locomotion	0.026
GO:0051239	regulation of multicellular organismal process	0.026
GO:0007165	signal transduction	0.026
GO:0090527	actin filament reorganization	0.026
GO:0048583	regulation of response to stimulus	0.026
GO:0009893	positive regulation of metabolic process	0.026
GO:0048522	positive regulation of cellular process	0.027
GO:0048856	anatomical structure development	0.027
GO:0030335	positive regulation of cell migration	0.027
GO:0051272	positive regulation of cellular component movement	0.028
GO:0048869	cellular developmental process	0.028
GO:0048518	positive regulation of biological process	0.029
GO:0032501	multicellular organismal process	0.030
GO:0007596	blood coagulation	0.032
GO:0050817	coagulation	0.033
GO:0007599	hemostasis	0.034
GO:0050789	regulation of biological process	0.035
GO:0065007	biological regulation	0.047

In our study population (V90+, n=90), paternal lifespan was not correlated with maternal lifespan (Spearman's rho=0.159, p=0.135) or with paternal age at conception (data available only for a subset of the population, n=21, Spearman's rho=-0.252, p=0.271). However, the offspring of longest-living fathers had more long-living siblings (siblings aged 85 or over) compared to progeny of the shortest-living fathers (Mann-Whitney U-test p=0.004). There was no association between maternal lifespan and the number of long-living siblings (p=0.148).

6 Discussion

6.1 Ageing-associated gene expression changes (I)

Ageing-associated gene expression changes have been analysed in whole blood, PBMCs and in purified leukocyte subsets (Cao et al., 2010; Harries et al, 2011; Irizar et al., 2015; Nakamura et al., 2012; Passtoors et al., 2012; Peters et al., 2015; Remondini et al., 2010; Reynolds et al., 2015). Such changes have also been studied in different solid tissues, such as skin, adipose tissue, brain and kidney (Erraji-Benchekroun et al., 2005; Glass et al., 2013; Hong et al., 2008; Rodwell et al., 2004; Zahn et al., 2006).

In general, genes identified as ageing-associated do not overlap a great deal between tissues or studies, indicating that tissues age differently, at least in terms of gene expression (Glass et al., 2013; Gheorghe et al., 2014; Hong et al., 2008; de Magalhães et al., 2009b; Reynolds et al., 2015; Rodwell et al., 2004). However, a few genes have been reported to be differentially expressed with age in several studies. *LRRN3* (leucine rich repeat neuronal 3), *CCR7* (chemokine (C-C motif) receptor 7) and *LEF1* (lymphoid enhancer-binding factor 1) have been identified as significantly down-regulated with age in our study I (Table 3) as well as in other studies using blood or PBMCs (Harries et al., 2011; Irizar et al., 2015; Jylhävä et al., 2010; Passtoors et al., 2012; Peters et al., 2015) and other tissues (Cao et al., 2010; Hong et al., 2008). Interestingly, our DNA methylation analysis (study II) showed that all three, *LRRN3*, *CCR7* and *LEF1,* contain CpG sites that are hypermethylated with increasing age and that the level of expression is inversely correlated with the level of DNA methylation in these genes (Table 15).

The genes that we identified as differentially expressed by age were almost exclusively associated with immune functions (Tables 4, 5 and 6). These results are concordant with those reported by others (Irizar et al., 2015; Passtoors et al., 2012; Remondini et al., 2010; Reynolds et al., 2015). This is a somewhat expected result given that the studies were performed using whole blood or a blood cell subpopulation and therefore were based on immune system cells. However, the ageing-associated changes in the expression of immune system associated genes has been identified also in kidney and brain tissue (Erraji-Benchekroun et al.,

2005; Rodwell et al., 2004; Zahn et al., 2006). The ageing-associated changes in the immune system, including immunosenescence and inflamm-aging, are extensively characterised (Arnold et al., 2011; Pawelec et al., 2010).

In addition to gene expression changes in the immune system, ageing-associated gene expression changes have also been reported in genes associated with RNA processing, chromatin remodelling and oxidative phosphorylation in mitochondria (Cao et al., 2010; Harries et al., 2011; Irizar et al., 2015 Reynolds et al., 2015). These results indicate that the very basic maintenance and metabolism processes in the cell are affected by ageing, and it can be speculated that these results represent the overall decline in functionality in ageing tissue. In our analysis, we did not identify marked expression changes in genes involved in these basic mechanisms. This difference is possibly because, in our analysis, we used a strict preprocessing pipeline, with only the top 5% of transcripts with the largest CV included in the analysis (2367 transcripts in total). It can be assumed that genes responsible for the basic functions of the cell do not change their expression to the same degree as those associated with immune functions and thus were excluded from our analysis.

It has also been reported that ageing-associated gene expression changes are not linear throughout a person's lifespan. There are certain ages where the rate of change is accelerated or decelerated and these differ between different tissues (Gheorghe et al., 2014; Remondini et al., 2010).

According to our results, gene expression changes associated with ageing differ between the sexes. Compared with males, slightly more genes and significantly more pathways were identified as differentially expressed in females. Pathways associated with ageing only in females were most prominently associated with nitric oxide production and signalling in T cells (Table 6). Our results support the idea of sexual dimorphism in the immune system but also in ageing. In general, females are considered to be more immunocompetent compared with males. Females are more resilient to different infections and certain cancers but are also more prone to autoimmune diseases (McClelland & Smith, 2011; Nunn et al., 2009).

To our knowledge, others studying ageing-associated gene expression changes have not assessed the role of sex. However, Jansen et al. (2014) studied gene expression differences between the sexes without the effect of ageing. According to their results, the genes overexpressed in females compared with males are enriched in immune system-associated GO term categories. These authors showed that the expression differences decrease when males are compared to

post-menopausal females and increase when compared to females using oral contraceptives, indicating that oestrogen is an important regulator of the sex-biased genes. As the levels of sex hormones, specifically that of oestrogen, decrease significantly during ageing (Horstman et al., 2012), it remains to be speculated what factors contribute to the perceived ageing-associated difference in gene expression between the sexes. As shown by the results of our DNA methylation analysis (study II), differences in DNA methylation are not a major contributor to sexual dimorphism of ageing-associated gene expression changes.

6.2 Ageing-associated DNA methylation changes (II & III)

DNA methylation changes associated with ageing are currently under extensive research, and the number of published studies is rapidly increasing. The great majority of recently reported results are obtained with Illumina Infinium HumanMethylation450 or 27 BeadChip and the most common tissues investigated are blood or different blood cell subpopulations (Bacalini et al., 2015; Bell et al., 2012; Florath et al., 2013; Garagnani et al., 2012; Gentilini et al., 2015; Hannum et al., 2013; Heyn et al., 2012; Johansson et al., 2013; McClay et al., 2014; Rakyan et al., 2010; Reynolds et al., 2014; Steegenga et al., 2014; Teschendorff et al., 2010; Tserel et al., 2015; Weidner et al., 2014; Xu & Taylor, 2013; Yuan et al., 2015). In addition, ageing-associated DNA methylation changes have been studied in muscle and brain tissue (Farré et al., 2015; Hernandez et al., 2011; Zykovich et al., 2014).

Of the methylation sites identified as ageing-associated in different studies, only a very limited set has been identified in several of the studies. Steegenga et al. (2014) compared the results of eight different studies in which 7477 different CpG sites have been identified as ageing-associated (Bell et al., 2012; Florath et al., 2013; Garagnani et al., 2012; Hannum et al., 2012; Heyn et al., 2012; Rakyan et al., 2010; Teschendorff et al., 2010; Xu & Taylor, 2014). Of these 7477 CpG sites, only 529 were reported by two or more groups. Our analysis shows a stronger concordance; of the 1202 a-CpGs identified in the YFS data (study III), 987 were also identified in the V90+ data (study II) when analysed with comparable methods (regression model only). Notably, the analysis methods in studies II and III were the same, but a different tissue was used (PBMCs in study II and whole blood, which includes also granulocytes, in study III). This agreement between the two studies implies that a large portion of the variation in

results previously reported by others may be due to differences in data processing and analysis methods, including the lack of adjustment for differences in cell type proportions.

Despite the lack of concordance in previous results, a handful of top hits have been reported in several of the studies. The single most often reported a-CpG, showing a large change in methylation level with age and high statistical significance, is cg16867657 in the *ELOVL2* gene (fatty acid elongase 2), which is hypermethylated with ageing (Bacalini et al., 2015; Hannum et al., 2012; Heyn et al., 2012; Florath et al., 2013; Garagnani et al., 2012; Johansson et al., 2013; Reynolds et al., 2014; Steegenga et al., 2014; Tserel et al., 2015; Zykovich et al., 2011). In both of our datasets, YFS (study III) and V90+ (study II), this locus was the most strongly ageing-associated CpG site. Despite its strong association with ageing, no functional role or association with ageing related impairments for the ageing-associated methylation change in *ELOVL2* have been reported. Other repeatedly reported ageing-associated CpG sites that were also identified in our studies include those in genes *FHL2*, *OTUD7A*, *PENK*, *KLF14* and *EDARADD* (Tables 1, 7 and 8). All of these genes, except *EDARADD*, are hypermethylated with ageing (Bacalini et al., 2015; Hannum et al., 2012; Heyn et al., 2012; Florath et al., 2013; Garagnani et al., 2012; Steegenga et al., 2014; Teschendorff et al., 2010; Xu & Taylor, 2013; Zykovich et al., 2011). As for *ELOVL2*, no functional role for these methylation changes in relation to the ageing process have been reported, and these genes are not associated with pathways linked to ageing-associated methylation changes. In our results (study II), the methylation level in these sites was not associated with gene expression levels of these genes. This result suggests that these sites are a part of a possible epigenetic clock mechanism that is associated purely with chronological age.

Ageing-associated CpG sites identified in our studies were not distributed uniformly across the genome. In both the V90+ and YFS datasets, the sites that were hypermethylated with age were predominantly located within 200 bp of the TSS or the 1st exon and in CGIs. In contrast, sites that were hypomethylated with age were located in gene bodies or outside of genes in non-CGIs (Figure 10). These data are in line with other reported results (Florath et al., 2013; Hernandez et al., 2011; Heyn et al., 2012; Johansson et al., 2013; McClay et al., 2014; Reynolds et al., 2014; Steegenga et al., 2014; Tserel et al., 2015; Yuan et al., 2015). As the majority of CGIs overlapping TSSs are unmethylated at baseline (D'Aquila et al., 2013), the observed pattern was unsurprising. However, this

result indicates that both the *de novo* methylation mechanisms as well as methylation maintenance are affected with ageing.

The genes containing CpG sites that were hypermethylated with age in our studies were associated with development and morphogenesis as well as DNA binding and gene expression (Tables 9, 10 and 11, Figure 11). The enrichment of ageing-associated hypermethylation in these processes has also been reported by others (Florath et al., 2013; Hernandez et al., 2011; Johansson et al., 2013; Rakyan et al., 2010). Interestingly, RNA processing was also reported to be affected in studies of ageing-associated gene expression changes (Cao et al., 2010; Harries et al., 2011; Reynolds et al., 2015). In contrast to ageing-associated hypermethylated sites and genes, no hypomethylation-associated GO terms or pathways were identified in the YFS data, and there were very few hypomethylation-associated GO terms and pathways in the V90+ data compared with hypermethylation-associated terms. This discrepancy was not due to different numbers of hyper- and hypomethylated sites; in fact, the hypomethylated sites were slightly more numerous (54% and 52% of all identified a-CpGs were hypomethylated in the V90+ and YFS data, respectively) (Table 12). Reynolds et al. (2014) and Yuan et al. (2015) also reported that the CpG sites hypermethylated with age are enriched in common processes and exhibit shared features, whereas hypomethylated a-CpGs are a less homogenous group. In addition, age-associated hypermethylation interactome hotspots have been reported (West et al., 2013).

Our findings along with those reported previously suggest that ageing-associated hyper- and hypomethylation are distinct processes, with possibly different causes and effects. We suggest that ageing-associated hypermethylation is primarily caused by programmed changes, whereas ageing-associated hypomethylation is primarily caused by stochastic and environmental effects. This hypothesis would imply that hypermethylation could be a strong predictor of chronological age, whereas hypomethylation could better serve as a marker of biological age, reflecting the environmental insults and/or allostatic load experienced by the individual.

The ageing-associated hypermethylation changes were enriched in common processes, such as development and morphogenesis (Table 10, Figure 11), implying a common driving force responsible for these changes. Also, the majority of the ageing-associated methylation changes reported by more than one study are hypermethylation events, including that of *ELOVL2*. While the strongest associations with ageing in our data are observed for hypermethylated

sites, hypomethylation events are more numerous, especially when the data are examined with a less-strict significance threshold. The excess of ageing-associated hypomethylation events has also been reported by others (Heyn et al., 2012; Johansson et al., 2013). Hypermethylation is by necessity an active, energy-consuming process, whereas hypomethylation may occur either actively or passively (Bhutani et al., 2011; Jones & Liang, 2009). Additionally, global hypomethylation has been associated with an increased risk of frailty (Bellizzi et al., 2012).

The need to adjust for differences in the proportions of different cell types when analysing DNA methylation in tissues composed of multiple cell types has been discussed in the literature. Reports advocating adjustment (Farré et al., 2015; Lam et al., 2012; Yuan et al., 2015) and those declaring it unnecessary (Bell et al., 2012; Tserel et al., 2015; Weidner et al., 2014; Xu & Taylor, 2014) have both been published. In these reports, the main purpose of the study has usually been to analyse the effects of ageing on DNA methylation, and the stance on cell type adjustment is usually a by-product on the obtained results. A report that directly considered this matter (Jaffe & Irizarry, 2014) concluded that differences in the proportions should be adjusted for in samples consisting of mixed cell types. This conclusion is reinforced by results showing that different blood cell types show distinctive DNA methylation profiles (Reinius et al., 2012; Zilbauer et al., 2013). Our PCA results also show that cell type proportions explain a significant part of the variation in the methylation data (Tables 13 and 14). Many studies claiming that differences in cell type proportions need not to be adjusted for have concentrated on a small number of CpGs that are very significantly associated with age. Strongly ageing-associated sites may represent ageing-associated programmed changes that are present in all tissues and thus are unaffected by differences in cell composition. However, the stochastic ageing-associated changes may be more modest, and thus the role of cell type heterogeneity becomes more evident and should be adjusted for.

In general, the results from the V90+ and YFS analysis (studies II & III) regarding ageing-associated DNA methylation changes were concordant in terms of the direction of change in the methylation levels, the location of a-CpGs and the associated cellular processes and functions of the a-CpGs. Fewer a-CpGs were identified in the YFS than in the V90+ data, which was expected given the narrow age range of the YFS population. In general, ageing-associated DNA methylation changes most likely include both linearly changing sites and sites where the change in the methylation level is increased at either end of the lifespan,

analogous to the findings regarding ageing-associated gene expression changes (Gheorghe et al., 2014; Remondini et al., 2010). It is possible that the sites identified in both datasets, V90+ and YFS, represent the former type, whereas sites identified only in the V90+ data represent the latter (see Figure 13).

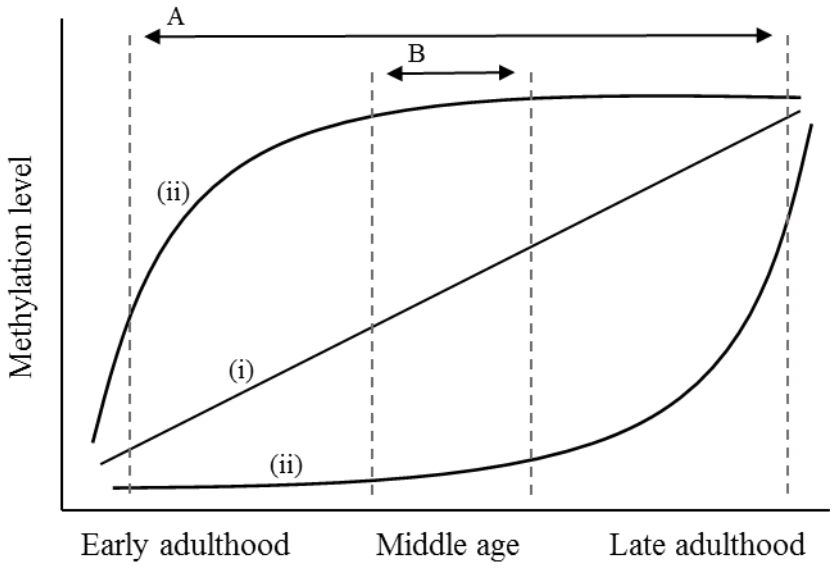


Figure 13. Ageing-associated DNA methylation changes may occur in a linear fashion (i) throughout lifespan, or they may be accelerated (ii) at either end of the lifespan. Both types of changes can be identified in a population with a broad age range (A), whereas in a population with a narrow age range (B) only the linearly occurring changes can be identified.

One notable exception between the V90+ and YFS data is the lack of overrepresentation of hypermethylation in chromosome 19 in YFS that was observed in the V90+ data. The hypermethylated sites in chromosome 19 were primarily in genes encoding zinc finger proteins. It has been proposed that the function of zinc finger proteins on chromosome 19 is to repress endogenous retroviruses (ERVs) (Lukic et al., 2014). Ageing is associated with genomic instability (López-Otín et al., 2013), and the possible expression of ERVs could contribute to this instability. In mice, ageing has been associated with changes in the expression of ERVs (Gaubatz et al., 1991; Wada et al., 1993). In our data (study II), we identified a correlation between DNA methylation and gene expression level for *ZNF154* (Table 16), indicating a functional role for the DNA methylation change. It appears that the change in the methylation level in

chromosome 19 zinc finger genes does not occur linearly with age but is possibly accelerated in the oldest old.

6.3 The relationship between ageing-associated DNA methylation changes and gene expression (I & II)

Of the 8540 a-CpGs identified in V90+ data (study II), only 377 showed an association with the expression of a corresponding gene (a-CpG annotated to the given gene). As one gene can contain several CpGs, there were 233 individual genes showing a correlation between expression and methylation. Of these, only 28 were identified as differentially expressed between nonagenarians and young controls in study I (Table 15). The majority of studies assessing the correlation between methylation and gene expression with array-based methods have reported this poor correlation (Hannum et al., 2012; Horvath, 2013; Johansson et al., 2013; Lam et al., 2012; Xu & Taylor, 2014; Yuan et al., 2015; Zykovich et al., 2014). However in contrast to our data, in several of the mentioned studies, expression and methylation levels were not obtained from the same samples; rather, one or the other was based on database information, leading to technical explanations to the poor correlation between gene expression and DNA methylation.

The poor correlation between DNA methylation and gene expression observed in our study is probably due to both biological and technical factors (see also Limitations, 6.6.2). The regulation of gene expression is a complex process, involving TFs, chromatin remodelling and histone modifications in addition to DNA methylation. Even the effects of CpG methylation alone are complex given that intergenic CGIs possibly represent alternative TSSs or regulate the expression of ncRNAs (Illingworth & Bird, 2009; Rauch et al., 2009; Rinn et al., 2007). In addition, we analysed the correlation between single CpG sites and single genes and only for CpGs sites located within the given gene. Thus, regulatory effects of long-distance CpGs were not detected and neither were regulatory effects of larger blocks of CpG sites. In a methylation array, both methylated and unmethylated CpGs are registered; in contrast, in the gene expression array, a non-expressed gene gives no signal. If hypermethylation completely prevents the expression of a given gene, this gene is omitted from the analysis, and the existing correlation between expression and DNA methylation is not detected.

The genes that displayed an association with DNA methylation levels were enriched in immunological processes as well as cytoskeletal remodelling and endocytosis (Table 17). The latter are also linked to immune function as they are required for leukocyte activation, migration and phagocytosis (Fenteany & Glogauer, 2004). In addition, in purified immune system cells (monocytes and CD8+ T cells), genes whose expression level was associated with DNA methylation level have been determined to be associated with immune functions (Reynolds et al., 2014; Tserel et al., 2015). Our results and those reported by others imply that ageing-associated altered functionality of the immune system may be partly mediated via changes in DNA methylation. However, our results also reflect the fact that the majority of genes with detectable expression levels in the samples were immune system associated.

6.4 Effects of parental lifespan on the DNA methylation profile of the progeny (IV)

In our study population (V90+), only paternal lifespan was associated with the DNA methylation profile of their nonagenarian progeny. No corresponding association was identified for maternal lifespan. It is noteworthy that the number of long-living siblings was also associated with paternal, not maternal, lifespan. These data support our results and suggest that different results regarding the association between paternal or maternal lifespan and the DNA methylation profile of the progeny are not merely artefacts caused by the study population. Ageing is known to exhibit sexual dimorphism (Seifarth et al., 2012), and the evidence also points to sexual dimorphism in the case of heritability of lifespan and transgenerational inheritance (Grossniklaus et al., 2013; You et al., 2010), although conflicting reports exist regarding the heritability of lifespan (Dutta et al., 2013). Our results support the notion of sexual dimorphism in heritability of lifespan and lifespan-associated features.

However, the probability of surviving to very old ages differs greatly between males and females. While females survive to extreme old ages due to heritable factors and/or lifestyle factors, males tend to survive to extreme old ages due to heritable factors only (Evert et al., 2003). It would be interesting to analyse whether the effect of maternal lifespan on DNA methylation profile of the progeny can be identified in a population consisting of individuals from families with a history of long lifespan. It is possible that the maternal effect on DNA

profile of the progeny cannot be identified in our population, as it consists of both familial and sporadic cases of longevity.

The CpG sites associated with paternal lifespan in our study were primarily located outside of CGIs (Figure 12). It has been demonstrated in mice that CpG sites associated with paternal environmental effects are primarily located in CpG-poor regions of the genome. In the mice studies, the location of the CpG sites associated with different exposures (jet fuel, plastics, ethanol) was similar in the genome, but there was no overlap between genes associated with these exposures (Skinner & Guerrero-Bosagna, 2014). CpG-poor regions of the genome have been proposed to have important regulatory roles (Hahn et al., 2011; Irizarry et al., 2009). It remains to be determined whether the methylation state of CpGs on CGIs is somehow more protected against environmental exposures compared with CpGs outside of CGIs due to, e.g., differences in the chromatin state.

Many genes and pathways associated with paternal lifespan in our study were associated with cell signalling (Table 19). The majority of CpG sites in these pathways were hypomethylated with advancing paternal age. Gentilini et al. (2013) previously showed that hypomethylated genes in the progeny of centenarians (as compared to progeny of non-long-lived parents) are associated with cell signalling. Cell signalling pathways function as an intertwined network, and changes in signalling intensities of these pathways has been associated with ageing (Carlson et al., 2008). Our results and those previously reported by others (Gentilini et al., 2013) suggest that ageing-associated changes in signalling pathways may be partially mediated via DNA methylation and affected by parental lifespan.

In addition to the cell signalling, genes with CpGs that were associated with paternal lifespan were associated with development and morphogenesis (Table 19). In our study on ageing-associated DNA methylation changes, we identified these processes to be associated with genes that were hypermethylated with ageing (studies II & III, Table 10, Figure 11). The association between developmental processes and morphogenesis and ageing-associated hypermethylation has been reported by others, as well (Florath et al., 2013; Gentilini et al., 2013; Johansson et al., 2013; Rakyan et al., 2010).

We speculate that the identified CpGs represent intergenerational epigenetic inheritance. However, we cannot exclude the possibility that the heritable component could be another epigenetic feature (i.e., not DNA methylation) or a genetic element that produces the observed DNA methylation profile. The molecular mechanism of epigenetic inheritance is still somewhat controversial as

the epigenome goes through two major reprogramming steps, one in primordial germs cells and another in the embryo after implantation (Heard & Martienssen et al., 2014; Szyf, 2015). However, epigenetic inheritance remains at least plausible, as certain loci show parent-of-origin-dependent expression (Hanna & Kelsey, 2014). In addition, certain single-copy loci avoid reprogramming in mice (Hackett et al., 2012; Seisenberger et al., 2012).

These sites may also represent one of the components that mediates the heritability of lifespan. Traditional genetic elements alone do not explain the 20-30% heritability of lifespan identified in epidemiological studies (Brooks-Wilson, 2013). Epigenetic components are possibly a part of this "missing" heritability.

6.5 Role of developmental pathways in ageing and longevity (II, III & IV)

Genes that are hypermethylated with increasing age in both the V90+ and YFS data were associated with organismal growth, development and morphogenesis, as shown by pathway and GO term analysis (see Table 10 and Figure 11). Our results are in line with those published by other groups (Florath et al., 2013; Hernandez et al., 2011; Johansson et al., 2013; Rakyan et al., 2010). Importantly, compared with ageing-associated hypomethylation, ageing-associated hypermethylation appears to be a regulated process (Table 12). In addition, genes where the methylation level is associated with paternal lifespan are also associated with organismal growth, development and morphogenesis (Table 19).

The hyperfunction theory of ageing proposes that ageing, in its essence, is the aimless continuation of developmental growth programmes (Blagosklonny, 2013a). This theory is supported by the reported association between the speed of development to maturity and age at death (Blagosklonny, 2013b) and also by long-lived mutant organisms that show impaired development (Kirkwood & Melov, 2011). In addition, the only known lifespan-extending intervention, calorie restriction, acts through mTOR, which modulates anabolic versus catabolic processes in response to nutrients, growth signals and the energy status of the cell (Johnson et al., 2013).

DNA methylation and other epigenetic mechanisms are important regulators of cell fate and differentiation (Boland et al., 2014; Cedar & Bergman, 2012). We speculate that the perceived methylation changes associated with ageing in our studies are parallel to those needed to control differentiation and lineage

commitment during developmental and differentiation phases of an individual. Thus, these changes represent the aimless continuation proposed by the hyperfunction theory of ageing. This suggestion is complemented by the identified differences in ageing-associated hyper- and hypomethylation. As the fidelity of development is of utmost importance, the small evolutionary pressure caused by the possible detrimental effects of these changes in old age are insufficient to select against them in the course of evolution.

6.6 Limitations of the study

6.6.1 Study population

The V90+ study population is a population-based sample, including both home-dwelling and institutionalised individuals. However, as it consists of the oldest of the old, the most frail and dependent on the help of others probably are underrepresented in the population. There are fewer males than females in the V90+ study population, which may have affected the results on sex differences. Additionally, the number of controls in the V90+ is smaller than that of nonagenarians, which affects the power of statistical testing to identify differentially expressed genes and differentially methylated CpG sites.

Our study population is cross-sectional; thus, some observed differences between nonagenarians and young controls may be due to birth cohort effects. However, longitudinal studies spanning the 70 years between our control and study subjects are challenging to execute in human populations.

Data on the age or age at death of the parents and siblings of the V90+ study participants were collected in an interview of the study subjects. As a portion of the subjects show impaired cognitive functions (MMSE <23), these data may contain some inaccuracies.

6.6.2 Gene expression analysis

For technical reasons, array-based gene expression analysis methods cannot distinguish a truly non-expressed gene from a gene with no expression signal. For this reason, genes that are not expressed in a given proportion of samples are often omitted from analysis. Therefore, the most drastic gene expression changes, i.e.,

genes that are completely non-expressed in one or the other study group, are omitted from the analysis.

In addition, recent reports have suggested that the transcriptome and proteome become increasingly discordant with age (Janssens et al., 2015). This further complicates the interpretation on gene expression array results.

Our gene expression analysis (study I) suffers from somewhat poorly chosen significance thresholds. In the preprocessing step the threshold was relatively strict, as only 5% of the probes (2367 probes), based on CV, were included in the final analysis, possibly excluding true differentially expressed transcripts from the final results. In contrast, the significance threshold used for statistical testing in the pathway analysis was relaxed (BH-corrected p-value <0.25), possibly leading to false positive pathways in the list of significant pathways.

6.6.3 DNA methylation analysis

The Illumina Infinium HumanMethylation450 BeadChip includes 485000 probes for CpG sites covering 99% of RefSeq genes and 96% of CGIs in the human genome (Illumina), making it the best available tool for high-throughput analysis of genome-wide DNA methylation. However, there are 28 million CpG sites in the human genome (D'Aquila et al., 2013), meaning that the methylation level of only 1.7% of the CpG sites can be identified. It has been suggested that to circumvent this issue and to identify truly biologically meaningful results, DNA methylation data should be analysed as larger blocks, i.e., as groups of probes that are adjacent to each other and show concordant methylation levels (Bacalini et al., 2015). As our data was analysed as individual CpG sites, some biologically significant ageing-associated DNA methylation changes may have been excluded from the results.

When interpreting DNA methylation analysis results, it should be taken into account that in each cell, a given CpG site can only be methylated or unmethylated, i.e., the methylation value in a given cell is 0 or 1. When the methylation level of a population of cells is measured, it is a continuous variable between 0 and 1. By adjusting for differences in the proportions of different cell types in samples of mixed cell types, this issue can be addressed. However, it is often the case that not all different cell types in the given sample can be accounted for. In addition, the different methylation status between different cells of a given cell type cannot be accounted for. As our analysis were not adjusted for all cell

types present in the sample (for example B cells in V90+ (studies II & IV) or CD4CD28- T cells in YFS (study III)), some results obtained may be artefacts caused by differences in these cell populations between the individuals.

In addition, the Illumina Infinium HumanMethylation450 BeadChip cannot distinguish between 5mC and 5hmC; instead, both modifications are detected as 5mC. 5hmC has been proposed to be a distinct epigenetic modification with specific functions (Breiling & Lyko, 2015), and the 5hmC content of whole blood DNA was recently characterised and associated with ageing (Xiong et al., 2015). Xiong et al. (2015) showed that ageing-associated decreases in the level of 5hmC were more pronounced compared with 5mC, necessitating further analysis of the role of 5hmC in ageing.

7 Summary and Conclusions

Here, we have characterised ageing-associated gene expression and DNA methylation changes in the oldest old and DNA methylation changes in middle-aged individuals. In addition, we investigated the possible intergenerational epigenetic inheritance of lifespan effects. The main results of these studies can be summarised as follows:

1. Ageing-associated gene expression changes show sexual dimorphism. However, ageing-associated DNA methylation changes fail to explain the majority of the observed gene expression changes, and specifically, the observed sexual dimorphism.
2. Ageing-associated hyper- and hypomethylation show distinct characteristics in terms of location (in relation to genes and CGIs) and associated genes, suggesting different causes and consequences for these processes. Ageing-associated hypermethylation appears to be a regulated process, whereas ageing-associated hypomethylation is due to environmental and stochastic effects.
3. Ageing-associated hypermethylation is associated with developmental pathways, supporting the proposed hyperfunction theory of ageing.
4. Paternal lifespan is associated with certain DNA methylation features in the progeny, implying the possibility of intergenerational epigenetic inheritance of lifespan.
5. As a technical note, our results affirm the need to adjust for differences in cell type proportions when analysing DNA methylation in samples composed of mixed cell types.

Our results add to the existing knowledge on ageing-associated gene expression and DNA methylation changes. Specifically, these results not only contribute to the understanding of these changes in the oldest old but also highlight the parallels between ageing-associated changes in middle-aged individuals and those at the end of adulthood. In addition, study IV shows the

effect of the full range of parental lifespan on progeny DNA methylation for the first time.

The interpretation of studies regarding ageing-associated gene expression and DNA methylation studies is complex due to both the methods used and the nature of ageing itself. Ageing is not a well-defined disease state that has a clear onset; ageing is a continuum that starts after maturation and progresses at different speeds in different individuals. Additionally, not all individuals are affected by all ageing-associated changes. Changes in the level of gene expression may be both cause and effect, thus complicating the analysis. DNA methylation is only one of the epigenetic mechanisms regulating gene expression, and it is clear that different epigenetic mechanisms function as an intertwined network. However, DNA methylation is currently the only epigenetic mechanism that can be analysed efficiently from patient samples with high-throughput methods. Thus, it will be important to study ageing in well-described cohorts and to combine different types of array data to gain a more comprehensive understanding of the molecular mechanisms of ageing.

We need to characterise the molecular mechanisms of ageing to fully understand ageing-associated phenotypic changes. Thus, it could be possible to pinpoint the causative mechanisms behind these detrimental changes and possibly develop interventions to prevent and postpone these effects. However, it is important to note that the ultimate goal of ageing research should not be the crude extension of lifespan but the preservation of youthful functionality under the cruel assault of passing time.

8 Acknowledgements

The studies of this PhD thesis were carried out at the School of Medicine, University of Tampere, at the department of Microbiology and Immunology during the years 2010–2016. My work was financially supported by Doctoral Programme in Biomedicine and Biotechnology, School of Medicine of University of Tampere and the Academy of Finland. The studies were financially supported by the Academy of Finland, the Competitive Research Fund of the Pirkanmaa Hospital District, Yrjö Jahnsson Foundation, Tampere Tuberculosis Foundation and Jenny and Antti Wihuri Foundation. The Science Fund of the City of Tampere is acknowledged for a grant to cover the printing expenses of this thesis.

I want to thank my supervisor Professor Mikko Hurme, MD, PhD, for giving me the opportunity to work in his group. I appreciate that in addition to the thesis work, I've been given time to participate in other projects as well. It has been wonderful to be able to immerse oneself in the world of ageing research.

The official reviewers of the thesis manuscript, Docent Antero Salminen, PhD and Docent Cilla Söderhäll, PhD are greatly acknowledged.

Professor Marja Jylhä, MD, PhD and Professor Terho Lehtimäki, MD, PhD, co-authors and members of my thesis committee are thanked for their help during my thesis work. In addition, Professor Jylhä deserves my thanks for the inspiring lectures given over the years, which provided me with another interesting angle on ageing research.

Docent Juulia Jylhävä, PhD, is warmly thanked for all the invaluable help I've received during the years. Although not an official supervisor of my thesis work, she has contributed greatly to my development from a rookie scientist to a not-quite-so-rookie scientist.

I want to thank my fellow PhD students, co-authors and co-workers Laura Kananen, MSc, and Tapio Nevalainen, MSc, for their contribution in these studies. Their contribution to the extension of my lifespan in the form of shared jokes is also appreciated.

My co-authors not mentioned before, Professor Matti Nykter, PhD, Professor Antti Hervonen, MD, PhD, Professor Mika Kähönen, MD, PhD, Professor Olli Raitakari, MD, PhD, Professor Pärt Peterson, PhD, Nina Mononen, PhD, Liina

Tserel, PhD and Sergei Häyrynen, are acknowledged for their important contributions to the original publications of this thesis.

I want to thank all the former co-workers in the department of Microbiology and Immunology, and also the people who participated in the collection of Vitality 90+ samples, Carita Eklund, PhD, Taru Kuparinen, MSc, Elina Nummenmaa, MSc, Tapio Kotipelto, MSc, Atte Haarala MD, PhD, Markus Teittinen, Katri Välimaa, Sanna Tuominen, Eeva Timonen and Janette Hinkka for their help and contribution to these studies. Sinikka Repo-Koskinen deserves a special thank-you, not only for participation in data collection and laboratory work, but also for making the day-to-day work so easy by taking care of all the little things that nobody truly wants to do, but somebody still has to.

I also wish to thank all the participants of Vitality 90+ study and Young Finns Study, without whom these studies would not have been possible.

I want to thank my friend and partner in crime Emma for enjoying the odd and broad world of biological sciences with me. It is soothing to now that I'm not the only one genuinely interested in the teeth of an aardvark, life of a short-eared elephant shrew, the epigenetic regulation in honey bees or the genetics of cat coat coloration. Elina, my almost-sister, who reminds me that there is a world outside of science, with whom I can enjoy all those things beginning with "black" or "folk" also deserves a big thank-you.

Isä ja äiti ansaitsevat ison kiitoksen kaikesta siitä tuesta, jota olen saanut aina koulutien alusta tähän päivään asti. En ole koskaan epäillyt pystymistäni, selviytymistäni tai pärjäämistäni, koska tekään ette ole. Äidille erityiskiitos villasukista (kirjoitan tätä ne mustaharmaat jalassa) ja isälle sarvista (ne eivät ole nyt päässä).

Tuomas, my sun and stars, my *hellren*, my *mellon*. I still agree with Jessica Rabbit. You make me laugh.

Tampere, February 2016

Saara Marttila

9 References

Abbott A. (2004) Ageing: growing old gracefully. Nature 428(6979):116-118.

Ademokun A, Wu YC, Dunn-Walters D. (2010) The ageing B cell population: composition and function. Biogerontology 11(2):125-137.

Ahuja N, Li Q, Mohan AL, Baylin SB, Issa JP. (1998) Aging and DNA methylation in colorectal mucosa and cancer. Cancer Res 58(23):5489-5494.

Anderson RM, Bitterman KJ, Wood JG, Medvedik O, Sinclair DA. (2003) Nicotinamide and PNC1 govern lifespan extension by calorie restriction in Saccharomyces cerevisiae. Nature 423(6936):181-185.

Anway MD, Cupp AS, Uzumcu M, Skinner MK. (2005) Epigenetic transgenerational actions of endocrine disruptors and male fertility. Science 308(5727):1466-1469.

Anway MD, Memon MA, Uzumcu M, Skinner MK. (2006) Transgenerational effect of the endocrine disruptor vinclozolin on male spermatogenesis. J Androl 27(6):868-879.

Anway MD, Rekow SS, Skinner MK. (2008) Transgenerational epigenetic programming of the embryonic testis transcriptome. Genomics 91(1):30-40.

Armanios M, Blackburn EH. (2012) The telomere syndromes. Nat Rev Genet 13(10):693-704.

Arnold CR, Wolf J, Brunner S, Herndler-Brandstetter D, Grubeck-Loebenstein B. (2011) Gain and loss of T cell subsets in old age--age-related reshaping of the T cell repertoire. J Clin Immunol 31(2):137-146.

Aspinall R, Andrew D. (2000) Thymic involution in aging. J Clin Immunol 20(4):250-256.

Atamna H, Cheung I, Ames BN. (2000) A method for detecting abasic sites in living cells: age-dependent changes in base excision repair. Proc Natl Acad Sci U S A 97(2):686-691.

Atzmon G, Schechter C, Greiner W, Davidson D, Rennert G, Barzilai N. (2004) Clinical phenotype of families with longevity. J Am Geriatr Soc 52(2):274-277.

Bacalini MG, Boattini A, Gentilini D, Giampieri E, Pirazzini C, Giuliani C, Fontanesi E, Remondini D, Capri M, Del Rio A, Luiselli D, Vitale G, Mari D, Castellani G, Di Blasio AM, Salvioli S, Franceschi C, Garagnani P. (2015) A meta-analysis on age-associated changes in blood DNA methylation: results from an original analysis pipeline for Infinium 450k data. Aging (Albany NY) 7(2):97-109.

Bahar R, Hartmann CH, Rodriguez KA, Denny AD, Busuttil RA, Dollé ME, Calder RB, Chisholm GB, Pollock BH, Klein CA, Vijg J. (2006) Increased cell-to-cell variation in gene expression in ageing mouse heart. Nature 441(7096):1011-1014.

Baker DJ, Dawlaty MM, Wijshake T, Jeganathan KB, Malureanu L, van Ree JH, Crespo-Diaz R, Reyes S, Seaburg L, Shapiro V, Behfar A, Terzic A, van de Sluis B, van Deursen JM. (2013) Increased expression of BubR1 protects against aneuploidy and cancer and extends healthy lifespan. Nat Cell Biol 15(1):96-102.

Barres R, Osler ME, Yan J, Rune A, Fritz T, Caidahl K, Krook A, Zierath JR. (2009) Non-CpG methylation of the PCGC-1alpha promoter through DNMT3B controls mitochondrial density. Cell Metab 10:189–198.

Bártová E, Krejcí J, Harnicarová A, Galiová G, Kozubek S. (2008) Histone modifications and nuclear architecture: a review. J Histochem Cytochem 56(8):711-721.

Barzilai N, Huffman DM, Muzumdar RH, Bartke A. (2012) The critical role of metabolic pathways in aging. Diabetes 61(6):1315-1322.

Beekman M, Nederstigt C, Suchiman HE, Kremer D, van der Breggen R, Lakenberg N, Alemayehu WG, de Craen AJ, Westendorp RG, Boomsma DI, de Geus EJ, Houwing-Duistermaat JJ, Heijmans BT, Slagboom PE. (2010) Genome-wide association study (GWAS)-identified disease risk alleles do not compromise human longevity. Proc Natl Acad Sci U S A 107(42):18046-18049.

Behl C, Skutella T, Lezoualc'h F, Post A, Widmann M, Newton CJ, Holsboer F. (1997) Neuroprotection against oxidative stress by estrogens: structure-activity relationship. Mol Pharmacol 51(4):535-541.

Bektas A, Zhang Y, Lehmann E, Wood WH 3rd, Becker KG, Madara K, Ferrucci L, Sen R. (2014) Age-associated changes in basal NF-κB function in human CD4+ T lymphocytes via dysregulation of PI3 kinase. Aging (Albany NY) 6(11):957-974.

Bell AC, Felsenfeld G. (2000) Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene. Nature 405(6785):482-485.

Bell JT, Tsai PC, Yang TP, Pidsley R, Nisbet J, Glass D, Mangino M, Zhai G, Zhang F, Valdes A, Shin SY, Dempster EL, Murray RM, Grundberg E, Hedman AK, Nica A, Small KS; MuTHER Consortium, Dermitzakis ET, McCarthy MI, Mill J, Spector TD, Deloukas P. (2012) Epigenome-wide scans identify differentially methylated regions for age and age-related phenotypes in a healthy ageing population. PLoS Genet 8(4):e1002629.

Bellizzi D, D'Aquila P, Montesanto A, Corsonello A, Mari V, Mazzei B, Lattanzio F, Passarino G. (2012) Global DNA methylation in old subjects is correlated with frailty. Age (Dordr) 34(1):169-179.

Belsky DW, Caspi A, Houts R, Cohen HJ, Corcoran DL, Danese A, Harrington H, Israel S, Levine ME, Schaefer JD, Sugden K, Williams B, Yashin AI, Poulton R, Moffitt TE. (2015) Quantification of biological aging in young adults. Proc Natl Acad Sci U S A 112(30):E4104-10.

Ben-Avraham D. (2015) Epigenetics of aging. Adv Exp Med Biol 847:179-191.

Bhutani N, Burns DM, Blau HM. (2011) DNA demethylation dynamics. Cell 146(6):866-872.

Biagi E, Candela M, Franceschi C, Brigidi P. (2011) The aging gut microbiota: new perspectives. Ageing Res Rev 10(4):428-429.

Bird A, Taggart M, Frommer M, Miller OJ, Macleod D. (1985) A fraction of the mouse genome that is derived from islands of nonmethylated, CpG-rich DNA. Cell 40(1):91-99.

Birky CW Jr. (2001) The inheritance of genes in mitochondria and chloroplasts: laws, mechanisms, and models. Annu Rev Genet 35:125-148.

Blagosklonny MV. (2012) Answering the ultimate question "what is the proximal cause of aging?" Aging (Albany NY) 4(12):861-877.

