Healthy at 100:
Genetic factors and phenotypes associated with healthy aging and longevity

by

Lauren C. Tindale

B.Sc., Simon Fraser University, 2012

Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of
Doctor of Philosophy

in the
Department of Biomedical Physiology and Kinesiology
Faculty of Science

© Lauren C. Tindale 2018
SIMON FRASER UNIVERSITY
Summer 2018

Copyright in this work rests with the author. Please ensure that any reproduction or re-use is done in accordance with the relevant national copyright legislation.

Approval

Name:	**Lauren C. Tindale**
Degree:	**Doctor of Philosophy (Biomedical Physiology and Kinesiology)**
Title:	**Healthy at 100: Genetic factors and phenotypes associated with healthy aging and longevity**
Examining Committee:	**Chair:** Sam Doesburg Associate Professor
	Angela Brooks-Wilson Senior Supervisor Professor
	Marco Marra Supervisor Professor Department of Medical Genetics University of British Columbia
	Ryan Morin Supervisor Assistant Professor Department of Molecular Biology and Biochemistry
	Miriam Rosin Internal Examiner Professor Department of Biomedical Physiology and Kinesiology
	Bradley Willcox External Examiner Professor Department of Geriatric Medicine University of Hawaii at Manoa
Date Defended/Approved:	July 27, 2018

Ethics Statement

The author, whose name appears on the title page of this work, has obtained, for the research described in this work, either:

a. human research ethics approval from the Simon Fraser University Office of Research Ethics

or

b. advance approval of the animal care protocol from the University Animal Care Committee of Simon Fraser University

or has conducted the research

c. as a co-investigator, collaborator, or research assistant in a research project approved in advance.

A copy of the approval letter has been filed with the Theses Office of the University Library at the time of submission of this thesis or project.

The original application for approval and letter of approval are filed with the relevant offices. Inquiries may be directed to those authorities.

Simon Fraser University Library
Burnaby, British Columbia, Canada

Update Spring 2016

Abstract

Background. Healthy aging and longevity are distinct phenotypes that describe an individual's health span and life span, respectively. Both are complex phenotypes that are influenced by lifestyle, environment, and genetics. Using Super-Seniors, individuals aged 85 years and older and free of major chronic disease, and mid-life controls, I investigated whether Super-Seniors have an overall decreased genetic susceptibility and increased resilience to age- related diseases.

Results. Using lifestyle, health, and genetic data, I compared the phenotypic and genetic characteristics of Super-Seniors to mid-life controls. Female Super-Seniors had an older age of last fertility, and were more likely to have had a child at ≥40 years. Super-Senior parents also exceeded the life expectancy for their era by a decade. Super-Seniors performed well on geriatric tests when compared to other long-lived populations. As well, an even more select group of Super-Seniors who survived 10 years after their initial recruitment still exhibited high function and good health.

Exome sequencing of two centenarian brothers suggested that long-lived individuals do not carry a decreased burden of common complex disease variants. Instead, protective buffering variants may play a role in healthy aging. In an analysis of buffering candidates, Super-Seniors were less likely than controls to carry an *APOEε4* allele or a haptoglobin *HP2* allele, and I identified 3 potential gene-gene interactions. In a network analysis of candidate buffering genes, lipid and cholesterol metabolism was a common theme. Further exploring the potential for protective factors in the Super-Seniors, I looked for evidence of allele-specific expression among disease genes, however no differences between groups were found.

Conclusions. Super-Seniors are cognitively and physically high functioning individuals who have evaded major age-related chronic diseases into old age. The familiality of long lifespan of the parents of Super-Seniors supports the hypothesis that heritable factors contribute to this desirable phenotype. Although Super-Seniors did not appear to carry a decreased disease burden, there was some evidence that they may carry protective factors. The Super-Seniors are a phenotypically healthy group in which to look for further genetic markers of healthy aging and longevity.

Acknowledgements

First of all, thank you to my supervisor Angie. She took me on as an undergraduate student when I knew nothing about genetics and helped me come out the other side with a PhD. A big thank you to our lab guru Steve who taught me everything I know in the lab and listened to my many rants and crises. And thank you to everyone who kept me (relatively) sane through the ups and downs of PhD life.