Blagosklonny MV. (2013a) Aging is not programmed: genetic pseudo-program is a shadow of developmental growth. Cell Cycle 12(24):3736-3742.
Blagosklonny MV. (2013b) Big mice die young but large animals live longer. Aging (Albany NY) 5(4):227-233.
Blagosklonny MV. (2014) Koschei the immortal and anti-aging drugs. Cell Death Dis 5:e1552.
Bocklandt S, Lin W, Sehl ME, Sánchez FJ, Sinsheimer JS, Horvath S, Vilain E. (2011) Epigenetic predictor of age. PLoS One 6(6):e14821.
Bodkin NL, Alexander TM, Ortmeyer HK, Johnson E, Hansen BC. (2003) Mortality and morbidity in laboratory-maintained Rhesus monkeys and effects of long-term dietary restriction. J Gerontol A Biol Sci Med Sci 58(3):212-219.
Bogdanović O, Veenstra GJ. (2009) DNA methylation and methyl-CpG binding proteins: developmental requirements and function. Chromosoma 118(5):549-565.
Boland MJ, Nazor KL, Loring JF. (2014) Epigenetic regulation of pluripotency and differentiation. Circ Res 115(2):311-324.
Bolduc V, Chagnon P, Provost S, Dubé MP, Belisle C, Gingras M, Mollica L, Busque L. (2008) No evidence that skewing of X chromosome inactivation patterns is transmitted to offspring in humans. J Clin Invest 118(1):333-341.
Bollati V, Schwartz J, Wright R, Litonjua A, Tarantini L, Suh H, Sparrow D, Vokonas P, Baccarelli A. (2009) Decline in genomic DNA methylation through aging in a cohort of elderly subjects. Mech Ageing Dev 130(4):234-239.
Boonekamp JJ, Simons MJ, Hemerik L, Verhulst S. (2013) Telomere length behaves as biomarker of somatic redundancy rather than biological age. Aging Cell 12(2):330-332.
Borgel J, Guibert S, Li Y, Chiba H, Schübeler D, Sasaki H, Forné T, Weber M. (2010) Targets and dynamics of promoter DNA methylation during early mouse development. Nat Genet 42(12):1093-1100.
Borrás C, Sastre J, García-Sala D, Lloret A, Pallardó FV, Viña J. (2003) Mitochondria from females exhibit higher antioxidant gene expression and lower oxidative damage than males. Free Radic Biol Med 34(5):546-552.
Breiling A, Lyko F. (2015) Epigenetic regulatory functions of DNA modifications: 5-methylcytosine and beyond. Epigenetics Chromatin 8:24.
Brooks-Wilson AR. (2013) Genetics of healthy aging and longevity. Hum Genet 132(12):1323-1338.
Brykczynska U, Hisano M, Erkek S, Ramos L, Oakeley EJ, Roloff TC, Beisel C, Schübeler D, Stadler MB, Peters AH. (2010) Repressive and active histone methylation mark distinct promoters in human and mouse spermatozoa. Nat Struct Mol Biol 17(6):679-687.
Bygren LO, Tinghög P, Carstensen J, Edvinsson S, Kaati G, Pembrey ME, Sjöström M. (2014) Change in paternal grandmothers' early food supply influenced cardiovascular mortality of the female grandchildren. BMC Genet 15:12.
Campanero MR, Armstrong MI, Flemington EK. (2000) CpG methylation as a mechanism for the regulation of E2F activity. Proc Natl Acad Sci U S A 97(12):6481-6486.
Campisi J, d'Adda di Fagagna F. (2007) Cellular senescence: when bad things happen to good cells. Nat Rev Mol Cell Biol 8(9):729-740.

Campisi J, Robert L. (2014) Cell senescence: role in aging and age-related diseases. Interdiscip Top Gerontol 39:45-61.
Candore G, Balistreri CR, Listì F, Grimaldi MP, Vasto S, Colonna-Romano G, Franceschi C, Lio D, Caselli G, Caruso C. (2006) Immunogenetics, gender, and longevity. Ann N Y Acad Sci 1089:516-37.
Cao JN, Gollapudi S, Sharman EH, Jia Z, Gupta S. (2010) Age-related alterations of gene expression patterns in human CD8+ T cells. Aging Cell 9(1):19-31.
Cao-Lei L, Massart R, Suderman MJ, Machnes Z, Elgbeili G, Laplante DP, Szyf M, King S. (2014) DNA methylation signatures triggered by prenatal maternal stress exposure to a natural disaster: Project Ice Storm. PLoS One 9(9):e107653.
Carlson ME, Silva HS, Conboy IM. (2008) Aging of signal transduction pathways, and pathology. Exp Cell Res 314(9):1951-1961.
Cedar H, Bergman Y. (2012) Programming of DNA methylation patterns. Annu Rev Biochem 81:97-117.
Challen GA, Sun D, Jeong M, Luo M, Jelinek J, Berg JS, Bock C, Vasanthakumar A, Gu H, Xi Y, Liang S, Lu Y, Darlington GJ, Meissner A, Issa JP, Godley LA, Li W, Goodell MA. (2011) Dnmt3a is essential for hematopoietic stem cell differentiation. Nat Genet 44(1):23-31.
Chen YA, Lemire M, Choufani S, Butcher DT, Grafodatskaya D, Zanke BW, Gallinger S, Hudson TJ, Weksberg R. (2013) Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. Epigenetics 8(2):203-209.
Chondrogianni N, Stratford FL, Trougakos IP, Friguet B, Rivett AJ, Gonos ES. (2003) Central role of the proteasome in senescence and survival of human fibroblasts: induction of a senescence-like phenotype upon its inhibition and resistance to stress upon its activation. J Biol Chem 278(30):28026-28037.
Colman RJ, Beasley TM, Kemnitz JW, Johnson SC, Weindruch R, Anderson RM. (2014) Caloric restriction reduces age-related and all-cause mortality in rhesus monkeys. Nat Commun 5:3557.
Cosgrove MS, Wolberger C. (2005) How does the histone code work? Biochem Cell Biol 83(4):468-476.
Cossetti C, Lugini L, Astrologo L, Saggio I, Fais S, Spadafora C. (2014) Soma-to-germline transmission of RNA in mice xenografted with human tumour cells: possible transport by exosomes. PLoS One 9(7):e101629.
Crews D, Gillette R, Miller-Crews I, Gore AC, Skinner MK. (2014) Nature, nurture and epigenetics. Mol Cell Endocrinol 398(1-2):42-52.
Cribari-Neto F, Zeileis A. (2010) Beta regression in R. JSS 34:1-24.
Cribbs A, Feldmann M, Oppermann U. (2015) Towards an understanding of the role of DNA methylation in rheumatoid arthritis: therapeutic and diagnostic implications. Ther Adv Musculoskelet Dis 7(5):206-219.
Dahl RE, Siegel SF, Williamson DE, Lee PA, Perel J, Birmaher B, Ryan ND. (1992) Corticotropin releasing hormone stimulation test and nocturnal cortisol levels in normal children. Pediatr Res 32(1):64-68.
D'Aquila P, Rose G, Bellizzi D, Passarino G. (2013) Epigenetics and aging. Maturitas 74(2):130-136.

Deaton AM, Bird A. (2011) CpG islands and the regulation of transcription. Genes Dev 25:1010-1022.

Deelen J, Uh HW, Monajemi R, van Heemst D, Thijssen PE, Böhringer S, van den Akker EB, de Craen AJ, Rivadeneira F, Uitterlinden AG, Westendorp RG, Goeman JJ, Slagboom PE, Houwing-Duistermaat JJ, Beekman M. (2013) Gene set analysis of GWAS data for human longevity highlights the relevance of the insulin/IGF-1 signaling and telomere maintenance pathways. Age (Dordr) (1):235-49.

Denis H, Ndlovu MN, Fuks F. (2011) Regulation of mammalian DNA methyltransferases: a route to new mechanisms. EMBO Rep 12(7):647-656.

Devanapally S, Ravikumar S, Jose AM. (2015) Double-stranded RNA made in C. elegans neurons can enter the germline and cause transgenerational gene silencing. (2015) Proc Natl Acad Sci U S A 112(7):2133-2138.

Dobyns WB, Filauro A, Tomson BN, Chan AS, Ho AW, Ting NT, Oosterwijk JC, Ober C. (2004) Inheritance of most X-linked traits is not dominant or recessive, just X-linked. Am J Med Genet A 129A(2):136-143.

Dobzhansky, T. (1973) nothing in biology makes sense except in the light of evolution. Am Biol Teatcher 35:125-129.

Drinkwater RD, Blake TJ, Morley AA, Turner DR. (1989) Human lymphocytes aged in vivo have reduced levels of methylation in transcriptionally active and inactive DNA. Mutat Res 219(1):29-37.

Du P, Kibbe WA, Lin SM. (2008) lumi: a pipeline for processing Illumina microarray. Bioinformatics 24(13):1547-1548.

Dutta A, Henley W, Robine JM, Langa KM, Wallace RB, Melzer D. (2013) Longer lived parents: protective associations with cancer incidence and overall mortality. J Gerontol A Biol Sci Med Sci 68(11):1409-1418.

Dutta A, Henley W, Robine JM, Llewellyn D, Langa KM, Wallace RB, Melzer D. (2014) Aging children of long-lived parents experience slower cognitive decline. Alzheimers Dement 10(5 Suppl):S315-322.

Eden E, Lipson D, Yogev S, Yakhini Z. (2007) Discovering motifs in ranked lists of DNA sequences. PLoS Comput Biol 3(3):e39.

Eden E, Navon R, Steinfeld I, Lipson D, Yakhini Z. (2009) GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. BMC Bioinformatics 10:48.

ElSharawy A, Keller A, Flachsbart F, Wendschlag A, Jacobs G, Kefer N, Brefort T, Leidinger P, Backes C, Meese E, Schreiber S, Rosenstiel P, Franke A, Nebel A. (2012) Genome-wide miRNA signatures of human longevity. Aging Cell 11(4):607-616.

Erraji-Benchekroun L, Underwood MD, Arango V, Galfalvy H, Pavlidis P, Smyrniotopoulos P, Mann JJ, Sibille E. (2005) Molecular aging in human prefrontal cortex is selective and continuous throughout adult life. Biol Psychiatry 57(5):549-558.

Evert J, Lawler E, Bogan H, Perls T. (2003) Morbidity profiles of centenarians: survivors, delayers, and escapers. J Gerontol A Biol Sci Med Sci 58(3):232-237.

Farré P, Jones MJ, Meaney MJ, Emberly E, Turecki G, Kobor MS. (2015) Concordant and discordant DNA methylation signatures of aging in human blood and brain. Epigenetics Chromatin 8:19.

Fenteany G, Glogauer M. (2004) Cytoskeletal remodeling in leukocyte function. Curr Opin Hematol 11(1):15-24.
Ferrari C. (2004) Beta regression for modelling rates and proportions. J Appl Stat 31:799-815.
Florath I, Butterbach K, Müller H, Bewerunge-Hudler M, Brenner H. (2014) Cross-sectional and longitudinal changes in DNA methylation with age: an epigenome-wide analysis revealing over 60 novel age-associated CpG sites. Hum Mol Genet 23(5):1186-1201.
Folstad I, Karter J. (1992) Parasites, bright males, and the immunocompetence handicap. Am Nat 139:603-622.
Fontana L, Partridge L. (2015) Promoting health and longevity through diet: from model organisms to humans. Cell 161(1):106-118.
Fraga MF, Ballestar E, Paz MF, Ropero S, Setien F, Ballestar ML, Heine-Suñer D, Cigudosa JC, Urioste M, Benitez J, Boix-Chornet M, Sanchez-Aguilera A, Ling C, Carlsson E, Poulsen P, Vaag A, Stephan Z, Spector TD, Wu YZ, Plass C, Esteller M. (2005) Epigenetic differences arise during the lifetime of monozygotic twins. Proc Natl Acad Sci U S A 102(30):10604-10609.
Fraga MF, Esteller M. (2007) Epigenetics and aging: the targets and the marks. Trends Genet 23(8):413-418.
Franceschi C, Bonafè M, Valensin S, Olivieri F, De Luca M, Ottaviani E, De Benedictis G. (2000a) Inflamm-aging. An evolutionary perspective on immunosenescence. Ann N Y Acad Sci 908:244-254.
Franceschi C, Valensin S, Bonafè M, Paolisso G, Yashin AI, Monti D, De Benedictis G. (2000b) The network and the remodeling theories of aging: historical background and new perspectives. Exp Gerontol 35(6-7):879-896.
Franceschi C, Campisi J. (2014) Chronic inflammation (inflammaging) and its potential contribution to age-associated diseases. J Gerontol A Biol Sci Med Sci 69 Suppl 1:S4-9.
Franklin TB, Russig H, Weiss IC, Gräff J, Linder N, Michalon A, Vizi S, Mansuy IM. (2010) Epigenetic transmission of the impact of early stress across generations. Biol Psychiatry 68(5):408-415.
Freitas AA, de Magalhães JP. (2011) A review and appraisal of the DNA damage theory of ageing. Mutat Res 728(1-2):12-22.
Fried LP, Tangen CM, Walston J, Newman AB, Hirsch C, Gottdiener J, Seeman T, Tracy R, Kop WJ, Burke G, McBurnie MA; Cardiovascular Health Study Collaborative Research Group. (2001) Frailty in older adults: evidence for a phenotype. J Gerontol A Biol Sci Med Sci 56(3):M146-156.
Ganna A, Rivadeneira F, Hofman A, Uitterlinden AG, Magnusson PK, Pedersen NL, Ingelsson E, Tiemeier H. (2013) Genetic determinants of mortality. Can findings from genome-wide association studies explain variation in human mortality? Hum Genet 132(5):553-561.
Gapp K, Jawaid A, Sarkies P, Bohacek J, Pelczar P, Prados J, Farinelli L, Miska E, Mansuy IM. (2014) Implication of sperm RNAs in transgenerational inheritance of the effects of early trauma in mice. Nat Neurosci 17(5):667-669.
Garagnani P, Bacalini MG, Pirazzini C, Gori D, Giuliani C, Mari D, Di Blasio AM, Gentilini D, Vitale G, Collino S, Rezzi S, Castellani G, Capri M, Salvioli S,

Franceschi C. (2012) Methylation of ELOVL2 gene as a new epigenetic marker of age. Aging Cell 11(6):1132-1134.
Garinis GA, van der Horst GT, Vijg J, Hoeijmakers JH. (2008) DNA damage and ageing: new-age ideas for an age-old problem. Nat Cell Biol 10(11):1241-1247.
Gaubatz JW, Arcement B, Cutler RG. (1991) Gene expression of an endogenous retrovirus-like element during murine development and aging. Mech Ageing Dev 57(1):71-85.
Gentilini D, Castaldi D, Mari D, Monti D, Franceschi C, Di Blasio AM, Vitale G. (2012) Age-dependent skewing of X chromosome inactivation appears delayed in centenarians' offspring. Is there a role for allelic imbalance in healthy aging and longevity? Aging Cell 11:277-283.
Gentilini D, Mari D, Castaldi D, Remondini D, Ogliari G, Ostan R, Bucci L, Sirchia SM, Tabano S, Cavagnini F, Monti D, Franceschi C, Di Blasio AM, Vitale G. (2013) Role of epigenetics in human aging and longevity: genome-wide DNA methylation profile in centenarians and centenarians' offspring. Age (Dordr) 35(5):1961-1973.
Gentilini D, Garagnani P, Pisoni S, Bacalini MG, Calzari L, Mari D, Vitale G, Franceschi C, Di Blasio AM. (2015) Stochastic epigenetic mutations (DNA methylation) increase exponentially in human aging and correlate with X chromosome inactivation skewing in females. Aging (Albany NY) 7(8):568-578.
Gheorghe M, Snoeck M, Emmerich M, Bäck T, Goeman JJ, Raz V. (2014) Major aging-associated RNA expressions change at two distinct age-positions. BMC Genomics 15:132.
Gibson KL, Wu YC, Barnett Y, Duggan O, Vaughan R, Kondeatis E, Nilsson BO, Wikby A, Kipling D, Dunn-Walters DK. (2009) B-cell diversity decreases in old age and is correlated with poor health status. Aging Cell 8(1):18-25.
Glass D, Viñuela A, Davies MN, Ramasamy A, Parts L, Knowles D, Brown AA, Hedman AK, Small KS, Buil A, Grundberg E, Nica AC, Di Meglio P, Nestle FO, Ryten M; UK Brain Expression consortium; MuTHER consortium, Durbin R, McCarthy MI, Deloukas P, Dermitzakis ET, Weale ME, Bataille V, Spector TD. (2013) Gene expression changes with age in skin, adipose tissue, blood and brain. Genome Biol 14(7):R75.
Goebeler S, Jylhä M, Hervonen A. (2003) Medical history, cognitive status and mobility at the age of 90. A population-based study in Tampere, Finland. Aging Clin Exp Res 15(2):154-161.
Goldberg EL, Romero-Aleshire MJ, Renkema KR, Ventevogel MS, Chew WM, Uhrlaub JL, Smithey MJ, Limesand KH, Sempowski GD, Brooks HL, Nikolich-Žugich J. (2015) Lifespan-extending caloric restriction or mTOR inhibition impair adaptive immunity of old mice by distinct mechanisms. Aging Cell 14(1):130-138.
Gombar S, Jung HJ, Dong F, Calder B, Atzmon G, Barzilai N, Tian XL, Pothof J, Hoeijmakers JH, Campisi J, Vijg J, Suh Y. (2012) Comprehensive microRNA profiling in B-cells of human centenarians by massively parallel sequencing. BMC Genomics 13:353.
Gooren LJ, Giltay EJ, Bunck MC. (2008) Long-term treatment of transsexuals with cross-sex hormones: extensive personal experience. J Clin Endocrinol Metab 93(1):19-25.

Greer EL, Maures TJ, Ucar D, Hauswirth AG, Mancini E, Lim JP, Benayoun BA, Shi Y, Brunet A. (2011) Transgenerational epigenetic inheritance of longevity in Caenorhabditis elegans. Nature 479(7373):365-371.

Grossniklaus U, Kelly WG, Ferguson-Smith AC, Pembrey M, Lindquist S. (2013) Transgenerational epigenetic inheritance: how important is it? Nat Rev Genet 14(3):228-235.

Guerra-Silveira F, Abad-Franch F. (2013) Sex bias in infectious disease epidemiology: patterns and processes. PLoS One 8(4):e62390.

Guibert S, Forné T, Weber M. (2012) Global profiling of DNA methylation erasure in mouse primordial germ cells. Genome Res 22(4):633-641.

Hackett JA, Zylicz JJ, Surani MA. (2012) Parallel mechanisms of epigenetic reprogramming in the germline. Trends Genet 28(4):164-174.

Hackett JA, Sengupta R, Zylicz JJ, Murakami K, Lee C, Down TA, Surani MA. (2013) Germline DNA demethylation dynamics and imprint erasure through 5-hydroxymethylcytosine. Science 339(6118):448-452.

Hahn MA, Wu X, Li AX, Hahn T, Pfeifer GP. (2011) Relationship between gene body DNA methylation and intragenic H3K9me3 and H3K36me3 chromatin marks. PLoS One 6(4):e18844.

Hammoud SS, Nix DA, Zhang H, Purwar J, Carrell DT, Cairns BR. (2009) Distinctive chromatin in human sperm packages genes for embryo development. Nature 460(7254):473-478.

Han S, Brunet A. (2012) Histone methylation makes its mark on longevity. Trends Cell Biol 22(1):42-49.

Hanna CW, Kelsey G. (2014) The specification of imprints in mammals. Heredity (Edinb) 113(2):176-183.

Hannum G, Guinney J, Zhao L, Zhang L, Hughes G, Sadda S, Klotzle B, Bibikova M, Fan JB, Gao Y, Deconde R, Chen M, Rajapakse I, Friend S, Ideker T, Zhang K. (2013) Genome-wide methylation profiles reveal quantitative views of human aging rates. Mol Cell 49(2):359-367.

Harman D. (1965) The free radical theory of aging: Effect of age on serum copper levels. J Gerontol 20:151-153.

Harries LW, Hernandez D, Henley W, Wood AR, Holly AC, Bradley-Smith RM, Yaghootkar H, Dutta A, Murray A, Frayling TM, Guralnik JM, Bandinelli S, Singleton A, Ferrucci L, Melzer D. (2011) Human aging is characterized by focused changes in gene expression and deregulation of alternative splicing. Aging Cell 10(5):868-878.

Hashimoto H, Liu Y, Upadhyay AK, Chang Y, Howerton SB, Vertino PM, Zhang X, Cheng X. (2012) Recognition and potential mechanisms for replication and erasure of cytosine hydroxymethylation. Nucleic Acids Res 40(11):4841-4849.

Hatakeyama C, Anderson CL, Beever CL, Peñaherrera MS, Brown CJ, Robinson WP. (2004) The dynamics of X-inactivation skewing as women age. Clin Genet 66(4):327-332.

Heard E, Martienssen RA. (2014) Transgenerational epigenetic inheritance: myths and mechanisms. Cell 157(1):95-109.

Heijmans BT, Tobi EW, Stein AD, Putter H, Blauw GJ, Susser ES, Slagboom PE, Lumey LH. (2008) Persistent epigenetic differences associated with prenatal exposure to famine in humans. Proc Natl Acad Sci U S A 105(44):17046-17049.

Hekimi S, Lapointe J, Wen Y. (2011) Taking a "good" look at free radicals in the aging process. Trends Cell Biol 21(10):569-576.

Hellman A, Chess A. (2007) Gene body-specific methylation on the active X chromosome. Science 315(5815):1141-1143.

Henikoff S, Furuyama T. (2012) The unconventional structure of centromeric nucleosomes. Chromosoma 121(4):341-352.

Hernandez DG, Nalls MA, Gibbs JR, Arepalli S, van der Brug M, Chong S, Moore M, Longo DL, Cookson MR, Traynor BJ, Singleton AB. (2011) Distinct DNA methylation changes highly correlated with chronological age in the human brain. Hum Mol Genet 20(6):1164-1172.

Heyn H, Li N, Ferreira HJ, Moran S, Pisano DG, Gomez A, Diez J, Sanchez-Mut JV, Setien F, Carmona FJ, Puca AA, Sayols S, Pujana MA, Serra-Musach J, Iglesias-Platas I, Formiga F, Fernandez AF, Fraga MF, Heath SC, Valencia A, Gut IG, Wang J, Esteller M. (2012) Distinct DNA methylomes of newborns and centenarians. Proc Natl Acad Sci U S A 109(26):10522-10527.

High KP. (2004) Infection as a cause of age-related morbidity and mortality. Ageing Res Rev 3(1):1-14.

Hill PW, Amouroux R, Hajkova P. (2014) DNA demethylation, Tet proteins and 5-hydroxymethylcytosine in epigenetic reprogramming: an emerging complex story. Genomics 104(5):324-333.

vB Hjelmborg J, Iachine I, Skytthe A, Vaupel JW, McGue M, Koskenvuo M, Kaprio J, Pedersen NL, Christensen K. (2006) Genetic influence on human lifespan and longevity. Hum Genet 119(3):312-321.

Holliday R. (2006) Epigenetics: a historical overview. Epigenetics 1(2):76-80.

Holliday R, Pugh JE. (1975) DNA modification mechanisms and gene activity during development. Science 187(4173):226-232.

Hong MG, Myers AJ, Magnusson PK, Prince JA. (2008) Transcriptome-wide assessment of human brain and lymphocyte senescence. PLoS One 3(8):e3024.

Horstman AM, Dillon EL, Urban RJ, Sheffield-Moore M. (2012) The role of androgens and estrogens on healthy aging and longevity. J Gerontol A Biol Sci Med Sci 67(11):1140-1152.

Horvath S. (2013) DNA methylation age of human tissues and cell types. Genome Biol 14(10):R115.

Horvath S, Garagnani P, Bacalini MG, Pirazzini C, Salvioli S, Gentilini D, Di Blasio AM, Giuliani C, Tung S, Vinters HV, Franceschi C. (2015) Accelerated epigenetic aging in Down syndrome. Aging Cell 14(3):491-495.

Hu H, Brittain GC, Chang JH, Puebla-Osorio N, Jin J, Zal A, Xiao Y, Cheng X, Chang M, Fu YX, Zal T, Zhu C, Sun SC. (2013) OTUD7B controls non-canonical NF-κB activation through deubiquitination of TRAF3. Nature 494(7437):371-374.

Hubbard RE, Woodhouse KW. (2010) Frailty, inflammation and the elderly. Biogerontology 11(5):635-641.

Hublitz P, Albert M, Peters AH. (2009) Mechanisms of transcriptional repression by histone lysine methylation. Int J Dev Biol 53(2-3):335-354.

Hunt KJ, Walsh BM, Voegeli D, Roberts HC. (2010) Inflammation in aging part 2: implications for the health of older people and recommendations for nursing practice. Biol Res Nurs 11(3):253-260.
Iguchi-Ariga SM, Schaffner W. (1989) CpG methylation of the cAMP-responsive enhancer/promoter sequence TGACGTCA abolishes specific factor binding as well as transcriptional activation. Genes Dev 3(5):612-619.
Illingworth RS, Bird AP. (2009) CpG islands – 'A rough guide'. FEBS Lett 583:1713-1720.
Inawaka K, Kawabe M, Takahashi S, Doi Y, Tomigahara Y, Tarui H, Abe J, Kawamura S, Shirai T. (2009) Maternal exposure to anti-androgenic compounds, vinclozolin, flutamide and procymidone, has no effects on spermatogenesis and DNA methylation in male rats of subsequent generations. Toxicol Appl Pharmacol 237(2):178-187.
Ingram DK, Roth GS. (2015) Calorie restriction mimetics: can you have your cake and eat it, too? Ageing Res Rev 20:46-62.
Inukai S, de Lencastre A, Turner M, Slack F. (2012) Novel microRNAs differentially expressed during aging in the mouse brain. PLoS One 7(7):e40028.
Irizar H, Goñi J, Alzualde A, Castillo-Triviño T, Olascoaga J, de Munain AL, Otaegui D. (2015) Age gene expression and coexpression progressive signatures in peripheral blood leukocytes. Exp Gerontol 72:50-56.
Irizarry RA, Ladd-Acosta C, Wen B, Wu Z, Montano C, Onyango P, Cui H, Gabo K, Rongione M, Webster M, Ji H, Potash JB, Sabunciyan S, Feinberg AP. (2009) The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. Nat Genet 41(2):178-186.
Issa JP. (2011) Epigenetic variation and cellular Darwinism. Nat Genet 43(8):724-726.
Issa JP, Vertino PM, Boehm CD, Newsham IF, Baylin SB. (1996) Switch from monoallelic to biallelic human IGF2 promoter methylation during aging and carcinogenesis. Proc Natl Acad Sci U S A 93(21):11757-11762.
Issa JP, Ahuja N, Toyota M, Bronner MP, Brentnall TA. (2001) Accelerated age-related CpG island methylation in ulcerative colitis. Cancer Res 61(9):3573-3577.
Jaffe AE, Irizarry RA. (2014) Accounting for cellular heterogeneity is critical in epigenome-wide association studies. Genome Biol 15(2):R31.
Jakobsson A, Westerberg R, Jacobsson A. (2006) Fatty acid elongases in mammals: their regulation and roles in metabolism. Prog Lipid Res 45(3):237-249.
Jansen R, Batista S, Brooks AI, Tischfield JA, Willemsen G, van Grootheest G, Hottenga JJ, Milaneschi Y, Mbarek H, Madar V, Peyrot W, Vink JM, Verweij CL, de Geus EJ, Smit JH, Wright FA, Sullivan PF, Boomsma DI, Penninx BW. (2014) Sex differences in the human peripheral blood transcriptome. BMC Genomics 15:33.
Janssens GE, Meinema AC, González J, Wolters JC, Schmidt A, Guryev V, Bischoff R, Wit EC, Veenhoff LM, Heinemann M. (2015) Protein biogenesis machinery is a driver of replicative aging in yeast. Elife 4:e08527.
Jia D, Jurkowska RZ, Zhang X, Jeltsch A, Cheng X. (2007) Structure of Dnmt3a bound to Dnmt3L suggests a model for de novo DNA methylation. Nature 449(7159):248-251.
Jin K. (2010) Modern Biological Theories of Aging. Aging Dis 1(2):72-74.

Jintaridth P, Mutirangura A. (2010) Distinctive patterns of age-dependent hypomethylation in interspersed repetitive sequences. Physiol Genomics 41(2):194-200.
Johansson A, Enroth S, Gyllensten U. (2013) Continuous Aging of the Human DNA Methylome Throughout the Human Lifespan. PLoS One 8(6):e67378.
Johnson WE, Li C, Rabinovic A. (2007) Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics 8(1):118-127.
Johnson SC, Rabinovitch PS, Kaeberlein M. (2013) mTOR is a key modulator of ageing and age-related disease. Nature 493(7432):338-345.
Jones MJ, Goodman SJ, Kobor MS. (2015) DNA methylation and healthy human aging. Aging Cell 14(6):924-932.
Jones PA, Liang G. (2009) Rethinking how DNA methylation patterns are maintained. Nat Rev Genet 10(11):805-811.
Jones PA. (2012) Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nat Rev Genet 13(7):484-492.
Joop G, Mitschke A, Rolff J, Siva-Jothy MT. (2006) Immune function and parasite resistance in male and polymorphic female Coenagrion puella. BMC Evol Biol 6:19.
Jylhävä J. (2014) Determinants of longevity: genetics, biomarkers and therapeutic approaches. Curr Pharm Des 20(38):6058-6070.
Jylhävä J, Eklund C, Jylhä M, Hervonen A, Hurme M. (2010) Expression profiling of immune-associated genes in peripheral blood mononuclear cells reveals baseline differences in co-stimulatory signalling between nonagenarians and younger controls: the vitality 90+ study. Biogerontology 11(6):671-677.
Kaeberlein M. (2013) Longevity and aging. F1000Prime Rep 5:5.
Kallio MA, Tuimala JT, Hupponen T, Klemelä P, Gentile M, Scheinin I, Koski M, Käki J, Korpelainen EI. (2011) Chipster: user-friendly analysis software for microarray and other high-throughput data. BMC Genomics 12:507.
Kanfi Y, Naiman S, Amir G, Peshti V, Zinman G, Nahum L, Bar-Joseph Z, Cohen HY. (2012) The sirtuin SIRT6 regulates lifespan in male mice. Nature 483(7388):218-221.
Kapahi P, Chen D, Rogers AN, Katewa SD, Li PW, Thomas EL, Kockel L. (2010) With TOR, less is more: a key role for the conserved nutrient-sensing TOR pathway in aging. Cell Metab 11(6):453-465.
Karrer U, Sierro S, Wagner M, Oxenius A, Hengel H, Koszinowski UH, Phillips RE, Klenerman P. (2003) Memory inflation: continuous accumulation of antiviral CD8+ T cells over time. J Immunol 170(4):2022-2029.
Kawahara TL, Michishita E, Adler AS, Damian M, Berber E, Lin M, McCord RA, Ongaigui KC, Boxer LD, Chang HY, Chua KF. (2009) SIRT6 links histone H3 lysine 9 deacetylation to NF-kappaB-dependent gene expression and organismal life span. Cell 136(1):62-74.
Kelly TK, Miranda TB, Liang G, Berman BP, Lin JC, Tanay A, Jones PA. (2010) H2A.Z maintenance during mitosis reveals nucleosome shifting on mitotically silenced genes. Mol Cell 39(6):901-911.
Kirkwood TB. (2005) Understanding the odd science of aging. Cell 120(4):437-447.

Kirkwood TB. (2008) Understanding ageing from an evolutionary perspective. J Intern Med 263(2):117-127.
Kirkwood TB, Melov S. (2011) On the programmed/non-programmed nature of ageing within the life history. Curr Biol 21(18):R701-707.
Kirschbaum C, Wüst S, Faig HG, Hellhammer DH. (1992) Heritability of cortisol responses to human corticotropin-releasing hormone, ergometry, and psychological stress in humans. J Clin Endocrinol Metab 75(6):1526-1530.
Koga H, Kaushik S, Cuervo AM. (2011) Protein homeostasis and aging: The importance of exquisite quality control. Ageing Res Rev 10(2):205-215.
Kolehmainen M, Ulven SM, Paananen J, de Mello V, Schwab U, Carlberg C, Myhrstad M, Pihlajamäki J, Dungner E, Sjölin E, Gunnarsdottir I, Cloetens L, Landin-Olsson M, Akesson B, Rosqvist F, Hukkanen J, Herzig KH, Dragsted LO, Savolainen MJ, Brader L, Hermansen K, Risérus U, Thorsdottir I, Poutanen KS, Uusitupa M, Arner P, Dahlman I. (2015) Healthy Nordic diet downregulates the expression of genes involved in inflammation in subcutaneous adipose tissue in individuals with features of the metabolic syndrome. Am J Clin Nutr 101(1):228-239.
Kolovou GD, Kolovou V, Mavrogeni S. (2014) We are ageing. Biomed Res Int 2014:808307.
Kriaucionis S, Heintz N. (2009) The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. Science 324(5929):929-930.
Krishnan V, Chow MZ, Wang Z, Zhang L, Liu B, Liu X, Zhou Z. (2011) Histone H4 lysine 16 hypoacetylation is associated with defective DNA repair and premature senescence in Zmpste24-deficient mice. Proc Natl Acad Sci U S A 108(30):12325-12330.
Kristan DM. (2007) Chronic calorie restriction increases susceptibility of laboratory mice (Mus musculus) to a primary intestinal parasite infection. Aging Cell 6(6):817-825.
Kumar A, Gibbs JR, Beilina A, Dillman A, Kumaran R, Trabzuni D, Ryten M, Walker R, Smith C, Traynor BJ, Hardy J, Singleton AB, Cookson MR. (2013) Age-associated changes in gene expression in human brain and isolated neurons. Neurobiol Aging 34(4):1199-1209.
Lam LL, Emberly E, Fraser HB, Neumann SM, Chen E, Miller GE, Kobor MS. (2012) Factors underlying variable DNA methylation in a human community cohort. Proc Natl Acad Sci U S A 109 Suppl 2:17253-17260.
Lardenoije R, Iatrou A, Kenis G, Kompotis K, Steinbusch HW, Mastroeni D, Coleman P, Lemere CA, Hof PR, van den Hove DL, Rutten BP. (2015) The epigenetics of aging and neurodegeneration. Prog Neurobiol 131:21-64.
Larson K, Yan SJ, Tsurumi A, Liu J, Zhou J, Gaur K, Guo D, Eickbush TH, Li WX. (2012) Heterochromatin formation promotes longevity and represses ribosomal RNA synthesis. PLoS Genet 8(1):e1002473.
Laurent L, Wong E, Li G, Huynh T, Tsirigos A, Ong CT, Low HM, Kin Sung KW, Rigoutsos I, Loring J, Wei CL. (2010) Dynamic changes in the human methylome during differentiation. Genome Res 20(3):320-331.
Lee JH, Flaquer A, Costa R, Andrews H, Cross P, Lantigua R, Schupf N, Tang MX, Mayeux R. (2004) Genetic influences on life span and survival among elderly

African-Americans, Caribbean Hispanics, and Caucasians. Am J Med Genet A 128A(2):159-164.
Lee GD, Wilson MA, Zhu M, Wolkow CA, de Cabo R, Ingram DK, Zou S. (2006) Dietary deprivation extends lifespan in Caenorhabditis elegans. Aging Cell 5(6):515-524.
Lemieux S, Prud'homme D, Bouchard C, Tremblay A, Després JP. (1993) Sex differences in the relation of visceral adipose tissue accumulation to total body fatness. Am J Clin Nutr 58(4):463-467.
Li E, Bestor TH, Jaenisch R. (1992) Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. Cell 69(6):915-926.
Liker A, Székely T. (2005) Mortality costs of sexual selection and parental care in natural populations of birds. Evolution 59(4):890-897.
Liu N, Pan T. (2015) RNA epigenetics. Transl Res 165(1):28-35.
Lock LF, Takagi N, Martin GR. (1987) Methylation of the Hprt gene on the inactive X occurs after chromosome inactivation. Cell 48(1):39-46.
López-Otín C, Blasco MA, Partridge L, Serrano M, Kroemer G. (2013) The hallmarks of aging. Cell 153(6):1194-1217.
Lu WD, Liu T, Li S, Woods VL Jr, Hook V. (2012) The prohormone proenkephalin possesses differential conformational features of subdomains revealed by rapid H-D exchange mass spectrometry. Protein Sci 21(2):178-187.
Lukic S, Nicolas JC, Levine AJ. (2014) The diversity of zinc-finger genes on human chromosome 19 provides an evolutionary mechanism for defense against inherited endogenous retroviruses. Cell Death Differ 21(3):381-387.
de Magalhães JP. (2012) Programmatic features of aging originating in development: aging mechanisms beyond molecular damage? FASEB J 26(12):4821-4826.
de Magalhães JP, Budovsky A, Lehmann G, Costa J, Li Y, Fraifeld V, Church GM. (2009a) The Human Ageing Genomic Resources: online databases and tools for biogerontologists. Aging Cell 8(1):65-72.
de Magalhães JP, Curado J, Church GM. (2009b) Meta-analysis of age-related gene expression profiles identifies common signatures of aging. Bioinformatics 25(7):875-881.
Manikkam M, Haque MM, Guerrero-Bosagna C, Nilsson EE, Skinner MK. (2014) Pesticide methoxychlor promotes the epigenetic transgenerational inheritance of adult-onset disease through the female germline. PLoS One 9(7):e102091.
Marioni RE, Shah S, McRae AF, Ritchie SJ, Muniz-Terrera G, Harris SE, Gibson J, Redmond P, Cox SR, Pattie A, Corley J, Taylor A, Murphy L, Starr JM, Horvath S, Visscher PM, Wray NR, Deary IJ. (2015a) The epigenetic clock is correlated with physical and cognitive fitness in the Lothian Birth Cohort 1936. Int J Epidemiol 44(4):1388-1396.
Marioni RE, Shah S5, McRae AF, Chen BH, Colicino E, Harris SE, Gibson J, Henders AK, Redmond P, Cox SR, Pattie A, Corley J, Murphy L, Martin NG, Montgomery GW, Feinberg AP, Fallin MD, Multhaup ML, Jaffe AE, Joehanes R, Schwartz J, Just AC, Lunetta KL, Murabito JM, Starr JM, Horvath S, Baccarelli AA, Levy D, Visscher PM, Wray NR, Deary IJ. (2015b) DNA methylation age of blood predicts all-cause mortality in later life. Genome Biol 16:25.
Markle JG, Fish EN. (2014) SeXX matters in immunity. Trends Immunol 35(3):97-104.

Marmorstein R, Roth SY. (2001) Histone acetyltransferases: function, structure, and catalysis. Curr Opin Genet Dev 11(2):155-161.
Martos SN, Tang WY, Wang Z. (2015) Elusive inheritance: Transgenerational effects and epigenetic inheritance in human environmental disease. Prog Biophys Mol Biol 118(1-2):44-54.
Mayhew LA, Smith D. (2014) Gender Convergence in Human Survival and the Postponement of Death. N Am Actuar J 18(1):194-216.
McCay CM, Crowell MF, Maynard LA. (1935) The effect of retarded growth upon the length of life span and upon the ultimate body size. J Nutr 10:63-79.
McClay JL, Aberg KA, Clark SL, Nerella S, Kumar G, Xie LY, Hudson AD, Harada A, Hultman CM, Magnusson PK, Sullivan PF, Van Den Oord EJ. (2014) A methylome-wide study of aging using massively parallel sequencing of the methyl-CpG-enriched genomic fraction from blood in over 700 subjects. Hum Mol Genet 23(5):1175-1185.
McClelland EE, Smith JM. (2011) Gender specific differences in the immune response to infection. Arch Immunol Ther Exp (Warsz) 59(3):203-213.
McCord RA, Michishita E, Hong T, Berber E, Boxer LD, Kusumoto R, Guan S, Shi X, Gozani O, Burlingame AL, Bohr VA, Chua KF. (2009) SIRT6 stabilizes DNA-dependent protein kinase at chromatin for DNA double-strand break repair. Aging (Albany NY) 1(1):109-121.
Mecocci P, Fanó G, Fulle S, MacGarvey U, Shinobu L, Polidori MC, Cherubini A, Vecchiet J, Senin U, Beal MF. (1999) Age-dependent increases in oxidative damage to DNA, lipids, and proteins in human skeletal muscle. Free Radic Biol Med 26(3-4):303-308.
Mevissen TE, Hospenthal MK, Geurink PP, Elliott PR, Akutsu M, Arnaudo N, Ekkebus R, Kulathu Y, Wauer T, El Oualid F, Freund SM, Ovaa H, Komander D. (2013) OTU deubiquitinases reveal mechanisms of linkage specificity and enable ubiquitin chain restriction analysis. Cell 154(1):169-84.
Min JN, Whaley RA, Sharpless NE, Lockyer P, Portbury AL, Patterson C. (2008) CHIP deficiency decreases longevity, with accelerated aging phenotypes accompanied by altered protein quality control. Mol Cell Biol 28(12):4018-4025.
Minks J, Robinson WP, Brown CJ. (2008) A skewed view of X chromosome inactivation. J Clin Invest 118(1):20-23.
Moskalev AA, Shaposhnikov MV, Plyusnina EN, Zhavoronkov A, Budovsky A, Yanai H, Fraifeld VE. (2013) The role of DNA damage and repair in aging through the prism of Koch-like criteria. Ageing Res Rev 12(2):661-684.
Moskalev AA, Aliper AM, Smit-McBride Z, Buzdin A, Zhavoronkov A. (2014) Genetics and epigenetics of aging and longevity. Cell Cycle 13(7):1063-1077.
Mostoslavsky R, Chua KF, Lombard DB, Pang WW, Fischer MR, Gellon L, Liu P, Mostoslavsky G, Franco S, Murphy MM, Mills KD, Patel P, Hsu JT, Hong AL, Ford E, Cheng HL, Kennedy C, Nunez N, Bronson R, Frendewey D, Auerbach W, Valenzuela D, Karow M, Hottiger MO, Hursting S, Barrett JC, Guarente L, Mulligan R, Demple B, Yancopoulos GD, Alt FW. (2006) Genomic instability and aging-like phenotype in the absence of mammalian SIRT6. Cell 124(2):315-329.

Murabito JM, Yuan R, Lunetta KL. (2012) The search for longevity and healthy aging genes: insights from epidemiological studies and samples of long-lived individuals. J Gerontol A Biol Sci Med Sci 67(5):470-479.

Nakamura S, Kawai K, Takeshita Y, Honda M, Takamura T, Kaneko S, Matoba R, Matsubara K. (2012) Identification of blood biomarkers of aging by transcript profiling of whole blood. Biochem Biophys Res Commun 418(2):313-318.

Newman AB, Murabito JM. (2013) The epidemiology of longevity and exceptional survival. Epidemiol Rev 35:181-197.

Newman AB, Walter S, Lunetta KL, Garcia ME, Slagboom PE, Christensen K, Arnold AM, Aspelund T, Aulchenko YS, Benjamin EJ, Christiansen L, D'Agostino RB Sr, Fitzpatrick AL, Franceschini N, Glazer NL, Gudnason V, Hofman A, Kaplan R, Karasik D, Kelly-Hayes M, Kiel DP, Launer LJ, Marciante KD, Massaro JM, Miljkovic I, Nalls MA, Hernandez D, Psaty BM, Rivadeneira F, Rotter J, Seshadri S, Smith AV, Taylor KD, Tiemeier H, Uh HW, Uitterlinden AG, Vaupel JW, Walston J, Westendorp RG, Harris TB, Lumley T, van Duijn CM, Murabito JM. (2010) A meta-analysis of four genome-wide association studies of survival to age 90 years or older: the Cohorts for Heart and Aging Research in Genomic Epidemiology Consortium. J Gerontol A Biol Sci Med Sci 65(5):478-487.

Nielsen S, Guo Z, Johnson CM, Hensrud DD, Jensen MD. (2004) Splanchnic lipolysis in human obesity. J Clin Invest 113(11):1582-1588.

Nuotio J, Oikonen M, Magnussen CG, Jokinen E, Laitinen T, Hutri-Kähönen N, Kähönen M, Lehtimäki T, Taittonen L, Tossavainen P, Jula A, Loo BM, Viikari JS, Raitakari OT, Juonala M. (2014) Cardiovascular risk factors in 2011 and secular trends since 2007: the Cardiovascular Risk in Young Finns Study. Scand J Public Health 42(7):563-571.

Nunn CL, Lindenfors P, Pursall ER, Rolff J. (2009) On sexual dimorphism in immune function. Philos Trans R Soc Lond B Biol Sci 364(1513):61-69.

Oberdoerffer P, Michan S, McVay M, Mostoslavsky R, Vann J, Park SK, Hartlerode A, Stegmuller J, Hafner A, Loerch P, Wright SM, Mills KD, Bonni A, Yankner BA, Scully R, Prolla TA, Alt FW, Sinclair DA. (2008) SIRT1 redistribution on chromatin promotes genomic stability but alters gene expression during aging. Cell 135(5):907-918.

Ohm JE, McGarvey KM, Yu X, Cheng L, Schuebel KE, Cope L, Mohammad HP, Chen W, Daniel VC, Yu W, Berman DM, Jenuwein T, Pruitt K, Sharkis SJ, Watkins DN, Herman JG, Baylin SB. (2007) A stem cell-like chromatin pattern may predispose tumor suppressor genes to DNA hypermethylation and heritable silencing. Nat Genet 39(2):237-242.

Okano M, Bell DW, Haber DA, Li E. (1999) DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell 99(3):247-257.

Orlich MJ, Singh PN, Sabaté J, Jaceldo-Siegl K, Fan J, Knutsen S, Beeson WL, Fraser GE. (2013) Vegetarian dietary patterns and mortality in Adventist Health Study 2. JAMA Intern Med 173(13):1230-1238.

Ozacmak VH, Sayan H. (2009) The effects of 17beta estradiol, 17alpha estradiol and progesterone on oxidative stress biomarkers in ovariectomized female rat brain subjected to global cerebral ischemia. Physiol Res 58(6):909-912.

Paganelli R, Quinti I, Fagiolo U, Cossarizza A, Ortolani C, Guerra E, Sansoni P, Pucillo LP, Scala E, Cozzi E, Bertollo L, Monti D, Franceschi C. (1992) Changes in circulating B cells and immunoglobulin classes and subclasses in a healthy aged population. Clin Exp Immunol 90(2):351-354.

Painter RC, Roseboom TJ, Bleker OP. (2005) Prenatal exposure to the Dutch famine and disease in later life: an overview. Reprod Toxicol 20(3):345-352.

Partridge L. (2010) The new biology of ageing. Philos Trans R Soc Lond B Biol Sci 365(1537):147-154.

Partridge L, Piper MD, Mair W. (2005) Dietary restriction in Drosophila. Mech Ageing Dev 126(9):938-950.

Paska AV, Hudler P. (2015) Aberrant methylation patterns in cancer: a clinical view. Biochem Med (Zagreb) 25(2):161-176.

Passtoors WM, Boer JM, Goeman JJ, Akker EB, Deelen J, Zwaan BJ, Scarborough A, Breggen Rv, Vossen RH, Houwing-Duistermaat JJ, Ommen GJ, Westendorp RG, van Heemst D, de Craen AJ, White AJ, Gunn DA, Beekman M, Slagboom PE. (2012) Transcriptional profiling of human familial longevity indicates a role for ASF1A and IL7R. PLoS One 7(1):e27759.

Pastor WA, Aravind L, Rao A. (2013) TETonic shift: biological roles of TET proteins in DNA demethylation and transcription. Nat Rev Mol Cell Biol 14(6):341-356.

Pawelec G, Akbar A, Caruso C, Effros R, Grubeck-Loebenstein B, Wikby A. (2004) Is immunosenescence infectious? Trends Immunol 25(8):406-410.

Pawelec G, Larbi A, Derhovanessian E. (2010) Senescence of the human immune system. J Comp Pathol 142 Suppl 1:S39-44.

Peleg S, Sananbenesi F, Zovoilis A, Burkhardt S, Bahari-Javan S, Agis-Balboa RC, Cota P, Wittnam JL, Gogol-Doering A, Opitz L, Salinas-Riester G, Dettenhofer M, Kang H, Farinelli L, Chen W, Fischer A. (2010) Altered histone acetylation is associated with age-dependent memory impairment in mice. Science 328(5979):753-756.

Pembrey ME. (2010) Male-line transgenerational responses in humans. Hum Fertil (Camb) 13(4):268-271.

Pembrey ME, Bygren LO, Kaati G, Edvinsson S, Northstone K, Sjöström M, Golding J; ALSPAC Study Team. (2006) Sex-specific, male-line transgenerational responses in humans. Eur J Hum Genet 14(2):159-166.

Perls TT, Wilmoth J, Levenson R, Drinkwater M, Cohen M, Bogan H, Joyce E, Brewster S, Kunkel L, Puca A. (2002) Life-long sustained mortality advantage of siblings of centenarians. Proc Natl Acad Sci U S A 99(12):8442-8447.

Peters MJ, Joehanes R, Pilling LC, Schurmann C, Conneely KN, Powell J, Reinmaa E, Sutphin GL, Zhernakova A, Schramm K, Wilson YA, Kobes S, Tukiainen T; NABEC/UKBEC Consortium, Ramos YF, Göring HH, Fornage M, Liu Y, Gharib SA, Stranger BE, De Jager PL, Aviv A, Levy D, Murabito JM, Munson PJ, Huan T, Hofman A, Uitterlinden AG, Rivadeneira F, van Rooij J, Stolk L, Broer L, Verbiest MM, Jhamai M, Arp P, Metspalu A, Tserel L, Milani L, Samani NJ, Peterson P, Kasela S, Codd V, Peters A, Ward-Caviness CK, Herder C, Waldenberger M, Roden M, Singmann P, Zeilinger S, Illig T, Homuth G, Grabe HJ, Völzke H, Steil L, Kocher T, Murray A, Melzer D, Yaghootkar H, Bandinelli S, Moses EK, Kent JW, Curran JE, Johnson MP, Williams-Blangero S, Westra HJ,

McRae AF, Smith JA, Kardia SL, Hovatta I, Perola M, Ripatti S, Salomaa V, Henders AK, Martin NG, Smith AK, Mehta D, Binder EB, Nylocks KM, Kennedy EM, Klengel T, Ding J, Suchy-Dicey AM, Enquobahrie DA, Brody J, Rotter JI, Chen YD, Houwing-Duistermaat J, Kloppenburg M, Slagboom PE, Helmer Q, den Hollander W, Bean S, Raj T, Bakhshi N, Wang QP, Oyston LJ, Psaty BM, Tracy RP, Montgomery GW, Turner ST, Blangero J, Meulenbelt I, Ressler KJ, Yang J, Franke L, Kettunen J, Visscher PM, Neely GG, Korstanje R, Hanson RL, Prokisch H, Ferrucci L, Esko T, Teumer A, van Meurs JB, Johnson AD. (2015) The transcriptional landscape of age in human peripheral blood. Nat Commun 6:8570.

Pfister G, Savino W. (2008) Can the immune system still be efficient in the elderly? An immunological and immunoendocrine therapeutic perspective. Neuroimmunomodulation 15(4-6):351-364.

Pidsley R, Y Wong CC, Volta M, Lunnon K, Mill J, Schalkwyk LC. (2013) A data-driven approach to preprocessing Illumina 450K methylation array data. BMC Genomics 14:293.

Pincus Z, Smith-Vikos T, Slack FJ. (2011) MicroRNA predictors of longevity in Caenorhabditis elegans. PLoS Genet 7(9):e1002306.

Pinney SE (2014) Mammalian non-CpG methylation: Stem cells and beyond. Biology 3:739-751.

Pollack Y, Stein R, Razin A, Cedar H. (1980) Methylation of foreign DNA sequences in eukaryotic cells. Proc Natl Acad Sci U S A 77(11):6463-6467.

Poulsen P, Esteller M, Vaag A, Fraga MF. (2007) The epigenetic basis of twin discordance in age-related diseases. Pediatr Res 61(5 Pt 2):38R-42R.

Power ML, Schulkin J. (2008) Sex differences in fat storage, fat metabolism, and the health risks from obesity: possible evolutionary origins. Br J Nutr 99(5):931-940.

Prestridge DS, Burks C. (1993) The density of transcriptional elements in promoter and non-promoter sequences. Hum Mol Genet 2(9):1449-1453.

Raitakari OT, Juonala M, Rönnemaa T, Keltikangas-Järvinen L, Räsänen L, Pietikäinen M, Hutri-Kähönen N, Taittonen L, Jokinen E, Marniemi J, Jula A, Telama R, Kähönen M, Lehtimäki T, Akerblom HK, Viikari JS. (2008) Cohort profile: the cardiovascular risk in Young Finns Study. Int J Epidemiol 37(6):1220-1226.

Rakyan VK, Down TA, Maslau S, Andrew T, Yang TP, Beyan H, Whittaker P, McCann OT, Finer S, Valdes AM, Leslie RD, Deloukas P, Spector TD. (2010) Human aging-associated DNA hypermethylation occurs preferentially at bivalent chromatin domains. Genome Res 20(4):434-439.

Ramirez-Carrozzi VR, Braas D, Bhatt DM, Cheng CS, Hong C, Doty KR, Black JC, Hoffmann A, Carey M, Smale ST. (2009) A unifying model for the selective regulation of inducible transcription by CpG islands and nucleosome remodeling. Cell 138(1):114-128.

Ramsey MJ, Moore DH 2nd, Briner JF, Lee DA, Olsen La, Senft JR, Tucker JD. (1995) The effects of age and lifestyle factors on the accumulation of cytogenetic damage as measured by chromosome painting. Mutat Res 338(1-6):95-106.

Rauch TA, Wu X, Zhong X, Riggs AD, Pfeifer GP. (2009) A human B cell methylome at 100-base pair resolution. Proc Natl Acad Sci U S A 106(3):671-678.

Ravussin E, Redman LM, Rochon J, Das SK, Fontana L, Kraus WE, Romashkan S, Williamson DA, Meydani SN, Villareal DT, Smith SR, Stein RI, Scott TM,

Stewart TM, Saltzman E, Klein S, Bhapkar M, Martin CK, Gilhooly CH, Holloszy JO, Hadley EC, Roberts SB; CALERIE Study Group. (2015) A 2-Year Randomized Controlled Trial of Human Caloric Restriction: Feasibility and Effects on Predictors of Health Span and Longevity. J Gerontol A Biol Sci Med Sci 70(9):1097-1104.

Reed T, Dick DM. (2003) Heritability and validity of healthy physical aging (wellness) in elderly male twins. Twin Res 6(3):227-234.

Reinius LE, Acevedo N, Joerink M, Pershagen G, Dahlén SE, Greco D, Söderhäll C, Scheynius A, Kere J. (2012) Differential DNA methylation in purified human blood cells: implications for cell lineage and studies on disease susceptibility. PLoS One 7(7):e41361.

Remondini D, Salvioli S, Francesconi M, Pierini M, Mazzatti DJ, Powell JR, Zironi I, Bersani F, Castellani G, Franceschi C. (2010) Complex patterns of gene expression in human T cells during in vivo aging. Mol Biosyst ;6(10):1983-1992.

Reynolds LM, Ding J, Taylor JR, Lohman K, Soranzo N, de la Fuente A, Liu TF, Johnson C, Barr RG, Register TC, Donohue KM, Talor MV, Cihakova D, Gu C, Divers J, Siscovick D, Burke G, Post W, Shea S, Jacobs DR Jr, Hoeschele I, McCall CE, Kritchevsky SB, Herrington D, Tracy RP, Liu Y. (2015) Transcriptomic profiles of aging in purified human immune cells. BMC Genomics 16:333.

Reynolds LM, Taylor JR, Ding J, Lohman K, Johnson C, Siscovick D, Burke G, Post W, Shea S, Jacobs DR Jr, Stunnenberg H, Kritchevsky SB, Hoeschele I, McCall CE, Herrington DM, Tracy RP, Liu Y. (2014) Age-related variations in the methylome associated with gene expression in human monocytes and T cells. Nat Commun 5:5366.

Riggs AD. (1975) X inactivation, differentiation, and DNA methylation. Cytogenet Cell Genet 14(1):9-25.

Rinn JL, Kertesz M, Wang JK, Squazzo SL, Xu X, Brugmann SA, Goodnough LH, Helms JA, Farnham PJ, Segal E, Chang HY. (2007) Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. Cell 129(7):1311-1323.

Rodwell GE, Sonu R, Zahn JM, Lund J, Wilhelmy J, Wang L, Xiao W, Mindrinos M, Crane E, Segal E, Myers BD, Brooks JD, Davis RW, Higgins J, Owen AB, Kim SK. (2004) A transcriptional profile of aging in the human kidney. PLoS Biol 2(12):e427.

Rose MR, Flatt T, Graves JL, Greer LF, Martinez DE, Matos M, Mueller LD, Shmookler Reis RJ, Shahrestani P. (2012) What is Aging? Front Genet 3:134.

Sadier A, Viriot L, Pantalacci S, Laudet V. (2014) The ectodysplasin pathway: from diseases to adaptations. Trends Genet 30(1):24-31.

Salminen A, Haapasalo A, Kauppinen A, Kaarniranta K, Soininen H, Hiltunen M. (2015) Impaired mitochondrial energy metabolism in Alzheimer's disease: Impact on pathogenesis via disturbed epigenetic regulation of chromatin landscape. Prog Neurobiol 131:1-20.

Salthouse TA. (2009) When does age-related cognitive decline begin? Neurobiol Aging 30(4):507-514.

Sandovici I, Naumova AK, Leppert M, Linares Y, Sapienza C. (2004) A longitudinal study of X-inactivation ratio in human females. Hum Genet 115(5):387-392.

Sarkar DK. (2016) Male germline transmits fetal alcohol epigenetic marks for multiple generations: a review. Addict Biol 21(1):23-34.
Schlesinger Y, Straussman R, Keshet I, Farkash S, Hecht M, Zimmerman J, Eden E, Yakhini Z, Ben-Shushan E, Reubinoff BE, Bergman Y, Simon I, Cedar H. (2007) Polycomb-mediated methylation on Lys27 of histone H3 pre-marks genes for de novo methylation in cancer. Nat Genet 39(2):232-236.
Schmidl C, Klug M, Boeld TJ, Andreesen R, Hoffmann P, Edinger M, Rehli M. (2009) Lineage-specific DNA methylation in T cells correlates with histone methylation and enhancer activity. Genome Res 19(7):1165-1174.
Schneider S, Kaufmann W, Buesen R, van Ravenzwaay B. (2008) Vinclozolin--the lack of a transgenerational effect after oral maternal exposure during organogenesis. Reprod Toxicol 25(3):352-360.
Schneider S, Marxfeld H, Gröters S, Buesen R, van Ravenzwaay B. (2013) Vinclozolin--no transgenerational inheritance of anti-androgenic effects after maternal exposure during organogenesis via the intraperitoneal route. Reprod Toxicol 37:6-14.
Sebastiani P, Perls TT. (2012) The genetics of extreme longevity: lessons from the new England centenarian study. Front Genet 3:277.
Sebastiani P, Solovieff N, Dewan AT, Walsh KM, Puca A, Hartley SW, Melista E, Andersen S, Dworkis DA, Wilk JB, Myers RH, Steinberg MH, Montano M, Baldwin CT, Hoh J, Perls TT. (2012) Genetic signatures of exceptional longevity in humans. PLoS One 7(1):e29848.
Seifarth JE, McGowan CL, Milne KJ. (2012) Sex and life expectancy. Gend Med 9(6):390-401.
Seisenberger S, Andrews S, Krueger F, Arand J, Walter J, Santos F, Popp C, Thienpont B, Dean W, Reik W. (2012) The dynamics of genome-wide DNA methylation reprogramming in mouse primordial germ cells. Mol Cell 48(6):849-862.
Selamat SA, Chung BS, Girard L, Zhang W, Zhang Y, Campan M, Siegmund KD, Koss MN, Hagen JA, Lam WL, Lam S, Gazdar AF, Laird-Offringa IA. (2012) Genome-scale analysis of DNA methylation in lung adenocarcinoma and integration with mRNA expression. Genome Res 22(7):1197-1211.
Sena LA, Chandel NS. (2012) Physiological roles of mitochondrial reactive oxygen species. Mol Cell 48(2):158-167.
Shadyab AH, LaCroix AZ. (2015) Genetic factors associated with longevity: a review of recent findings. Ageing Res Rev 19:1-7.
Sharma A. (2015) Transgenerational epigenetic inheritance requires a much deeper analysis. Trends Mol Med 21(5):269-270.
Skinner MK, Manikkam M, Tracey R, Guerrero-Bosagna C, Haque M, Nilsson EE. (2013) Ancestral dichlorodiphenyltrichloroethane (DDT) exposure promotes epigenetic transgenerational inheritance of obesity. BMC Med 11:228.
Skinner MK, Guerrero-Bosagna C. (2014) Role of CpG deserts in the epigenetic transgenerational inheritance of differential DNA methylation regions. BMC Genomics 15:692.
Skulachev VP. (2011) Aging as a particular case of phenoptosis, the programmed death of an organism (a response to Kirkwood and Melov "On the programmed/non-

programmed nature of ageing within the life history"). Aging (Albany NY) 3(11):1120-1123.
Smith ZD, Chan MM, Mikkelsen TS, Gu H, Gnirke A, Regev A, Meissner A. (2012) A unique regulatory phase of DNA methylation in the early mammalian embryo. Nature 484(7394):339-344.
Smyth GK. (2004) Linear models and empirical bayes methods for assessing differential expression in microarray experiments. Stat Appl Genet Mol Biol 3:Article3.
So K, Tamura G, Honda T, Homma N, Waki T, Togawa N, Nishizuka S, Motoyama T. (2006) Multiple tumor suppressor genes are increasingly methylated with age in non-neoplastic gastric epithelia. Cancer Sci 97(11):1155-1158.
Sohal RS, Forster MJ. (2014) Caloric restriction and the aging process: a critique. Free Radic Biol Med 73:366-382.
Sommer M, Poliak N, Upadhyay S, Ratovitski E, Nelkin BD, Donehower LA, Sidransky D. (2006) DeltaNp63alpha overexpression induces downregulation of Sirt1 and an accelerated aging phenotype in the mouse. Cell Cycle 5(17):2005-2011.
Song WH, Ballard JW2 Yi YJ, Sutovsky P. (2014) Regulation of mitochondrial genome inheritance by autophagy and ubiquitin-proteasome system: implications for health, fitness, and fertility. Biomed Res Int 2014:981867.
Steegenga WT, Boekschoten MV, Lute C, Hooiveld GJ, de Groot PJ, Morris TJ, Teschendorff AE, Butcher LM, Beck S, Müller M. (2014) Genome-wide age-related changes in DNA methylation and gene expression in human PBMCs. Age (Dordr) 36(3):9648.
Stein R, Sciaky-Gallili N, Razin A, Cedar H. (1983) Pattern of methylation of two genes coding for housekeeping functions. Proc Natl Acad Sci U S A 80(9):2422-2426.
Suderman M, Borghol N, Pappas JJ, Pinto Pereira SM, Pembrey M, Hertzman C, Power C, Szyf M. (2014) Childhood abuse is associated with methylation of multiple loci in adult DNA. BMC Med Genomics 7:13.
Suderman M, McGowan PO, Sasaki A, Huang TC, Hallett MT, Meaney MJ, Turecki G, Szyf M. (2012) Conserved epigenetic sensitivity to early life experience in the rat and human hippocampus. Proc Natl Acad Sci U S A 109 Suppl 2:17266-17272.
Suzuki H, Toyota M, Kondo Y, Shinomura Y. (2009) Inflammation-related aberrant patterns of DNA methylation: detection and role in epigenetic deregulation of cancer cell transcriptome. Methods Mol Biol 512:55-69.
Szyf M. Nongenetic inheritance and transgenerational epigenetics. Trends Mol Med 21(2):134-144.
Talens RP, Christensen K, Putter H, Willemsen G, Christiansen L, Kremer D, Suchiman HE, Slagboom PE, Boomsma DI, Heijmans BT. (2012) Epigenetic variation during the adult lifespan: cross-sectional and longitudinal data on monozygotic twin pairs. Aging Cell 11(4):694-703.
Terry DF, Wilcox MA, McCormick MA, Pennington JY, Schoenhofen EA, Andersen SL, Perls TT. (2004) Lower all-cause, cardiovascular, and cancer mortality in centenarians' offspring. J Am Geriatr Soc 52(12):2074-2076.
Teschendorff AE, Menon U, Gentry-Maharaj A, Ramus SJ, Weisenberger DJ, Shen H, Campan M, Noushmehr H, Bell CG, Maxwell AP, Savage DA, Mueller-Holzner E, Marth C, Kocjan G, Gayther SA, Jones A, Beck S, Wagner W, Laird PW, Jacobs

IJ, Widschwendter M. (2010) Age-dependent DNA methylation of genes that are suppressed in stem cells is a hallmark of cancer. Genome Res 20(4):440-446.
Teschendorff AE, Marabita F, Lechner M, Bartlett T, Tegner J, Gomez-Cabrero D, Beck S. (2013) A beta-mixture quantile normalization method for correcting probe design bias in Illumina Infinium 450 k DNA methylation data. Bioinformatics 29(2):189-196.
Tower J. (2006) Sex-specific regulation of aging and apoptosis. Mech Ageing Dev 127(9):705-718.
Tserel L, Kolde R, Limbach M, Tretyakov K, Kasela S, Kisand K, Saare M, Vilo J, Metspalu A, Milani L, Peterson P. (2015) Age-related profiling of DNA methylation in CD8+ T cells reveals changes in immune response and transcriptional regulator genes. Sci Rep 5:13107.
Valinluck V, Sowers LC. (2007) Endogenous cytosine damage products alter the site selectivity of human DNA maintenance methyltransferase DNMT1. Cancer Res 67(3):946-950.
Vaquero A. (2009) The conserved role of sirtuins in chromatin regulation. Int J Dev Biol 53(2-3):303-322.
Veenendaal MV, Painter RC, de Rooij SR, Bossuyt PM, van der Post JA, Gluckman PD, Hanson MA, Roseboom TJ. (2013) Transgenerational effects of prenatal exposure to the 1944-45 Dutch famine. BJOG 120(5):548-553.
Verset L, Feys L, Trépant AL, De Wever O, Demetter P. (2015) FHL2: a scaffold protein of carcinogenesis, tumour-stroma interactions and treatment response. Histol Histopathol 17:11709.
Vijg J, Dollé ME. (2007) Genome instability: cancer or aging? Mech Ageing Dev 128(7-8):466-468.
Voltz JW, Card JW, Carey MA, Degraff LM, Ferguson CD, Flake GP, Bonner JC, Korach KS, Zeldin DC. (2008) Male sex hormones exacerbate lung function impairment after bleomycin-induced pulmonary fibrosis. Am J Respir Cell Mol Biol 39(1):45-52.
Wada Y, Tsukada M, Kamiyama S, Koizumi A. (1993) Retroviral gene expression as a possible biomarker of aging. Int Arch Occup Environ Health 65(1 Suppl):S235-239.
Waddington CH. (1942) The epigenotype. Endeavour 18-20.
Wagner W, Horn P, Bork S, Ho AD. (2008) Aging of hematopoietic stem cells is regulated by the stem cell niche. Exp Gerontol 43(11):974-980.
Walker GA, Lithgow GJ. (2003) Lifespan extension in C. elegans by a molecular chaperone dependent upon insulin-like signals. Aging Cell 2(2):131-139.
Walter S, Atzmon G, Demerath EW, Garcia ME, Kaplan RC, Kumari M, Lunetta KL, Milaneschi Y, Tanaka T, Tranah GJ, Völker U, Yu L, Arnold A, Benjamin EJ, Biffar R, Buchman AS, Boerwinkle E, Couper D, De Jager PL, Evans DA, Harris TB, Hoffmann W, Hofman A, Karasik D, Kiel DP, Kocher T, Kuningas M, Launer LJ, Lohman KK, Lutsey PL, Mackenbach J, Marciante K, Psaty BM, Reiman EM, Rotter JI, Seshadri S, Shardell MD, Smith AV, van Duijn C, Walston J, Zillikens MC, Bandinelli S, Baumeister SE, Bennett DA, Ferrucci L, Gudnason V, Kivimaki M, Liu Y, Murabito JM, Newman AB, Tiemeier H, Franceschini N. (2011) A genome-wide association study of aging. Neurobiol Aging 32(11):2109.e15-28.

Wang C, Jurk D, Maddick M, Nelson G, Martin-Ruiz C, von Zglinicki T. (2009) DNA damage response and cellular senescence in tissues of aging mice. Aging Cell 8(3):311-323.

Wang L, Wang FS, Gershwin ME. (2015) Human autoimmune diseases: a comprehensive update. J Intern Med ;278(4):369-395.

Weidner CI, Lin Q, Koch CM, Eisele L, Beier F, Ziegler P, Bauerschlag DO, Jöckel KH, Erbel R, Mühleisen TW, Zenke M, Brümmendorf TH, Wagner W. (2014) Aging of blood can be tracked by DNA methylation changes at just three CpG sites. Genome Biol 15(2):R24.

Weindruch R, Walford RL, Fligiel S, Guthrie D. (1986) The retardation of aging in mice by dietary restriction: longevity, cancer, immunity and lifetime energy intake. J Nutr 116(4):641-654.

van de Werken C, van der Heijden GW, Eleveld C, Teeuwssen M, Albert M, Baarends WM, Laven JS, Peters AH, Baart EB. (2014) Paternal heterochromatin formation in human embryos is H3K9/HP1 directed and primed by sperm-derived histone modifications. Nat Commun 5:5868.

West J, Beck S, Wang X, Teschendorff AE. (2013) An integrative network algorithm identifies age-associated differential methylation interactome hotspots targeting stem-cell differentiation pathways. Sci Rep 3:1630.

White GP, Watt PM, Holt BJ, Holt PG. (2009) Differential patterns of methylation of the IFN-gamma promoter at CpG and non-CpG sites underlie differences in IFN-gamma gene expression between human neonatal and adult CD45RO- T cells. J Immunol 168:2820–2827.

Widschwendter M, Fiegl H, Egle D, Mueller-Holzner E, Spizzo G, Marth C, Weisenberger DJ, Campan M, Young J, Jacobs I, Laird PW. (2007) Epigenetic stem cell signature in cancer. Nat Genet 39(2):157-158.

Wigler M, Levy D, Perucho M. (1981) The somatic replication of DNA methylation. Cell 24(1):33-40.

Willcox BJ, Willcox DC, He Q, Curb JD, Suzuki M. (2006) Siblings of Okinawan centenarians share lifelong mortality advantages. J Gerontol A Biol Sci Med Sci 61(4):345-354.

Willcox DC, Scapagnini G, Willcox BJ. (2014) Healthy aging diets other than the Mediterranean: a focus on the Okinawan diet. Mech Ageing Dev 136-137:148-162.

Wilson VL, Smith RA, Ma S, Cutler RG. (1987) Genomic 5-methyldeoxycytidine decreases with age. J Biol Chem 262(21):9948-9951.

Wu J, Wang SH, Potter D, Liu JC, Smith LT, Wu YZ, Huang TH, Plass C. (2007) Diverse histone modifications on histone 3 lysine 9 and their relation to DNA methylation in specifying gene silencing. BMC Genomics 24;8:131.

Xiong J, Jiang HP, Peng CY, Deng QY, Lan MD, Zeng H, Zheng F, Feng YQ, Yuan BF. (2015) DNA hydroxymethylation age of human blood determined by capillary hydrophilic-interaction liquid chromatography/mass spectrometry. Clin Epigenetics 7(1):72.

Xu Z, Taylor JA. (2014) Genome-wide age-related DNA methylation changes in blood and other tissues relate to histone modification, expression and cancer. Carcinogenesis 35(2):356-364.

Xu X, Su S, Barnes VA, De Miguel C, Pollock J, Ownby D, Shi H, Zhu H, Snieder H, Wang X. (2013) A genome-wide methylation study on obesity: differential variability and differential methylation. Epigenetics 8(5):522-533.

Yang X, Stedra J, Cerny J. (1996) Relative contribution of T and B cells to hypermutation and selection of the antibody repertoire in germinal centers of aged mice. J Exp Med 183(3):959-970.

Yashin AI, Wu D, Arbeev KG, Ukraintseva SV. (2010) Joint influence of small-effect genetic variants on human longevity. Aging (Albany NY) 2(9):612-620.

You D, Gu D, Yi Z. (2010) Familial Transmission of Human Longevity Among the Oldest-Old in China. J Appl Gerontol 29:308-332.

Yuan T, Jiao Y, de Jong S, Ophoff RA, Beck S, Teschendorff AE. (2015) An integrative multi-scale analysis of the dynamic DNA methylation landscape in aging. PLoS Genet 11(2):e1004996.

Zahn JM, Sonu R, Vogel H, Crane E, Mazan-Mamczarz K, Rabkin R, Davis RW, Becker KG, Owen AB, Kim SK. (2006) Transcriptional profiling of aging in human muscle reveals a common aging signature. PLoS Genet 2(7):e115.

Zambelli F, Pesole G, Pavesi G. (2009) Pscan: finding over-represented transcription factor binding site motifs in sequences from co-regulated or co-expressed genes. Nucleic Acids Res 37(Web Server issue):W247-52.

Zampieri M, Ciccarone F, Calabrese R, Franceschi C, Bürkle A, Caiafa P. (2015) Reconfiguration of DNA methylation in aging. Mech Ageing Dev 151:60-70.

Zane L, Sharma V, Misteli T. (2014) Common features of chromatin in aging and cancer: cause or coincidence? Trends Cell Biol 24(11):686-694.

Zbeida M, Goldsmith R, Shimony T, Vardi H, Naggan L, Shahar DR. (2014) Mediterranean diet and functional indicators among older adults in non-Mediterranean and Mediterranean countries. J Nutr Health Aging 18(4):411-418.

Zhang Z, Zhang R. (2015) Epigenetics in autoimmune diseases: Pathogenesis and prospects for therapy. Autoimmun Rev 14(10):854-863.

Zilbauer M, Rayner TF, Clark C, Coffey AJ, Joyce CJ, Palta P, Palotie A, Lyons PA, Smith KG. (2013) Genome-wide methylation analyses of primary human leukocyte subsets identifies functionally important cell-type-specific hypomethylated regions. Blood 122(25):e52-60.

Zimniak P. (2012) What is the Proximal Cause of Aging? Front Genet 3:189.

Zykovich A, Hubbard A, Flynn JM, Tarnopolsky M, Fraga MF, Kerksick C, Ogborn D, MacNeil L, Mooney SD, Melov S. (2014) Genome-wide DNA methylation changes with age in disease-free human skeletal muscle. Aging Cell 13(2):360-366.

Ørstavik KH. (2009) X chromosome inactivation in clinical practice. Hum Genet 126(3):363-373.

Other references:
Gene expression omnibus (GEO) database: http://www.ncbi.nlm.nih.gov/geo/
Gerontology Research Group, http://www.grg.org/
Illumina, http://www.illumina.com/products/methylation_450_beadchip_kits.html 8.1.2016
SVT 2015a: (Suomen virallinen tilasto), 2015: Kuolleet ISSN=1798-2529. 01 2013. Helsinki: Tilastokeskus 3.9.2015.
SVT 2015b: (Suomen virallinen tilasto), 2015: Väestöennuste ISSN=1798-5137. 2012. Helsinki: Tilastokeskus 14.7.2015.
WHO, 2011: http://www.who.int/ageing/publications/global_health/en/ 14.7.2015
QIAGEN's Ingenuity® pathway analysis, www.ingenuity.com

10 Original Communications

OPEN ACCESS Freely available online

PLOS ONE

Transcriptional Analysis Reveals Gender-Specific Changes in the Aging of the Human Immune System

Saara Marttila[1,2]*, Juulia Jylhävä[1,2], Tapio Nevalainen[1,2], Matti Nykter[3], Marja Jylhä[2,4], Antti Hervonen[2,4], Liina Tserel[5], Pärt Peterson[5], Mikko Hurme[1,2,6]

1 Department of Microbiology and Immunology, School of Medicine, University of Tampere, Tampere, Finland, **2** Gerontology Research Center, Tampere, Finland, **3** Institute of Biomedical Technology, University of Tampere, Tampere, Finland, **4** School of Health Sciences, University of Tampere, Tampere, Finland, **5** Molecular Pathology, University of Tartu, Tartu, Estonia, **6** Centre for Laboratory Medicine, Tampere University Hospital, Tampere, Finland

Abstract

Aging and gender have a strong influence on the functional capacity of the immune system. In general, the immune response in females is stronger than that in males, but there is scant information about the effect of aging on the gender difference in the immune response. To address this question, we performed a transcriptomic analysis of peripheral blood mononuclear cells derived from elderly individuals (nonagenarians, n = 146) and young controls (aged 19–30 years, n = 30). When compared to young controls, we found 339 and 248 genes that were differentially expressed ($p<0.05$, fold change >1.5 or <-1.5) in nonagenarian females and males, respectively, 180 of these genes were changed in both genders. An analysis of the affected signaling pathways revealed a clear gender bias: there were 48 pathways that were significantly changed in females, while only 29 were changed in males. There were 24 pathways that were shared between both genders. Our results indicate that female nonagenarians have weaker T cell defenses and a more prominent pro-inflammatory response as compared to males. In males significantly fewer pathways were affected, two of which are known to be regulated by estrogen. These data show that the effects of aging on the human immune system are significantly different in males and females.

Citation: Marttila S, Jylhävä J, Nevalainen T, Nykter M, Jylhä M, et al. (2013) Transcriptional Analysis Reveals Gender-Specific Changes in the Aging of the Human Immune System. PLoS ONE 8(6): e66229. doi:10.1371/journal.pone.0066229

Editor: Gualtiero Colombo, Centro Cardiologico Monzino IRCCS, Italy

Received October 18, 2012; **Accepted** May 2, 2013; **Published** June 11, 2013

Funding: This work was financially supported by grants from the Academy of Finland (132704) (http://www.aka.fi/eng), the Competitive Research Fund of Pirkanmaa Hospital District (9M017 and 9N013) (http://www.pshp.fi/default.aspx?nodeid = 10066&contentlan = 2) (MH), European Regional Development Fund (http://www.communities.gov.uk/regeneration/regenerationfunding/europeanregionaldevelopment/), Archimedes Foundation (http://archimedes.ee/), Centre for Translational Genomics of Tartu University (http://www.ctg.ut.ee/) and Estonian targeted grant SF0180021s07 (LT, PP). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: saara.marttila@uta.fi

Introduction

Old age is associated with a higher risk of inflammatory diseases, autoimmune disorders and malignancies. This increased risk is due to the decreased function of the immune system, with immunosenescence and chronic low-grade inflammation, termed inflamm-aging, representing the key changes [1]. With advancing age, the number of naïve CD4+ and CD8+ T cells declines, while the number of memory and effector cells increases. One prominent feature of old age is the increased proportion of late-stage differentiated CD8+ T cell clones that lack the expression of the costimulatory molecule CD28. Additionally, T cell function is modulated with advancing age; older individuals show a restricted T cell receptor (TCR) repertoire and defects in TCR-mediated signaling [2]. Similar to T cells, the number of naïve B cells is decreased, while the number of memory B cells is increased [3]. Inflamm-aging is another hallmark of aging. In the elderly, the blood levels of pro-inflammatory cytokines (IL-6, TNF-α and CRP) are increased, but the cellular sources and inductive signals underlying this expression are still largely unknown [4].

The immune system shows strong sexual dimorphism. Generally, females are more immunocompetent, meaning that they show increased resilience to various infections and some non-infectious diseases, such as cancer [5]. However, as a result, females are more prone to autoimmune disorders. Sex hormones are correlated with some of these differences, but other physical, and possibly social, factors may have a role in the sexual dimorphism of immune functions [5], [6]. In general, females and males age differently, as most clearly observed in the variance of morbidity and mortality rates between the genders [7]. However, the combined effects of aging and gender on the human immune system have not been analyzed previously.

Results

To better understand the combinatorial effects of age and gender on the immune system, we analyzed the global gene expression profile of peripheral blood mononuclear cells (PBMCs) from nonagenarians (n = 146, 103 females, 43 males) and young controls (n = 30, aged 19–30 years, 21 females, 9 males) using an Illumina Human HT12v4 BeadChip array. The data were analyzed with the Chipster program [8] (IT Center for Science Ltd (CSC), Espoo, Finland). Using a cut-off of $p<0.05$ and a fold change (FC) of below -1.5 or above 1.5, we identified 339 genes that were differentially expressed in female nonagenarians compared to female controls, and 248 genes that were differen-

tially expressed in male nonagenarians compared to male controls. Of these genes, 180 were common to both genders (Figure 1). The top 10 up- and down-regulated genes are shown in Table 1, and all differentially expressed genes are listed in Tables S1 and S2. The expression levels of four transcripts were verified with qPCR. The transcripts verified included both up- and down-regulated transcript as well as transcripts with high and low FC. The results acquired through qPCR were positively correlated with the microarray results. The expression of CD83, IL8 and LRRN3 were measured, with FCs (microarray/qPCR) in males of 1.73/1.90, 3.46/7.26 and −4.68/−5.65, respectively. In females, the fold changes (microarray/qPCR) for CD83, IL8, LRRN3 and PLCG1 were 1.70/1.71, 4.85/6.15, −5.64/−7.81 and −1.63/−1.98, respectively.

To identify the biological pathways affected, Ingenuity Pathway Analysis software (IPA) (Ingenuity® Systems, www.ingenuity.com) was used. Of the pathways by the Ingenuity Knowledge Base, our analysis revealed 48 pathways that were significantly affected in females ($p<0.05$, FDR<0.25 and at least 3 genes from the pathway were up- or down-regulated) and 29 pathways that were affected in males. Of these pathways, 24 were common to both genders. *B cell development* was the most significantly affected pathway in both genders. Other pathways that were significantly affected included the *Dendritic cell maturation* pathway and *T helper cell differentiation* pathway (Table 2, Table S3). Furthermore, changes in a significant number of pathways were found to be age-dependent in only one of the genders. In females, there were 24 gender-specific pathways (i.e. pathways that were only affected in females), and the most significantly affected signaling pathway was *CTLA4 signaling in cytotoxic T lymphocytes*. In males, there were fewer gender-specific pathways (5 in total), and the *Estrogen mediated S-phase entry* pathway was most affected (Tables 3 and 4). The proportions of different T cell subpopulations in the study subjects were determined with FACS analysis (Table S4) and no statistically significant differences were found between the genders. Thus, unequal representation of different T cell subsets can be excluded as an explanation for these gender-specific differences.

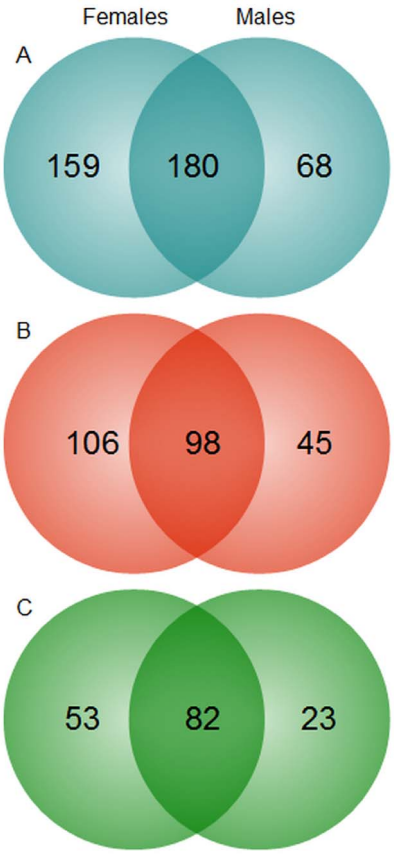


Figure 1. Genes differentially expressed in nonagenarians. We found 339 genes that were differentially expressed in female nonagenarians, compared to young female controls, and 248 genes that were differentially expressed in male nonagenarians, compared to young male controls ($p<0.05$, $-1.5> FC >1.5$). A total of 180 of these genes were common to both genders. Slightly more genes were up-regulated (1b) than were down-regulated (1c) in the nonagenarians of both genders.
doi:10.1371/journal.pone.0066229.g001

Discussion

In summary, the data presented here suggests that the effect of aging on the function of the human immune system is different between males and females. Aging-associated changes in gene expression have previously been studied in PBMCs [9], [10], but these studies did not consider the effect of gender. Several genes identified in our study have previously been associated with aging or advanced age, including LEF1 [9], [10], [11], [12], VPREB3 [10], NR4A2 [11], LRRN3 [9], [10], [12], CCR7 [12], [13] and CD19 [13]. All of these genes were affected in both genders. All of the pathways that were found to be significantly affected by aging in both genders of nonagenarians in this study have been reported associated with aging in the literature. We also identified one novel pathway, *TREM1 signaling*, that has not been previously associated with aging. *TREM1 signaling* has a role in acute inflammation; it is expressed in blood neutrophils and monocytes, and its expression is induced by pathogens (LPS, bacteria and fungi [14], [15]). It appears that the *TREM1 signaling* pathway contributes to the pro-inflammatory state in elderly populations.

In addition to the pro-inflammatory pathways that are affected in nonagenarians of both genders, several pro-inflammatory pathways were affected only in females (Table 3). This result is not surprising, because females generally to have stronger inflammatory reactions [6]. One reason for the muted inflammatory response in males may be testosterone, which is known to have anti-inflammatory effects [16]. NF-κB signaling is affected in both genders, but in females, there were more genes with this pathway that were affected. In addition, two NO synthesis-associated pathways were affected only in females, which indicates more potent NF-κB signaling and an elevated pro-inflammatory response in females because iNOS induction and NO synthesis are induced by NF-κB and other pro-inflammatory cytokines [17]. The *p38 MAPK signaling* pathway, which was significantly affected in females only, can also be activated by cell stressors other than pro-inflammatory cytokines [18]. This result indicates that females may have a more potent stress response than males.

Several of the age dependent, female-specific pathways are involved in the activation of T lymphocytes (Table 3), suggesting that gender may play an important regulatory role in T cell-mediated defense. Because the expression levels of several genes in these pathways were either up- or down regulated, it is difficult to reconstruct the functional end result. However, because the *CTLA4 signaling in cytotoxic T lymphocytes* pathway, including elevated expression of CTLA4, and the *iCOS-iCOSL signaling in T helper cells* pathway were significantly affected in females only, it appears that the T cell-mediated defense is weaker in female nonagenarians than in male nonagenarians.