Table of Contents

Approval ... ii
Ethics Statement ... iii
Abstract ... iv
Acknowledgements ... v
Table of Contents ... vi
List of Tables ... ix
List of Figures ... x
List of Acronyms ... xi
Preface ... xiii

Chapter 1. Background ... 1
1.1. Aging and disease in society ... 1
1.2. Genetics of healthy aging and longevity ... 2
1.3. Hallmarks of aging ... 6
1.4. Model organism genetics ... 9
1.5. Next-generation sequencing ... 10
1.6. Genetic studies of aging and longevity ... 11
1.7. The Super-Senior Study ... 14

Chapter 2. The Super-Seniors Study: Phenotypic characterization of a healthy 85+ population ... 17
2.1. Introduction ... 18
2.2. Design and Methods ... 19
2.2.1. Recruitment ... 19
2.2.2. Data collection ... 20
2.2.3. Phenotypic review ... 20
2.2.4. Statistical analysis ... 21
2.3. Results ... 21
2.3.1. Recruitment ... 21
2.3.2. Descriptive statistics ... 21
2.3.3. Functional tests ... 22
2.3.4. Lifestyle ... 22
2.3.5. Number of offspring and age of fertility ... 23
2.3.6. The parents of Super-Seniors lived longer than their contemporaries ... 23
2.4. Discussion ... 24

Chapter 3. 10-year follow-up of the Super-Seniors Study: Compression of morbidity and genetic factors in survivors ... 31
3.1. Introduction ... 32
3.2. Design and Methods ... 33
3.3. Results ... 33
3.4. Discussion ... 35

Chapter 4. Burden of common complex disease variants in the exomes of two healthy centenarian brothers 42
4.1. Introduction 43
4.2. Design and Methods 43
4.2.1. Subjects 43
4.2.2. GWAS variants 44
4.2.3. High penetrance variants 44
4.2.4. Mitochondrial analysis 45
4.3. Results 45
4.3.1. GWAS variants 45
4.3.2. High penetrance variants 45
4.3.3. Mitochondrial analysis 46
4.4. Discussion 46

Chapter 5. The search for protective buffering variants: Lipid and Alzheimer's Disease genes associated with healthy aging and longevity in healthy oldest-old 51
5.1. Introduction 51
5.2. Design and Methods 52
5.2.1. Subjects 52
5.2.2. Literature search 53
5.2.3. Genotyping 53
5.2.4. Association tests of individual variants 54
5.2.5. Gene-gene interaction analysis 54
5.2.6. Network analysis 54
5.3. Results 55
5.3.1. Candidate variants 55
5.3.2. Genotypes and Quality Control 55
5.3.3. Association tests of individual variants 55
5.3.4. Gene-gene interaction analysis 56
5.3.5. Network analysis 57
5.4. Discussion 58

Chapter 6. Allele-specific expression: A pilot study in healthy centenarians 67
Introduction 68
6.1. Design and Methods 69
6.1.1. Subjects 69
6.1.2. Sample preparation 69
6.1.3. Cell fraction prediction 69
6.1.4. Variant filtering 70
6.1.5. Analysis of ASE 70
6.2. Results 70
6.2.1. Cell fraction prediction 70
6.2.2. Variant filtering 70
6.2.3. Distribution and function of variants with ASE 71
6.2.4. Disease association of genes and variants with ASE 71
6.3. Discussion 72

Chapter 7. Discussion and Conclusions 79
7.1. The contribution of this thesis to the field of healthy aging 79
7.2. Future directions 82
7.2.1. Sample size and power 82
7.2.2. Sick elderly as a control group 84
7.2.3. Gut microbiome of the Super-Seniors 84
7.2.4. Epigenetics 85
7.2.5. Translation of genetics to the clinic 86