Table 1. Ten most up- and down-regulated genes found in females and males.

Gender	Symbol	Description	p	FC
Female	LRRN3*	leucine rich repeat neuronal 3	<0.000001	−5.639
	CCR7	chemokine (C-C motif) receptor 7	<0.000001	−3.613
	LOC652694	similar to Ig kappa chain V-I region HK102 precursor	<0.000001	−2.877
	IGJ	immunoglobulin J polypeptide, linker protein for immunoglobulin alpha and mu polypeptides	<0.000001	−2.848
	CD27	CD27 molecule	<0.000001	−2.747
	CD79A	CD79a molecule, immunoglobulin-associated alpha	<0.000001	−2.680
	CD19	CD19 molecule	<0.000001	−2.651
	IGLL1	immunoglobulin lambda-like polypeptide 1	<0.000001	−2.643
	SGK223	homolog of rat pragma of Rnd2	<0.000001	−2.643
	FCRLA	Fc receptor-like A	<0.000001	−2.628
	IL8*	interleukin 8	<0.000001	4.851
	PTGS2	prostaglandin-endoperoxide synthase 2 (prostaglandin G/H synthase and cyclooxygenase)	<0.000001	3.792
	NR4A2	nuclear receptor subfamily 4, group A, member 2	<0.000001	3.169
	RHOB	ras homolog gene family, member B	<0.000001	2.951
	CDKN1A	cyclin-dependent kinase inhibitor 1A (p21, Cip1)	<0.000001	2.909
	IL1B	interleukin 1, beta	<0.000001	2.840
	RGS1	regulator of G-protein signaling 1	<0.000001	2.766
	EGR1	early growth response 1	<0.000001	2.609
	CCL3L3	chemokine (C-C motif) ligand 3-like 3	<0.000001	2.600
	HBEGF	heparin-binding EGF-like growth factor	<0.000001	2.557
Male	LRRN3*	leucine rich repeat neuronal 3	<0.000001	−4.678
	CD79A	CD79a molecule, immunoglobulin-associated alpha	0.000034	−3.073
	CCR7	chemokine (C-C motif) receptor 7	<0.000001	−3.045
	CD19	CD19 molecule	0.000005	−2.963
	FCRLA	Fc receptor-like A	0.000010	−2.874
	CD79B*	CD79b molecule, immunoglobulin-associated beta	<0.000001	−2.713
	NELL2	NEL-like 2 (chicken)	<0.000001	−2.709
	LOC652694	similar to Ig kappa chain V-I region HK102 precursor	0.000169	−2.423
	LEF1	lymphoid enhancer-binding factor 1	<0.000001	−2.414
	VPREB3	pre-B lymphocyte 3	<0.000001	−2.413
	IL8	interleukin 8	0.007899	3.459
	NR4A2	nuclear receptor subfamily 4, group A, member 2	0.000001	3.422
	JUN	jun proto-oncogene	0.005935	3.046
	RGS1	regulator of G-protein signaling 1	0.000015	2.935
	CDKN1A	cyclin-dependent kinase inhibitor 1A (p21, Cip1)	<0.000001	2.868
	OSM	oncostatin M	0.000784	2.716
	PTGS2	prostaglandin-endoperoxide synthase 2 (prostaglandin G/H synthase and cyclooxygenase)	0.003393	2.684
	HBEGF	heparin-binding EGF-like growth factor	0.000422	2.615
	IL1B	interleukin 1, beta	0.005170	2.561
	ADM	adrenomedullin	0.000034	2.433

The 10 most down- and up-regulated transcripts in female and male nonagenarians compared to young controls. All differentially expressed transcripts are listed in supplementary Tables S1 (males) and S2 (females).
*There are several transcripts of this gene, only the one with largest FC is shown.
doi:10.1371/journal.pone.0066229.t001

The *Cytotoxic T lymphocyte mediated apoptosis of target cells* pathway was also found to be affected by age in females only. Chronic viral infections, e.g. cytomegalovirus (CMV) and Epstein-Barr virus, affect a large majority of the nonagenarian population. In our nonagenarian study population 96% of the females and 95% of the males were seropositive for CMV. Cytotoxic CD8+ cells are the primary cell type that controls these viral infections, but chronic viral infections can also induce the generation of atypical

Table 2. The canonical pathways that were most affected in nonagenarians of both genders.

Canonical pathway		−logFDR	−logP	Ratio	Rank	Molecules
B Cell Development	Females	6.04	8.40	0.276	1	PTPRC ↑, CD19 ↓, HLA-DOA ↓, CD79B ↓, HLA-DQA1 ↑, CD86 ↑, HLA-DOB ↓, CD79A ↓
	Males	5.43	7.86	0.241	1	PTPRC ↑, CD19 ↓, HLA-DOA ↓, SPN ↑, CD79B ↓,HLA-DOB ↓,CD79A ↓
B Cell Receptor Signaling	Females	2.04	3.29	0.062	12	PTPRC ↑, CD19 ↓, JUN ↑, CD79B ↓, FCGR2A ↑, EGR1 ↑, PIK3AP1 ↑, CREB5 ↑, BCL6 ↑, CD79A ↓
	Males	1.06	2.32	0.043	11	PTPRC ↑, CD19 ↓, JUN ↑, CD79B ↓, EGR1 ↑, BCL6 ↑, CD79A ↓
Communication between Innate and Adaptive Immune Cells	Females	3.48	5.15	0.097	4	IL8 ↑, TLR10 ↓, IL15 ↑, CCL3L1/CCL3L3 ↑, CD86 ↑, IL1B ↑, CD83 ↑, CCL3 ↑, CCR7 ↓
	Males	1.75	3.34	0.065	5	IL8 ↑, TLR10 ↓, CCL3L1/CCL3L3 ↑, IL1B ↑, CD83 ↑, CCR7 ↓
Crosstalk between Dendritic Cells and Natural Killer Cells	Females	2.36	3.72	0.089	9	LTA ↓, CD69 ↑, IL15 ↑, LTB ↓, CD86 ↑, CD83 ↑, CCR7 ↓, FAS ↑
	Males	0.62	1.52	0.044	21	KIR3DL2 ↑, LTB ↓, CD83 ↑, CCR7 ↓
Dendritic Cell Maturation	Females	6.04	8.24	0.089	2	HLA-DOA ↓, LEP ↑, FCGR2A ↑, IL15 ↑, HLA-DQA1 ↑, PLCG1 ↓, LTB ↓, CD83 ↑, CREB5 ↑, FCGR1A ↑, LTA ↓, HLA-DOB ↓, CD86 ↑, IL1B ↑, STAT1 ↑, CCR7 ↓, FCGR1B ↑
	Males	1.84	3.49	0.047	4	HLA-DOA ↓, LTB ↓, HLA-DOB ↓, IL1B ↑, CD83 ↑, STAT1 ↑, FCGR1A ↑, CCR7 ↓, FCGR1B ↑
IL-8 Signaling	Females	2.48	3.90	0.063	8	IL8 ↑, GNG11 ↑, JUN ↑, RHOB ↑, DEFA1 ↑, VEGFB ↓, HBEGF ↑, PTGS2 ↑, IRAK3 ↑, ITGB5 ↑, MMP9 ↑, GNG7 ↓
	Males	0.62	1.44	0.031	26	IL8 ↑, JUN ↑, RHOB ↑, HBEGF ↑, PTGS2 ↑, GNG7 ↓
ILK Signaling	Females	2.07	3.34	0.060	11	MYC ↓, FLNB ↓, JUN ↑, RHOB ↑, VEGFB ↓, LEF1 ↓, HIF1A ↑, PTGS2 ↑, CREB5 ↑, ITGB5 ↑, MMP9 ↑
	Males	0.62	1.46	0.032	25	MYC ↓, JUN ↑, RHOB ↑, LEF1 ↓, HIF1A ↑, PTGS2 ↑
Oncostatin M Signaling	Females	1.47	2.50	0.118	21	MT2A ↑, IL6ST ↓, OSM ↑, STAT1 ↑
	Males	1.45	2.98	0.118	6	MT2A ↑, IL6ST ↓, OSM ↑, STAT1 ↑
OX40 Signaling Pathway	Females	1.70	2.77	0.082	19	HLA-DOA ↓, JUN ↑, HLA-DQA1 ↑, HLA-DOB ↓, TRAF5 ↓
	Males	1.11	2.42	0.066	9	HLA-DOA ↓, JUN ↑, HLA-DOB ↓, TRAF5 ↓
PI3K Signaling in B Lymphocytes	Females	2.64	4.11	0.078	7	BLK ↓, PTPRC ↑, CD19 ↓, IL4R ↓, JUN ↑, ATF3 ↑, CD79B ↓, PLCG1 ↓, PIK3AP1 ↑, CD79A ↓
	Males	1.45	2.89	0.055	7	BLK ↓, PTPRC ↑, CD19 ↓, JUN ↑, ATF3 ↑, CD79B ↓, CD79A ↓
Role of JAK family kinases in IL-6-type Cytokine Signaling	Females	1.87	3.01	0.154	16	IL6ST ↓, SOCS3 ↑, OSM ↑, STAT1 ↑
	Males	1.84	3.51	0.154	3	IL6ST ↓, SOCS3 ↑, OSM ↑, STAT1 ↑
Role of NFAT in Regulation of the Immune Response	Females	3.48	5.14	0.070	5	HLA-DOA ↓, CD79B ↓, FCGR2A ↑, HLA-DQA1 ↑, PLCG1 ↓, FCGR1A ↑, GNG7 ↓, CD79A ↓, GNG11 ↑, JUN ↑, CD86 ↑, HLA-DOB ↓, FCGR1B ↑
	Males	1.45	2.84	0.043	8	HLA-DOA ↓, JUN ↑, CD79B ↓, HLA-DOB ↓, FCGR1A ↑, GNG7 ↓, CD79A ↓, FCGR1B ↑
T Helper Cell Differentiation	Females	4.00	5.80	0.130	3	IL6ST ↓, IL4R ↓, HLA-DOA ↓, HLA-DQA1 ↑, CD86 ↑, HLA-DOB ↓, CXCR5 ↓, STAT1 ↑, BCL6 ↑
	Males	1.93	3.76	0.087	2	IL6ST ↓, HLA-DOA ↓, HLA-DOB ↓, CXCR5 ↓, STAT1 ↑, BCL6 ↑
TREM1 Signaling	Females	2.70	4.25	0.123	6	IL8 ↑, TLR10 ↓, PLCG1 ↓, CD86 ↑, IL1B ↑, CD83 ↑, CCL3 ↑
	Males	1.01	2.21	0.070	12	IL8 ↑, TLR10 ↓, IL1B ↑, CD83 ↑

P-values were derived from Fisher's exact test, -log(0.05) = 1.3, and the Benjamini-Hochberg-corrected t-test (FDR), –log(0.25) = 0.61. The ratio is the number of differentially expressed genes in the data set divided by the total number of genes in the given pathway. The rank indicates the position of the pathway in the gender-specific pathway list. A complete list of the pathways affected in both genders is shown in supplementary Table S3 .

doi:10.1371/journal.pone.0066229.t002

cytotoxic CD4+ cells, which express granzyme B [19]. Because chronic viral infections are thought to be a driving force behind age-associated changes in the immune system, and because there is no difference in seroprevalence between the genders, it is of great interest to determine whether the immune systems of males and females control these infections in different ways. Additionally, the production of IL-15, a major homeostatic cytokine, was affected only in females (the *IL-15 production* pathway). Previously, it was shown that the blood levels of this cytokine are elevated in centenarians [20], [21], but our data show that this phenomenon is restricted to female nonagenarians, at least at the transcriptional level (FC = 1.6).

Significantly fewer pathways were affected in males than in females (Table 4). Males also had fewer differentially expressed transcripts, but this does not completely explain the difference in the number of pathways, as the canonical pathway analysis in IPA takes into account the number of input transcripts. The *Estrogen mediated S-phase entry* pathway was most significantly changed. Generally, estrogen is known to have an effect on inflammation, and to possibly have a protective role against oxidative stress. The

Table 3. The canonical pathways that were affected in nonagenarian females.

Canonical pathways	−logFDR	−logP	Ratio	Molecules
Prostanoid Biosynthesis	2.07	3.34	0.333	PTGS1 ↑, PTGS2 ↑, PTGDS ↑
CTLA4 Signaling in Cytotoxic T Lymphocytes	1.90	3.06	0.074	HLA-DOA ↓, HLA-DQA1 ↑, TRAT1 ↓, PLCG1 ↓, CD86 ↑, HLA-DOB ↓, CTLA4 ↑
CCR5 Signaling in Macrophages	1.90	3.06	0.074	GNG11 ↑, JUN ↑, PLCG1 ↓, CCL3 ↑, GNG7 ↓, FAS ↑
IL-15 Production	1.77	2.88	0.138	TXK ↓, IL15 ↑, STAT1 ↑, IRF1 ↑
IL-10 Signaling	1.76	2.85	0.083	IL1R2 ↑, SOCS3 ↑, IL4R ↓, JUN ↑, FCGR2A ↑, IL1B ↑
p38 MAPK Signaling	1.36	2.33	0.060	IL1R2 ↑, MYC ↓, IL1B ↑, IRAK3 ↑, CREB5 ↑, STAT1 ↑, FAS ↑
P2Y Purigenic Receptor Signaling Pathway	1.30	2.25	0.056	MYC ↓, ITGA2B ↑, GNG11 ↑, JUN ↑, PLCG1 ↓, CREB5 ↑, GNG7 ↓
iNOS Signaling	1.26	2.17	0.087	JUN ↑, IRAK3 ↑, STAT1 ↑, IRF1 ↑
Cytotoxic T Lymphocyte-mediated Apoptosis of Target Cells	1.24	2.13	0.077	HLA-DOA ↓, HLA-DQA1 ↑, HLA-DOB ↓, FAS ↑
Differential Regulation of Cytokine Production in Intestinal Epithelial Cells by IL-17A and IL-17F	1.23	2.10	0.130	LCN2 ↑, IL1B ↑, CCL3 ↑
iCOS-iCOSL Signaling in T Helper Cells	1.21	2.06	0.054	PTPRC ↑, HLA-DOA ↓, HLA-DQA1 ↑, TRAT1 ↓, PLCG1 ↓, HLA-DOB ↓
IL-4 Signaling	1.19	2.04	0.067	IL4R ↓, HLA-DOA ↓, HLA-DQA1 ↑, HLA-DOB ↓, FCER2 ↓
Nur77 Signaling in T Lymphocytes	1.19	2.03	0.070	HLA-DOA ↓, HLA-DQA1 ↑, CD86 ↑, HLA-DOB ↓
PKCθ Signaling in T Lymphocytes	1.06	1.87	0.047	HLA-DOA ↓, JUN ↑, HLA-DQA1 ↑, PLCG1 ↓, CD86 ↑, HLA-DOB ↓
TNFR2 Signaling	1.06	1.86	0.094	JUN ↑, LTA ↓, TNFAIP3 ↑
Calcium-induced T Lymphocyte Apoptosis	1.02	1.82	0.066	HLA-DOA ↓, HLA-DQA1 ↑, PLCG1 ↓, HLA-DOB ↓
G Protein Signaling Mediated by Tubby	0.97	1.74	0.081	GNG11 ↑, PLCG1 ↓, GNG7 ↓
Glucocorticoid Receptor Signaling	0.97	1.74	0.036	IL1R2 ↑, IL8 ↑, JUN ↑, SGK1 ↑, CDKN1A ↑, IL1B ↑, PTGS2 ↑, STAT1 ↑,CCL3 ↑, FCGR1A ↑
Inhibition of Angiogenesis by TSP1	0.93	1.67	0.091	JUN ↑, THBS1 ↑, MMP9 ↑
Role of JAK1 and JAK3 in γc Cytokine Signaling	0.91	1.64	0.064	SOCS3 ↑, IL4R ↓, IL15 ↑, STAT1 ↑
ERK5 Signaling	0.87	1.59	0.063	IL6ST ↓, MYC ↓, SGK1 ↑, CREB5 ↑
Antigen Presentation Pathway	0.86	1.57	0.075	HLA-DOA ↓, HLA-DQA1 ↑, HLA-DOB ↓
Production of Nitric Oxide and Reactive Oxygen Species in Macrophages	0.68	1.35	0.038	JUN ↑, RHOB ↑, PLCG1 ↓, PCYOX1 ↑, STAT1 ↑, SPI1 ↑, IRF1 ↑
Phospholipase C Signaling	0.66	1.33	0.033	GNG11 ↑, CD79B ↓, RHOB ↑, FCGR2A ↑, PLCG1 ↓, CREB5 ↑, GNG7 ↓, CD79A ↓

The pathways that were significant only in females are shown. The p-values were derived from Fisher's exact test, −log(0.05) = 1.3 and the Benjamini-Hochberg-corrected t-test (FDR), −log(0.25) = 0.61. The ratio is the number of differentially expressed genes in the data set divided by the total number of genes in the given pathway.

doi:10.1371/journal.pone.0066229.t003

levels of estrogen and androgens decrease in aging males and low levels of estrogen are associated with a risk of fracture. However, the relative contribution of estrogens versus androgens in aging males is unclear [16]. *PDGF signaling*, which was also affected in males, has also been shown to be affected by a lack of estrogen [22].

The study presented here has some limitations. We have shown that the proportions of different T cell subpopulations or the proportion of monocyte-macrophage lineage cells do not differ between the genders (Table S4). However, there may be differences in the proportion of other cell populations that may explain some of the observed differences in gene expression.

Table 4. The canonical pathways that were affected in nonagenarian males.

Canonical pathways	−logFDR	−logP	Ratio	Molecules
Estrogen-mediated S-phase Entry	1.11	2.41	0.111	MYC ↓, CDKN1A ↑, E2F5 ↓
PDGF Signaling	0.77	1.71	0.051	MYC ↓, JUN ↑, STAT1 ↑, PDGFRB ↑
CD27 Signaling in Lymphocytes	0.62	1.52	0.055	JUN ↑, TRAF5 ↓, CD27 ↓
PPAR Signaling	0.62	1.49	0.040	JUN ↑, IL1B ↑, PTGS2 ↑, PDGFRB ↑
Role of Pattern Recognition Receptors in Recognition of Bacteria and Viruses	0.62	1.48	0.042	NLRP3 ↑, C5AR1 ↑, IL1B ↑, OAS3 ↑

The pathways that were significant only in males are shown. The p-values were derived from Fisher's exact test, −log(0.05) = 1.3 and the Benjamini-Hochberg-corrected t-test (FDR), −log(0.25) = 0.61. The ratio is the number of differentially expressed genes in the data set divided by the total number of genes in the given pathway.

doi:10.1371/journal.pone.0066229.t004

The number of healthy young controls used is relatively small in comparison to nonagenarian group. Thus, the small sample size will have an effect on the power of statistical testing to identify differentially expressed genes. To address this limitation, we have used statistical test specifically designed for small sample sizes. Data interpretation through pathway enrichment also mitigates this limitation as we do not need to observe all, only a significant fraction of genes belonging to a given pathway.

Additionally, we have previously shown that aging-related changes are affected by the CMV serostatus [23]. Because of the high seroprevalence of CMV in the nonagenarian study population (96% of females and 95% of males are seropositive for CMV), we cannot assess the combinatorial effect of gender and CMV on the age-associated changes in transcription.

This study focused on the effect of aging on the immune systems of males and females. It has been known for decades that gender has an influence on the function of the immune system, with females generally having a stronger immune response. Gender differences in the immune response are also detectable at transcriptomic levels [24]. Sex steroids, estrogen and testosterone, clearly play a role in driving gender differences in the immune response. Presently, there is no biological explanation for these aging-induced differences, and we can only speculate based on the available evidence. For example, aging strongly influences the levels of sex steroids, but during menopause estrogen levels decrease more rapidly than testosterone levels do during andropause [16]. Specifically, the positive effects of estrogens on the immune system stop at about age 45–55.

Another interesting possibility involves potential changes in the X-chromosome. The X-chromosome contains the largest number of immune-related genes in the genome [25], and aging may modify the function of genes on the X-chromosome. In females, X-chromosome is inactivated at random during an early embryonic stage (i.e. there is a 50/50 ratio of the maternal and paternal X-chromosomes). However, in elderly individuals, this ratio may be skewed. During a 13-year follow-up it was recently shown that this skewing is associated with survival [26]. It remains to be established whether this skewing has an influence on the expression of immune-related genes.

Materials and Methods

Ethics Statement

All participants in this study provided their written, informed consent. This study was conducted according to the principles expressed in the declaration of Helsinki, and the study protocol was approved by the ethics committee of the city of Tampere (Study protocol number SOTE 1592/402/96).

Population

The study population consisted of 146 nonagenarians (females n = 103, males n = 43) who were participating in the Vitality 90+ study, and 30 young, healthy controls (aged 19–30 years, median 22.5 years; females n = 21, males n = 9). All of the study subjects were of Western European descent. The Vitality 90+ study is an ongoing prospective population-based study that includes both home-dwelling and institutionalized individuals aged 90 years or more who live in the city of Tampere, Finland. The recruitment and characterization of the participants were performed as previously reported for earlier Vitality 90+ study cohorts [27]. In this study, we included only individuals born in 1920, and the samples used in this study were collected in the year 2010. The nonagenarians included in the study had not had any infections or received any vaccinations in the 30 days prior to the blood sample collection. The young controls consisted of healthy laboratory personnel who did not have any medically diagnosed chronic illnesses, who were non-smokers and who had not had any infections or received any vaccinations within the two weeks prior to the blood sample collection.

Sample Collection

The blood samples were collected into EDTA-containing tubes by a trained medical student during a home visit. All of the blood samples were drawn between 8 am and 12 am. The samples were directly subjected to leucocyte separation with a Ficoll-Paque density gradient (Ficoll-Paque^{TM} Premium, cat.no. 17-5442-03, GE Healthcare Bio-Sciences AB, Uppsala, Sweden). The PBMC layer was collected and a subset of cells was suspended in 150 µl of RNAlater solution (Ambion Inc., Austin, TX, USA) for use in a microarray analysis. The cells that were to be used for FACS analysis were suspended in 1 ml of a Freezing solution (5/8 FBS, 2/8 RPMI-160 medium, 1/8 DMSO) (FBS cat. no. F7524, Sigma-Aldrich, MO, USA; RPMI: cat. no. R0883, Sigma-Aldrich, MO, USA; DMSO: cat. no. 1.02931.0500, VWR, Espoo, Finland).

RNA Extraction and Transcriptomic Analysis

For RNA extraction, equal amounts of PBS and RNAlater were added to the cell suspension and then removed by centrifugation, leaving only the cell pellet. RNA was purified using an miRNeasy mini kit (Qiagen, CA, USA) according to the manufacturer's protocol using on-column DNase digestion (AppliChem GmbH, Darmstadt, Germany). The concentration and quality of the RNA were assessed with an Agilent RNA 6000 Nano Kit on Agilent 2100 Bioanalyzer (Agilent Technologies, CA, USA).

Labeled cRNA was prepared from 330 ng of total RNA using an Illumina TotalPrep RNA Amplification Kit (Ambion Inc., TX, USA) with overnight incubation according to the manufacturer's protocol. The quality of the labeled cRNA was determined using a 2100 Bioanalyzer (Agilent Technologies). In total, 1,500 ng of labeled cRNA was hybridized to a HumanHT-12 v4 Expression BeadChip (Cat no. BD-103-0204, Illumina Inc., CA, USA) overnight according to the Illumina protocol in the Core Facility at the Department of Biotechnology, University of Tartu. The chips were scanned using a Beadscan (Illumina Inc.). The microarray data are available in the GEO database (http://www.ncbi.nlm.nih.gov/geo/), accession number GSE40366.

Data Preprocessing and Statistical Analysis

The preprocessing, filtering and analysis of the data were performed with the Chipster v2.3 program [8] (CSC, Espoo, Finland). The lumi pipeline was used for data preprocessing and normalization [28]. The Array_Address_ID was used as a probe identifier, background correction was performed with the bgAdjust.affy package, and the data were transformed with the vst (variance stabilizing transformation) method and normalized with the rsn (robust spline normalization) method. The vst and rsn methods were chosen, because they are recommended in the literature and are designed to take in to account the technical replicates in each Illumina chip (the bead array technology) [28], [29], [30]. The quality control was performed by using box blot, density blot and PCA analysis.

To filter out the non-expressed probes and probes whose expression did not change between study groups, we filtered the data based on the coefficient of variation (CV, standard deviation/mean). We included the 5% of probes (2367) with the highest CV, i.e., the highest variation between nonagenarian and control samples. The nonagenarian samples and control samples were

compared with an empirical Bayes two-group test from the limma package [31] using the Benjamini-Hochberg false discovery rate (FDR) for multiple testing correction. The threshold for significance for p-values was set to 0.05. From these genes, we classified those with a linear fold change above 1.5 or below -1.5 as differentially expressed. This classification was performed to obtain comparable groups of genes from both genders. Because more females were included in the study, performing the analysis without a fold change limit produced almost three times more genes for the females compared with the males.

IPA

To identify canonical pathways associated with aging, we analyzed the gene sets with IPA software (Ingenuity® Systems, www.ingenuity.com). According to the manufacturer, the canonical pathways are well-characterized metabolic and cell signaling pathways that have been curated and hand-drawn by PhD-level scientists. The information used to construct the canonical pathways is derived from specific journal articles, review articles, text books, and the KEGG Ligand database. The canonical pathways are directional. All of the data sources provided by the Ingenuity Knowledge Base were included in the IPA analysis. For the association of molecules, only experimentally observed results were accepted and only human data were considered. The HumanHT-12 v 4.0 was used as a reference set to generate p-values for the pathways, and Fisher's exact test and Benjamini-Hochberg multiple testing correction (FDR) were used to calculate p-values for the pathways. With these parameters, we obtained 293 and 213 analysis-ready molecules for females and males, respectively, out of a total of 339 and 248 genes that were differentially expressed. We considered a canonical pathway to be significantly affected at $p<0.05$ ($-\log P>1.3$), at $FDR<0.25$ and when the pathway contained a minimum of 3 genes. Pathways associated with cancer and other disease, as defined by Ingenuity Systems®, were excluded from the analysis. The IPA analysis was performed on 6.3.2013.

qPCR Verification

In total, 300 ng of RNA was converted to cDNA using a High Capacity cDNA Reverse Transcription Kit (Part No. 4368814, Applied Biosystems, CA, USA). Because the amount of cDNA was limited, we performed a pre-amplification step using TaqMan® PreAmp Master Mix (Part No. 4348266, Applied Biosystems). This protocol amplifies small amounts of cDNA without introducing bias to the sample. In brief, 15 ng of cDNA was amplified for 10 cycles according to the manufacturer's instructions using the same assays with which the actual qPCRs were performed (CD83 Hs01077168_g1, IL8 Hs00174103_m1, LRRN3 Hs00539582_s1, PLCG1 Hs01008225_m1 and GUSB Hs00939627_m1 as endogenous control). GUSB was selected as the control gene because GUSB was shown to be the most stable transcript in gene expression studies that used aging PBMCs [32]. The pre-amplified product was diluted 1:5 with TE- buffer.

Transcript levels were determined with the single gene assays described above, using TaqMan® Gene Expression Master Mix (Part No. 4369016, Applied Biosystems). To determine whether the transcripts were differentially expressed between the nonagenarians and the young controls, we calculated the RQ values with RQ Manager Software (Applied Biosystems).

FACS

The proportions of different lymphocyte populations were determined using FACS analysis (BD FACSCanto II) and the results were analyzed with BD FACS Diva, version 6.1.3 (BD Biosciences, Franklin Lakes, NJ, USA). The antibodies used were FITC-CD14 (cat. no. 11-0149), PerCP-Cy5.5-CD3 (45-0037), APC-CD28 (17-0289) (eBioscience, San Diego, CA, USA), PE-Cy™7-CD4 (cat. no. 557852) and APC-Cy™7-CD8 (557834) (BD Biosciences).

Supporting Information

Table S1 Differentially expressed transcripts in male nonagenarians compared to male controls.
(XLSX)

Table S2 Differentially expressed transcripts in female nonagenarians compared to female controls.
(XLSX)

Table S3 Affected canonical pathways in both genders.
(XLSX)

Table S4 Proportions of T cell populations in nonagenarians and young controls of both genders.
(XLSX)

Acknowledgments

We would like to thank Sinikka Repo-Koskinen, Janette Hinkka, Katri Välimaa, Sanna Tuominen and Mario Saare for their skillful technical assistance.

Author Contributions

Conceived and designed the experiments: SM JJ MH. Performed the experiments: SM JJ TN LT. Analyzed the data: SM JJ MN. Contributed reagents/materials/analysis tools: MJ AH PP MH. Wrote the paper: SM JJ MH.

References

1. Pawelec G, Larbi A, Derhovanessian E (2010) Senescence of the human immune system. J Comp Pathol 142: S39–S44.
2. Arnold CR, Wolf J, Brunner S, Herndler-Brandstetter D, Grubeck-Loebenstein B (2011) Gain and loss of T cell subsets in old age – Age-related reshaping of the T cell repertoire. J Clin Immunol 31: 137–146.
3. Ademokun A, Wu Y-C, Dunn-Walters D (2010) The ageing B cell population: Composition and function. Biogerontology 11: 125–137.
4. Franceschi C, Capri M, Monti D, Giunta S, Olivieri F, et al. (2007) Inflammaging and anti-inflammaging: a systemic perspective on aging and longevity emerged from studies in humans. Mech Ageing Dev 128: 92–105.
5. Nunn CL, Lindenfors P, Pursall ER, Rolff J (2009) On sexual dimorphism in immune function. Phil Trans Soc B 364: 61–69.
6. McClelland EE, Smith JM (2011) Gender specific differences in the immune response to infection. Arch Immunol Ther Exp 59: 203–213.
7. Graves BM, Strand M, Lindsay AR (2006) A reassessment of sexual dimorphism in human senescence: theory, evidince and causation. Am J Hum Biol 18: 161–168.
8. Kallio MA, Tuimala JT, Hupponen T, Klemelä P, Gentile M, et al. (2011) Chipster: user-friendly analysis software for microarray and other high-throughput data. MBC Genomics 12: 507.
9. Harries LW, Hernandez D, Henley W, Wood AR, Holly AC, et al. (2010) Human aging is characterized by focused changes in gene expression and deregulation of alternative splicing. Aging Cell 10: 868–878.
10. Passtoors WM, Boer JM, Goeman JJ, Akker EB, Deelen J, et al. (2012) Transcriptional profiling of human familial longevity indicates a role for ASF1A and IL7R. PLoS one 7; e27759.
11. Cao J-N, Gollapudi S, Sharman EH, Jia Z, Gupta S (2010) Age-related alterations of gene-expression patterns in human CD8+ T cells. Aging Cell 9: 19–31.
12. Hong M-G, Myers AJ, Magnusson PKE, Prince JA (2008) Transcriptome-wide assessment of human brain and lymphocyte senescence. PLoS One 3: e3024.
13. Jylhävä J, Eklund C, Jylhä M, Hervonen A, Hurme M (2010) Expression profiling of immune-associated genes in peripheral blood mononuclear cells

reveals baseline differences in co-stimulatory signalling between nonagenarians and younger controls: the vitality 90+ study. Biogerontology 11: 671–677.
14. Bouchon A, Dietrich J, Colonna M (2000) Cutting Edge: Inflammatory Responses Can Be Triggered by TREM-1, a Novel Receptor Expressed on Neutrophils and Monocytes. J Immunol 164: 4991–4995.
15. Schenk M, Bouchon A, Seibold F, Mueller C (2007) TREM-1–expressing intestinal macrophages crucially amplify chronic inflammation in experimental colitis and inflammatory bowel diseases. J Clin Invest 117: 3097–3106.
16. Horstman AM, Dillon EL, Urban RJ, Sheffield-Moore M (2012) The role of androgens and estrogens on healthy aging and longevity. J Gerontol A Biol Sci Med Sci 67: 1140–1152.
17. Pautz A, Hahn S, Nowag S, Voss C, Kleinert H (2010) Regulation of the expression of inducible nitric oxide synthase. Nitric Oxide 23: 75–93.
18. Coulthard LR, White DE, Jones DL, McDermott F, Burchill SA (2009) $p38^{MAPK}$: stress responses from molecular mechanisms to therapeutics. Trends Mol Med 15: 369–379.
19. Brown DM (2010) Cytolytic CD4 cells: Direct mediators in infectious disease and malignancy. Cell Immunol 262: 89–95.
20. Waldman TA (2006) The Biology of interleukin-2 and inteleukin-15: implications for cancer therapy and vaccine design. Nat Rev Immunol 6: 595–601.
21. Gangemi S, Basile G, Monti D, Merendino RA, Di Pasquale G, et al. (2005) Age-related modifications in circulating IL-15 levels in humans. Mediators Inflamm 4: 245–247.
22. Weigel MT, Ghazoui Z, Dunbier A, Pancholi S, Dowsett M, et al. (2012) Preclinical and clinical studies of estrogen deprivation support the PDGF/Abl pathway as a novel therapeutic target for overcoming endocrine resistance in breast cancer. Breast Cancer Res 14: R78.
23. Kuparinen T, Marttila S, Jylhävä J, Tserel L, Peterson P, et al. (2013) Cytomegalovirus (CMV)-dependent and -independent changes in the aging of the human immune system: a transcriptomic analysis. Exp Gerontol 48: 305–312.
24. Hewagama A, Patel D, Yarlagadda S, Stricland FM, Richardson BC (2009) Stronger inflammatory/cytotoxic T-cell response in women identified by microarray analysis. Genes Immunol 10: 509–516.
25. Bianchi I, Lleo A, Gershwin ME, Invernizzi P (2012) The X chromosome and immune associated genes. J Autoimmun 38: J187–192.
26. Mengel-From J, Thinggaard M, Christiansen L, Vaupel JW, Orstavik KH, et al. (2012) Skewed X inactivation and survival: a 13-year follow-up study of elderly twins and singletons. Eur J Hum Genet 20: 361–364.
27. Goebeler S, Jylhä M, Hervonen A (2003) Medical history, cognitive status and mobility at the age of 90. A population based study in Tampere, Finland. Aging Clin Exp Res 15: 154–161.
28. Du P, Kibbe WA, Lin SM (2008) lumi: a pipeline for processing Illumina microarray. Bioinformatics 24: 1547–1548.
29. Lin SM, Du P, Huber W, Kibbe WA (2008) Model-based variance-stabilizing transformation for Illumina microarray data. Nucleic Acids Res 36: e11.
30. Dunning MJ, Ritchie ME, Barbosa-Morais NL, Tavaré S, Lynch AG (2008) Spike-in validation of an illumina-specific variance-stabilizing transformation. BMC Res Notes 1: 18.
31. Smyth GK (2004). Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. Stat Appl Genet Mol Biol 3: Article 3.
32. Zampieri M, Ciccarone F, Guastafierro T, Bacalini MG, Calabrese R, et al. (2010) Validation of suitable internal control genes for expression studies in aging. Mech Ageing Dev 131: 89–95.

Marttila et al. BMC Genomics (2015) 16:179
DOI 10.1186/s12864-015-1381-z

BMC Genomics

RESEARCH ARTICLE Open Access

Ageing-associated changes in the human DNA methylome: genomic locations and effects on gene expression

Saara Marttila[1,2*], Laura Kananen[1,2], Sergei Häyrynen[3], Juulia Jylhävä[1,2], Tapio Nevalainen[1,2], Antti Hervonen[2,4], Marja Jylhä[2,4], Matti Nykter[3] and Mikko Hurme[1,2,5]

Abstract

Background: Changes in DNA methylation are among the mechanisms contributing to the ageing process. We sought to identify ageing-associated DNA methylation changes at single-CpG-site resolution in blood leukocytes and to ensure that the observed changes were not due to differences in the proportions of leukocytes. The association between DNA methylation changes and gene expression levels was also investigated in the same individuals.

Results: We identified 8540 high-confidence ageing-associated CpG sites, 46% of which were hypermethylated in nonagenarians. The hypermethylation-associated genes belonged to a common category: they were predicted to be regulated by a common group of transcription factors and were enriched in a related set of GO terms and canonical pathways. Conversely, for the hypomethylation-associated genes only a limited set of GO terms and canonical pathways were identified. Among the 8540 CpG sites associated with ageing, methylation level of 377 sites was also associated with gene expression levels. These genes were enriched in GO terms and canonical pathways associated with immune system functions, particularly phagocytosis.

Conclusions: We find that certain ageing-associated immune-system impairments may be mediated via changes in DNA methylation. The results also imply that ageing-associated hypo- and hypermethylation are distinct processes: hypermethylation could be caused by programmed changes, whereas hypomethylation could be the result of environmental and stochastic processes.

Keywords: Epigenetics, Methylome, DNA methylation, Ageing, PBMCs, Gene expression, Molecular ageing, Hypermethylation, Hypomethylation

Background

Ageing can be described as a functional decline that leads to a diminished ability to respond to stress, increased homeostatic instability and an increased risk of diseases such as cancer and inflammatory diseases. Ultimately, these changes lead to death [1]. The molecular basis of ageing is multifactorial, including changes in energy metabolism, alterations in DNA repair mechanisms, increased inflammation and changes in leukocyte proportions (changes in CD4+/CD8+ ratio, increase of costimulatory CD28 receptor-deficient T cells [2]). Consequently, several theories exist regarding the mechanisms underlying ageing. Whether the ageing process itself consists of the accumulation of molecular damage due to environmental and stochastic effects or is a truly programmed or pseudo-programmed process that stems from development remains to be determined, yet a process as complex as ageing most likely involves aspects of all these phenomena [3-5].

Ageing leads to both global and local changes in the DNA methylation profile. Global hypomethylation has been shown to occur across tissues, and promoter-specific hypermethylation has been demonstrated for various tissues and genes [6]. Several ageing-relates diseases, such as cancer, Alzheimer's disease and type 2 diabetes, have also been shown to be associated with changes in DNA methylation [7]. The role of epigenetics in ageing-associated processes

* Correspondence: saara.marttila@uta.fi
[1]Department of Microbiology and Immunology, School of Medicine, University of Tampere, Tampere, Finland
[2]Gerontology Research Center, Tampere, Finland
Full list of author information is available at the end of the article

could be significant, as genetics appears to explain only a small portion of the observed variation in lifespan and healthspan [8]. As the epigenome is modified throughout life by varying environmental conditions, the accumulated effects of these changes could be most prominent in the aged population.

DNA methylation was suggested to control the activity of genes as early as 1975 [9,10] and has since been demonstrated to control the expression of single genes and the silencing of large sections of chromatin. DNA methylation mainly occurs on CpG-dinucleotides, which form CpG islands containing above-average CpG content. These CpG islands overlap the transcription start sites (TSSs) of the majority of human genes, and the classical role of DNA methylation is transcriptional inhibition, with the methylation of TSSs preventing the initiation of transcription [11,12]. The role of methylation in the gene body is less clear; methylation does not appear to block transcriptional elongation but may actually enhance it, and methylation may have a role in alternative splicing. Furthermore, DNA methylation is required for the suppression of transposable elements [13]. DNA methylation controls gene expression by directly inhibiting the binding of transcription factors (TFs), by recruiting methyl-binding proteins that prevent TFs from binding to DNA [14], or by affecting the conformation of the surrounding chromatin [15].

The relationship between ageing and DNA methylation has been studied previously by measuring the DNA methylation level of repetitive elements (global DNA methylation [6]) as well as with Illumina Golden Gate array [16] and the Infinium HumanMethylation27 BeadChip (27 K array) [17-22]. These arrays included a severely biased set of CpGs located in known cancer-associated genes and CpGs located almost exclusively in CpG island promoter regions, respectively. The Illumina Infinium HumanMethylation450 BeadChip (450K array) offers an improvement in this area, as the probes span 99% of the RefSeq genes and are distributed more evenly across the genome, such as on the shores and shelves of CpG islands and in non-CpG islands (non-CGIs), as well as in gene bodies and untranslated regions (UTRs) [23-28]. However, the majority of previous studies did not take into consideration the prominent ageing-associated changes in the proportions of leukocytes, thereby introducing possible bias into the analyses [29].

In this study, our aim was to identify ageing-associated DNA methylation changes that are independent of changes in leukocyte proportions. We also examined gene expression data from the same individuals from whom the methylation data were obtained, and we were therefore able to explore the relationship between gene expression and DNA methylation in these elderly individuals.

Results

Ageing-associated DNA methylation changes

Our study population consisted of the Vitality 90+ study population: there was a total of 146 nonagenarians and 30 young controls, from whom we extracted peripheral blood mononuclear cells (PBMCs). The methylation data were produced with the 450K array, and the expression data were obtained with the Illumina HumanHT12v4 BeadChip. Our aim was to identify ageing-associated changes in the level of DNA methylation. Our approach was two-sided, as we sought to concentrate on CpG sites that showed a large enough difference in the level of methylation to have a plausible biological significance but also to ensure that the identified differences were not due to changes in the proportions of leukocyte populations.

The proportions of different leukocytes differed between the nonagenarians and young controls in our study population, as reported previously [30]. A principal component analysis (PCA) revealed that the first principal component accounted for 20.5% of the observed variation in methylation levels detected in our data (Figure 1). This component was strongly associated with leukocyte proportions, indicating that the analysis needs to be adjusted for the proportions of leukocytes.

First, we compared the methylation levels at individual CpGs in the nonagenarian group (n = 122) with those in the young control group (n = 21) using the Wilcoxon rank-sum test and identified 10083 CpG sites that were differentially methylated between these two groups (with a Benjamini-Hochberg-corrected p-value <0.05 and a difference between absolute M-value medians >1). Second, age group-associated methylation sites were identified with a beta regression model, with sex and different leukocyte populations (the ratio of CD4+ and CD8+ T cells and the proportions of CD4 + CD28-, CD8 + CD28- and CD14+ cells) as covariates. This method identified 45507 CpG sites for which age group was a significant covariate (Bonferroni-corrected p-value <0.05). The 10083 CpG sites identified via the group comparison were enriched at the top of the list of the 45507 ageing-associated CpGs. However, 1543 of the 10083 CpG sites showed no statistical significance in the regression analysis, indicating that the perceived difference in methylation was due to differences in leukocyte proportions rather than ageing *per se*. We now report the 8540 CpG sites, which exhibit a large, statistically significant difference in the level of methylation between the nonagenarians and young controls and remain significant after adjusting for differences in leukocyte populations in the regression analysis, as truly ageing-associated methylation changes (for a list of all ageing-associated CpGs, see Additional file 1). Sex chromosomes were excluded from the analysis.