References. 88

Appendices. 112
Appendix A. 113
A.1. Life expectancy in Canada 113
A.2. Collection of Super-Seniors and controls from BC Medical Services Plan lists 114
A.3. Age distribution of participants in the Super-Seniors Study 115
A.4. Distribution of ethnicity in Super-Seniors and controls 116
A.5. Contingency table of smoking status in Super-Seniors and controls 117
A.6. Contingency table of the number of offspring in Super-Seniors and controls 118
Appendix B. 119
B.1. Follow-up of participants in the Super-Seniors Study 119
Appendix C. 120
C.1. Supplementary methods: Exome sequencing 120
C.2. Supplementary methods: Mitochondrial sequencing analysis 121
C.3. List of 73 filtered GWAS variants from the National Human Genome Research Institute catalogue 122
C.4. Homoplasmic variants found in the mitochondrial genomes of the brothers 123
Appendix D. 125
D.1. *p*-values for main effects examined 125
D.2. *p*-values for interaction tests identified a priori 126
D.3. Interaction odds ratios and 95% confidence intervals between variants 127
Appendix E. 128
E.1. Proportion of immune cell types as estimated by CIBERSORT by sample. 128
E.2. Proportion of variants showing allele-specific expression by predicted effect. 129
E.3. Estimated average cell fractions of immune cell types from CIBERSORT. The signature gene file of 22 distinct immune cell types provided by CIBERSORT was used as a reference for gene expression signatures. 130
E.4. List of 35 top biased centenarian and 23 top biased control genes. Top biased centenarian genes were defined as genes were there was evidence of ASE for a particular gene in 4 centenarians and only 1 or no controls, or in 3 centenarians and no controls; and vice versa for top biased control genes 131

List of Tables

Table 1-1. Genetic studies of healthy aging and/or longevity. 16

Table 2-1. Characteristics of the study participants. 28

Table 3-1. Characteristics of Super-Seniors at two interviews approximately 10 years apart 39

Table 3-2. Genotype comparison between Super-Senior survivors, and the original Phase 1 collection of Super-Seniors and controls. 40

Table 4-1. Predicted damaging SNPs found in the brothers in ACMG genes 49

Table 5-1. Candidate genes and candidate epistatic variants 62

Table 5-2. Genotype counts and frequencies in Super-Seniors and controls 63

Table 5-3. Odds ratios and 95% confidence intervals for the association between variants in *APOE* and *HP* and healthy aging 64

Table 5-4. The top 20 functions and diseases represented in a candidate gene network in IPA® 65

Table 6-1. NHGRI-EBI GWAS catalog SNPs with evidence of allele-specific expression in centenarians of controls. 75

Table 6-2. Top canonical pathways and top disease categories among genes with allele-specific expression as determined by QIAGEN Ingenuity® Pathway Analysis. 76

Table 6-3. ACMG genes with evidence of allele-specific expression in centenarians of controls.. 77

List of Figures

Figure 2-1. Recruitment of Super-Seniors and controls.. 29

Figure 2-2. Distribution of functional test scores of the Super-Seniors. 30

Figure 4-1. Proportion of risk genotypes and alleles in the brothers and other exome data sets.. 50

Figure 5-1. A network including 15 candidate epistatic longevity genes 66

Figure 6-1. A) Proportion of major and minor allele bias by group. B) Coding region variants by major and minor allele bias (n = 3700 coding variants with ASE ≥ 0.7:0.3). C) Coding region variants by major and minor allele bias with an alternate allele frequency > 0.9 or < 0.1 (n = 360 coding variants with ASE ≥ 0.9:0.1). 78

List of Acronyms

Aβ	Amyloid-beta
ACMG	American College of Medical Genetics and Genomics
AD	Alzheimer disease
ASE	Allele-specific expression
BWA	Burrows-Wheeler aligner
cDNA	Complementary DNA
dTTP	Deoxythymidine triphosphate
dUTP	Deoxyuridine Triphosphate
CEU	Utah Residents with Northern and Western European Ancestry
CHD	Coronary heart disease
CI	Confidence interval
CNS	Central nervous system
CVD	Cardiovascular disease
DNA	Deoxyribonucleic acid
DNAm	DNA methylation
EBV	Epstein-Barr virus
FDA	Food and drug administration
FDR	False discovery rate
GDS	Geriatric depression scale
GM	Gut microbiome
GWAS	Genome-wide association scan
HDL	High density lipoprotein
IADL	Instrumental activities of daily living
INDEL	Insertion or deletion
IPA	Ingenuity® Pathway Analysis
LCL	Lymphoblastoid cell lines
LDL	Low density lipoprotein
LOS	Length of stay
MAF	Minor allele frequency
MMSE	Mini-mental status exam
mRNA	Messenger RNA
NMN	Nicotinamide mononucleotide