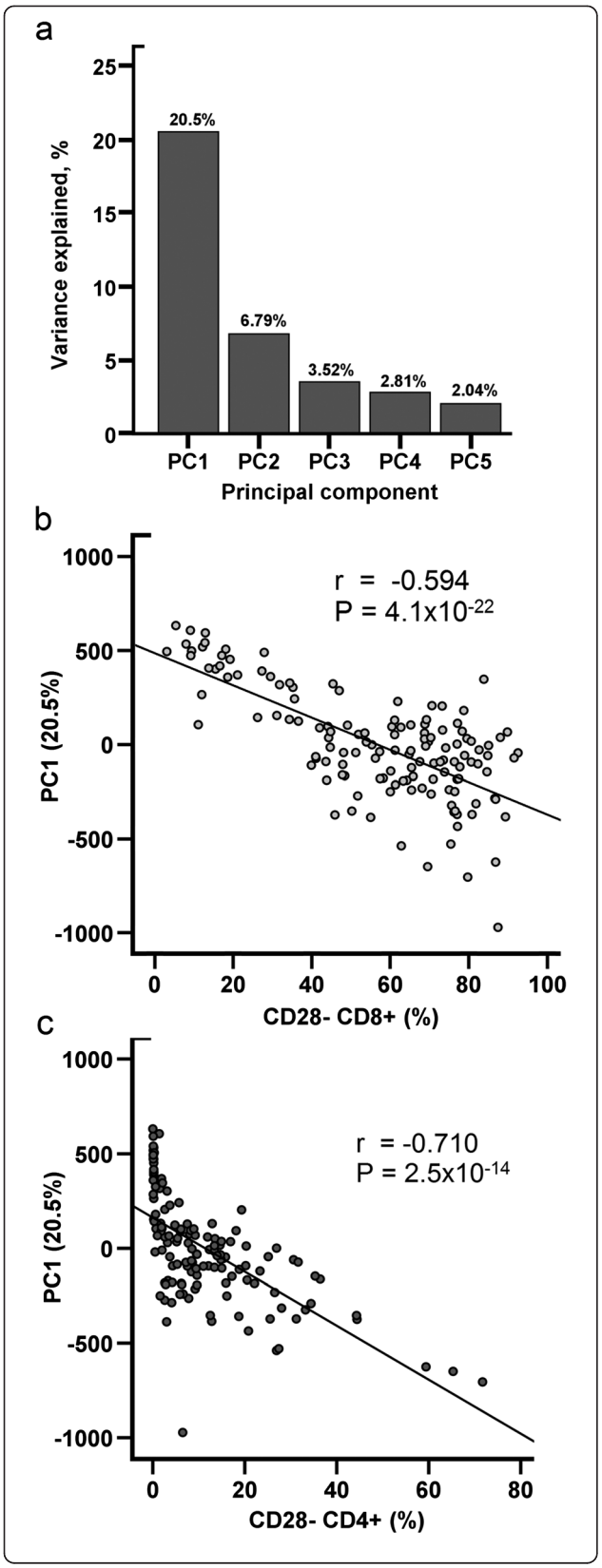


Figure 1 The association of cell type proportions with DNA methylation. The global DNA methylation was decomposed into a set of linearly independent principal component (PC) patterns. Components were used to examine the relationships between global DNA methylation and biological or non-biological covariates (e.g., gender, the batch effect and cell types). **(a)** The top 5 components (PC1-PC5) with the largest proportion of explained variance from the data. The percentages of explained variance are shown above the bars. **(b)** The association of the proportion of CD8 + CD28- cells with the first principal component (Spearman's rank correlation coefficient -0.594 (p = 4.1e-22)) and **(c)** the association of the proportion of CD4 + CD28- cells with the first principal component (Spearman's rank correlation coefficient -0.710 (p = 2.5e-14)).

Among the 8540 ageing-associated CpG sites, 3925 (46%) were hypermethylated, while 4615 (54%) were hypomethylated, in the nonagenarians. The most significant hits, based on the p-values obtained using the site-specific regression models, were cg16867657 (*ELOVL2*), cg16762684 (*MBP*), cg11344352 (*ERCC1*), cg17110586 and cg04875128 (*OTUD7A*). The largest differences in the level of methylation were observed for cg07211259 (*PDCD1LG2*), cg18826637 and cg26063719 (*VIM*), which were hypomethylated in the nonagenarians, and for cg06352730, cg00674365 (*ZNF471*) and cg21402921 (*GABRA5*), which were hypermethylated in the nonagenarians. The top-ranking hits are presented in Tables 1, 2 and Figure 2.

Genomic location of the ageing-associated methylation sites

The ageing-associated CpGs were not uniformly distributed across chromosomes, CpG islands or genes. Chromosomes 2, 3, 4, 5 and 18 contained more ageing-associated methylation sites than expected, whereas chromosomes 16, 17, 19 and 22 had fewer ageing-associated methylation sites than expected (Hypergeometric test $p < 0.05$, Additional file 2). On the majority of these chromosomes, the proportion of hypermethylated sites compared with the proportion of hypomethylated sites was roughly equal or was slightly biased towards an excess of hypomethylated sites, as in the overall data. Interestingly, on chromosomes 18 and 19, there were considerably more hypermethylated sites than expected: among the identified ageing-associated methylation sites on these chromosomes, 72% and 75% were hypermethylated, constituting a clear overrepresentation compared with the 46% of hypermethylated sites identified in the total data.

The CpG sites that were hypermethylated with advancing age were enriched at CpG islands, rather than on island shores or shelves or in non-CGIs. By contrast, the hypomethylated CpG sites were enriched in non-CGIs; their absence from CpG islands was striking, as only 1.2% of all hypomethylated sites were located in CpG islands, whereas 31.5% of the total probes were located

Table 1 Top 10 age-group associated CpG sites from the regression model

ProbeID	Gene	betareg estimate	betareg p-value	Δβ	Wilcoxon p-value
cg16867657	*ELOVL2*	1.023	6.38E-66	0.243	1.53E-10
cg16762684	*MBP*	-1.486	4.74E-64	-0.168	1.53E-10
cg11344352	*ERCC1*	-1.202	9.15E-63	-0.153	1.53E-10
cg17110586	na	0.895	1.46E-59	0.200	1.53E-10
cg04875128	*OTUD7A*	1.514	7.2E-58	0.279	1.53E-10
cg08262002	*LDB2*	-0.710	2.72E-55	-0.197	1.53E-10
cg18618815	*COL1A1*	-0.941	1.78E-52	-0.225	1.53E-10
cg00748589	na	0.864	1.36E-51	0.179	1.53E-10
cg15416179	*MAP2K3*	-1.131	2.38E-51	-0.187	1.53E-10
cg12065799	*RRAGC*	-0.823	8.15E-51	-0.088	1.53E-10
cg23479922	*MARCH11*	0.940	4.07E-49	0.263	1.53E-10
cg07544187	*CILP2*	1.541	2.35E-48	0.252	1.53E-10
cg09038267	*C10orf26*	1.227	1.48E-47	0.150	1.53E-10
cg13033938	*IP6K1*	-0.699	7.54E-47	-0.061	1.53E-10
cg19283806	*CCDC102B*	-1.253	9.82E-47	-0.267	1.53E-10
cg07547549	*SLC12A5*	0.900	5.02E-46	0.245	1.53E-10
cg01949403	*APOL3*	0.807	7.53E-46	0.111	1.53E-10
cg01243823	*NOD2*	-1.280	7.9E-46	-0.232	1.53E-10
cg22242842	na	-0.952	1.99E-44	-0.206	1.53E-10
cg06007201	*FAM38A*	-0.932	5.65E-44	-0.156	1.53E-10

CpG sites with most significant association to age group in the beta regression models (betareg). To clarify, Δβ refers to difference in the median of DNA methylation values between nonagenarians and young controls (difference in β-value), whereas betareg estimate refers to the estimate obtained from a regression model termed beta regression. Thus the absolute value of betareg estimate and the absolute value of Δβ for a given CpG site are not directly comparable, only the signs of the values are.

in CpG islands (Figure 3). In regard to gene regions, hypermethylated CpGs were enriched in regions near the TSSs and the 1st exons of genes, whereas hypomethylated sites were scarce in these areas and were enriched in the gene body and, more strongly, in the regions outside of genes (Figure 3).

Functional annotation of the ageing-associated methylation sites

The locations of ageing-associated hyper- and hypomethylation differ, thus it can be assumed that their origins and/or functions also differ. Therefore, we performed the functional analyses separately for hypermethylated and hypomethylated sites and genes harbouring these ageing-associated methylation sites. The 3925 hypermethylated sites were annotated to 1832 different genes, and the 4615 hypomethylated CpG sites were annotated to 2057 different genes.

GOrilla (Gene Ontology enRIchment anaLysis and visuaLizAtion tool) [31,32] was used to identify the GO functions and processes associated with ageing-associated hyper- and hypomethylation-associated genes. For both categories, we identified more significant GO terms for hypermethylation-associated genes, even though there were fewer hypermethylation-associated genes compared with hypomethylation-associated genes. For the hypermethylation-associated genes, 36 enriched GO function terms were identified (Bonferroni corrected $p < 0.05$), whereas for the hypomethylation-associated genes, 27 enriched GO function terms were identified; 11 of these terms were common to the two groups (Additional file 3). The top GO function terms for the hypermethylation-associated genes were unique to these genes; these terms were associated with sequence-specific DNA binding and transcription factor binding (also presented as a diagram in Additional file 4). The GO terms that were enriched only for hypomethylated sites did not reveal similar enrichment for a common process (Additional file 4). The GO function terms that were common to hypermethylation- and hypomethylation-associated genes also formed a group and were clustered around channel function-associated GO terms. The results for GO process terms was similar to that for GO functions, as we identified 265 significant GO terms for hypermethylation-associated genes, whereas for hypomethylation-associated genes, we identified only 53 significant GO terms; 41 of these terms were common to hyper- and hypomethylation-associated genes. (Additional file 5). The top-ranking hypermethylation-

Table 2 Top 10 CpG sites with the largest Δβ between nonagenarians and young controls

ProbeID	Gene	betareg estimate	betareg p-value	Δβ	Wilcoxon p-value
cg07211259	*PDCD1LG2*	-1.086	3.24E-30	-0.290	1.53E-10
cg18826637	na	-1.280	2.23E-32	-0.289	1.53E-10
cg26063719	*VIM*	-1.036	6.09E-25	-0.284	1.53E-10
cg08548498	*SLPI*	-0.767	1.24E-15	-0.278	1.66E-10
cg19283806	*CCDC102B*	-1.253	9.82E-47	-0.267	1.53E-10
cg13591783	*ANXA1*	-0.826	5.17E-22	-0.266	1.53E-10
cg27192248	na	-1.246	2.57E-20	-0.265	1.59E-10
cg03274391	na	-1.263	1.18E-15	-0.264	1.54E-10
cg23654401	*VOPP1*	-0.781	2.94E-16	-0.262	1.54E-10
cg26269881	*BHLHE40*	-1.005	4.25E-25	-0.261	1.53E-10
cg18952796	*NPTX2*	1.121	6.89E-26	0.264	1.56E-10
cg17688525	*L3MBTL4*	0.865	1.36E-11	0.265	5.86E-10
cg27526665	*THRB*	0.940	2.64E-22	0.266	2.0E-10
cg09555124	*IGF2R*	0.944	4.34E-23	0.277	1.53E-10
cg23160016	*GABRA2*	1.041	1.01E-17	0.277	2.49E-10
cg10568066	*RNF39*	0.973	4.68E-13	0.278	1.60E-8
cg04875128	*OTUD7A*	1.514	7.2E-58	0.279	1.53E-10
cg21402921	*GABRA5*	0.868	4.90E-17	0.285	5.58E-10
cg00674365	*ZNF471*	1.033	6.27E-24	0.288	3.65E-10
cg06352730	na	1.437	1.26E-23	0.288	1.76E-10

CpG sites with the largest difference in the methylation level (Δβ) between nonagenarians and controls. To clarify, Δβ refers to difference in the median of DNA methylation values between nonagenarians and young controls (difference in β-value), whereas betareg estimate refers to the estimate obtained from a regression model termed beta regression. Thus the absolute value of betareg estimate and the absolute value of Δβ for a given CpG site are not directly comparable, only the signs of the values are.

specific GO terms were clustered around two types of processes: development and morphogenesis; and metabolic processes, gene expression and nucleotide metabolism (Additional file 6). Again, the significant GO terms associated with hypomethylation did not belong to a specific group (Additional file 6). The hypermethylation-specific GO terms that formed specific clusters are presented in Tables 3, 4 and 5.

PScan [33] was used to identify transcription factors that could be common regulators of the identified genes. For the hypermethylation-associated genes, 24 common transcription factors were identified (Additional file 7), whereas for the hypomethylation-associated genes, only one TF (EWSR1-FLI1, p-value 1.502e-5), was identified. Among the 24 transcription factors that were common to hypermethylation-associated genes, half (12) were zinc-coordinating transcription factors.

We also identified canonical pathways related to hypo- and hypermethylation-associated genes through Ingenuity pathway analysis (IPA) [34]. For the hypermethylation-associated genes, we identified 19 affected canonical pathways (Benjamini-Hochberg-corrected p-value <0.05), whereas for the hypomethylation-associated genes, 3 pathways were identified, 1 of which was common to both groups of genes (Additional file 8). The canonical pathways associated with hypermethylation in nonagenarians belonged to signalling pathway categories such as *Organismal growth & development*, *Cellular growth* and *Proliferation and development* (Figure 4).

Effect of sex on ageing-associated DNA methylation changes

Among the 8540 ageing-associated, high-confidence CpG sites, only 7 showed a statistically significant association with sex in our beta regression analysis in which age group, sex and leukocyte proportions were included

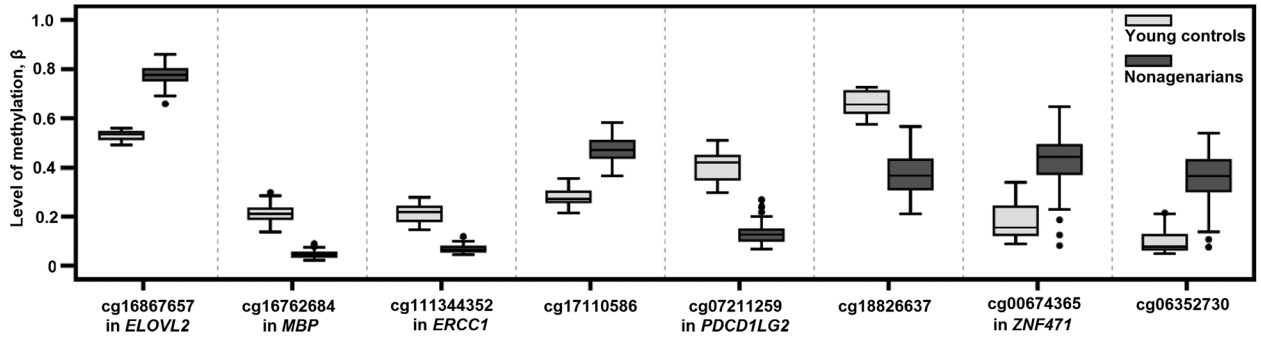


Figure 2 The top ageing-associated CpG sites. The level of DNA methylation presented as a box plot in the control and nonagenarian groups and in CpG sites with the strongest association to age group (cg16867657 (*ELOVL2*), cg16762684 (*MBP*), cg111344352 (*ERCC1*) and cg17110586) and in CpG sites with the largest methylation differences (cg07211259 (*PDCD1LG2*), cg18826637, cg00674365 (*ZNF471*) and cg06352730). Gene annotation is shown where applicable. See also Tables 1 and 2.

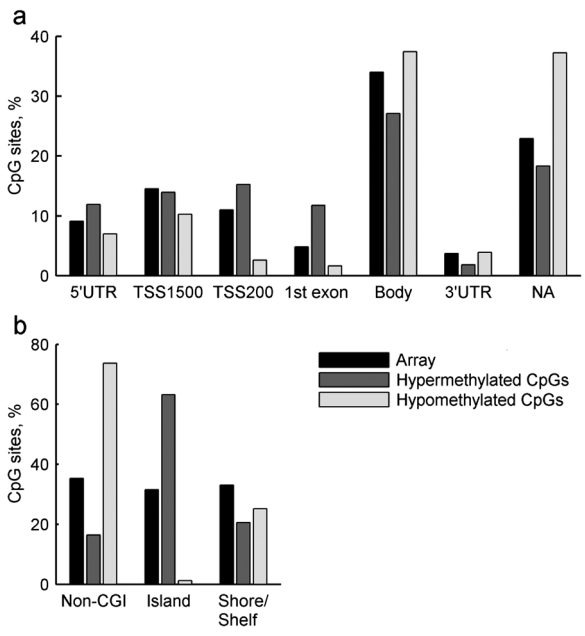


Figure 3 Locations of the ageing-associated methylation sites identified in the nonagenarians. Ageing-associated hyper- and hypomethylated probes are distributed differently across the genome. The distribution of ageing-associated CpGs in relation to **(a)** genes and **(b)** CpG islands. Ageing-associated hypermethylation is mainly located in CpG islands, TSSs and the 1st exons of genes, whereas ageing-associated hypomethylation occurs mainly in non-CGIs, gene bodies and areas outside of genes. In the figure, array denotes the distribution of probes in the 450K array.

as covariates. The sex-associated sites that were also ageing-associated are listed in Additional file 9.

Association between ageing-associated DNA methylation changes and gene expression

We performed a correlation analysis between the level of methylation at ageing-associated CpGs and the expression level of genes in which these CpG sites were located. In nonagenarians, we identified 422 correlation pairs (Pearson correlation, Benjamini-Hochberg-corrected p-value < 0.05) that consisted of 377 individual CpG sites and 233 individual genes (Additional file 10). The apparent discrepancy in these numbers is because a single CpG can be located in a region of overlapping transcripts, and there can be several CpGs within the coding region of a single transcript. In the young controls, we identified 50 expression-methylation correlation pairs (Pearson correlation, Benjamini-Hochberg-corrected p-value < 0.05), consisting of 43 individual CpGs and 37 individual genes. In nonagenarians, 255 (60%) of these correlated CpG-gene pairs showed an inverse correlation, and 167 (40%) exhibited a direct correlation. In the young controls, these numbers were 46 (92%) and 4 (8%), respectively. Among the genes whose expression level was correlated with the level of DNA methylation, 23 were common to the nonagenarians and young controls, and in all cases, the direction of correlation was the same. We previously showed that 14 of the 233 genes identified in the present study were differentially expressed with age in both sexes and that an additional 14 were differentially expressed with age in either sex [35] (For details, see Additional file 10).

The correlated CpGs did not exhibit a similar distribution in the genome to the ageing-associated methylation sites. Those CpG sites whose methylation level correlated with the level of gene expression were concentrated in non-CGIs and on CpG island shores and shelves, whereas only a few (14.6%) were located in CpG islands. In the non-CGIs, the majority of correlations were direct, whereas the opposite situation was observed in CpG islands and on island shores and shelves. With regard to regions within genes, the correlated CpGs were relatively evenly distributed. However, we identified an abundance of correlated CpGs within gene bodies (55% of all correlated sites), where the majority of sites were directly correlated. In regions near a TSS (TSS200, from TSS

Table 3 Hypermethylation-specific GO function terms in nonagenarians

GO term	Description	P-value	FDR q-value	Rank (out of 36)
GO:0043565	Sequence-specific DNA binding	1.18E-32	4.65E-29	1
GO:0001071	Nucleic acid binding transcription factor activity	9.38E-31	1.85E-27	2
GO:0003700	Sequence-specific DNA binding transcription factor activity	2.22E-30	2.92E-27	3
GO:0003677	DNA binding	6.48E-16	6.38E-13	4
GO:0000981	Sequence-specific DNA binding RNA polymerase II transcription factor activity	2.6E-15	2.05E-12	5
GO:0000976	Transcription regulatory region sequence-specific DNA binding	4.8E-13	2.7E-10	7
GO:0044212	Transcription regulatory region DNA binding	7.92E-12	3.47E-9	9
GO:0000975	Regulatory region DNA binding	2.22E-11	8.75E-9	10
GO:0001067	Regulatory region nucleic acid binding	2.22E-11	7.96E-9	11

This table includes only the hypermethylation-specific GO function terms that form a common cluster, associated with DNA binding and transcription. The presented p-values are unadjusted and the threshold for significance is 1.27e-5 (Bonferroni). The rank denotes the placement of a given GO term in the list of all significant GO terms. For all statistically significant GO function terms, see Additional file 3.

Table 4 Hypermethylation-specific GO process terms in nonagenarians

GO term	Description	P-value	FDR q-value	Rank (out of 265)
GO:0048598	Embryonic morphogenesis	1.25E-22	8.85E-20	17
GO:0048729	Tissue morphogenesis	2.99E-19	1.64E-16	22
GO:0002009	Morphogenesis of an epithelium	6.94E-18	3.22E-15	26
GO:0001763	Morphogenesis of a branching structure	1.84E-17	7.65E-15	29
GO:0048754	Branching morphogenesis of an epithelial tube	1.26E-15	3.09E-13	49
GO:0048562	Embryonic organ morphogenesis	6.12E-14	1.1E-11	67
GO:0009887	Organ morphogenesis	1.13E-13	1.92E-11	71
GO:0035107	Appendage morphogenesis	2.17E-12	2.85E-10	92
GO:0035108	Limb morphogenesis	2.17E-12	2.82E-10	93
GO:0030326	Embryonic limb morphogenesis	7.07E-12	8.7E-10	98
GO:0035113	Embryonic appendage morphogenesis	7.07E-12	8.61E-10	99
GO:0048704	Embryonic skeletal system morphogenesis	2.25E-11	2.56E-9	106
GO:0048705	Skeletal system morphogenesis	1.04E-10	1.11E-8	113
GO:0048732	Gland development	2.37E-10	2.3E-8	124

This table includes only the hypermethylation-specific GO process terms that form a common cluster, associated with development and morphogenesis. The presented p-values are unadjusted and the threshold for significance is 4.15e-6 (Bonferroni). The rank denotes the placement of a given GO term in the list of all significant GO terms. For all statistically significant GO process terms, see Additional file 5.

to -200 nucleotides upstream of TSS), directly correlated CpGs were almost completely absent (Additional file 11).

To analyse the processes associated with the genes that displayed a correlation between expression and methylation levels, we performed GO term analysis and IPA for the nonagenarians. We identified 20 GO process terms (Bonferroni-corrected p-value <0.05), of which 6 (30%) were immune system associated. Numerous immune system pathways were also identified when considering GO process terms that were more loosely associated with these genes (Benjamini-Hochberg-corrected p-value <0.05), where 39 of 121 (32%) statistically significant GO process terms were immune system associated (Additional file 12). Only one GO function term (GO:0005515 Protein binding) was associated with the correlated CpGs. In addition to the immune system, pathways related to the reaction to the environment were affected. Ingenuity canonical pathway analysis revealed 15 canonical pathways (Benjamini-Hochberg-corrected p-value <0.05) (Table 6), the majority of which were directly immune system associated (*Crosstalk between Dendritic Cells and Natural Killer Cells*, *Antigen Presentation Pathway*, *Fcγ Receptor-mediated Phagocytosis in Macrophages and Monocytes*, *T Helper Cell Differentiation*) or associated with cytoskeleton remodelling and endocytosis (*Integrin Signalling*, *Actin Cytoskeleton Signalling*, *Tec Kinase Signalling*, *Paxillin Signalling*, *Caveolar-mediated Endocytosis Signalling*).

Discussion

Ageing-associated DNA methylation changes; single CpG sites and their location and function

Here, we present the results of our ageing-associated DNA methylation analysis. In summary, our results were obtained with a 450K array using PBMCs collected from nonagenarians and young controls. The study subjects were analysed as two age groups, and we used two different statistical methods to verify that the ageing-associated methylation sites identified had a prominent difference in the level of methylation between the age groups and that this difference was not due to changes in leukocyte proportions. The proportions of leukocytes were measured via FACS. We also added a layer of information by including gene expression data from the same individuals. The small number of young controls is a potential limitation in our study; thus, the results should be interpreted accordingly.

In previous ageing-methylation studies, the age range of the oldest study subjects has typically been from 70 to 80 years of age [17,19-22,25,26], and the youngest age group to be included has ranged from new-borns [22,28] to 50 years of age [20,25]. In studies in which subjects over 90 years old have been analysed, these individuals represented a minority of the study population or the overall sample size has been small [18,27,28,36]. Hence, a strength of our study is the large number of the oldest-old individuals homogenous in terms of age. In addition, our study population represents the two extremities of adulthood, and as age was used as a dichotomous variable we were able to identify both changes occurring linearly with age as well as changes that occur in either end of the spectrum. The DNA methylation studies performed with 27K arrays [17-22] fail to capture methylation changes outside gene promoters, yet our results, as well as those of others [25,28,36], show that ageing-associated changes are not restricted to gene promoters. In contrast to our study, previous reports combining methylation and expression

Table 5 Hypermethylation-specific GO process terms in nonagenarians

GO term	Description	P-value	FDR q-value	Rank (out of 265)
GO:0045935	Positive regulation of nucleobase-containing compound metabolic process	7.07E-17	2.37E-14	36
GO:0051173	Positive regulation of nitrogen compound metabolic process	1.06E-16	3.44E-14	37
GO:0031328	Positive regulation of cellular biosynthetic process	2.74E-16	8.47E-14	39
GO:0009891	Positive regulation of biosynthetic process	3.61E-16	1.06E-13	41
GO:0045893	Positive regulation of transcription, DNA-templated	5.05E-16	1.35E-13	45
GO:0019219	Regulation of nucleobase-containing compound metabolic process	7.72E-16	2.02E-13	46
GO:0010628	Positive regulation of gene expression	1.11E-15	2.78E-13	48
GO:0006357	Regulation of transcription from RNA polymerase II promoter	5.05E-15	1.15E-12	53
GO:0031326	Regulation of cellular biosynthetic process	1.07E-14	2.27E-12	57
GO:0051171	Regulation of nitrogen compound metabolic process	1.28E-14	2.66E-12	58
GO:0009889	Regulation of biosynthetic process	1.45E-14	2.96E-12	59
GO:0051254	Positive regulation of RNA metabolic process	1.8E-14	3.56E-12	61
GO:0006355	Regulation of transcription, DNA-templated	1.98E-14	3.85E-12	62
GO:1902680	Positive regulation of RNA biosynthetic process	2.06E-14	3.93E-12	63
GO:0045944	Positive regulation of transcription from RNA polymerase II promoter	2.94E-14	5.45E-12	65
GO:0010557	Positive regulation of macromolecule biosynthetic process	6.72E-14	1.19E-11	68
GO:0031323	Regulation of cellular metabolic process	1.09E-13	1.87E-11	70
GO:2001141	Regulation of RNA biosynthetic process	1.37E-13	2.29E-11	72
GO:0031325	Positive regulation of cellular metabolic process	2.22E-13	3.57E-11	75
GO:0045934	Negative regulation of nucleobase-containing compound metabolic process	2.8E-13	4.39E-11	77
GO:0031327	Negative regulation of cellular biosynthetic process	3.75E-13	5.8E-11	78
GO:0009893	Positive regulation of metabolic process	3.84E-13	5.85E-11	79
GO:0009890	Negative regulation of biosynthetic process	3.84E-13	5.78E-11	80
GO:0000122	Negative regulation of transcription from RNA polymerase II promoter	5.25E-13	7.72E-11	82
GO:0051252	Regulation of RNA metabolic process	9.69E-13	1.39E-10	84
GO:0051172	Negative regulation of nitrogen compound metabolic process	1.07E-12	1.52E-10	85
GO:2000112	Regulation of cellular macromolecule biosynthetic process	1.14E-12	1.6E-10	86
GO:0080090	Regulation of primary metabolic process	1.39E-12	1.92E-10	87
GO:0010629	Negative regulation of gene expression	2.02E-12	2.68E-10	91
GO:0010556	Regulation of macromolecule biosynthetic process	3.09E-12	3.96E-10	94
GO:0045892	Negative regulation of transcription, DNA-templated	4.17E-12	5.29E-10	95
GO:1902679	Negative regulation of RNA biosynthetic process	4.83E-12	6.07E-10	96
GO:0019222	Regulation of metabolic process	1.56E-11	1.84E-9	102
GO:0051253	Negative regulation of RNA metabolic process	1.94E-11	2.25E-9	104
GO:0010468	Regulation of gene expression	1.35E-10	1.39E-8	117
GO:0010558	Negative regulation of macromolecule biosynthetic process	1.39E-10	1.41E-8	119
GO:0010604	Positive regulation of macromolecule metabolic process	1.49E-10	1.46E-8	123
GO:2000113	Negative regulation of cellular macromolecule biosynthetic process	3.62E-10	3.38E-8	129

This table includes only the hypermethylation-specific GO process terms that form a common cluster, associated with nucleotide metabolism, RNA metabolism and transcription. The presented p-values are unadjusted and the threshold for significance is 4.15e-6 (Bonferroni). The rank denotes the placement of a given GO term in the list of all significant GO terms. For all statistically significant GO process terms, see Additional file 5.

data have relied on individuals from different study cohorts [27,36]. A group of studies have also tried to identify a small set of methylation sites that could be used to construct an ageing signature [22,23,25,27]. However, by focusing on a broader set of ageing-associated methylation sites, the mechanisms of ageing can be more thoroughly examined. Given that published ageing-methylation studies have been conducted using various age ranges and

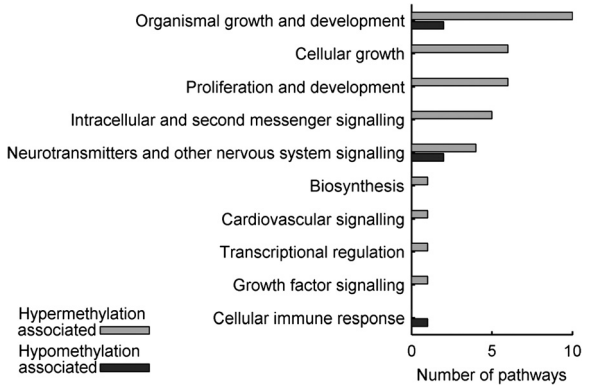


Figure 4 Canonical pathway categories associated with differentially methylated genes in the nonagenarians. Hypomethylation-associated genes are enriched in only three canonical pathways, thus corresponding to only a few pathway categories. Hypermethylation-associated genes are enriched in canonical pathways associated mainly with organismal and cellular growth and development. One canonical pathway can belong to several categories; for the individual pathways, see Additional file 8.

statistical methods, discrepancies in the results are most likely due to both biological and statistical factors.

The main characteristics of the ageing-associated methylation sites identified in the present study are presented in Table 7. We identified 8540 high-confidence CpG sites that show a large difference in methylation levels between nonagenarians and young controls and that present high statistical significance in a regression model adjusted for the leukocyte proportion. A slight majority (54%) of the identified sites were hypomethylated in the nonagenarians. Among the top-ranking ageing-associated methylation changes that have been reported with a high frequency, *ELOVL2* (cg16867657, cg24724428), *PENK* (cg04598121), *FHL2* (cg22454769, cg24079702, cg06639320) *EDARADD* (cg09809672), *KLF14* (cg04528819, cg07955995) and *OTUD7A* (cg04875128) were also identified in our study. Of these genes, only *EDARADD* was hypomethylated in the nonagenarians compared with the controls. As reported by Steegenga et al. [24], among 8 previous studies analysing the association of ageing and DNA methylation changes in PBMCs [17,19-21,25-28], only 529 probes were reported to be affected by age by more than one research group. Of these probes, our analysis identified 105. Interestingly, the majority of frequently reported CpG sites are hypermethylated with increasing age. Among the 151 CpG sites (148 of which are present in 450K) reported to be associated with ageing by more than 3 groups [24], 77% were hypermethylated. Of the 105 CpG sites that are frequently reported and were identified in our study, 79% (83/105) were hypermethylated.

The functional roles of the 10 most frequently reported ageing-associated methylation sites are currently unclear, as they are not associated with a common, ageing-related process. According to our results, the genes associated with these sites are not enriched under a common GO term or

Table 6 Canonical pathways associated with genes whose expression levels correlate with the level of DNA methylation in nonagenarians

Ingenuity canonical pathways	p-value (B-H corrected)	Ratio	Molecules
Integrin signalling	0.016	0.054	ITGB1,PTK2,RAP2A,FYN,PAK1,RALA,ACTA2,ITGA6,CAPN2,ITGAL,ACTN1
Actin cytoskeleton signalling	0.017	0.048	ITGB1,PTK2,TIAM1,PAK1,F2R,ACTA2,TRIO,PDGFD,GSN,ARHGAP24,ACTN1
Tec kinase signalling	0.019	0.051	STAT4,ITGB1,PTK2,FYN,GNAI3,GNB4,PAK1,ACTA2,HCK
Agrin interactions at neuromuscular junction	0.019	0.090	ITGB1,PTK2,PAK1,ACTA2,ITGA6,ITGAL
Paxillin signalling	0.020	0.064	ITGB1,PTK2,PAK1,ACTA2,ITGA6,ITGAL,ACTN1
Reelin signalling in neurons	0.026	0.073	ITGB1,FYN,HCK,ITGA6,ARHGEF11,ITGAL
Phospholipase C signalling	0.030	0.041	ITGB1,FYN,GNB4,RALA,AHNAK,SYK,MEF2C,ARHGEF11,PLD6,CREB5
Germ cell-sertoli cell junction signalling	0.030	0.051	ITGB1,PTK2,TGFBR2,PAK1,ACTA2,ITGA6,GSN,ACTN1
Crosstalk between dendritic cells and natural killer cells	0.030	0.066	IFNG,ACTA2,CD86,HLA-F,ITGAL,CCR7
Protein kinase A signalling	0.030	0.035	TGFBR2,PTK2,GNB4,GNAI3,TCF4,PTPN7,YWHAG,DUSP10, RYR1,LEF1,CREB5,PTPRM,SIRPA
Antigen presentation pathway	0.030	0.100	PSMB9,IFNG,HLA-F,HLA-DPB1
Fcγ receptor-mediated phagocytosis in macrophages and monocytes	0.030	0.063	FYN,PAK1,SYK,ACTA2,HCK,PLD6
T helper cell differentiation	0.037	0.073	STAT4,TGFBR2,IFNG,IFNGR2,CD86
Ephrin receptor signalling	0.038	0.041	ITGB1,PTK2,FYN,GNAI3,GNB4,PAK1,PDGFD,CREB5
Caveolar-mediated endocytosis signalling	0.044	0.062	ITGB1,FYN,ACTA2,ITGA6,ITGAL

Canonical pathways (IPA [34]) associated with genes whose expression levels correlate with the level of DNA methylation. P-values are Benjamini-Hochberg corrected. Ratio = number of identified genes/number of genes in the pathway. Molecules refer to genes affected in our analysis present in the given pathway.

Table 7 Characteristics of ageing-associated methylation sites

	Hypermethylated	Hypomethylated
n	3925	4615
CpG island location	CpG islands	Non-CGI
Genomic location	TSS, 1st exon	Gene body, outside genes
Associated genes	1832	2057
GO function terms	36	27
GO process terms	265	53
Canonical pathways (IPA)	19	3
Transcription factors	24	1

The CpG island location and genomic location refer to the sites where hyper- and hypomethylated sites are most abundant. Notably, there are more hypomethylated CpG sites compared with hypermethylated CpG sites and therefore also more hypomethylation-associated genes, yet the hypermethylation-associated genes are enriched in more GO terms and canonical pathways, and they share more common transcription factors.

in common canonical pathways. Only *FHL2*, *PENK* and *OTUD7A* are included in any identified GO term, and none of them are included in affected canonical pathways. The methylation levels of frequently reported CpGs are not correlated with the expression levels of the corresponding genes. For *EDARADD*, we identified an additional CpG site (cg18964582) that was differentially methylated between nonagenarians and young controls, located within TSS1500, where there is an inverse correlation between the methylation level and the expression level. However, based on previous findings and our results, it appears that the frequently reported ageing-associated CpG sites are not strongly associated with known ageing-related mechanisms but could instead represent a cellular chronological clock mechanism.

Our results revealed an enrichment of ageing-associated hypermethylation at CpG islands, whereas hypomethylation was enriched in non-CGIs and was almost totally absent from CpG islands. These findings are in line with previously reported results [16,24,25,28,36,37]. The majority of CpG sites are not initially methylated in CpG islands, and the change observed during ageing is hypermethylation. The opposite is true for regions with few CpG sites that initially are heavily methylated, and the non-CGIs are associated with hypomethylation. These results support the notion that the normal maintenance of DNA methylation patterns is disrupted with ageing [38]. As both hypomethylation and hypermethylation occur with ageing, it appears that both *de novo* methylation processes, mediated by *DNMT3A* and *DNMT3B* methyltransferases, and the maintenance of existing DNA methylation, mediated by *DNMT1*, are disrupted with ageing. Interestingly, our results identified 4 CpG sites in *DNMT3A* that were ageing associated (cg00050692, which was hypomethylated, and cg15302376, cg15843262 and cg26544247, which were hypermethylated in the nonagenarians). However, there was no correlation between the level of methylation and *DNMT3A* expression.

Our results showed that not only are ageing-associated hyper- and hypomethylation found at different genomic sites but that these changes are also found in genes associated with different functions. Our findings further revealed that ageing-associated hypermethylation is concentrated in genes associated with developmental processes as well as DNA-binding and transcription of genes, whereas hypomethylation is not enriched among a specific set of genes. Johansson et al. [36], Rakyan et al. [19] and Florath et al. [25] previously reported the association of hypermethylation with developmental processes and DNA binding. As DNA methylation regulates DNA transcription, it is interesting that the genes required during this process are differentially methylated with ageing. In comparison with ageing-associated hypomethylation, hypermethylation appears to be a more regulated process, as no strongly hypomethylation-specific functions or processes were identified in this study.

It is notable that while the individual sites reported to be ageing associated differ to some extent between studies, the results regarding their locations in the genome and the molecular functions with which they are associated are more uniform. Single highly significant CpG sites have also been reported in various studies, including sites located in the *ELOVL2* and *FLH2* genes. Common ageing-associated DNA methylation changes can also be observed across different tissues [6,23]. Thus, it appears that at least some fraction of ageing-associated DNA methylation changes is caused by programmed or pseudo-programmed changes that occur in a similar manner across tissues and individuals. As certain processes and sites are reported frequently, it can be hypothesised that these sites and processes represent clock-like changes associated with ageing. For example, a strong association with chronological age has been shown for *ELOVL2* (cg16867657) [25,26,36]. However, it remains to be investigated whether these sites are only associated with chronological age or if there are also associations with phenotypic changes related to (successful) ageing. If these frequently reported sites are only markers of chronological age, markers of biological age are yet to be identified.

The role of cell proportions in DNA methylation studies

The majority of DNA methylation and expression studies are performed with whole blood or PBMCs due to the accessibility of these tissue types. However, PBMCs consist of various cell types, and different individuals can exhibit differences in the proportions of different cell populations. Ageing is known to be associated with changes in the proportions of T cells [2,39]. Furthermore, the different leukocyte subtypes show differences in their DNA methylation levels [40], and changes in

DNA methylation are known to be one of the factors regulating lineage development in leukocytes [41].

Previous reports have claimed that differences in the proportions of leukocytes do not cause bias in methylation analyses [17,21,22]. However, contradictory reports have also been published [42], and recently it has been systematically shown that differences in leukocyte proportions should be taken into consideration when analysing ageing-associated methylation differences [29]. Our PCA revealed that the largest percentage of the variation in our methylation data was associated with the proportions of different leukocyte subtypes (Figure 1).

The role of sex in ageing-associated DNA methylation studies

According to our results, sex does not have a large effect on ageing-associated DNA methylation changes in autosomes, as we identified only 7 CpG sites for which sex, in addition to age group, was a significant covariate in the regression model. However, the small number of male samples in our control population may have precluded the identification of ageing-associated sex differences. Nevertheless, Johansson et al. [36] and McClay et al. [37] previously reported similar findings in studies focusing on individual sites associated with ageing. In studies where methylation profiles have been used to predict age, however, the methylome has been shown to age more rapidly in men than in women [22,27]. DNA methylation is believed to mediate the long-term regulation of gene expression [13], and it is therefore interesting to note that sex differences appear to be mediated via mechanisms other than DNA methylation. Apparently, the effects of sex observed in methylome studies predicting age are small global effects rather than large changes at a limited number of sites. We have previously reported [35] that there are sex-specific differences in the gene expression changes associated with ageing, but based on the results of the present study, these expression differences are not regulated by DNA methylation.

The role of zinc-associated proteins in ageing

We observed a clear enrichment of hypermethylation on chromosome 19, which seems to be due to the abundance of zinc finger proteins on this chromosome. The increased methylation of zinc finger genes on chromosome 19 has previously been observed in oropharyngeal squamous cell carcinoma [43], and similarities between the methylation changes that occur in ageing and cancer have been demonstrated in multiple studies [20,21,23]. It has recently been proposed that the zinc finger proteins on chromosome 19 have specifically evolved to repress endogenous retroviruses (ERVs) [44]. On the other hand, the expression of ERVs has been associated with ageing in mice [45,46]. Hence, the hypermethylation of zinc finger genes observed with ageing offers an explanation for why ERVs are able to be expressed with advancing age. One of the zinc finger genes predicted to repress ERVs by Lukic et al. [44] was *ZNF154*. We identified 10 CpGs within this gene as being hypermethylated in the nonagenarians, and we identified a strong negative correlation between the level of methylation and the expression of this gene, indicating that its expression is truly downregulated in the aged individuals. Both ageing and cancer are associated with genomic instability [1], and the role of active ERVs in inducing this genomic instability with increasing age could be analogous to that proposed in cancer [47].

Zinc-coordinating transcription factors were also enriched among the TFs predicted to regulate hypermethylation-associated genes in this study, as 12 out of the 24 identified TFs were zinc coordinating. Zinc has been associated with various processes that are known to be regulated during ageing, such as immune function, DNA repair mechanisms, cell proliferation, apoptosis and transcription [48,49].

The association between ageing-associated DNA methylation changes and gene expression

We sought to examine the relationship between ageing-associated DNA methylation changes and gene expression levels. Compared with previous studies, a key asset of our study is that methylation and gene expression data were available from the same samples. Those ageing-associated methylation sites in which the level of methylation is associated with the level of gene expression are concentrated in non-CGIs and on shores and shelves, as well as in gene body regions. Similar findings have been reported by Zilbauer et al. [40]. Gene-body methylation has been proposed to affect gene expression via splicing and alternative start site usage [13,50]. It is important to note that many previous studies examining DNA methylation changes during ageing have been performed using the Illumina 27K array, where the majority of the probes are located in promoter regions. In these studies, the effects of gene-body methylation on gene expression levels remained unidentified.

The identified genes that display expression-methylation correlations are strongly enriched in immunological processes and in cytoskeletal remodelling and endocytosis. Cytoskeletal remodelling is required for leukocyte activation, migration and phagocytosis [51]. The results imply that some fraction of ageing-associated immune system changes may be regulated by DNA methylation. Defects in the immune system are a hallmark of ageing, leading to increased susceptibility to infectious diseases, cancer and ultimately death [1]. DNA methylation typically regulates long-term trends in gene expression [11,13], and the possibility that immune system-related processes may be locked in a particular state by DNA methylation could

offer one explanation as to why the immune system of elderly individuals is not able to respond appropriately to various insults.