RBC	Red blood cell
RD	Read depth
RNA	Ribonucleic acid
RNA-seq	Methodology for sequencing RNA by NGS
NGS	Next-generation sequencing
NHLBI	National Heart, Lung, and Blood Institute
OR	Odds ratio
PCR	Polymerase chain reaction
RD	Read depth
REB	Research ethics board
RNA	Ribonucleic acid
RNA-seq	RNA sequencing
ROS	Reactive oxygen species
SD	Standard deviation
SNP	Single nucleotide polymorphism
TAME	Targeting Aging with Metformin
TUG	Timed up and go
UTR	Untranslated region
WB	Whole blood
WBC	White blood cell

Preface

Chapter 2 has been published in near-identical form as: Halaschek-Wiener J*, Tindale LC*, Collins JA, Leach S, McManus B, Madden K, Meneilly G, Le ND, Connors JM, Brooks-Wilson AR. The Super-Seniors Study: Phenotypic characterization of a healthy 85+ population. PLOS One. 2018; 13(5): e0197578. PLOS applies the Creative Commons Attribution (CC BY) license to works they publish.

A version of **Chapter 3** has been submitted for publication as: Tindale LC, Salema D, Brooks-Wilson AR. 10-year follow-up of the Super-Seniors Study: Compression of morbidity and genetic factors in survivors.

Chapter 4 has been published in near identical form as: Tindale LC, Zeng A, Bretherick KL, Leach S, Thiessen N and Brooks-Wilson AR. Burden of Common Complex Disease Variants in the Exomes of Two Healthy Centenarian Brothers. Gerontology. 2016; 62(1):58-62 DOI:10.1159/000430462. Written permission was received from S. Karger AG, Basel to republish this article in this thesis.

Chapter 5 has been published in near identical form as: Tindale LC et al. Lipid and Alzheimer's Disease genes associated with healthy aging and longevity in healthy oldest-old. Oncotarget. 2017; 8(13):pp: 20612-20621 DOI:10.18632/oncotarget.15296. Oncotarget applies the Creative Commons Attribution 3.0 License (CC BY 3.0) to all works they publish.

Chapters 2-6 are part of the Super-Seniors Study, which was funded and collected by Angela Brooks-Wilson. She was also involved in study design and editorial production of all chapters.

For **Chapter 2** I reviewed the aging database for the Super-Senior Study and requested query parameters from our study coordinators, Ruth Thomas and Diane Salema. I also examined the original hard-copy documents of the questionnaires when there were discrepancies in the database or to better understand how the questions were asked. Using information from the database regarding health and personal history, I characterized the study participants and ran comparisons between the Super-Seniors and controls. The

original collection of the Super-Seniors Study was done by Dr. Julius Halaschek-Wiener and the Brooks-Wilson laboratory.

For **Chapter 3**, Diane Salma coordinated the follow-up interviews. I compiled interview results, performed the statistical analysis, and wrote the manuscript.

For **Chapter 4** I completed the analysis of the two centenarian brothers' exomes, including alignment, variant calling, and examination of findings, and wrote the manuscript. Bioinformatics guidance was provided by Nina Thiessen. Andy Zeng performed the mitochondrial analysis.

In **Chapter 5** I conducted the literature search for buffering and epistatic variants, which were then sent for genotyping at the Genome Quebec Innovation Centre. For variants that could not be genotyped by this method, I genotyped them by TaqMan assay or PCR. I performed the statistical analysis with the help of Dr. John Spinelli and wrote the manuscript.

For **Chapter 6** I did the sample preparation, including RNA extraction and library construction, and analysis of sequencing results. Alignment and variant calling was performed by the bioinformatics team at Canada's Michael Smith Genome Science Centre.

Chapter 1.