We found that only a minority of ageing-associated CpG sites showed an association between methylation and expression levels. Furthermore, only a minority of these genes have been identified as differentially expressed between nonagenarians and young individuals [35]. Previous studies have also found a limited number of associations between ageing-associated DNA methylation changes and gene expression levels [21,23,27,36,42,52]. Due to the methods applied in the present study, not all the effects of DNA methylation on gene expression could be detected; this limitation is also true for previously reported results. The textbook case of DNA methylation regulating gene expression (the methylation of a promoter and silencing of a gene) remains undetected in many cases because in an array analysis, an unexpressed gene shows no signal that can be distinguished from background and is therefore typically omitted from the analysis. Additionally, in the present study, the methylation level of each CpG was correlated separately with gene expression. In CpG island regions in particular, the effect of DNA methylation changes on gene expression could be observed when a cluster of closely located CpG sites were analysed as a whole. The effects of CpG sites that are not located in the regulated gene itself also remain unidentified. The short list of methylation-gene expression associations linked to ageing reported herein and previously by others should be interpreted as a defined set of one type of methylation-gene expression associations, and it should be assumed that other types of mechanisms exist and require different methodologies to be identified.

Conclusions

Based on the results presented here, it appears that ageing-associated hyper- and hypomethylation are distinct processes, both in terms of their causes and consequences. We suggest that hypermethylation is an active process, caused by programmed or pseudo-programmed ageing processes, and that hypermethylation is strongly associated with chronological age. Ageing-associated hypomethylation, however, is a passive process caused by stochastic or environmental effects and is associated with biological age, i.e., the phenotype of the individual. Whether the underlying cause of ageing is programmed, pseudo-programmed or due to the accumulation of molecular damage has been widely discussed in the literature. Given that evidence supporting each theory can be found, it is plausible that these mechanisms all contribute to the ageing process but possibly affect different aspects [3-5].

First, hypermethylation is an active process that consumes energy as new methyl groups are added to DNA by DNA methyltransferases. Hypomethylation can also be an active process in some cases, but contrary to hypermethylation, it may occur passively as well [53,54]. The most frequently reported ageing-associated DNA methylation changes (for example in *ELOVL2*) that are repeated across tissues and study populations, thus implying programmed changes, are hypermethylation events. In studies where chronological age has been explained in association with DNA methylation levels, it has been found that at sites showing the strongest correlation with chronological age, methylation increases with age [25,26]. The ageing-associated hypermethylated sites form common groups with regard to cellular processes and functions. According to the results of the present study, hypermethylation-associated genes are predicted to be regulated by a common group of transcription factors and are also enriched in common GO terms, whereas hypomethylation-associated genes do not to appear to form common groups. The top-ranking ageing-associated sites are hypermethylated, but hypomethylated sites are more numerous. This difference becomes more significant when the threshold of significance is lowered; of the 8540 sites identified here, 54% were hypomethylated, but among the 45507 sites identified with the regression model, 64% were hypomethylated. Johansson et al. [36] also reported an excess of hypomethylation over hypermethylation with ageing.

Global hypomethylation has been associated with an increasing risk of frailty [55], but very few other associations between phenotype and DNA methylation have been reported [17]. However, this may be due to technical concerns, as the study by Bell et al. [17] was performed with the 27K array, which almost exclusively contains promoter-associated probes that are not methylated at baseline and can therefore primarily acquire hypermethylation. Phenotype association studies performed with the 450K array or using sequencing techniques are necessary to clarify if hypomethylation is associated with typical ageing-associated phenotypes.

The role of DNA methylation is known to differ depending on its location in the genome. Thus, it would not be surprising if different DNA methylation changes in the genome are affected by different ageing mechanisms. As DNA methylation analyses are complicated by the different effects of methylation sites at different genomic positions and by the cumulative effects of nearby CpG sites, all possible known biases, such as the proportions of leukocytes, should be accounted for in DNA methylation analyses.

Methods

Study population

The study population consisted of 146 nonagenarians (females n = 103, males n = 43) participating in the Vitality 90+ study and 30 young, healthy controls (aged 19-30 years, median 22.5 years; females n = 21, males

n = 9). Gene expression data were available for all the individuals, and methylation data were available for 122 nonagenarians (n = 89 females and n = 33 males) and 21 young controls (n = 14 females and n = 7 males), and data on cell proportions were available for 115 nonagenarians (n = 84 females and n = 31 males) and all 30 of the young controls. All the study subjects were of Western European descent. The Vitality 90+ study is an ongoing prospective population-based study that includes both home-dwelling and institutionalised individuals aged 90 years or more who live in the city of Tampere, Finland. The recruitment and characterisation of the participants were performed as previously reported for earlier Vitality 90+ study cohorts [56]. In this study, we included only individuals born in 1920, and the evaluated samples were collected in the year 2010. The nonagenarians included in the study had not had any infections or received any vaccinations in the 30 days prior to blood sample collection. The young controls consisted of healthy laboratory personnel who did not have any medically diagnosed chronic illnesses, were non-smokers and had not had any infections or received any vaccinations within the two weeks prior to blood sample collection. The study participants provided their written informed consent. The study has been conducted according to the principles expressed in the declaration of Helsinki, and the study protocol was approved by the ethics committee of the city of Tampere (1592/403/1996).

Sample collection

The blood samples were collected into EDTA-containing tubes by a trained medical student during a home visit. All the blood samples were drawn between 8 am and 12 am. The samples were directly subjected to leucocyte separation on a Ficoll-Paque density gradient (Ficoll-Paque™ Premium, cat. no. 17-5442-03, GE Healthcare Bio-Sciences AB, Uppsala, Sweden). The PBMC layer was collected, and a subset of the cells was suspended in 150 μl of RNAlater solution (Ambion Inc., Austin, TX, USA) for use in a gene expression microarray analysis. Cells that were to be subjected to FACS analysis and DNA extraction were suspended in 1 ml of a freezing solution (5/8 FBS, 2/8 RPMI-160 medium, 1/8 DMSO) (FBS cat. no. F7524, Sigma-Aldrich, MO, USA; RPMI: cat. no. R0883, Sigma-Aldrich, MO, USA; DMSO: cat. no. 1.02931.0500, VWR, Espoo, Finland).

DNA extraction

DNA was extracted from PBMCs using the QIAamp DNA Mini kit (Qiagen, CA, USA), following the manufacturer's instructions for the spin protocol. The DNA was eluted in 60 μl of AE elution buffer and stored at -20°C. The concentration and quality of the DNA was assessed with the Qubit dsDNA HS Assay (Invitrogen, Eugene, OR, USA).

RNA extraction

For RNA extraction, equal amounts of PBS and RNAlater were added to the cell suspension and then removed via centrifugation, leaving only the cell pellet. RNA was purified using an miRNeasy mini kit (Qiagen, CA, USA), according to the manufacturer's protocol, with on-column DNase digestion (AppliChem GmbH, Darmstadt, Germany). The concentration and quality of the RNA were assessed with the Agilent RNA 6000 Nano Kit on an Agilent 2100 Bioanalyzer (Agilent Technologies, CA, USA).

FACS

The proportions of different lymphocyte populations were determined through FACS analysis (BD FACSCanto II), and the results were analysed with BD FACS Diva, version 6.1.3 (BD Biosciences, Franklin Lakes, NJ, USA). The antibodies employed in this analysis were FITC-CD14 (cat. no. 11-0149), PerCP-Cy5.5-CD3 (45-0037), APC-CD28 (17-0289) (eBioscience, San Diego, CA, USA), PE-Cy™7-CD4 (cat. no. 557852) and APC-Cy™7-CD8 (557834) (BD Biosciences).

Expression array

Labelled cRNA was prepared from 330 ng of total RNA using the Illumina TotalPrep RNA Amplification Kit (Ambion Inc., TX, USA) with overnight incubation according to the manufacturer's protocol. The quality of the labelled cRNA was determined using a 2100 Bioanalyzer (Agilent Technologies). In total, 1500 ng of labelled cRNA was hybridised overnight to a HumanHT-12 v4 Expression BeadChip (Cat no. BD-103-0204, Illumina Inc., CA, USA), according to the Illumina protocol, in the Core Facility of the Department of Biotechnology of the University of Tartu. Samples were assigned to the arrays in a randomised order. The chips were scanned using Beadscan (Illumina Inc.).

Methylation array

Genome-wide DNA methylation profiling was performed at the Institute for Molecular Medicine Finland (FIMM) Technology Centre of the University of Helsinki in two batches (time interval, 6 months). Bisulfite conversion of 1 μg of DNA was performed using the EZ-96 DNA Methylation Kit (Zymo Research, Irvine, CA, USA) according to manufacturer's instructions. A 4-μl aliquot of bisulphite-converted DNA was subjected to whole-genome amplification and then enzymatically fragmented and hybridised to the Infinium HumanMethylation450 BeadChip (Illumina, San Diego, CA, USA) according to manufacturer's protocol. Samples were assigned to the arrays in a randomised order. The BeadChips were scanned with the iScan reader (Illumina).

Preprocessing of the methylation microarray data

The methylation data were preprocessed as a *methylumiset* object using R software with the *wateRmelon* array-specific package from Bioconductor [57]. The annotation information was based on the GRCh37/hg19 genome assembly from February 2009. Prior to any processing, all unspecific or polymorphic sites (n = 76775) were removed based on database information [58]. Samples and target sites of a technically poor quality were filtered out by excluding sites with a beadcount of <3 in 5% of the samples (n = 526) and sites for which 1% of the samples showed a detection p-value >0.05 (n = 740). Background correction and quantile normalisation via the *dasen* method were conducted individually for the two applied chemistries (Infinium I and II) as well as for the intensities of methylation (m) and un-methylation (u). After *dasen* treatment, the u and m intensities were transformed to beta (β) and M values. β is the ratio of the methylated probe (m) intensities to the overall intensities (m + u + α), where α is the constant offset, 100. Thus, β ranges linearly from 0 (non-methylated, 0%) to 1 (completely methylated, 100%). The β values were further transformed into M values using the equation log2(β/(1- β)). Next, the batch effect of the chemistries was adjusted using the BMIQ method, which is based on beta mixture models and the EM algorithm [59]. Several visualisation styles were used to verify the quality of the preprocessed data, such as boxplots from the raw intensities, Kernel density plots in the chemistry correction procedure and PCA plots (see Additional file 13). The batch effect of two laboratory days (time interval of 6 months) was confirmed via PCA (PC2 6.8%) to be a cause of severe bias in the data. Thus, the bias was corrected using an algorithm based on Empirical Bayes methods, as implemented in the R package *Combat* [60].

Preprocessing of the gene expression microarray data

The gene expression microarray data were preprocessed as a *Lumibatch* object with the *lumi* pipeline using R software [61]. Background correction was performed with the *bgAdjust.affy* package. The gene expression values were then transformed with vst and normalised using the *rsn* method. Transcripts with transformed expression values of greater than 7.5 in 20% of the samples were included in the analysis. Visualisations, boxplots and PCA plots were used in the pipeline to verify the quality of the data.

Comparison of age groups

To detect CpG sites showing substantial differences in DNA methylation between nonagenarians and young adults, the sites displaying the largest difference in the absolute value of the methylation level were included in the analysis (-1 > ΔM > 1, threshold for ΔM based on [61]). The rank-sums of the methylation values of the two groups were further compared with the Wilcoxon rank-sum test, and the nominal Benjamini-Hochberg-adjusted p-value was set to 0.05.

Multiple regression analyses

To assess the relationship between age- and site-specific methylation levels in greater detail, a generalised regression model referred to as variable dispersion beta regression was utilised in an iterative manner (n = 407 646). Age group was employed as a predictor of the site-specific methylation outcome, in the form of β values (ranging from 0 to 1), in each equation of the mean model with a linker function of *logit*. Furthermore, as it was observed through PCA that the DNA methylation levels fluctuated based on the composition of blood cell subtypes, the proportions of CD28-/CD4+ and CD28-/CD8+ cells showed especially clear correlations with principal component 1, which explained 20% of the overall variance in DNA methylation. Therefore, the variables corresponding to cell type proportions (the CD4+ to CD8+ ratio and the proportions of CD28-/CD4+, CD28-/CD8+ and CD14+ cells) were set as adjustments in the analysis to determine leukocyte proportions independent of genome-wide ageing-associated DNA methylation changes. Sex was used as an additional covariate. The regression analyses were performed using R software and with algorithms implemented in the *betareg* package [62,63]. The nominal Bonferroni-adjusted p-value was set to 0.05. See Additional file 14 for a flow chart summary of the analysis steps to identify the high-confidence ageing-associated CpG sites.

Correlations with gene expression levels

The associations between gene expression and DNA methylation levels were separately examined through bivariate correlation (Pearson) analyses for young and old individuals. The correlation analyses were designed for each transcript and CpG site pair showing identical annotation for a gene. Thus, multiple CpG sites were paired with the same gene, and several genes were matched with the same CpG site. In total, 2461 expression-methylation pairs were tested. The nominal Benjamini-Hochberg-adjusted p-value was set to 0.05.

Pathway analyses

All the pathway analyses were performed on differentially methylated genes, i.e., genes that harbour at least one ageing-associated CpG site. There were 1832 hypermethylation-associated genes (3925 CpG sites) and 2057 hypomethylation-associated genes (4615 CpG sites) included in the dataset. Of the hypomethylated CpG sites, 1719 were not associated with any known gene, and of the hypermethylated CpG sites, 720 were not associated with any known gene.

IPA [34] was used to identify canonical pathways associated with our differentially methylated genes. According to the manufacturer, these canonical pathways are well-characterised metabolic and cell signalling pathways that have been curated and hand-drawn by PhD-level scientists. All the data sources provided by the Ingenuity Knowledge Base were included in the IPA, and the Ingenuity Knowledge Base was used as the reference set in all analyses. For the association of molecules, only experimentally observed results were accepted, and only human data were considered. Benjamini-Hochberg multiple testing correction (FDR) was employed to calculate the p-values for the pathways. Canonical pathways were considered significant at $p < 0.05$ ($-\log P > 1.3$) and when the pathway contained a minimum of 3 genes. Pathways associated with cancer and other disease, as defined by Ingenuity Systems®, were excluded from the analysis. The IPA for hyper- and hypomethylation-associated genes was performed on 14.3.2014, and the IPA for genes showing a correlation between methylation and expression levels was performed on 12.3.2014.

GOrilla [31,32] was used to identify the enriched GO terms for the hyper- and hypomethylation-associated genes and for genes showing a correlation between the levels of methylation and expression. GO terms were searched based on two unranked lists (target and background), and all genes with at least one probe in the 450K array were used as the background list. A Bonferroni-corrected p-value of <0.05 was used as the threshold for significance.

PScan [33] can be used to predict if a group of genes is regulated by a common transcription factor. The analysis was performed with the default settings, i.e., using the Jaspar database and the -450 - +50 bp region around the TSS. PScan was able to identify 1811 and 2020 of the total hyper- and hypomethylation-associated transcripts, respectively. This analysis was performed on 11.3.2014. A Bonferroni-corrected p-value of <0.05 was used as a threshold for significance.

Array data

The array data are available in the GEO database (http://www.ncbi.nlm.nih.gov/geo/) under the accession numbers GSE40366 for the gene expression data and GSE58888 for methylation data.

Additional files

Additional file 1: All 8540 ageing-associated CpG sites.

Additional file 2: Distribution of ageing-associated methylation sites across chromosomes.

Additional file 3: Enriched GO function terms. A table of GO function terms associated with identified ageing-associated methylation sites. GO function terms associated with hypermethylated sites in sheet a and GO function terms associated with hypomethylation in sheet b.

Additional file 4: Diagram of enriched GO function terms. A visualisation of Additional file 3.

Additional file 5: Enriched GO process terms. A table of GO process terms associated with identified ageing-associated methylation sites. GO process terms associated with hypermethylated sites in sheet a and GO process terms associated with hypomethylation in sheet b.

Additional file 6: Diagram of enriched GO process terms. A visualisation of Additional file 5.

Additional file 7: Transcription factors predicted to regulate hypermethylation-associated genes.

Additional file 8: Ingenuity canonical pathways associated with differentially methylated genes. Canonical pathways associated with hypermethylation are in sheet a and canonical pathways associated with hypomethylation are in sheet b.

Additional file 9: Sex- and age-group -associated methylation sites.

Additional file 10: CpG sites showing a correlation to the expression level of the corresponding gene.

Additional file 11: Genomic locations of the CpG sites where the level of DNA methylation correlates with the expression level of the corresponding gene.

Additional file 12: GO process terms associated with genes whose expression levels were correlating with the level of DNA methylation.

Additional file 13: PCA plots for quality control of methylation data.

Additional file 14: Flow chart of the analysis steps used to identify ageing-associated methylation sites.

Abbreviations

27K array: Infinium HumanMethylation27 BeadChip; 450K array: Infinium HumanMethylation450 BeadChip; Betareg: Beta regression model; ERV: Endogenous retrovirus; Non-CGI: Non-CpG island; PBMC: Peripheral blood mononuclear cell; TF: Transcription factor; TSS: Transcription start site; UTR: Untranslated region.

Competing interests

The authors declare they have no competing interests.

Authors' contributions

SM performed the experiments and statistical analyses and was primarily responsible for writing the manuscript. LK and SH processed the data and performed statistical analyses. JJ and TN performed the experiments. AH and MJ were responsible for recruiting the study population. MN performed statistical analyses. MH provided reagents and materials for the study. SM, LK, JJ an MH contributed to the design of the study. All authors contributed to the writing of the manuscript and read and approved the final manuscript.

Acknowledgements

We would like to thank Sinikka Repo-Koskinen, Janette Hinkka, Katri Välimaa and Sanna Tuominen for their skillful technical assistance.

Author details

[1]Department of Microbiology and Immunology, School of Medicine, University of Tampere, Tampere, Finland. [2]Gerontology Research Center, Tampere, Finland. [3]Institute of Biosciences and Medical Technology, University of Tampere, Tampere, Finland. [4]School of Health Sciences, University of Tampere, Tampere, Finland. [5]Fimlab Laboratories, Tampere, Finland.

Received: 11 September 2014 Accepted: 21 February 2015
Published online: 14 March 2015

References

1. López-Otín C, Blasco MA, Partridge L, Serrano M, Kroemer G. The hallmarks of aging. Cell. 2013;153:1194–217.

2. Arnold CR, Wolf J, Brunner S, Herndler-Brandstetter D, Grubeck-Loebenstein B. Gain and loss of T cell subsets in old age – age-related reshaping of the T cell repertoire. J Clin Immunol. 2011;31:137–46.
3. Blagosklonny MV. Aging is not programmed: genetic pseudo-program is a shadow of developmental growth. Cell Cycle. 2013;12:3736–42.
4. de Magalhães JP. Programmatic features of aging originating in development: aging mechanisms beyond molecular damage? FASEB J. 2012;26:4821–6.
5. Kirkwood TB, Melov S. On the programmed/non-programmed nature of ageing within the life history. Curr Biol. 2011;21:R701–7.
6. Bacalini MG, Friso S, Olivieri F, Pirazzini C, Giuliani C, Capri M, et al. Present and future of anti-ageing epigenetic diets. Mech Ageing Dev. 2014;136:101–15.
7. Johnson AA, Akman K, Calimport SR, Wuttke D, Stolzing A, de Magalhães JP. The role of DNA methylation in aging, rejuvenation, and age-related disease. Rejuvenation Res. 2012;15:483–94.
8. Brooks-Wilson AR. Genetics of healthy aging and longevity. Hum Genet. 2013;132:1323–38.
9. Riggs AD. X inactivation, differentiation, and DNA methylation. Cytogenet Cell Genet. 1975;14:9–25.
10. Holliday R, Pugh JE. DNA modification mechanisms and gene activity during development. Science. 1975;187:226–32.
11. Deaton AM, Bird A. CpG islands and the regulation of transcription. Genes Dev. 2011;25:1010–22.
12. Illingworth RS, Bird AP. CpG islands–'a rough guide'. FEBS Lett. 2009;583:1713–20.
13. Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nat Rev Genet. 2012;13:484–92.
14. Bogdanović O, Veenstra GJ. DNA methylation and methyl-CpG binding proteins: developmental requirements and function. Chromosoma. 2009;118:549–65.
15. Cedar H, Bergman Y. Programming of DNA methylation patterns. Annu Rev Biochem. 2012;81:97–117.
16. Christensen BC, Houseman EA, Marsit CJ, Zheng S, Wrensch MR, Wiemels JL, et al. Aging and environmental exposures alter tissue-specific DNA methylation dependent upon CpG island context. PLoS Genet. 2009;5:e1000602.
17. Bell JT, Tsai PC, Yang TP, Pidsley R, Nisbet J, Glass D, et al. Epigenome-wide scans identify differentially methylated regions for age and age-related phenotypes in a healthy ageing population. PLoS Gene. 2012;8:e1002629.
18. Gentilini D, Mari D, Castaldi D, Remondini D, Ogliari G, Ostan R, et al. Role of epigenetics in human aging and longevity: genome-wide DNA methylation profile in centenarians and centenarians' offspring. Age (Dordr). 2013;35:1961–73.
19. Rakyan VK, Down TA, Maslau S, Andrew T, Yang TP, Beyan H, et al. Human aging-associated DNA hypermethylation occurs preferentially at bivalent chromatin domains. Genome Res. 2010;20:434–9.
20. Teschendorff AE, Menon U, Gentry-Maharaj A, Ramus SJ, Weisenberger DJ, Shen H, et al. Age-dependent DNA methylation of genes that are suppressed in stem cells is a hallmark of cancer. Genome Res. 2010;20:440–6.
21. Xu Z, Taylor JA. Genome-wide age-related DNA methylation changes in blood and other tissues relate to histone modification, expression and cancer. Carcinogenesis. 2014;35:356–64.
22. Weidner CI, Lin Q, Koch CM, Eisele L, Beier F, Ziegler P, et al. Aging of blood can be tracked by DNA methylation changes at just three CpG sites. Genome Biol. 2014;15:R24.
23. Horvath S. DNA methylation age of human tissues and cell types. Genome Biol. 2013;14:R115.
24. Steegenga WT, Boekschoten MV, Lute C, Hooiveld GJ, de Groot PJ, Morris TJ, et al. Genome-wide age-related changes in DNA methylation and gene expression in human PBMCs. Age (Dordr). 2014;36:9648.
25. Florath I, Butterbach K, Muller H, Bewerunge-Hudler M, Brenner H. Cross-sectional and longitudinal changes in DNA methylation with age: an epigenome-wide analysis revealing over 60 novel age-associated CpG sites. Hum Mol Genet. 2013;23:1186–201.
26. Garagnani P, Bacalini MG, Pirazzini C, Gori D, Giuliani C, Mari D, et al. Methylation of ELOVL2 gene as a new epigenetic marker of age. Aging Cell. 2012;11:1132–4.
27. Hannum G, Guinney J, Zhao L, Zhang L, Hughes G, Sadda S, et al. Genome-wide methylation profiles reveal quantitative views of human aging rates. Mol Cell. 2012;49:359–67.
28. Heyn H, Li N, Ferreira HJ, Moran S, Pisano DG, Gomez A, et al. Distinct DNA methylomes of newborns and centenarians. Proc Natl Acad Sci U S A. 2012;109:10522–7.
29. Jaffe AE, Irizarry RA. Accounting for cellular heterogeneity is critical in epigenome-wide association studies. Genome Biol. 2014;15:R31.
30. Marttila S, Jylhävä J, Pesu M, Hämäläinen S, Jylhä M, Hervonen A, et al. IL-7 concentration is increased in nonagenarians but is not associated with markers of T cell immunosenescence. Exp Gerontol. 2011;46:1000–2.
31. Eden E, Lipson D, Yogev S, Yakhini Z. Discovering motifs in ranked lists of DNA sequences. PLoS Comput Biol. 2007;3:e39.
32. Eden E, Navon R, Steinfeld I, Lipson D, Yakhini Z. GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. BMC Bioinformatics. 2009;10:48.
33. Zambelli F, Pesole G, Pavesi G. Pscan: finding over-represented transcription factor binding site motifs in sequences from co-regulated or co-expressed genes. Nucleic Acids Res. 2009;37:W247–52.
34. Ingenuity Systems®, www.ingenuity.com.
35. Marttila S, Jylhävä J, Nevalainen T, Nykter M, Jylhä M, Hervonen A, et al. Transcriptional analysis reveals gender-specific changes in the aging of the human immune system. PLoS One. 2013;8:e66229.
36. Johansson A, Enroth S, Gyllensten U. Continuous aging of the human DNA methylome throughout the human lifespan. PLoS One. 2013;8:e67378.
37. McClay JL, Aberg KA, Clark SL, Nerella S, Kumar G, Xie LY, et al. A methylome-wide study of aging using massively parallel sequencing of the methyl-CpG-enriched genomic fraction from blood in over 700 subjects. Hum Mol Genet. 2014;23:1175–85.
38. D'Aquila P, Rose G, Bellizzi D, Passarino G. Epigenetics and aging. Maturitas. 2013;74:130–6.
39. Pawelec G, Larbi A, Derhovanessian E. Senescence of the human immune system. J Comp Pathol. 2010;142 Suppl 1:S39–44.
40. Zilbauer M, Rayner TF, Clark C, Coffey AJ, Joyce CJ, Palta P, et al. Genome-wide methylation analyses of primary human leukocyte subsets identifies functionally important cell-type-specific hypomethylated regions. Blood. 2013;122:e52–60.
41. Kondilis-Mangum HD, Wade PA. Epigenetics and the adaptive immune response. Mol Aspects Med. 2013;34:813–25.
42. Lam LL, Emberly E, Fraser HB, Neumann SM, Chen E, Miller GE, et al. Factors underlying variable DNA methylation in a human community cohort. Proc Natl Acad Sci U S A. 2012;109 Suppl 2:17253–60.
43. Lleras RA, Adrien LR, Smith RV, Brown B, Jivraj N, Keller C, et al. Hypermethylation of a cluster of Krüppel-type zinc finger protein genes on chromosome 19q13 in oropharyngeal squamous cell carcinoma. Am J Pathol. 2011;178:1965–74.
44. Lukic S, Nicolas JC, Levine AJ. The diversity of zinc-finger genes on human chromosome 19 provides an evolutionary mechanism for defense against inherited endogenous retroviruses. Cell Death Differ. 2014;21:381–7.
45. Gaubatz JW, Arcement B, Cutler RG. Gene expression of an endogenous retrovirus-like element during murine development and aging. Mech Ageing Dev. 1991;57:71–85.
46. Wada Y, Tsukada M, Kamiyama S, Koizumi A. Retroviral gene expression as a possible biomarker of aging. Int Arch Occup Environ Health. 1993;65(1 Suppl):S235–9.
47. Romanish MT, Cohen CJ, Mager DL. Potential mechanisms of endogenous retroviral-mediated genomic instability in human cancer. Semin Cancer Biol. 2010;20:246–53.
48. Sharif R, Thomas P, Zalewski P, Fenech M. The role of zinc in genomic stability. Mutat Res. 2012;733:111–21.
49. Mocchegiani E, Costarelli L, Giacconi R, Cipriano C, Muti E, Malavolta M. Zinc-binding proteins (metallothionein and alpha-2 macroglobulin) and immunosenescence. Exp Gerontol. 2006;41:1094–107.
50. Maunakea AK, Nagarajan RP, Bilenky M, Ballinger TJ, D'Souza C, Fouse SD, et al. Conserved role of intragenic DNA methylation in regulating alternative promoters. Nature. 2010;466:253–7.
51. Fenteany G, Glogauer M. Cytoskeletal remodeling in leukocyte function. Curr Opin Hematol. 2004;11:15–24.
52. Zykovich A, Hubbard A, Flynn JM, Tarnopolsky M, Fraga MF, Kerksick C, et al. Genome-wide DNA methylation changes with age in disease-free human skeletal muscle. Aging Cell. 2014;13:360–6.
53. Bhutani N, Burns DM, Blau HM. DNA demethylation dynamics. Cell. 2011;146:866–72.
54. Jones PA, Liang G. Rethinking how DNA methylation patterns are maintained. Nat Rev Genet. 2009;10:805–11.

55. Bellizzi D, D'Aquila P, Montesanto A, Corsonello A, Mari V, Mazzei B, et al. Global DNA methylation in old subjects is correlated with frailty. Age (Dordr). 2012;34:169–79.
56. Goebeler S, Jylhä M, Hervonen A. Medical history, cognitive status and mobility at the age of 90. A population based study in Tampere, Finland. Aging Clin Exp Res. 2003;15:154–61.
57. Pidsley R, Wong CC Y, Volta M, Lunnon K, Mill J, Schalkwyk LC. A data-driven approach to preprocessing Illumina 450K methylation array data. BMC Genomics. 2013;14:293.
58. Chen YA, Lemire M, Choufani S, Butcher DT, Grafodatskaya D, Zanke BW, et al. Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. Epigenetics. 2013;8:203–9.
59. Teschendorff AE, Marabita F, Lechner M, Bartlett T, Tegner J, Gomez-Cabrero D, et al. A beta-mixture quantile normalization method for correcting probe design bias in Illumina Infinium 450 k DNA methylation data. Bioinformatics. 2013;29:189–96.
60. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics. 2007;8:118–27.
61. Du P, Kibbe WA, Lin SM. lumi: a pipeline for processing Illumina microarray. Bioinformatics. 2008;24:1547–8.
62. Cribari-Neto F, Zeileis A. Beta regression in R. JSS. 2010;34:1–24.
63. Ferrari C. Beta regression for modelling rates and proportions. J Appl Stat. 2004;31:799–15.

Kananen et al. BMC Genomics (2016) 17:103
DOI 10.1186/s12864-016-2421-z

BMC Genomics

RESEARCH ARTICLE Open Access

Aging-associated DNA methylation changes in middle-aged individuals: the Young Finns study

L. Kananen[1,2*], S. Marttila[1,2], T. Nevalainen[1,2], J. Jylhävä[1,2], N. Mononen[3], M. Kähönen[4], O. T. Raitakari[5], T. Lehtimäki[3,6] and M. Hurme[1,2,3]

Abstract

Background: Chronological aging-associated changes in the human DNA methylome have been studied by multiple epigenome-wide association studies (EWASs). Certain CpG sites have been identified as aging-associated in multiple studies, and the majority of the sites identified in various studies show common features regarding location and direction of the methylation change. However, as a whole, the sets of aging-associated CpGs identified in different studies, even with similar tissues and age ranges, show only limited overlap. In this study, we further explore and characterize CpG sites that show close relationship between their DNA methylation level and chronological age during adulthood and which bear the relationship regardless of blood cell type heterogeneity.

Results: In this study, with a multivariable regression model adjusted for cell type heterogeneity, we identified 1202 aging-associated CpG sites (a-CpGs, FDR < 5 %), in whole blood in a population with an especially narrow age range (40 - 49 years). Repeatedly reported a-CpGs located in genes *ELOVL2*, *FHL2*, *PENK* and *KLF14* were also identified. Regions with aging-associated hypermethylation were enriched regarding several gene ontology (GO) terms (especially in the cluster of developmental processes), whereas hypomethylated sites showed no enrichment. The genes with higher numbers of a-CpG hits were more often hypermethylated with advancing age. The comparison analysis revealed that of the 1202 a-CpGs identified in the present study, 987 were identified as differentially methylated also between nonagenarians and young adults in a previous study (The Vitality 90+ study), and importantly, the directions of changes were identical in the previous and in the present study.

Conclusions: Here we report that aging-associated DNA methylation features can be identified in a middle-aged population with an age range of only 9 years. A great majority of these sites have been previously reported as aging-associated in a population aged 19 to 90 years. Aging is associated with different types of changes in DNA methylation, clock-like as well as random. We speculate that the a-CpGs identified here in a population with a narrow age-range represent clock-like changes, as they showed concordant methylation behavior in population spanning whole adulthood as well.

Keywords: Aging-associated, DNA methylation, EWAS, CpG sites, Adulthood, Hypermethylation, Blood cell type heterogeneity

* Correspondence: laura.kananen@uta.fi
[1]Department of Microbiology and Immunology, School of Medicine, University of Tampere, Tampere, Finland
[2]Gerontology Research Center, Tampere, Finland
Full list of author information is available at the end of the article

Background

The epigenome includes DNA methylation (DNAmet), post-translational histone modifications and chromatin remodeling. Tens of millions of nucleotides referred to as CpG sites, which are prone to DNAmet, exist in the haploid human genome. Furthermore, the genome-wide DNAmet profile is maintained through cell divisions. DNA methyltransferases apply methyl groups on CpG sites to form 5-methylcytosine, whereas demethylation may occur either passively due to dysfunction of the transferring enzyme or actively through 5-hydromethylcytosine formation. Genomic regions spanning approximately 0.5 kilobases with a high density of CpG sites are called CpG islands, and these are commonly localized near transcription start sites. CpG sites in such islands are often less methylated; thus, the genes are available for initiation of transcription. Moreover, DNAmet plays crucial roles in gene expression by not only blocking the promoter region but also altering the activities of regulatory elements, such as enhancers and insulators. Alternatively, gene body methylation may influence alternative splicing [1, 2]. Thus, the cell identity is in part determined and maintained by a cell type-specific genome-wide methylation pattern, which may therefore be used in the laboratory as a marker to characterize the cell types [3–5].

The genome-wide DNAmet profile of the cell changes; DNAmet patterns are altered in diseases, such as Alzheimer disease, cancer and type 2 diabetes, and are also influenced by the accumulating effects of environmental factors such as toxin exposure and diet [1, 6, 7]. Single CpG sites undergo hypo- and hypermethylation either randomly by stochastic factors or via more systematic mechanisms [1]. For example, exposure to environmental factors such as smoking induces hypomethylation of a well-characterized single CpG site in the gene *F2RL3;* this represents an example of a non-random change in DNAmet because the magnitude of the change is dose and exposure-time dependent [8, 9].

Furthermore, the epigenome is modified by the biological aging process. As also Heyn et al. [10] reported and Zampieri et al. [1] reviewed, in general, aging induces a decrease in average DNA methylation level genome-wide (global hypomethylation). This was demonstrated by whole-genome bisulfite sequencing of newborns and centenarians with as high as ~90 % genomic coverage. The comparison of methylation states between the two extremes of the human lifespan also revealed how the systematic methylation patterns of the CpG sites are eventually lost and how inter-individual differences increase with advanced age. In addition, hypermethylation in regions near promoters can cause down-regulation of essential genes that influence vitally important pathways; Heyn et al. [10] reported that aging-accelerated hypermethylation events occurred in 13 % of the CpG sites among the millions of sites in the genome. Therefore, methylation alterations may be considered as one important factor in the development of aging-associated diseases [1, 10].

Many studies have addressed the aging-associated DNAmet changes in blood cells using Illumina array technology-based methods, which cover 27000 or 485000 CpG sites in the genome [1]. The methylation levels of specific CpG sites are known to be associated with chronological aging in a wide variety of tissues [11–13]. However, as a whole, the sets of aging-associated CpGs identified in different studies, even with comparable tissues and age ranges, show limited overlap. Only few EWASs on age have taken the cell type heterogeneity into account [14–17]. We and others [4] hypothesize that lack of cell type adjustment may have potentially distorted the results obtained, and this may have contributed to the lack of concordance observed between the studies.

In this study, we aimed to discover and characterize regions where the DNAmet levels are associated with chronological age (a-CpGs) in a middle-aged population (aged 40–49 years) through analysis where the cell type heterogeneity was adjusted for. Middle-aged individuals were selected from the Young Finns Study (YFS) [18] follow-up in 2011; the selection in the present study is a balanced sample (i.e. the number of subjects in each age group was equal and the groups had similar sex-distribution), and it therefore provides an excellent opportunity to inspect the effects of aging on DNA methylome. Furthermore, this sample comprises individuals in an extremely narrow age range of only nine years. The subjects' DNA methylomes were characterized using Illumina Infinium HumanMethylation450 BeadChips and the cell type heterogeneity and sex were adjusted for in the analysis.

Additionally, our findings were interpreted together with compatible data obtained using the same 450BeadChip technology, including our previous results obtained from an EWAS on age (The Vitality 90 + Study, V90+), in which the subjects' ages ranged from 19 to 90 years [15], as well as other results compiled by Steegenga et al. [19]. The results from the YFS were interpreted by considering that rates of aging-associated DNAmet changes fluctuate, especially during the growth period before adulthood and at the end of the lifespan [11, 20]. Accordingly, the a-CpGs found in the YFS that overlap with those established from adult samples with wider age ranges, such as V90+ study, may be speculated to be DNAmet regions with constant rate of change throughout adulthood. Thus, we aimed to explore the a-CpGs where level of methylation changes in a clocklike fashion throughout adulthood from those that show a more random aging-associated pattern.

Results

Aging-associated alterations in DNA methylation

In this study, the genome-wide DNAmet levels in whole blood samples of middle-aged individuals were measured using 450BeadChip technology. The sample heterogeneity (i.e., the proportions of CD8T and CD4T cells, monocytes, granulocytes, and NK and B cells) were estimated by comparing DNAmet profiles to the reference dataset [4] (Additional file 1: Figure S1). The cell type proportions were verified as important determinants of variation in DNAmet using Spearman's correlation analysis, in which the cell type proportions were correlated with the main principal components (PCs). The PCs were defined with principal component analysis (PCA) from the DNAmet data without cell subtype adjustment (Additional file 2: Table S1a). The analysis revealed that PC1 to PC6 together explained a large proportion (24 %) of the variance in the DNA methylome data. Among those PCs, several PCs had considerable large ($-0.5 > r > 0.5$) correlation coefficients; thus, adjustments for the cell type proportion in the analysis were mandatory. The hypothesis whether DNAmet level of a CpG site is associated with chronological age was tested at each CpG site using generalized linear regression analysis ('beta regression'), where sex and cell type proportions were adjusted for.

We found 1202 a-CpGs (i.e. CpG sites where age was a statistically significant variable in the multivariable regression model, FDR < 5 %) in middle-aged individuals (aged 40–49 years), of which 622 (52 %) were hypomethylated and 580 (48 %) were hypermethylated with advancing age. These hypo- and hypermethylated sites were annotated on 440 and 437 genes, respectively. Lists of the most significant aging-associations in YFS are shown in Tables 1 and 2 and in Additional file 3: Table S4. Frequently reported CpG sites (summarized by Steegenga et al. [19]) located in the *ELOVL2* (cg16867657, cg24724428 and cg21572722), three sites in the *FHL2* (cg06639320, cg22454769 and cg24079702), two sites in the *PENK* (cg16219603, cg16419235), and two sites in the *KLF14* (cg08097417, cg09499629 and cg07955995) were also identified as hypermethylated in the present study.

Interestingly, similar to correlation analysis results shown in Additional file 2: Table S1a, the cell type proportions were important determinants of variation in

Table 1 The top 20 hypermethylated a-CpGs in middle-aged individuals. The hypermethylated and hypomethylated a-CpGs are shown separately in Tables 1 and 2, respectively. The top-ranking hypermethylated a-CpGs were selected with the following criteria: 1) direction of the association based on the value of beta regression (denoted as 'betareg') estimate of age; 2) more than one hit identified per gene (q-value < 0.05 which corresponds to false discovery rate <5 %) and 3) the top-ranking *p*-values. The full list of a-CpGs is shown in Additional file 3: Table S4. The q-value denotes the Benjamini-Hochberg-corrected *p*-value

ProbeID	Gene name	CHR	Coordinate	Betareg estimate of age	q-value
cg16867657	*ELOVL2*	6	11152863	0.022	0.00E + 00
cg24724428	*ELOVL2*	6	11152874	0.021	4.80E-07
cg21572722	*ELOVL2*	6	11152880	0.013	3.46E-06
cg06639320	*FHL2*	2	105382171	0.018	3.46E-06
cg00059225	*GLRA1*	5	151284550	0.013	5.13E-06
cg08097417	*KLF14*	7	130069673	0.020	1.87E-05
cg22454769	*FHL2*	2	105382199	0.021	5.03E-05
cg07553761	*TRIM59*	3	161650671	0.016	6.12E-05
cg01588592	*ETV3L*	1	155335949	0.011	1.14E-04
cg11176990	*LOC375196*	2	39041037	0.014	1.54E-03
cg09499629	*KLF14*	7	130069676	0.018	1.54E-03
cg22158769	*LOC375196*	2	39041043	0.020	2.43E-03
cg18898125	*NEFM*	8	24826286	0.012	2.49E-03
cg21911021	*ZIK1*	19	62786823	0.020	3.07E-03
cg27217742	*RGS12*	4	3335078	0.013	3.07E-03
cg17737681	*DLX1*	2	172660382	0.015	3.29E-03
cg24079702	*FHL2*	2	105382203	0.015	5.99E-03
cg16219603	*PENK*	8	57523140	0.013	7.00E-03
cg23930856	*TFAP2B*	6	50919683	0.013	7.22E-03
cg11152943	*TRAPPC9*	8	141318170	0.013	7.57E-03

Table 2 The top 20 hypomethylated a-CpGs in middle-aged individuals. The hypermethylated and hypomethylated a-CpGs are shown separately in Tables 1 and 2, respectively. The top-ranking hypomethylated a-CpGs were selected with the following criteria: 1) direction of the association based on the value of beta regression (denoted as 'betareg') estimate of age; 2) more than one hit identified per gene (q-value < 0.05 which corresponds to false discovery rate < 5 %) and 3) the top-ranking *p*-values. The full list of a-CpGs is shown in Additional file 3: Table S4. The q-value denotes the Benjamini-Hochberg-corrected *p*-value

ProbeID	Gene name	CHR	Coordinate	Betareg estimate of age	q-value
cg00791074	*MTHFD1L*	6	151227862	-0.018	7.51E-04
cg18618815	*COL1A1*	17	45630323	-0.018	5.99E-03
cg14169886	*PRDM16*	1	3101709	-0.014	5.99E-03
cg01820374	*LAG3*	12	6752344	-0.014	9.24E-03
cg19421125	*LAG3*	12	6753117	-0.022	1.02E-02
cg14829066	*NTRK3*	15	86360145	-0.013	1.49E-02
cg03290281	*C6orf195*	6	2577606	-0.021	1.49E-02
cg05561193	*DCLK2*	4	151218492	-0.017	1.96E-02
cg20249566	*NWD1*	19	16691739	-0.024	1.97E-02
cg23928726	*PEX10*	1	2334858	-0.014	1.97E-02
cg20007894	*SCAND3*	6	28648421	-0.019	2.08E-02
cg16355231	*PEX10*	1	2334839	-0.019	2.14E-02
cg15058210	*HDAC4*	2	239861814	-0.018	2.16E-02
cg06030846	*TMEM108*	3	134581182	-0.011	2.16E-02
cg25994988	*UBASH3B*	11	122157592	-0.011	2.16E-02
cg18345924	*NCAM2*	21	21294102	-0.016	2.18E-02
cg00638021	*COL1A1*	17	45622061	-0.013	2.26E-02
cg19344626	*NWD1*	19	16691749	-0.024	2.36E-02
cg01288258	*ITFG2*	12	2792128	-0.011	2.41E-02
cg05221385	*TAF10*	11	6590080	-0.010	2.43E-02

DNAmet levels of the 1202 a-CpGs as well (Additional file 2: Table S1b). In this second correlation analysis, the PCs were defined with PCA from DNA methylation data of the 1202 a-CpGs (aging-associated CpG sites, FDR < 5 %); methylation data in PCA were not adjusted for the cell subtype heterogeneity. Correlation analysis revealed that PC1-PC6 determined more than 50 % of variance in methylation levels of these a-CpGs and these PCs correlated clearly with age and the cell counts. It is also worth of mentioning that of the 1202 a-CpGs in our initial aging-association analysis, there were 526 multivariable regression models (corresponding 526 CpG sites) where all cell count variables (monocytes, granulocytes, NK, CD8T and CD4T cells) were detected as statistically significant (FDR < 5 %) predictors of DNA methylation levels.

The importance of the cell count considerations was explored with an additional set of regression models, where the DNA methylation level in each CpG site genome-wide was explained with age and sex only while the cell counts were not adjusted for. In this analysis, only 56 sites were classified as aging-associated (FDR < 5 %) and these sites were all included to the original pool of 1202 a-CpGs. The 56 a-CpGs are pointed out in the Additional file 3: Table S4.

Aging-associated hypermethylation and hypomethylation differ in their features. The exploration of aging-associations in the YFS revealed that hypermethylation was more frequent within genes with more association hits as shown in Additional file 4: Table S5 and Fig. 1). Specifically, there were 70 genes in total either with more than one hypomethylated or more than one hypermethylated a-CpGs per gene. Of those, 22 genes comprised more than one hypomethylated a-CpGs per gene and 48 genes comprised more than one hypermethylated a-CpGs per gene as shown in Additional file 4: Table S5.

Next, the genomic locations of the a-CpGs were investigated, revealing that 388 of the 1202 a-CpGs were located on CpG islands rather than island shores, shelves or non-island regions, and a majority (N = 331) of those were hypermethylated (Additional file 1: Figure S2). The remaining sites were distributed to shores, shelves and non-island regions with opposite manner as shown in Additional file 1: Figure S2; the aging-associated hypomethylation was more abundant on those regions. The a-CpG locations on genes were also investigated; no

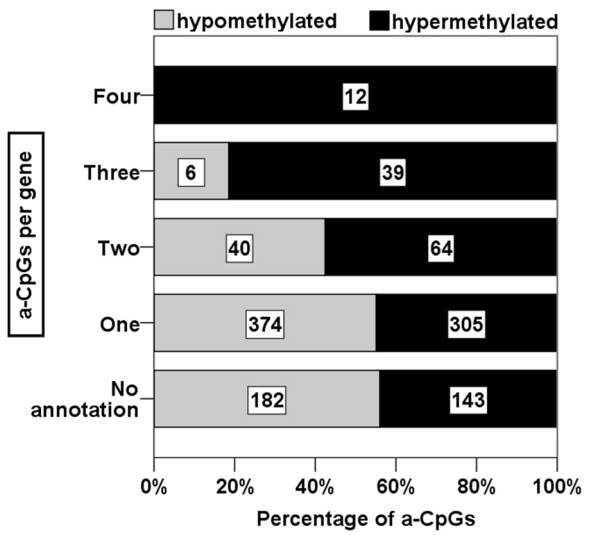


Fig. 1 Numbers of aging-associated CpG sites (hits) per gene in regard to hypermethylation and hypomethylation is visualized as bars. Aging-associated hypermethylation was more frequent within genes with more association hits. First, the genes were categorized into groups based on the number of hypermethylated or hypomethylated a-CpG hits per gene. Next, the frequencies of hypermethylated and hypomethylated a-CpGs within the groups were calculated. The number of a-CpGs for each group is shown inside each bar

enrichment of a-CpGs was detected in the regions of 3′ untranslated regions (UTRs), 5′UTRs or close distances to transcription start sites or gene bodies (Additional file 1: Figure S3a and b). The distributions of the a-CpGs on chromosomes were also investigated; hypermethylated a-CpGs were over-represented on chromosome 18, whereas hypomethylated sites were not enriched on any chromosome (hypergeometric test, nominal p-value of 0.05) (Additional file 1: Figure S3c). In addition, we ensured using visual examination that there were no spatial local cluster(s) of a-CpGs on Chr-18.

Sex specificity of the aging-associated CpG sites

To evaluate the sex specificity of the aging-associations, an interaction model with variables corresponding to sex, age and the interaction of sex and age (age*sex) was constructed. No sex-specific a-CpGs were identified, as analysis revealed that no interaction term had a false discovery rate (FDR) below 5 % (q-value < 0.05) in the interaction models. Furthermore, we analyzed women (N = 111) and men (N = 73) separately as well: sex-specific a-CpGs were explored among all CpG sites with an multivariable regression model ('beta regression') where age and cell type proportion variables were used to predict DNA methylation level in each CpG site. These analyses revealed that there were 105 and 173 a-CpGs (FDR < 5 %) among men and women, respectively; these CpG sites were all included to our original pool of 1202 a-CpGs which were detected using whole sample (N = 184). Importantly, as shown in Additional file 1: Figure S5, when the directions of change among the 1202 a-CpGs were cross-compared between men and women (without p-value cut-off), all sites, except one, showed concordant behavior regarding hypermethylation or hypomethylation during aging (i.e. whether the estimate of age variable in the regression model was negative or positive value). This behavior was also identical to the directions of change among the 1202 a-CpGs in the initial analysis (N = 184). As a conclusion, these results were in line with our interaction analysis: there were no significantly sex-specific a-CpGs among middle-aged individuals.

Functional roles of a-CpGs in the YFS

The gene ontology (GO) functions and processes of the genes with a-CpGs were investigated using the Gene Ontology enRIchment anaLysis and visuaLizAtion (GOrilla) tool [21]. The analysis was conducted separately for genes with hypermethylated a-CpGs and for hypomethylated a-CpGs (N = 440 and N = 437, respectively). The analysis revealed an unambiguous differences between hypo- and hypermethylated a-CpGs, as 73 GO process terms and to 8 GO function terms were enriched to genes with hypermethylated a-CpGs (Tables 3 and 4, respectively; Additional file 2: Table S2.), whereas there was no enrichment of terms among the genes with hypomethylated a-CpGs (Bonferroni-adjusted p-value threshold of 0.05). The most statistically significant processes were anatomical structure development (GO:0048856, $p = 1.02*10^{-11}$) and morphogenesis (GO:0009653, $p = 5.02*10^{-10}$), both of which cluster under the term 'developmental process'.

In addition, Pscan [22] was used to predict whether there were common regulators for groups of genes. The hypermethylation-associated genes were predicted to be regulated by 11 common transcription factors (Additional file 2: Table S3), several of which were zinc coordinating. For hypomethylation-associated genes, no common transcription factors were found. A large proportion of the 11 regulators of genes with hypermethylated a-CpGs in the YFS were zinc coordinating, and four (E2F1, EGR1, SP1, TFAP2A) were identical to those identified in the V90+ study [15].

Comparisons to other studies

In the explorative cross-comparison analysis, the a-CpGs identified in middle-aged individuals of the YFS were compared to aging-associated DNA methylome alterations between nonagenarians and 19–30-year-old individuals evidenced in our previous study (the V90+ study) [15]. The a-CpGs identified in the V90+ study were strongly associated with aging while the cell type heterogeneity was adjusted for in the analysis. A total of

Table 3 Several GO process terms were enriched within genes with hypermethylated a-CpGs in the analysis with GOrilla [21, 43]. This table represents the main clusters of processes (53 redundant GO terms were filtered out of 73 terms using REViGO [44]). The full list of processes is shown in Additional file 2: Table S2

GO term	Description of the process	*p*-value (-log10)
GO:0048856	Anatomical structure development	10.9914
GO:0050794	Regulation of cellular process	8.9788
GO:0007389	Pattern specification process	8.2343
GO:0032502	Developmental process	8.2041
GO:0009893	Positive regulation of metabolic process	8.0511
GO:0044708	Single-organism behavior	7.5544
GO:0035108	Limb morphogenesis	7.5544
GO:0003002	Regionalization	7.3585
GO:0051239	Regulation of multicellular organismal process	7.301
GO:0006357	Regulation of transcription from RNA polymerase II promoter	7.2248
GO:0065007	Biological regulation	7.1675
GO:0007610	Behavior	7.08
GO:0048598	Embryonic morphogenesis	7.0778
GO:0048518	Positive regulation of biological process	6.8761
GO:0048519	Negative regulation of biological process	6.7122
GO:0008285	Negative regulation of cell proliferation	6.4921
GO:0048523	Negative regulation of cellular process	5.8827
GO:0010842	Retina layer formation	5.8041
GO:0051961	Negative regulation of nervous system development	5.7423
GO:0032774	RNA biosynthetic process	5.4225

the 1202 a-CpGs established in the YFS cohort, 999 a-CpGs were also aging-associated in the V90+ sample (FDR < 5 %, Additional file 3: Table S4). Of these 999 a-CpGs, 464 (46 %) were hypermethylated, and 535 (54 %) were hypomethylated with advancing age. Furthermore, in 987 of the overlapping 999 a-CpGs the direction of the aging-associated change was the same: in the present and in the V90+ study, 455 a-CpGs were hypermethylated, and 532 were hypomethylated with advancing age (Fig. 2).

Finally, a-CpGs that were characterized from whole blood samples as aging-associated using 450 BeadChip technology and previously reported by Hannum et al. (number of hits, 89) [13], Garagnani et al. (number of hits, 9) [12] and Florath et al. (number of hits, 162) [23] and presented as summary table in Steegenga et al [19] were further compared with our data. The corresponding age of the samples ranged between 19 -101, 9–83 and 50–75 years, respectively. The comparison revealed 21 common CpG sites out of the 999 a-CpGs in two or more studies in addition to the YFS and the V90+ study (Fig. 3).

Discussion

In this study, we identified 1202 a-CpGs where the DNAmet level was associated with aging in middle-aged individuals (i.e. with an age range of 40 to 49 years), in whom the growth and development of youth has ended yet old age and its associated diseases had not begun. Of the 1202 a-CpGs, 622 (52 %) were hypomethylated, and 580 (48 %) were hypermethylated with advancing age, with annotations on 440 and 437 different genes, respectively. In general, the functional features of these aging-associated sites are mostly similar to those identified from cohorts with larger age differences. Our study highlights also that a large number of sites undergo aging-associated DNAmet level changes throughout adulthood and we speculate that a great proportion of those probably change with a clock-like manner.

A large fraction of the DNAmet sites are altered during the lifespan, as shown by previous studies performed using 450BeadChip technology [15, 24] and whole-genome bisulfite sequencing [10]. Furthermore, the rates of these changes may fluctuate at different stages of the lifespan. Studies have shown that a-CpGs behave differently during the growth period before adulthood and at the end of the lifespan [11, 20]. Nonetheless, there are genes (*ELOVL2*, *SFMBT1*, *KLF14*, *PENK*, and *FHL2*) with CpG sites that are consistently detected as being aging-associated despite of differences in sample tissue

Table 4 GO function terms were enriched within genes with hypermethylated a-CpGs in the analysis with GOrilla. Table contains the full list of enriched GO function terms (Bonferroni-adjusted $p < 0.05$) obtained from analysis with GOrilla [21, 43]

GO term	Description of the function	*p*-value (-log10)
GO:0043565	Sequence-specific DNA binding	10.001
GO:0000981	Sequence-specific DNA binding RNA polymerase II transcription factor activity	7.322
GO:0001071	Nucleic acid binding transcription factor activity	6.721
GO:0003700	Sequence-specific DNA binding transcription factor activity	6.721
GO:0003677	DNA binding	6.625
GO:0005326	Neurotransmitter transporter activity	5.148
GO:0005488	Binding	4.967

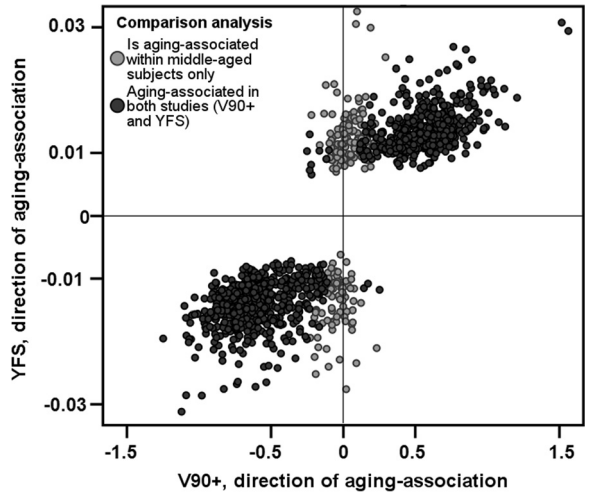


Fig. 2 The direction of aging-association in 1202 a-CpGs is visualized as scatterplot. Each dot corresponds to single a-CpG; directions of associations correspond to estimates of age which are fetched from the regression models. Of 1202 sites, 987 CpG sites were similarly associated with aging in both the YFS and in the V90+ study. The analyses in both studies were adjusted for leukocyte cell subtype proportions, and the studies consisted of the samples with distinct age ranges: the YFS comprised 40 to 49 years old subjects whereas the V90+ study consisted of 19–30-year-old individuals and nonagenarians. The corresponding data illustrated in the Fig. 2 is presented in Additional file 3: Table S4

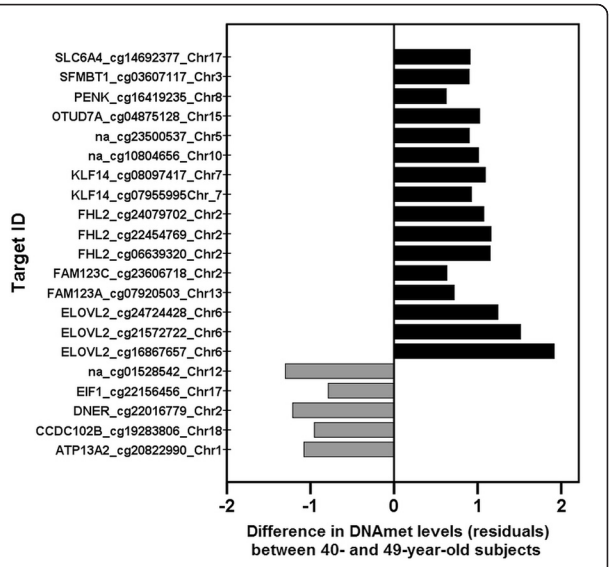


Fig. 3 The top 21 most commonly reported a-CpGs and their direction of association with aging. The top 21 a-CpGs were selected with following criteria: the a-CpG was identified in present study and in the V90+ study, as well as in two or more other studies (Hannum et al. [13], Garagnani et al. [12] or Florath et al. [23]); the sites were reported as aging-associated in blood samples and the data were obtained using 450 BeadChip technology. Methylation level differences in YFS between the highest and the lowest age groups (between 40- and 49-year-old individuals; calculated from the medians of residuals after adjusting for effects of sex and cell type proportions), are illustrated as bars. The bars are colored according to the hypomethylation or hypermethylation status (grey = hypomethylated, black = hypermethylated). Gene annotation is shown for each bar, where applicable (na = no gene annotation). The corresponding data is presented in Additional file 3: Table S4

types or age distributions [11–13, 15]; notably, these genes were also identified in the present study as being aging-associated (Tables 1 and 2; Additional file 3: Table S4). However, a recent meta-analysis on three DNAmet data sets obtained using 450BeadChip illustrated discrepancies in the lists of regions where DNAmet levels were altered during the entire human lifespan, ranging from 0 to 100 years of age [2]. Because blood sample heterogeneity has been shown to have a great impact on EWASs [4, 15], our speculation is that the discrepancies might be due to the presence of different cell types.

In the primary analysis, we aimed to identify a-CpGs in middle-aged individuals representing general population with age range of only one decade. Then, we cross-compared the results to those obtained with similar analysis pipeline from a population aged 19 to 90 years (Vitality 90+ study) [15]. Among the 1202 a-CpGs characterized from the YFS with an age range of nine years, 987 sites had an identical association direction as detected in the Vitality 90+ study, as shown in Fig. 2 and in Additional file 3: Table S4. We hypothesize that sites displaying aging-associated methylation changes in both populations possibly represent sites where the change in DNA methylation follows a clock-like pattern. We further speculate that the non-overlapping CpG sites identified in the population with a wider age range (19 to 90 years of age) may possibly represent sites where the aging-associated change is accelerated in either early or late adulthood; the a-CpGs identified only when comparing group of nonagenarians to young adults may represent changes that reflect e.g. aging-associated pathologies or accumulation of aging-associated impairments.

As aging influences the immune system of men and women differently and as the risk rates of several diseases between sexes are unequal [25, 26], 1) an interaction analysis was performed to address the sex specificity of a-CpGs, and 2) the aging-associations were also evaluated in separate analyses among men and women. These analyses revealed no sex-specific single a-CpGs; thus, the identified a-CpGs are universally altered in both men and women. These results are in accordance with our previous results from the V90+ study, in which the DNAmet states of nonagenarians were compared with 19–30-year-old individuals [15], and with results published by others [24, 27]. However, studies have shown that as a whole, the DNA methylomes of males age more rapidly than those of females [13, 28].

Aging-accelerated hypomethylation may be thought as an erosion-like event, whereas hypermethylation may be thought as an actively guided process. In practice, the difference between these features is manifested, for example, through the enrichment of GO terms for groups of genes and for signaling pathways [1, 15]. The distinct roles of the methylation status were demonstrated in the present study with the numbers of a-CpG hits in a gene, as we observed notable enrichment of hypermethylation events located in genes with more than one a-CpG (Fig. 1). The functional roles of genes with a-CpGs were established by GO term enrichment analysis, which revealed obvious difference between hypo- and hypermethylated a-CpGs, even though the analysis was conducted with an equal number of genes in the GO term analyses. A high number of GO terms were enriched to genes with hypermethylated a-CpGs (Tables 3 and 4; Additional file 2: Table S2), whereas there was no GO term enrichment within genes with hypomethylated a-CpGs. The most statistically significant processes enriched to genes with hypermethylated a-CpGs were 'anatomical structure development' and 'morphogenesis', both of which cluster under the term 'developmental process'. The enrichment of hypermethylated a-CpGs to these processes has been reported previously [14, 23, 24, 29]. Reynolds [30] and Yuan [16] reported also that the CpG sites hypermethylated during aging are enriched to common processes and exhibit shared features, whereas hypomethylated a-CpGs are a less homogenous group. Furthermore, age-associated hypermethylation interactome hotspots have been reported [31].

In addition to the details mentioned above, we observed other similar hypermethylation characteristics in the YFS, as those reported in previous studies [1, 15]. For example, the majority (85 % out of 388) of a-CpGs localized in CpG-islands (instead of shores, shelves or other regions) were hypermethylated, and an excess of hypermethylated a-CpGs were also found on chromosome 18. However, there was no enrichment of a-CpGs on chromosome 19. In the V90+ study, the hypermethylated a-CpGs located in the genes encoding zinc-associated proteins were more abundant on chromosome 19 [15], where zinc-finger genes are clustered. The zinc-finger genes (such as *ZNF154*) located in chromosome 19 are proposed to be repressors of endogenous retroviruses (ERVs) [32], and the repressor activity may be disturbed by hypermethylation. Interestingly, CpG sites located in the gene *ZNF154* and almost all other genes encoding zinc-fingers on chromosome 19 were absent from our pool of 1202 a-CpGs. Thus, as the hypermethylation of CpG sites located in genes encoding zinc-fingers was observed in the oldest age group, we hypothesize that rates of methylation level changes at the CpG sites located in ERV repressor genes (e.g. *ZNF154*) may fluctuate throughout the lifespan and that the rates may be enhanced in association with other senescence-related factors. Therefore, it is possible that DNAmet-based dysfunction of the repression system might explain the increased expression of ERVs in old age [33]. Future studies are required to address these questions.

To further inspect the roles of the genes with aging-accelerated DNAmet changes, analysis of the common regulators (transcription factors) of groups of genes with hypermethylated and hypomethylated a-CpGs was conducted with Pscan [22]. The results were again surprisingly concordant with those in the V90+ study. There were 11 regulators with unique identifiers for hypermethylated a-CpGs (Additional file 2: Table S3), whereas hypomethylated a-CpGs had no common regulators. A great proportion of the 11 regulators of genes with hypermethylated a-CpGs in the YFS were zinc coordinating, and four (E2F1, EGR1, SP1, and TFAP2A) were identical to those identified in the V90+ study results [15]. Overall, the results from analysis of the functional roles of the genes with a-CpGs were surprisingly well in line with the observations from the V90+ study and supported the proposition that aging-associated hypermethylation is a more tightly regulated process, whereas aging-associated hypomethylation is induced more by environmental effects and stochastic factors.

Finally, we demonstrated the lack of concordance in previously reported pools of a-CpGs by comparing three published lists of overlapping a-CpGs produced using 450BeadChips from whole blood samples from subjects with age ranges of 50–75, 19–101 and 9–83 [12, 13, 23]. Although 987 of the a-CpGs in the YFS showed similar association directions as in the V90+ study (Fig. 2 and Additional file 3: Table S4), we observed only 61 overlapping a-CpGs in the YFS and the V90+ study, which were also reported as aging-associated in one or more other robustly compatible studies (same sample type and array technology). Of these, only 21 a-CpGs were observed in two or more of the studies in the comparison (Fig. 3). To the best of our knowledge [4, 15], the main factor that contributes to the DNAmet profiles in blood cells is cell type heterogeneity; thus, we speculate that the lack of cell type adjustments may account for the majority of disparity in the cross-comparisons. The results of aging-association analysis and combined PCA-correlation analysis in this study supports our speculation. Cell type heterogeneity should be taken into account when analyzing samples composed of mixed cell types, but a limited number of such studies have been conducted [4, 14–17].

Notably, our study had an obvious limitation, it would substantially benefit from being a follow-up; therefore, future studies are needed. Nevertheless, the analysis is

powered by well-designed sample characteristics because each age group was matched by sex and sample size and because adjustments were made for cell type heterogeneity. Thus, the analysis was sensitive enough to detect DNAmet changes within an age range spanning nine years.

Conclusions

Here we report that aging-associated DNA methylation changes can be identified in a middle-aged population with a narrow age range of 9 years. Aging-associated DNAmet changes are not uniform, but occur due to different reasons, at different rates and directions in different parts of the genome and are not alike in all cell types. Thus, due to this diverse nature of aging-associated DNA methylation changes, all confounding factors should be accounted for in the analysis, in order to obtain comparable results. Our results support the notion that cell type heterogeneity should be adjusted for when analyzing tissues consisting of mixed cell types. Moreover, our results imply that considerable proportion of DNAmet changes show clock-like behavior throughout adulthood.

Methods

Study population

The Young Finns study (YFS) comprises a series of six cohorts, representing general population, born in 1962, 1965, 1968, 1971, 1974 and 1977 from five cities with university hospitals in Finland (Helsinki, Kuopio, Oulu, Tampere and Turku) [18]. A subsample of 184 individuals was randomly assigned from a follow-up in 2011. The sample collection in 2011 is described in more detail elsewhere [34]. The categories of age in the methylation analysis were 40, 43, 46 and 49 years old, with group sizes of 50, 44, 55 and 35, in which 58 %, 68.2, 56.4 and 60 % were women, respectively. All of the participants were of western European descent. The study followed the guidelines of the Declaration of Helsinki and was approved by the Ethical Review Committee of Turku University Hospital. All participants provided informed consent.

DNA methylome quantification

Sample preparations

Leukocyte DNA of the YFS cohort was obtained from EDTA-blood samples using a Wizard® Genomic DNA Purification Kit (Promega Corporation, Madison, WI, USA) according to the manufacturer's instructions. Genome-wide DNA methylation levels were obtained using Illumina Infinium HumanMethylation450 BeadChips [35–37] in the Core Facility at the Institute of Molecular Medicine Finland (FIMM), University of Helsinki according to the protocol by Illumina.

The methylation data set was preprocessed identically with a previously described analysis pipeline which was used in the DNA methylation analysis of the V90+ study samples [15, 38, 39]. Briefly, methylation signal data was preprocessed as a methylumiset object using R software ($R >= 2.15.3$) with array-specific algorithms implemented in the R package wateRmelon [40] and BMIQ [38]. The resulting β values ranged linearly from 0 (non-methylated, 0 %) to 1 (completely methylated, 100 %). The quality of DNA samples and methylation data was carefully ensured by standard examinations with principal component analysis (PCA) and visualizations with density plots, boxplots and dotplots. Three of the YFS samples were excluded due to atypically low probe intensities compared with control probe intensities.

The YFS sample was lacking leukocyte cell type characterizations; thus, the proportions were determined by the estimation algorithm implemented in the estimateCellCounts function of the minfi Bioconductor package [4] using R software ($R >= 2.15.3$). The algorithm utilizes the selection of 600 control probes that represents specific signatures of CD8T and CD4T cells, monocytes, granulocytes, and NK and B cells (Additional file 1: Figure S1). The reference data used in the estimation is available in the FlowSorted.Blood.450K Bioconductor package [4].

Quality control of the DNA methylome data

As the cell type proportions contribute to most of the variation in genome-wide DNAmet [4, 15], the significance of the estimated cell counts in the DNAmet data was investigated by PCA, and the main PCs of DNAmet were correlated with the cell counts (Additional file 2: Table S1a). Spearman's correlation analysis indicated a clear connection between methylation profiles and estimated cell proportions. Thus, the estimated cell counts as well as the genome-wide methylation data was shown to behave as expected.

As part of the quality control step, a well-known CpG site with phenotype association was selected. Smoking is strongly associated with the hypomethylation of cg03636183, located in the gene *F2RL3* [8, 9]; our data from the YFS replicated this finding, as we observed a difference between daily smokers and others (Wilcoxon rank sum-test, $P = 2.4*10^{-6}$; Additional file 1: Figure S4). Analysis with multivariable regression model (function lm() in R) revealed that the cell type heterogeneity, age or sex of the samples did not alter the finding of cg0363618.

Detection of aging-associated methylation regions

Aging-associated CpG sites, the a-CpGs, were explored using a generalized linear regression model, referred to as the 'variable dispersion beta regression' in an iterative

manner for each methylation locus (CpG site). The age (categories of 40, 43, 46 and 49) was employed as a variable to predict the site-specific methylation outcome in the form of a β value (ranging from 0 to 1); this was done in each equation using the mean model and a linker function of *logit*. The cellular heterogeneity was adjusted in the initial multivariable regression analyses: in addition to age and sex variables, variables corresponding to each estimated blood cell subtype proportion (CD8T and CD4T cells, monocytes, granulocytes, NK and B cells; all ranging linearly from 0 to 1) were included to the regression models as predictors of DNA methylation level. Additionally, sex-specific a-CpGs were explored among all CpG sites using two approaches: 1) with an interaction model where age, sex, sex*age and cell type proportion variables were used to predict DNA methylation level, and 2) with an regression model where age and cell type proportion variables were used to predict DNA methylation level separately for men and women. Furthermore, to explore the relevance of the cell count considerations in the regression analyses, an additional set of age-association analyses was performed. In these regression models, the DNA methylation level of each CpG site was explained with age and gender variables only and the cell proportions were not adjusted for. The analyses were performed using R software ($R >= 2.15.3$), and the regression analyses were mainly conducted with algorithms implemented in the betareg package [41]. The nominal Benjamini-Hochberg adjusted *p*-value (q-value) was set to 0.05. The a-CpGs were annotated based on the assembly provided by the R package, FDb.InfiniumMethylation.hg19 [42]. For the purpose of visualization in Fig. 3, standardized weighted residual values of the methylation levels were extracted for each CpG site from regression models in which only sex and cell type proportion variables were set as predictors.

Analysis of the functional roles of a-CpGs

The enriched gene ontology (GO) terms of the genes with a-CpGs were discovered using GOrilla [21, 43], and the significant terms were further clustered by REViGO [44]. The GOrilla analysis was performed for the process, function and component categories with two un-ranked lists, of which the first list comprised genes with hypomethylated or hypermethylated a-CpGs (Additional file 3: Table S4), and the second comprised the genes in the background (N = 20,902; analysis date, 9.3.2015). Furthermore, the prediction of common transcription factors of the groups of genes with either hypermethylated or hypomethylated a-CpGs (as two separate analyses) was conducted using Pscan with the default settings (JASPAR database; analysis date, 10.3.2015) [22]. The nominal *p*-value was set to at the Bonferroni-corrected value of 0.05 in each analysis.

Availability of supporting data

The methylation data presented in this manuscript have been submitted to the Gene Expression Omnibus (GEO) database (http://www.ncbi.nlm.nih.gov/geo/) under the accession number GSE69270.

Additional files

Additional file 1: Figures S1-S5. 1) A figure of estimated proportions of CD8T, CD4T, NK, B cell, monocyte and granulocyte cells of peripheral blood samples in YFS. Proportions are visualized as boxplots, categorized by age group and organized to separate panels by sex. 2) A figure of aging-associated CpG site locations in regard to CpG islands (CGIs). Number of aging-associated CpG sites are visualized with stacked bars. 3) A figure (a-c) presenting locations of a-CpGs. 4) A figure showing results for association of DNA methylation level in cg03636183 with smoking. 5) A figure presenting sex specificity of the aging-associated CpG sites (a-CpGs). (DOCX 357 kb)

Additional file 2: Tables S1-S3. 1) Two summary tables (a and b) of the results from Spearman correlation analyses between age, the cell counts and the first principal components (PCs). PCs were defined from either the whole methylation data or 1202 a-CpGs using PCA. 2) A table of the GO terms of the bio processes that are enriched to genes with aging-associated CpG-sites. 3) A table of common transcription factors for genes with hypermethylated a-CpGs characterized using Pscan. (DOCX 31 kb)

Additional file 3: Table S4. A full table of 1202 a-CpGs with detailed information. (XLSX 183 kb)

Additional file 4: Table S5. A summary table where the 70 genes with more than one hypomethylated or more than one hypermethylated a-CpGs per gene are presented. (XLSX 15 kb)

Abbreviations

a-CpGs: aging-associated CpG sites (FDR < 5 %); DNAmet: DNA methylation; ERV: endogenous retrovirus; EWAS: epigenome-wide association studies; PC: principal component; UTR: untranslated region; V90+: The Vitality 90+ Study; YFS: Young Finns study.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

LK processed the data and performed analyses and was responsible for writing the manuscript. NM, JJ and TN were responsible of the experiments. MH, MK and TL provided reagents and materials and LK, SM, JJ, TL and MH contributed to the design of the study. OTR, MK and TL were responsible for recruiting the subjects in the study. All authors contributed to the writing of manuscript and read and approved the final manuscript.

Acknowledgements

The Young Finns Study has been financially supported by the Academy of Finland: grants 286284 (T.L), 132704 (M.H.), 134309 (Eye), 126925, 121584, 124282, 129378 (Salve), 117787 (Gendi), and 41071 (Skidi); the Social Insurance Institution of Finland; Kuopio, Tampere and Turku University Hospital Medical Funds (grant X51001 for T.L.); Juho Vainio Foundation; Paavo Nurmi Foundation; Finnish Foundation of Cardiovascular Research (T.L.); Finnish Cultural Foundation; Tampere Tuberculosis Foundation (T.L., M.H.); Emil Aaltonen Foundation (T.L.); and Yrjö Jahnsson Foundation (T.L., M.H.) This work was also supported by grants from the Competitive Research Fund of Pirkanmaa Hospital District (9M017, 9N013 to M.H.). We thank Nina Peltonen and Sinikka Repo-Koskinen for their excellent technical assistance.

Author details
[1]Department of Microbiology and Immunology, School of Medicine, University of Tampere, Tampere, Finland. [2]Gerontology Research Center, Tampere, Finland. [3]Fimlab Laboratories, Tampere, Finland. [4]Departments of Clinical Physiology, Tampere University Hospital and University of Tampere School of Medicine, Tampere, Finland. [5]Research Centre of Applied and Preventive Cardiovascular Medicine and the Department of Clinical Physiology and Nuclear Medicine, University of Turku and Turku University Hospital, Turku, Finland. [6]Department of Clinical Chemistry, University of Tampere School of Medicine, Tampere, Finland.

Received: 5 June 2015 Accepted: 1 February 2016
Published online: 09 February 2016

References

1. Zampieri M, Ciccarone F, Calabrese R, Franceschi C, Burkle A, Caiafa P. Reconfiguration of DNA methylation in aging. Mech Ageing Dev. 2015.
2. Bacalini MG, Boattini A, Gentilini D, Giampieri E, Pirazzini C, Giuliani C, et al. A meta-analysis on age-associated changes in blood DNA methylation: results from an original analysis pipeline for Infinium 450k data. Aging (Albany NY). 2015;7(2):97–109.
3. Accomando WP, Wiencke JK, Houseman EA, Nelson HH, Kelsey KT. Quantitative reconstruction of leukocyte subsets using DNA methylation. Genome Biol. 2014;15(3):R50.
4. Jaffe AE, Irizarry RA. Accounting for cellular heterogeneity is critical in epigenome-wide association studies. Genome Biol. 2014;15(2):R31.
5. Houseman EA, Accomando WP, Koestler DC, Christensen BC, Marsit CJ, Nelson HH, et al. DNA methylation arrays as surrogate measures of cell mixture distribution. BMC Bioinformatics. 2012;13:86.
6. Johnson AA, Akman K, Calimport SR, Wuttke D, Stolzing A, de Magalhaes JP. The role of DNA methylation in aging, rejuvenation, and age-related disease. Rejuvenation Res. 2012;15(5):483–94.
7. Bacalini MG, Friso S, Olivieri F, Pirazzini C, Giuliani C, Capri M, et al. Present and future of anti-ageing epigenetic diets. Mech Ageing Dev. 2014;136–137:101–15.
8. Zhang Y, Yang R, Burwinkel B, Breitling LP, Brenner H. F2RL3 methylation as a biomarker of current and lifetime smoking exposures. Environ Health Perspect. 2014;122(2):131–7.
9. Breitling LP, Yang R, Korn B, Burwinkel B, Brenner H. Tobacco-smoking-related differential DNA methylation: 27K discovery and replication. Am J Hum Genet. 2011;88(4):450–7.
10. Heyn H, Li N, Ferreira HJ, Moran S, Pisano DG, Gomez A, et al. Distinct DNA methylomes of newborns and centenarians. Proc Natl Acad Sci U S A. 2012; 109(26):10522–7.
11. Horvath S. DNA methylation age of human tissues and cell types. Genome Biol. 2013;14(10):R115.
12. Garagnani P, Bacalini MG, Pirazzini C, Gori D, Giuliani C, Mari D, et al. Methylation of ELOVL2 gene as a new epigenetic marker of age. Aging Cell. 2012;11(6):1132–4.
13. Hannum G, Guinney J, Zhao L, Zhang L, Hughes G, Sadda S, et al. Genome-wide methylation profiles reveal quantitative views of human aging rates. Mol Cell. 2013;49(2):359–67.
14. Rakyan VK, Down TA, Maslau S, Andrew T, Yang TP, Beyan H, et al. Human aging-associated DNA hypermethylation occurs preferentially at bivalent chromatin domains. Genome Res. 2010;20(4):434–9.
15. Marttila S, Kananen L, Hayrynen S, Jylhava J, Nevalainen T, Hervonen A, et al. Ageing-associated changes in the human DNA methylome: genomic locations and effects on gene expression. BMC Genomics. 2015;16(1):179. Epub 2015 Mar 14.
16. Yuan T, Jiao Y, de Jong S, Ophoff RA, Beck S, Teschendorff AE. An integrative multi-scale analysis of the dynamic DNA methylation landscape in aging. PLoS Genet. 2015;11(2), e1004996.
17. Teschendorff AE, Menon U, Gentry-Maharaj A, Ramus SJ, Weisenberger DJ, Shen H, et al. Age-dependent DNA methylation of genes that are suppressed in stem cells is a hallmark of cancer. Genome Res. 2010; 20(4):440–6.
18. Raitakari OT, Juonala M, Ronnemaa T, Keltikangas-Jarvinen L, Rasanen L, Pietikainen M, et al. Cohort profile: the cardiovascular risk in Young Finns study. Int J Epidemiol. 2008;37(6):1220–6.
19. Steegenga WT, Boekschoten MV, Lute C, Hooiveld GJ, de Groot PJ, Morris TJ, et al. Genome-wide age-related changes in DNA methylation and gene expression in human PBMCs. Age (Dordr). 2014;36(3):9648. Epub 2014 May 2.
20. Alisch RS, Barwick BG, Chopra P, Myrick LK, Satten GA, Conneely KN, et al. Age-associated DNA methylation in pediatric populations. Genome Res. 2012;22(4):623–32.
21. Eden E, Navon R, Steinfeld I, Lipson D, Yakhini Z. GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. BMC Bioinformatics. 2009;10:48.
22. Zambelli F, Pesole G, Pavesi G. Pscan: finding over-represented transcription factor binding site motifs in sequences from co-regulated or co-expressed genes. Nucleic Acids Res. 2009;37(Web Server issue):W247–52.
23. Florath I, Butterbach K, Muller H, Bewerunge-Hudler M, Brenner H. Cross-sectional and longitudinal changes in DNA methylation with age: an epigenome-wide analysis revealing over 60 novel age-associated CpG sites. Hum Mol Genet. 2014;23(5):1186–201.
24. Johansson A, Enroth S, Gyllensten U. Continuous aging of the human DNA methylome throughout the human lifespan. PLoS One. 2013;8(6), e67378.
25. Gubbels Bupp MR. Sex, the aging immune system, and chronic disease. Cell Immunol. 2015.
26. Oertelt-Prigione S. The influence of sex and gender on the immune response. Autoimmun Rev. 2012;11(6-7):A479–85.
27. McClay JL, Aberg KA, Clark SL, Nerella S, Kumar G, Xie LY, et al. A methylome-wide study of aging using massively parallel sequencing of the methyl-CpG-enriched genomic fraction from blood in over 700 subjects. Hum Mol Genet. 2014;23(5):1175–85.
28. Weidner CI, Lin Q, Koch CM, Eisele L, Beier F, Ziegler P, et al. Aging of blood can be tracked by DNA methylation changes at just three CpG sites. Genome Biol. 2014;15(2):R24.
29. Hernandez DG, Nalls MA, Gibbs JR, Arepalli S, van der Brug M, Chong S, et al. Distinct DNA methylation changes highly correlated with chronological age in the human brain. Hum Mol Genet. 2011;20(6):1164–72.
30. Reynolds LM, Taylor JR, Ding J, Lohman K, Johnson C, Siscovick D, et al. Age-related variations in the methylome associated with gene expression in human monocytes and T cells. Nat Commun. 2014;5:5366.
31. West J, Beck S, Wang X, Teschendorff AE. An integrative network algorithm identifies age-associated differential methylation interactome hotspots targeting stem-cell differentiation pathways. Sci Rep. 2013;3: 1630.
32. Lukic S, Nicolas JC, Levine AJ. The diversity of zinc-finger genes on human chromosome 19 provides an evolutionary mechanism for defense against inherited endogenous retroviruses. Cell Death Differ. 2014;21(3):381–7.
33. Taruscio D, Mantovani A. Factors regulating endogenous retroviral sequences in human and mouse. Cytogenet Genome Res. 2004;105(2-4): 351–62.
34. Nuotio J, Oikonen M, Magnussen CG, Jokinen E, Laitinen T, Hutri-Kahonen N, et al. Cardiovascular risk factors in 2011 and secular trends since 2007: the cardiovascular risk in Young Finns study. Scand J Public Health. 2014; 42(7):563–71.
35. Bibikova M, Lin Z, Zhou L, Chudin E, Garcia EW, Wu B, et al. High-throughput DNA methylation profiling using universal bead arrays. Genome Res. 2006;16(3):383–93.
36. Bibikova M, Le J, Barnes B, Saedinia-Melnyk S, Zhou L, Shen R, et al. Genome-wide DNA methylation profiling using Infinium(R) assay. Epigenomics. 2009;1(1):177–200.
37. Bibikova M, Barnes B, Tsan C, Ho V, Klotzle B, Le JM, et al. High density DNA methylation array with single CpG site resolution. Genomics. 2011; 98(4):288–95.
38. Teschendorff AE, Marabita F, Lechner M, Bartlett T, Tegner J, Gomez-Cabrero D, et al. A beta-mixture quantile normalization method for correcting probe design bias in Illumina Infinium 450 k DNA methylation data. Bioinformatics. 2013;29(2):189–96.
39. Marabita F, Almgren M, Lindholm ME, Ruhrmann S, Fagerstrom-Billai F, Jagodic M, et al. An evaluation of analysis pipelines for DNA methylation profiling using the Illumina HumanMethylation450 BeadChip platform. Epigenetics. 2013;8(3):333–46.
40. Pidsley R, Wong CC Y, Volta M, Lunnon K, Mill J, Schalkwyk LC. A data-driven approach to preprocessing Illumina 450K methylation array data. BMC Genomics. 2013;14:293.
41. Cribari-Neto F, Zeileis A. Beta regression in R. JSS. 2010;34:1–24 [http://www.jstatsoft.org/v34/i02/].

42. Triche T, Jr. InfiniumMethylation.hg19: Annotation package for Illumina Infinium DNA methylation probes. R package version 2.2.0.
43. Eden E, Lipson D, Yogev S, Yakhini Z. Discovering motifs in ranked lists of DNA sequences. PLoS Comput Biol. 2007;3(3), e39.
44. Supek F, Bosnjak M, Skunca N, Smuc T. REVIGO summarizes and visualizes long lists of gene ontology terms. PLoS One. 2011;6(7), e21800.

Length of paternal lifespan is manifested in the DNA methylome of their nonagenarian progeny

Saara Marttila[1,2], Laura Kananen[1,2], Juulia Jylhävä[1,2], Tapio Nevalainen[1,2], Antti Hervonen[2,3], Marja Jylhä[2,3,4] and Mikko Hurme[1,2,5]

[1] Department of Microbiology and Immunology, School of Medicine, University of Tampere, Tampere, Finland

[2] Gerontology Research Center, Tampere, Finland

[3] School of Health Sciences, University of Tampere, Tampere, Finland

[4] Institute for Advanced Social Research, University of Tampere, Tampere, Finland

[5] Fimlab Laboratories, Tampere, Finland

***Correspondence to**: Saara Marttila, **email**: Saara.Marttila@staff.uta.fi*

Keywords: lifespan, longevity, DNA methylation, methylome, Intergenerational inheritance, Gerotarget

***Received**: August 17, 2015* ***Accepted**: August 27, 2015* ***Published**: September 30, 2015*

ABSTRACT

The heritability of lifespan is 20-30%, but only a few genes associated with longevity have been identified. To explain this discrepancy, the inheritance of epigenetic features, such as DNA methylation, have been proposed to contribute to the heritability of lifespan.

We investigated whether parental lifespan is associated with DNA methylation profile in nonagenarians. A regression model, adjusted for differences in blood cell proportions, identified 659 CpG sites where the level of methylation was associated with paternal lifespan. However, no association was observed between maternal lifespan and DNA methylation. The 659 CpG sites associated with paternal lifespan were enriched outside of CpG islands and were located in genes associated with development and morphogenesis, as well as cell signaling. The largest difference in the level of methylation between the progeny of the shortest-lived and longest-lived fathers was identified for CpG sites mapping to *CXXC5*. In addition, the level of methylation in three Notch-genes (*NOTCH1*, *NOTCH3* and *NOTCH4*) was also associated with paternal lifespan.