Background

1.1. Aging and disease in society

We live in an aging society that has a longer average life expectancy than ever before in human history. According to Canadian demographic data [1], the life expectancy at birth in 1921 (the earliest year this data is available) was 58.16 years for women and 55.95 years for men, and has steadily increased so that in 2011 the life expectancy at birth for women was 83.9 years for women and 79.52 years for men (Appendix A.1). Life expectancy has increased due to improvements in sanitation, health care, education, and scientific breakthroughs [2], however with this increase in aged individuals, there has also been an increase in chronic disease prevalence and the number of years spent with disability [3]. In the US, the Centre for Disease Control asked adults to report whether they had the following chronic conditions: hypertension, coronary heart disease, stroke, diabetes, cancer, arthritis, hepatitis, weak or failing kidneys, chronic obstructive pulmonary disorder, or current asthma [4]. Between 2002 and 2013, individuals 45-64 years old who reported having 0-1 chronic conditions decreased from 71.4% to 67.7%, while the incidence of having 2-3 chronic conditions increased from 24.2% to 26.6%, and the incidence of have 4+ chronic conditions increased from 4.4% to 5.6%. Over this same time period, the number of individuals 75 years and older who reported having 0-1 chronic conditions decreased from 41.4% to 34.4%, while the incidence of having 2-3 chronic conditions increased from 45.8% to 48.9%, and the incidence of have 4+ chronic conditions increased from 13.3% to 16.7%. It is also interesting that the survey does not differentiate between having 0 or 1 chronic condition, suggesting that having one chronic condition is still considered to be “good” health.

The increase in chronic disease prevalence is a burden on our health care system and on taxpayers. For example, the Canadian Institute for Health Information’s Patient Cost

Estimator [5] showed that among patients 80+ years in Canada from 2012-13, chronic obstructive pulmonary disease was the most treated patient condition, which cost an average of $7,527 per patient with an average acute length of stay (LOS) of 7.5 days, followed by heart failure without a coronary angiogram, which cost an average of $7,358 with an average acute LOS of 7.6 days. Other common patient conditions included dementia, which cost an average of $15,916 and resulted in an average acute LOS of 17 days, and myocardial infarction/shock/arrest without cardioangiogram, which cost an average of $8,328 with an average acute LOS of 7.0 days. Clearly, if we could pre-emptively reduce the number of patients who develop these types of chronic diseases, this would decrease hospital costs. The study of healthy aging and longevity not only aims to help people live longer, but more importantly to live healthy and disease-free into those years.

1.2. Genetics of healthy aging and longevity

Healthy aging and longevity are distinct phenotypes that describe an individual's health span and life span, respectively. Both are complex phenotypes that are influenced by lifestyle, environment, and genetics. Estimates for heritability range between 15% to 30% in twin and population-based studies [6]. The genetic contribution to survival is suggested to be minimal before the age of 60, but becomes greater beyond this age [7]. Many candidates for longevity associated genes and pathways are thought to enhance the capacity to handle stress-induced molecular damage, and conversely, mutations in these pathways may lead to age-related susceptibility to cellular damage and disease [8]. However, identifying specific longevity related genes and variants has proven to be a challenge. While genes that affect lifespan have been found in model organisms, few findings have been transferable to humans [9].

Polygenic traits can have large numbers of associated single nucleotide polymorphisms (SNPs) that each have a small but important contribution to the trait. For example, in type 2 diabetes, 38 associated variants that have odds ratios from 1.06 to 1.40 have been estimated to explain about 10% of the observed heritability of the disease [10]. As well, although it is generally estimated that ~80% of human height is heritable, one study found that using 180 SNPs explains only ~10% of the variability in height [11]. Another study used ~300,000 SNPs in their model for variation in height and found that they could explain

~45%, therefore suggesting that some “missing” heritability is a result of very small individual effects that were previously unobserved or that could not be found until sample sizes became sufficiently large [12].

The major replicable genetic association in human longevity is apolipoprotein E (*APOE*), which has been reported in multiple genome-wide association scans (GWAS) [13, 14]. There are two single nucleotide polymorphisms (SNPs) in *APOE* that result in three distinct isoforms. *APOE ε2* has been associated with decreased risk for mortality [15-17], *APOE ε3* is considered to be neutral, and *APOE ε4* has been associated with increased risk for mortality as well as Alzheimer Disease (AD) [18]. *APOE ε4* frequency ranges from 0.09 to 0.30 in different populations, and *APOE* haplotype risk is consistent with a semi-dominant inheritance pattern [19]. *APOE* and *FOXO3* variants have also been consistently seen in candidate gene studies for longevity [9, 20] and in a GWAS meta-analysis for longevity [21]. *FOXO3* is and evolutionarily conserved gene that is involved in the insulin/insulin-like growth factor 1 signaling pathway. Mutations in genes in this pathway have been seen to affect lifespan in model organisms [22-24] however *FOXO3* is the only gene in this pathway associated with longevity in humans to date [25, 26]. *FOXO3* variants were found to be associated with longevity in Japanese American men (aged 95 years and older) [25] and in German centenarians [27], and to have an effect on risk for stroke, cardiovascular mortality, and all-cause mortality in the Leiden 85-plus Study. The top SNPs of interest in *FOXO3* are located in introns, which have been associated with expression quantitative trait loci (eQTLs), indicating that the longevity association of *FOXO3* may be linked to expression level of the gene [28].