There are implications for the inheritance of acquired traits via epigenetic mechanisms in mammals. Here we describe DNA methylation features that are associated with paternal lifespan, and we speculate that the identified CpG sites may represent intergenerational epigenetic inheritance.

INTRODUCTION

The heritability of lifespan (age at death) has been estimated to be approximately 20-30%, and it has been shown to increase with advancing age. Healthy aging is also heritable, and the offspring of long-lived parents show delayed onset of aging-associated diseases [1, 2, 3, 4]. Much of the research studying the heritability of lifespan has focused on extreme age (nonagenarians, centenarians, supercentenarians), but recently it has been shown that every decade of parental age after the age of 65 reduces the mortality and incidence of cancer of their offspring [5].

Even though the heritability of the lifespan is acknowledged, only one genomic locus (on chromosome 3) and a few genetic variants, such as in *APOE* and *FOXO3,* have consistently been shown to be associated with longevity. Data regarding other genomic loci and genes, including *CETP*, *HSF2* and *MTP*, have been inconsistent between studies [3]. Therefore, in addition to disease susceptibility alleles, rare genetic variants and environment-genome interactions, epigenetic mechanisms such as DNA methylation may be mediating the heritability of lifespan.

Changes in DNA methylation are associated with

Table 1: Grouping of study population according to paternal and maternal lifespan.

	n	Age of father at death	Age of mother at death
Whole population	90	40-103 (67)	40-101 (79.5)
Group FI	32	40-60 (55)	
Group FII	30	61-75 (67.5)	
Group FIII	28	77-103 (83)	
Group MI	32		40-72 (58)
Group MII	32		75-83 (80)
Group MIII	26		84-101 (88.5)

For group comparisons, the population was divided according to paternal and maternal lifespan. Presented here are the age range and (median) age at death for fathers and mothers.

aging and many aging-associated diseases, such as cancer, Alzheimer's disease and type 2 diabetes [6]. These changes include global hypomethylation and site-specific hypermethylation [7, 8, 9], and both tightly regulated and environmental or stochastic effects have been reported [10, 11, 12]. Aging-associated hypomethylation has been shown to be delayed in the offspring of centenarians [13].

The role of transgenerational epigenetic inheritance in the heritability of acquired traits has been discussed in the literature. In mice and rats, there is evidence that environmental exposure (for example, to vinclozolin and ethanol) causes phenotypic effects and changes in somatic and sperm DNA methylation, and these effects have been shown to be transmitted to the F4 generation [14, 15, 16, 17]. In a majority of the studies, the inheritance of epigenetic features is affected by the sex of the parent or progeny [15, 16, 18].

In humans, the environmental conditions experienced during early childhood or fetal development have been shown to link to epigenetic features, typically DNA methylation, in adulthood. For example, the progeny of mothers who experienced famine during early pregnancy are more prone to obesity and raised blood lipids, and the methylation status of the *IGF2* gene is affected in these progeny [19, 20]. In addition, childhood abuse has been associated with alterations to DNA methylation in middle-aged men [21]. In some cases, the environmental factors (e.g. nutrition, tobacco smoking, and betel quid chewing) experienced by fathers or grandfathers have been shown to affect the phenotype of their sons or grandsons (e.g. increased risk of diabetic death and increased adiposity). It is suspected that these traits are inherited via epigenetic mechanisms [18].

The effect of length of parental lifespan on the DNA methylation profile of progeny has not been previously studied. Here, we sought to identify DNA methylation patterns that are associated with maternal or paternal lifespan (age at death) to determine whether this trait manifests in the DNA methylome of progeny. DNA methylation profiles that are common among the progeny of longer-living parents may be components that are partially responsible for the heritability of lifespan.

RESULTS

Long-living fathers, long-living siblings

The study population consisted of 90 nonagenarians who participated in the Vitality 90+ study cohort of 2010 [8, 22]. In the regression model used to identify CpG sites associated with parental age, parental age was used as a continuous variable. However, for group comparisons, the population was divided into three groups according to paternal (FI (shortest-living fathers), FII, FIII (longest-living fathers)) and maternal age (MI (shortest-living mothers), MII, MIII (longest-living mothers)). See Table 1 for distribution of parental ages.

We found that group FIII (progeny of the longest-living fathers) had more long-living siblings (siblings living over 85 years) compared to group FI (Mann-Whitney U-test $p = 0.004$). This difference remains statistically significant when considering siblings over 75 years ($p = 0.006$) or siblings over 80 years ($p = 0.006$). The lifespan of the mother had no effect on the number of long-living siblings (comparison between groups MIII and MI: for siblings over 85 years of age, $p = 0.148$, for siblings over 80 years, $p = 0.338$ and for siblings over 75 years, $p = 0.242$).

Paternal lifespan was not correlated with maternal lifespan (Spearman's rho = 0.159, $p = 0.135$) or with paternal age at conception (data on paternal age at conception available only for a subset of the population ($n = 21$), Spearman's rho = -0.252, $p = 0.271$). In addition, paternal lifespan was not associated with the socio-economic status of offspring.

Association of paternal age with DNA methylation profile

The DNA methylation profile was determined with Illumina Infinium HumanMethylation450 BeadChip from peripheral blood mononuclear cells. We identified 659 CpG sites where the level of methylation was associated

Table 2: Genes with the largest number of CpG sites associated with paternal lifespan.

	n(CpG)	ID	*p*-value (BH-corrected)	Δβ
CXXC5	6	cg19628988	0.049	-0.082
		cg15165154	0.023	-0.072
		cg22885332	0.049	-0.042
		cg14871225	0.040	-0.034
		cg00906476	0.046	-0.015
		cg01008405	0.032	-0.012
COL11A2	4	cg13683990	0.042	-0.025
		cg21232625	0.042	-0.024
		cg25459558	0.028	-0.023
		cg02266086	0.046	-0.020
KCNS1	4	cg25353142	0.023	-0.033
		cg27634724	0.025	-0.023
		cg07589968	0.038	-0.021
		cg06193004	0.021	-0.017
BID	3	cg03433260	0.042	-0.017
		cg20234121	0.044	-0.014
		cg01280609	0.025	-0.013
FGR	3	cg09845000	0.029	-0.046
		cg09370867	0.030	-0.046
		cg13448978	0.046	-0.042
LOC283050	3	cg24658487	0.021	-0.041
		cg22890825	0.046	-0.035
		cg06891775	0.049	-0.018

In total, 42 genes harbored more than one affected CpG site (see Supplementary Table 2) and 6 genes contained three or more CpG sites associated with paternal lifespan. Δβ refers to the difference in methylation level between group FIII and group FI.

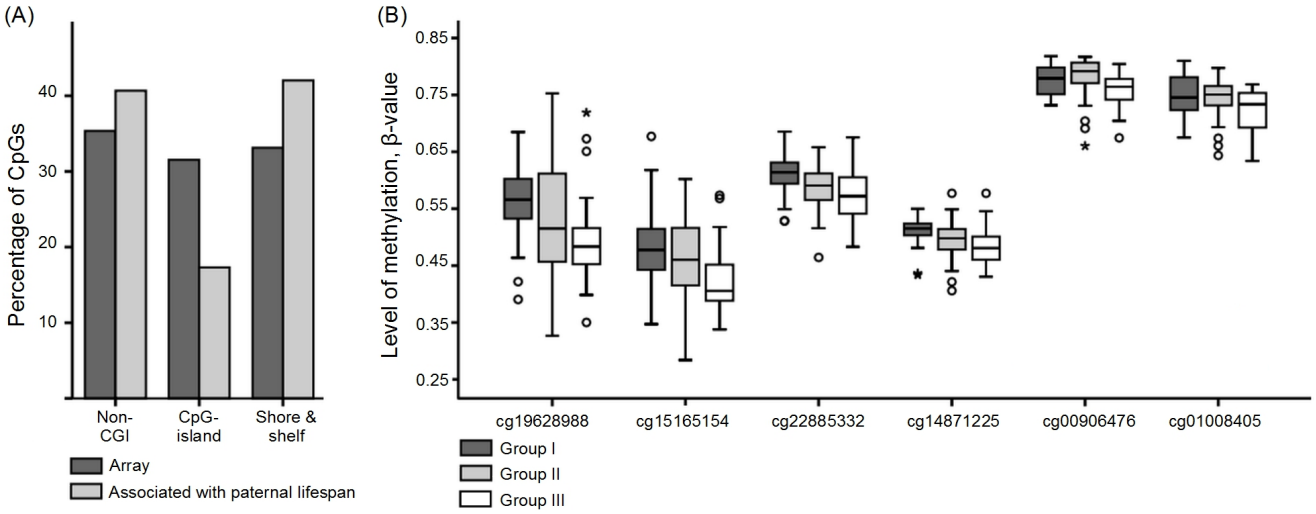


Figure 1: Location of CpG sites associated with paternal lifespan and methylation level of CXXC5. A. Location of CpG sites associated with paternal lifespan with regard to CpG islands. There were fewer than expected CpG sites found in CpG islands (hypergeometric test $p < 0.05$). **B.** Differences in the level of methylation in *CXXC5*. Level of methylation in each CpG site is presented for each group (Group I, progeny of shortest-lived fathers, Group III, progeny of longest-lived fathers).

Table 3: CpG sites associated with paternal lifespan that had the largest Δβ between group FIII and group FI.

Gene	ID	*p*-value (BH-corrected)	Δβ
CXXC5	cg19628988	0.048	-0.082
NOTCH1	cg12076931	0.032	-0.080
KRT27	cg10747531	0.032	-0.077
na	cg11284147	0.047	-0.077
CXXC5	cg15165154	0.023	-0.072
MPZL1	cg04846203	0.035	-0.067
NOTCH4	cg06023661	0.038	-0.066
UEVLD	cg15846482	0.033	-0.065
SORT1	cg02175308	0.028	-0.065
DAP	cg14129473	0.032	-0.064
MORC2	cg23825480	0.047	0.055
RRAD	cg06410849	0.032	0.056
RESP18	cg19020434	0.032	0.057
ITPKB	cg23717186	0.037	0.059
na	cg00248242	0.041	0.059
CPA5	cg22664614	0.039	0.059
na	cg14828411	0.040	0.060
GULP1	cg16947583	0.034	0.062
EPM2AIP1	cg24607398	0.023	0.069
na	cg23644389	0.045	0.072

For all CpG sites associated with paternal lifespan, see Supplementary Table 1. na = no gene annotation available.

with paternal lifespan (regression model p-value < 0.05 (BH-corrected), Δβ between group FIII and FI >1%, see Supplementary Table 1). Of the CpG sites associated with paternal lifespan, higher paternal age was associated with decreasing level of methylation in 423 (64%). There were no CpG sites where the level of methylation was associated with maternal lifespan. It is noteworthy that both the number of long-living siblings and the DNA methylation profile were associated with paternal lifespan, but not with maternal lifespan.

The CpG sites that were associated with paternal age were not enriched in any particular chromosome or gene location (hypergeometric test p>0.05). However, there were fewer than expected CpG sites in CpG islands (hypergeometric test $p < 0.05$, see Figure 1A).

Because a small number of DNA methylation changes in key genes that are involved in a given biological process can regulate the whole process or pathway (without DNA methylation changes to other genes), we wanted to investigate the identified top hits more closely. We defined the top hits as CpG sites with >5% difference in Δβ between groups FIII and FI or CpG sites that were located in a gene that harbored at least two CpG sites associated with paternal age. There were 65 CpG sites located in 46 different genes with Δβ>5%, and 31 additional genes had at least two CpG sites that were associated with paternal lifespan. Combined, there were 146 CpG sites and 77 genes that were further characterized (Supplementary Table 2). Among all CpG sites associated with paternal age, there were more sites where the level of methylation decreased as paternal age increased. This trend was even more pronounced among the top hits, where 116 out of 146 sites (79%) showed decreasing methylation levels with increasing paternal age.

CXXC5 (CXXC finger protein 5) was the most affected gene, harboring 6 CpG sites where the level of methylation was associated with paternal lifespan, and in all of these CpG sites, higher paternal age was associated with a decreased level of DNA methylation (Figure 1B). All genes with 3 or more CpG sites associated with paternal age are presented in Table 2. The largest Δβ between group FIII and group FI were observed at cg19628988 (*CXXC5*, Δβ = -0.082), cg12076931 (*NOTCH1*, Δβ = -0.080), cg23644389 (Δβ = 0.072) and cg24607398 (*EPM2AIP1*, Δβ = 0.069) (Table 3). In addition to *NOTCH1*, *NOTCH4* also included a CpG site with a Δβ>5% (cg06023661, Δβ = -0,066) and both genes harbored an additional CpG site (in *NOTCH1*, cg13861904, Δβ = -0.042 and in *NOTCH4*, cg06815976, Δβ = -0.042). *NOTCH3* also harbored two CpG sites where the level of methylation was associated with paternal lifespan (cg27320207, Δβ = -0.038 and cg26880200, Δβ = 0.020).

Table 4: GO process terms associated with genes where methylation level is associated with paternal lifespan.

GO Term	Description	*p*-value (BH-corrected)
GO:0048523	negative regulation of cellular process	0.011
GO:0010646	regulation of cell communication	0.012
GO:0022603	regulation of anatomical structure morphogenesis	0.013
GO:0023051	regulation of signaling	0.014
GO:0040012	regulation of locomotion	0.016
GO:0044767	single-organism developmental process	0.016
GO:0009966	regulation of signal transduction	0.017
GO:0032502	developmental process	0.019
GO:0048519	negative regulation of biological process	0.022
GO:0030154	cell differentiation	0.022
GO:0009653	anatomical structure morphogenesis	0.024
GO:0051270	regulation of cellular component movement	0.024
GO:0050878	regulation of body fluid levels	0.024
GO:0050794	regulation of cellular process	0.025
GO:2000147	positive regulation of cell motility	0.025
GO:0044707	single-multicellular organism process	0.025
GO:0031325	positive regulation of cellular metabolic process	0.025
GO:0040017	positive regulation of locomotion	0.026
GO:0051239	regulation of multicellular organismal process	0.026
GO:0007165	signal transduction	0.026
GO:0090527	actin filament reorganization	0.026
GO:0048583	regulation of response to stimulus	0.026
GO:0009893	positive regulation of metabolic process	0.026
GO:0048522	positive regulation of cellular process	0.027
GO:0048856	anatomical structure development	0.027
GO:0030335	positive regulation of cell migration	0.027
GO:0051272	positive regulation of cellular component movement	0.028
GO:0048869	cellular developmental process	0.028
GO:0048518	positive regulation of biological process	0.029
GO:0032501	multicellular organismal process	0.030
GO:0007596	blood coagulation	0.032
GO:0050817	coagulation	0.033
GO:0007599	hemostasis	0.034
GO:0050789	regulation of biological process	0.035
GO:0065007	biological regulation	0.047

The 659 CpG sites associated with paternal lifespan were located in 422 different genes, and these genes were enriched to 35 GO process terms (Benjamini-Hochberg multiple testing corrected *p*-value of < 0.05).

Pathways

The 659 CpG sites associated with paternal lifespan were located in 422 different genes. Cellular processes and signaling pathways associated with the identified genes were searched using QIAGEN's Ingenuity® pathway analysis (IPA) [23] and GOrilla [24,25]. We identified only one canonical pathway, B cell receptor signaling, that was associated with the identified genes when p-values were corrected for multiple testing (BH-corrected *p*-value < 0.05). Using GO term analysis, we identified 35 enriched GO process terms (BH-corrected *p*-value < 0.05) that were associated with genes harboring the CpG sites that were associated with paternal lifespan. The identified GO process terms were associated with development and morphogenesis and with cell signaling (Table 4).

In the GO term analysis for the top hits, no term reached multiple testing-corrected statistical significance (BH-corrected *p*-value < 0.05), but there was a trend toward developmental and signaling processes. Similarly, no significant canonical pathways were identified when multiple testing correction was used (BH-corrected *p*-value < 0.05). However, Notch-signaling was closest to the significance threshold (BH-corrected *p*-value = 0.084).

DISCUSSION

Here, we report the identification of 659 CpG sites where the level of methylation was associated with length of paternal lifespan. These results were adjusted for differences in blood cell type percentages. We speculate that these sites may represent intergenerational epigenetic inheritance and that these methylation sites could be associated with heritability of lifespan.

Cell signaling

CXXC5, a member of the small zinc finger protein family, contained 6 CpG sites where the level of methylation decreased as paternal lifespan increased. *CXXC5* negatively regulates Wnt/β-catenin signaling [26, 27, 28] and has been shown to be a mediator in BMP-signaling [29]. *CXXC5* has a role in normal and tumoral myelopoiesis [30] and in endothelial cell differentiation and migration and vessel formation [29]. The CXXC motif recognizes unmethylated CpG sites, and these proteins are involved in epigenetic modifications [31].

Six CpG sites in three Notch genes (*NOTCH1, 3* and *4*) were associated with paternal lifespan in our study. In addition, pathway analyses implied that Notch-signaling is associated with DNA methylation changes that are associated with paternal lifespan. The Notch-signaling pathway functions in various cell types and at various time points during development. Notch-signaling plays a role in development and organogenesis, and also in adult tissue maintenance and repair [32]. Notch-signaling has been associated with aging associated loss of muscle mass and function (sarcopenia). Impairments in Notch-signaling may be responsible for loss of myogenic potential in aged muscle. This association may also be due to an imbalance in Notch- and Wnt-signaling [32, 33, 34]. In addition, disruptions in Notch-signaling have been implicated in certain cancers and associated with Alzheimer's disease [32].

GO term analysis showed that signaling was affected, and B cell receptor signaling was also specifically identified as being associated with the identified genes in our study. In parallel with other changes in the immune system, the B cell pool goes through various changes during aging, and some of these changes have also been associated with adverse health outcomes [35, 36]. Our results imply that these changes can be partially regulated by DNA methylation. However, because we were unable to adjust the analysis for the proportion of B cells, this result may be due to differences in B cell proportions across study samples.

A previous study showed that genes that are hypomethylated in the offspring of nonagenarians (compared to progeny of non-long-lived parents) were also associated with signal transmission [13], and our own results show that GO process terms, such as regulation of cell communication (GO:0010646) and regulation of signaling (GO:0023051), are associated with genes that contain methylation sites where the level of methylation is associated with paternal lifespan (Table 4). A review by Carlson et al. [32] discussed the association between aging and changes in signaling intensities in various signaling pathways (for example Notch-, TGFβ- and Wnt-signaling). These pathways function in an intertwined network, and proper regulation is needed to balance signaling during development and adult tissue maintenance and repair.

Location of CpGs associated with paternal age

Of the identified CpG sites associated with paternal lifespan, fewer were located in CpG-dense CpG islands than expected. They were instead enriched outside of CpG islands and in shores and shelves. It has been shown, in mice, that methylation level is associated with paternal environmental effects at CpG sites that are located in low-CG areas of the genome [37]. DNA methylation at CpG islands, and particularly at transcription start sites, is usually considered to be the more important regulator of gene expression [38], but the CpG-poor regions of the genome have also recently been proposed to be important for regulation [39, 40].

Aging and longevity are linked with development

We found that methylation sites that were associated with paternal lifespan were enriched in genes associated with development and morphogenesis. Aging associated hypermethylation has also been shown to be enriched in genes associated with development and morphogenesis [8, 13, 41, 42, 43].

The roles of developmental or metabolic rates in aging and longevity have been extensively studied, and caloric restriction, body size, changes in insulin signaling and the mTOR-pathway also have implied associations with longevity [44, 45, 46]. Alterations in the epigenetic mechanisms that control developmental processes may also contribute to lifespan. The role of developmental programs in aging is also a component of the quasi-programmed hyperfunction theory of aging, which states that aging is the aimless continuation of a developmental program when it is no longer needed [47]. Because these developmental programs are needed early in life, large alterations to these pathways would likely be deleterious. Thus, we expect that the DNA methylation changes identified in this study have only small effects on lifespan.

Mechanism of epigenetic inheritance

Both human and mouse studies have implied that certain traits acquired by a parent can be inherited by progeny, at least for one or two generations, and that some of these cases involve DNA methylation (see Introduction). The molecular mechanism explaining how inheritance through DNA methylation patterns occurs is still lacking, because the DNA methylome goes through two major reprogramming steps, first in the embryo and then in primordial germ cells [15, 48]. However, imprinted genes do have parent-of-origin-dependent expression patterns [49], and it has recently been shown that in mice, certain genomic regions at least partially avoid the reprogramming of the DNA methylome [50, 51]. Thus, transgenerational epigenetic inheritance is at least plausible in humans.

The DNA methylation features associated with paternal lifespan that were identified in this study may be intergenerationally inherited. However, we cannot exclude that the hereditary component may be another epigenetic feature, rather than DNA methylation, or traditional genetic element that contributes to the perceived DNA methylation pattern.

Both transgenerational epigenetic inheritance and the heritability of longevity and lifespan appear to be dependent on the sex of the parent and/or progeny, although reported results are inconsistent in the case of longevity and lifespan [5]. Our results show that the DNA methylation landscape and the number of longer-living siblings are associated only with paternal, and not maternal, lifespan. Our results therefore support the notion that there are sex differences in the heritability of lifespan. Due to the small study population, we were unable to identify the effects of paternal lifespan on the DNA methylome of daughters and sons separately, although sex was included as a covariate in the regression model. There is a female advantage in longevity, and females have better survival at all ages. Various mechanisms, including hormonal effects and differences in immune function (role of estrogen and androgens, susceptibility to infections [52,53]) as well as the role of X chromosome (skewing of X chromosome inactivation [54]) have been speculated to play a role, but definitive proof is lacking. Similarly, sexual dimorphism in the heritability of factors contributing to lifespan remain to be speculated [55]. Our results also indicate that paternal lifespan is not associated with the socio-economic status of the progeny, suggesting that this observed effect is not due to a shared environment.

CONCLUSIONS

In summary, we show that length of paternal lifespan is associated with progeny DNA methylation profiles and that this effect can be identified in nonagenarians. To our knowledge, the effects of the full range of parental lifespan on DNA methylation have not been previously analyzed. However, Gentilini et al. [13] did study the effect of extreme longevity in women. The methylation sites associated with paternal lifespan reported in the current study were located in genes associated with development and morphogenesis, as well as cell signaling. These results imply that these processes may be epigenetically regulating lifespan.

These results suggest that part of the "missing" heritability of lifespan may be epigenetic in nature. In addition to epigenetics, rare genetic variants most likely contribute to the heritability of lifespan. Because the length of lifespan is also significantly affected by environmental effects, lifestyle factors and interaction effects between environment and genetics, further studies are needed to uncover the genetic and epigenetic features that provide minor contributions to the heritability of lifespan.

MATERIALS AND METHODS

Study population

The study population consisted of 90 individuals born in 1920 (females n = 66, males n = 24) who participated in the home examinations in the Vitality 90+ Study in the year 2010. The study subjects included in this study were selected from the Vitality 90+ study cohort of 2010 based on two criteria: (i) information on both maternal and paternal lifespan was available and (ii) both parents had a lifespan of 40 years or more. The Vitality 90+ study is an on-going, prospective, population based study that includes both home dwelling and institutionalized individuals who are aged 90 years or more, and who live in the city of Tampere, Finland. The recruitment and characterization of participants were performed as has been reported in earlier Vitality 90+ study cohorts [22]. The study subjects were all of Western European descent and had not had any infections or received any vaccinations in the 30 days prior to blood sample collection. The study participants provided their written informed consent. This study was conducted according to the principles expressed in the declaration of Helsinki, and the study protocol was approved by the ethics committee of the city of Tampere (1592/403/1996).

Sample collection

Blood samples were collected into EDTA-containing tubes by a trained medical student during a home visit. All blood samples were drawn between 8 am and 12 am. Samples were directly subjected to leucocyte separation on a Ficoll-Paque density gradient (Ficoll-Paque™ Premium,

cat. no. 17-5442-03, GE Healthcare Bio-Sciences AB, Uppsala, Sweden). The PBMC layer was collected and was suspended in 1 ml of a freezing solution (5/8 FBS, 2/8 RPMI-160 medium, 1/8 DMSO) (FBS cat. no. F7524, Sigma-Aldrich, MO, USA; RPMI: cat. no. R0883, Sigma-Aldrich, MO, USA; DMSO: cat. no. 1.02931.0500, VWR, Espoo, Finland) and stored in liquid nitrogen.

Information on the age of death of parents and siblings and the age of living siblings was collected with a questionnaire at the home visit.

DNA extraction

DNA was extracted from PBMCs using the QIAamp DNA Mini kit (Qiagen, CA, USA), following the manufacturer's instructions for the spin protocol. The DNA was eluted in 60 µl of AE elution buffer and stored at -20°C. The concentration and quality of the DNA was assessed with the Qubit dsDNA HS Assay (Invitrogen, Eugene, OR, USA).

FACS

The proportions of different lymphocyte populations were determined through FACS analysis (BD FACSCanto II), and the results were analyzed with BD FACS Diva, version 6.1.3 (BD Biosciences, Franklin Lakes, NJ, USA). The antibodies employed in this analysis were FITC-CD14 (cat. no. 11-0149), PerCP-Cy5.5-CD3 (45-0037), APC-CD28 (17-0289) (eBioscience, San Diego, CA, USA), PE-Cy™7-CD4 (cat. no. 557852) and APC-Cy™7-CD8 (557834) (BD Biosciences).

Methylation array

Genome-wide DNA methylation profiling was performed at the Institute for Molecular Medicine Finland (FIMM) Technology Centre of the University of Helsinki in two batches (time interval, 6 months). Bisulfite conversion of 1 µg of DNA was performed using an EZ-96 DNA Methylation Kit (Zymo Research, Irvine, CA, USA) according to the manufacturer's instructions. A 4 µl aliquot of bisulfite-converted DNA was subjected to whole genome amplification and then enzymatically fragmented and hybridized to the Infinium HumanMethylation450 BeadChip (Illumina, San Diego, CA, USA) according to the manufacturer's protocol. Samples were assigned to the arrays in a randomized order. The BeadChips were scanned using an iScan reader (Illumina).

Processing of the methylation data

The data were processed as described previously [8], and can be accessed in GEO database (GSE58888) [56]. Before any processing, all unspecific or polymorphic sites (n = 76775) with minor allele frequency higher than 5%, based on database information [57], and probes mapping to sex chromosomes (n = 11 648) were removed. Methylation data were preprocessed as a methylumiset object using R software (R> = 2.15.3) with the wateRmelon array-specific package [58]. Technically poor quality samples and target sites were filtered out by excluding sites with a beadcount of < 3 in 5% of the samples (n = 515) and sites for which 1% of the samples had a detection p-value > 0.05 (n = 698). Background correction and quantile normalization were conducted individually for the two chemistries (Infinium I and II) as well as for the intensities of methylation (m) and un-methylation (u) using the dasen method. After dasen normalization, the u and m intensities were transformed to values of beta (β). β is the ratio of methylated probe (m) intensities to overall intensities (m+u+α), where α is the constant offset, 100. Thus, β ranges linearly between 0 (non-methylated, 0%) and 1 (completely methylated, 100%). Next, the batch effect of the Infinium chemistries was adjusted using the BMIQ algorithm, which is based on beta mixture-models and the EM-algorithm [59]. Several visualization styles were used to verify the quality of the data, including boxplots from the raw intensities, Kernel density plots in the chemistry correction procedure and PCA (principal component analysis).

Detection of methylation sites associated with parental age

To assess the relationship between site-specific methylation level and the age of the father/mother at the time of death, a generalized regression model, referred to as variable dispersion beta regression [60, 61], was utilized on each CpG site. The age of the father and mother at the time of death (linear variable) and the gender (categories 0 and 1) of the subject were employed as predictors of the site-specific methylation outcome in the form of β-values (ranging from 0 to 1) in each equation, where the mean model with a linker function of logit was utilized. Furthermore, as was previously observed, because methylation levels fluctuate based on the composition of blood cell subtypes [8,62], the variables corresponding to cell type proportions (the CD4+ to CD8+ ratio and the proportions of CD28/CD4+ and CD28-/CD8+ and CD14+ cells) were included as linear covariates in the model. The bias caused by the batch effect of two laboratory days (time interval of 6 months) was also confirmed by PCA. Therefore, a variable corresponding to the batches (categories 0 and 1) was set as covariate in the model. The nominal Benjamini-Hochberg corrected p-value was set to 0.05. Next, the CpG sites with substantial differences in methylation level between the extreme age groups were extracted. The subjects were categorized to groups FI, FII,

FIII (and MI, MII, MIII), with equal group sizes according to the age of the father/mother at the time of death (see Table 1). The extraction procedure was conducted by calculating the difference in median values of methylation in each CpG site for groups I and III, and only sites with $-0.01>\Delta\beta>0.01$ were included for further analysis.

Pathway analyses

Pathway analyses were performed on genes harboring CpG sites where the level of methylation was associated with paternal lifespan. The 659 CpG sites were located in 422 different genes.

IPA [23] was used to identify canonical pathways associated with the identified genes. According to the manufacturer, these canonical pathways are well characterized metabolic and cell signaling pathways that have been curated and hand-drawn by PhD-level scientists. All of the data sources provided by the Ingenuity Knowledge Base were included in the IPA, and the Ingenuity Knowledge Base was used as the reference set in all analyses. For the association of molecules, only experimentally observed results were accepted, and only human data were considered. A Benjamini-Hochberg multiple testing corrected *p*-value of < 0.05 was used as the threshold for significance. The Ingenuity pathway analysis was performed on the 12th of March 2015.

GOrilla [24, 25] was used to identify the enriched GO terms for the identified genes. GO terms were searched based on two unranked lists (target and background), and all genes with at least one probe in the 450K array were used as the background list. A Benjamini-Hochberg multiple testing corrected *p*-value of < 0.05 was used as the threshold for significance. GOrilla analysis was performed on the 29th of April 2015.

ACKNOWLEDGMENTS

We would like to thank Sinikka Repo-Koskinen, Janette Hinkka, Katri Välimaa, Sanna Tuominen and Tuomas Marttila for their skillful technical assistance.

GRANT SUPPORT

This work was supported by grants from the Tampere Tuberculosis Foundation (M.H.), Yrjö Jahnsson Foundation (M.H.), the Academy of Finland (132704 to M.H) and the Competitive Research Fund of Pirkanmaa Hospital District (9M017, 9N013 to M.H.; 9N012 to A.H.).

CONFLICTS OF INTEREST

The authors declare no conflicts of interest.

REFERENCES

1. Reed T, Dick DM. Heritability and validity of healthy physical aging (wellness) in elderly male twins. Twin Res. 2003; 6: 227-234.
2. Murabito JM, Yuan R, Lunetta KL. The search for longevity and healthy aging genes: insights from epidemiological studies and samples of long-lived individuals. J. Gerontol. A. Biol. Sci. Med. Sci. 2012; 67: 470-479.
3. Brooks-Wilson AR. Genetics of healthy aging and longevity. Hum. Genet. 2013; 132: 1323-1338.
4. Shadyab AH, LaCroix AZ. Genetic factors associated with longevity: a review of recent findings. Ageing Res. Rev. 2015; 19:1-7.
5. Dutta A, Henley W, Robine JM, Langa KM, Wallace RB, Melzer D. Longer lived parents: protective associations with cancer incidence and overall mortality. J. Gerontol. A. Biol. Sci. Med. Sci. 2013; 68: 1409-1418.
6. Johnson AA, Akman K, Calimport SR, Wuttke D, Stolzing A, de Magalhães JP. The role of DNA methylation in aging, rejuvenation, and age-related disease. Rejuvenation Res. 2012; 15: 483-494.
7. Bacalini MG, Friso S, Olivieri F, Pirazzini C, Giuliani C, Capri M, Santoro A, Franceschi C, Garagnani P. Present and future of anti-ageing epigenetic diets. Mech. Ageing. Dev. 2014; 136-137: 101-115.
8. Marttila S, Kananen L, Häyrynen S, Jylhävä J, Nevalainen T, Hervonen A, Jylhä M, Nykter M, Hurme M. Ageing-associated changes in the human DNA methylome: genomic locations and effects on gene expression. BMC Genomics. 2015; 16: 179.
9. D'Aquila P, Rose G, Bellizzi D, Passarino G. Epigenetics and aging. Maturitas. 2013; 74: 130-136.
10. Horvath S. DNA methylation age of human tissues and cell types. Genome Biol. 2013; 14: R115.
11. Steegenga WT, Boekschoten MV, Lute C, Hooiveld GJ, de Groot PJ, Morris TJ, Teschendorff AE, Butcher LM, Beck S, Müller M. Genome-wide age-related changes in DNA methylation and gene expression in human PBMCs. Age (Dordr). 2014; 36: 9648.
12. Fraga MF, Ballestar E, Paz MF, Ropero S, Setien F, Ballestar ML, Heine-Suñer D, Cigudosa JC, Urioste M, Benitez J, Boix-Chornet M, Sanchez-Aguilera A, Ling C, et al. Epigenetic differences arise during the lifetime of monozygotic twins. Proc. Natl. Acad. Sci. U. S. A. 2005; 102: 10604-10609.
13. Gentilini D, Mari D, Castaldi D, Remondini D, Ogliari G, Ostan R, Bucci L, Sirchia SM, Tabano S, Cavagnini F, Monti D, Franceschi C, Di Blasio AM, Vitale G. Role of epigenetics in human aging and longevity: genome-wide DNA methylation profile in centenarians and centenarians' offspring. Age (Dordr) 2013; 35: 1961-1973.
14. Anway MD, Cupp AS, Uzumcu M, Skinner MK. Epigenetic transgenerational actions of endocrine disruptors and male

fertility. Science. 2005; 308: 1466-1469.

15. Szyf M. Nongenetic inheritance and transgenerational epigenetics. Trends. Mol. Med. 2015; 21: 134-144.
16. Martos SN, Tang WY, Wang Z. Elusive inheritance: Transgenerational effects and epigenetic inheritance in human environmental disease. Prog. Biophys. Mol. Biol. 2015; 118: 44-54.
17. Sarkar DK. Male germline transmits fetal alcohol epigenetic marks for multiple generations: a review. Addict Biol. 2015 Jan 12. doi: 10.1111/adb.12186. [Epub ahead of print].
18. Grossniklaus U, Kelly WG, Ferguson-Smith AC, Pembrey M, Lindquist S. Transgenerational epigenetic inheritance: how important is it? Nat. Rev. Genet. 2013; 14: 228-235.
19. Painter RC, Roseboom TJ, Bleker OP. Prenatal exposure to the Dutch famine and disease in later life: an overview. Reprod. Toxicol. 2005; 20: 345-352.
20. Heijmans BT, Tobi EW, Stein AD, Putter H, Blauw GJ, Susser ES, Slagboom PE, Lumey LH. Persistent epigenetic differences associated with prenatal exposure to famine in humans. Proc. Natl. Acad. Sci. U. S. A. 2008; 105: 17046-17049.
21. Suderman M, Borghol N, Pappas JJ, Pinto Pereira SM, Pembrey M, Hertzman C, Power C, Szyf M. Childhood abuse is associated with methylation of multiple loci in adult DNA. BMC Med. Genomics. 2014; 7: 13.
22. Goebeler S, Jylhä M, Hervonen A. Medical history, cognitive status and mobility at the age of 90. A population based study in Tampere, Finland. Aging Clin. Exp. Res. 2003; 15: 154-161.
23. IPA®, QIAGEN Redwood City, www.qiagen.com/ingenuity
24. Eden E, Lipson D, Yogev S, Yakhini Z. Discovering motifs in ranked lists of DNA sequences. PLoS Comput. Biol. 2007; 3: e39.
25. Eden E, Navon R, Steinfeld I, Lipson D, Yakhini Z. GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. BMC Bioinformatics. 2009; 10: 48.
26. Kim MS, Yoon SK, Bollig F, Kitagaki J, Hur W, Whye NJ, Wu YP, Rivera MN, Park JY, Kim HS, Malik K, Bell DW, Englert C, et al. A novel Wilms tumor 1 (WT1) target gene negatively regulates the WNT signaling pathway. J. Biol. Chem. 2010; 285: 14585-14593.
27. Andersson T, Södersten E, Duckworth JK, Cascante A, Fritz N, Sacchetti P, Cervenka I, Bryja V, Hermanson O. CXXC5 is a novel BMP4-regulated modulator of Wnt signaling in neural stem cells. J. Biol. Chem. 2009; 284: 3672-3681.
28. Kim HY, Yoon JY, Yun JH, Cho KW, Lee SH, Rhee YM, Jung HS, Lim HJ, Lee H, Choi J, Heo JN, Lee W, No KT, et al. CXXC5 is a negative-feedback regulator of the Wnt/β-catenin pathway involved in osteoblast differentiation. Cell Death Differ. 2015; 22: 912-920.
29. Kim HY, Yang DH, Shin SW, Kim MY, Yoon JH, Kim S, Park HC, Kang DW, Min D, Hur MW, Choi KY. CXXC5 is a transcriptional activator of Flk-1 and mediates bone morphogenic protein-induced endothelial cell differentiation and vessel formation. FASEB J. 2014; 28: 615-626.
30. Pendino F, Nguyen E, Jonassen I, Dysvik B, Azouz A, Lanotte M, Ségal-Bendirdjian E, Lillehaug JR. Functional involvement of RINF, retinoid-inducible nuclear factor (CXXC5), in normal and tumoral human myelopoiesis. Blood. 2009; 113: 3172-3181.
31. Long HK, Blackledge NP, Klose RJ. ZF-CxxC domain-containing proteins, CpG islands and the chromatin connection. Biochem. Soc. Trans. 2013; 41: 727-740.
32. Carlson ME, Silva HS, Conboy IM. Aging of signal transduction pathways, and pathology. Exp. Cell. Res. 2008; 314: 1951-1961.
33. Buas MF, Kadesch T. Regulation of skeletal myogenesis by Notch. Exp. Cell. Res. 2010; 316: 3028-3033.
34. Arthur ST, Cooley ID. The effect of physiological stimuli on sarcopenia; impact of Notch and Wnt signaling on impaired aged skeletal muscle repair. Int. J. Biol. Sci. 2012; 8: 731-760.
35. Kogut I, Scholz JL, Cancro MP, Cambier JC. B cell maintenance and function in aging. Semin. Immunol. 2012; 24: 342-349.
36. Scholz JL, Diaz A, Riley RL, Cancro MP, Frasca D. A comparative review of aging and B cell function in mice and humans. Curr. Opin. Immunol. 2013; 25: 504-510.
37. Skinner MK, Guerrero-Bosagna C. Role of CpG deserts in the epigenetic transgenerational inheritance of differential DNA methylation regions. BMC Genomics. 2014; 15: 692.
38. Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nat. Rev. Genet. 2012; 13: 484-492.
39. Hahn MA, Wu X, Li AX, Hahn T, Pfeifer GP. Relationship between gene body DNA methylation and intragenic H3K9me3 and H3K36me3 chromatin marks. PLoS One. 2011; 6: e18844.
40. Irizarry RA, Ladd-Acosta C, Wen B, Wu Z, Montano C, Onyango P, Cui H, Gabo K, Rongione M, Webster M, Ji H, Potash JB, Sabunciyan S, Feinberg AP. The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. Nat. Genet. 2009; 41: 178-186.
41. Johansson A, Enroth S, Gyllensten U. Continuous aging of the human DNA methylome throughout the human lifespan. PLoS One. 2013; 8: e67378.
42. Florath I, Butterbach K, Muller H, Bewerunge-Hudler M, Brenner H. Cross-sectional and longitudinal changes in DNA methylation with age: an epigenome-wide analysis revealing over 60 novel age-associated CpG sites. Hum. Mol. Genet. 2013; 23: 1186–1201.
43. Rakyan VK, Down TA, Maslau S, Andrew T, Yang TP, Beyan H, Whittaker P, McCann OT, Finer S, Valdes AM, Leslie RD, Deloukas P, Spector TD. Human aging-

associated DNA hypermethylation occurs preferentially at bivalent chromatin domains. Genome Res. 2010; 20: 434-439.

44. Bartke A. Healthy aging: is smaller better? - a mini-review. Gerontology. 2012; 58: 337-343.
45. Lamming DW. Diminished mTOR signaling: a common mode of action for endocrine longevity factors. Springerplus. 2014; 3: 735.
46. Ingram DK, Roth GS. Calorie restriction mimetics: can you have your cake and eat it, too? Ageing Res. Rev. 2015; 20: 46-62.
47. Blagosklonny MV. Aging is not programmed: genetic pseudo-program is a shadow of developmental growth. Cell Cycle. 2013; 12: 3736-3742.
48. Heard E, Martienssen RA. Transgenerational epigenetic inheritance: myths and mechanisms. Cell. 2014; 157: 95-109.
49. Hanna CW, Kelsey G. The specification of imprints in mammals. Heredity (Edinb) 2014; 113: 176-183.
50. Hackett JA, Zylicz JJ, Surani MA. Parallel mechanisms of epigenetic reprogramming in the germline. Trends Genet. 2012; 28: 164-174.
51. Seisenberger S, Andrews S, Krueger F, Arand J, Walter J, Santos F, Popp C, Thienpont B, Dean W, Reik W. The dynamics of genome-wide DNA methylation reprogramming in mouse primordial germ cells. Mol. Cell. 2012; 48: 849-862.
52. Nunn CL, Lindenfors P, Pursall ER, Rolff J. On sexual dimorphism in immune function. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 2009; 364: 61-69.
53. Seifarth JE, McGowan CL, Milne KJ. Sex and life expectancy. Gend. Med. 2012; 9: 390-401.
54. Gentilini D, Castaldi D, Mari D, Monti D, Franceschi C, Di Blasio AM, Vitale G. Age-dependent skewing of X chromosome inactivation appears delayed in centenarians' offspring. Is there a role for allelic imbalance in healthy aging and longevity? Aging Cell. 2012; 11: 277-283.
55. Newman AB, Murabito JM. The epidemiology of longevity and exceptional survival. Epidemiol. Rev. 2013; 35: 181-197.
56. Gene Expression Omnibus, http://www.ncbi.nlm.nih.gov/geo/
57. Chen YA, Lemire M, Choufani S, Butcher DT, Grafodatskaya D, Zanke BW, Gallinger S, Hudson TJ, Weksberg R. Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. Epigenetics. 2013; 8: 203-209.
58. Pidsley R, Y Wong CC, Volta M, Lunnon K, Mill J, Schalkwyk LC. A data-driven approach to preprocessing Illumina 450K methylation array data. BMC Genomics. 2013; 14: 293.
59. Teschendorff AE, Marabita F, Lechner M, Bartlett T, Tegner J, Gomez-Cabrero D, Beck S (2013) A beta-mixture quantile normalization method for correcting probe design bias in Illumina Infinium 450 k DNA methylation data. Bioinformatics. 2013; 29: 189-196.
60. Cribari-Neto F, Zeileis A. Beta regression in R. JSS. 2010; 34: 1-24.
61. Ferrari C (2004) Beta regression for modelling rates and proportions. J. Appl. Stat. 2004; 31: 799-815.
62. Jaffe AE, Irizarry RA. Accounting for cellular heterogeneity is critical in epigenome-wide association studies. Genome Biol. 2014; 15: R31.