Recently, whole exome sequencing was performed on 100 individuals between 98 and 100 years of age, where researchers did not find evidence of any rare protein-altering SNPs, nor did they find any genes with an increased burden of rare variants [29]. As well, additional longevity loci have been found in an informed GWAS that took into account information about previous disease associations [30], and a GWAS in Han Chinese that had a sample size with 2.7 times more centenarians than any other previously published GWAS [31]. As is the case in these studies, identifying additional common variants will likely require greater sample sizes or more elaborate methods of analyzing the data. However, since healthy aging is a polygenic trait, there may not be many more common individual variants with large effect sizes to find. Therefore future research may instead

want to focus on collections of common variants with small individual effect sizes that have a larger cumulative effect, rare variants with large individual effects, or epistatic variants that are difficult to detect on their own.

Epistasis occurs when one variant is influence by the presence of another variant. An undetected underlying gene-environment or gene-gene interaction could lead to an unaccounted for confounder. Thus, a significant effect seen in one study, may not to replicate in a subsequent population due to differences in the genetic background of the study populations [32]. These variants that fail to replicate are therefore candidates for having an underlying interacting mechanism. One proposed mechanism for protective variants in longevity is that buffering gene variants may attenuate the negative effects of unfavourable variants. A well-known candidate for a genetic buffer, or capacitor, is the heat shock protein Hsp90 in *Drosophila melanogaster*. Hsp90 is a molecular chaperone that promotes protein folding and stability, consequently masking genetic variation and increasing the ability to handle heat stress [33, 34].

In humans, buffering variants can be recognized by change in genotype frequency with age in a population. A protective buffering gene would show a monotonic increase in frequency with age when plotted from approximately age 65 years and older since it is favourable to survival in the population. A deleterious variant that is buffered, however, would be seen at high frequencies in young and very old ages, and at decreased frequency in early old age; thus forming a U-shaped curve [32]. This initial decrease in allele frequency is a result of the variant being deleterious and disease causing in the population. The inflection point of the U-shaped curve where allele frequency changes from decreasing to increasing occurs because the surviving carriers of the deleterious variant also carry a protective buffering variant thereby raising the allele frequency closer to the initial unselected frequency. Previous studies have identified a small number of potentially longevity- associated buffering variants, as well as potentially deleterious buffered variants in centenarians [32, 35]. For example, in a population of Ashkenazi Jews aged 50-110 years, it was found that there was a monotonic increase in the nonsynonymous variant *CETP VV* (rs5882), which appeared to attenuate the deleterious effect of the heterozygous variant rs1853021 in *LPA*. The frequency of heterozygotes at this position decreased in the population until approximately 80 years old, and then subsequently increased until nearly the frequency of the younger age group, forming a U-

shaped curve. When the population was subdivided for whether they carried the buffering *CETP VV* genotype, the subset with the favourable longevity variant exhibited a slight increase in the frequency of the deleterious *LPA* genotype, whereas the subset lacking it showed a monotonic decrease in frequency, similar to a non-buffered deleterious variant [32].

Allele-specific expression (ASE) refers to a difference in the expression between two alleles at a specific site; ASE can be investigated by using heterozygous SNPs as indicators to distinguish between the transcripts with each allele [36]. *Cis*-acting factors are in close physical proximity to the regulated gene, whereas *trans*-acting factors regulate distant genes; although it appears that both types of regulation act together to coordinate gene expression [37]. In the absence of *cis*-acting factors, equal expression of both copies of the allele would be expected [38]. The advantage of looking at ASE is that the other allele acts as an internal control, where both alleles share the same environment and other *trans*-acting factors that affect gene expression [38, 39]. Imbalances can occur due to allele-specific alternative splicing, variation in transcriptional start or stop sites, *cis*-acting regulatory variants that may alter regulation for just one allele through a change to promoter/enhancer regions (transcription factor binding sites), epigenetic differences such as DNA methylation or chromatin state, and differences in 3′ UTR mutations that can affect mRNA stability or microRNA binding [39]. An extreme case of ASE is X-chromosome inactivation which can occur due to gene imprinting where either the maternal or paternal copy of the gene is silenced [40].

ASE has been found to be relatively common throughout the genome with studies finding that 20-25% of human genes exhibit ASE [41, 42]. Another study found that 4.6% of heterozygous HapMap SNPs showed evidence of ASE at a read depth of 50 and an allelic ratio of 67:33 [36]. ASE has also been associated with disease phenotypes. ASE was found in the *ERCC5* gene, part of the nucleotide excision DNA repair pathway, in individuals with lung cancer compared to non-lung cancer controls [43]. ASE was also seen in the *MTTP* gene due to the minor alleles of two promoter SNPs, which may alter the susceptibility to ischemic heart disease [44]. As well, ASE was associated with phenotypic variability and penetrance of malignant hyperthermia [45].

Monozygotic twins have been found to have a more similar degree of ASE than unrelated controls [46], suggesting that ASE has a heritable component. In particular, allele-specific chromatin signatures have also been demonstrated to be heritable [47, 48], and allele-specific DNA methylation has been associated with disease risk [49, 50]. Epigenetic inheritance has been suggested to account for some of the "missing heritability" that is not explained by DNA mutations alone [51].

The field of genetics of healthy aging research aims to identify genetic variants and biological mechanisms that contribute to the healthy longevity of exceptionally long-lived individuals and families. This will hopefully lead to the development of therapies and drugs that will help those individuals not fortunate enough to carry "longevity genes" live healthily into old age.

1.3. Hallmarks of aging

Aging is a universal experience characterized by a physiological decline with time. Although many succumb to age-related diseases in old age, disease is not a normal part of aging. Nine hallmarks of aging were proposed in a seminal 2013 paper by López-Otín et al. [52]: genomic instability, telomere attrition, epigenetic alterations, loss of proteostasis, deregulated nutrient-sensing, mitochondrial dysfunction, cellular senescence, stem cell exhaustion, and altered intercellular communication. These hallmarks are further subdivided into primary hallmarks (genomic instability, telomere attrition, epigenetic alterations, loss of proteostasis) which are the cause of cellular damage and are intrinsically detrimental, antagonistic hallmarks (deregulated nutrient-sensing, mitochondrial dysfunction, cellular senescence) which have contradictory protective and deleterious effects, and integrative hallmarks (stem cell exhaustion, altered intercellular communication) which are the resulting effects of the previous hallmarks and are responsible for the deteriorating aging phenotype.

Genomic instability occurs through the accumulation of genetic damage that occurs with aging [52]. Nuclear DNA can accumulate somatic mutations, chromosomal aneuploidies and copy-number variations during aging. Mitochondrial DNA (mtDNA), which has less protective architecture and limited mDNA repair mechanisms, also incurs mutations. The resulting damage can a result in dysfunctional cells [52].

Telomere attrition is the progressive loss of telomeres that occurs due to the inability of the DNA replication process to replicate to the end of chromosomes [52]. The Hayflick limit refers to the number of times a cell can divide before telomeres shorten to a critical length, thus reaching cellular senescence [53, 54]. Telomere length has been found to have an unclear association as a biomarker of aging, likely due to the fact that cells are largely unaffected until a critical telomere length is reached [55, 56].

Epigenetic alterations arise during aging with changing patterns of DNA methylation, post-translational histone modifications, and chromatin remodelling [52]. The epigenetic clock of aging developed by Dr. Horvath [57] proposes that DNA methylation age is a biomarker of biological age that more accurately predicts the functional capability of a tissue or person over chronological age [56]. Epigenetic age uses a mathematical algorithm to estimate age based on methylation states of specific CpGs; and since methylation is reversible, this leads to the possibility of using DNA methylation age as a means of identifying or validating aging interventions [56].

Proteostasis is the maintenance of protein concentration, conformation, interaction, localization, and turnover within the cell [58]. With aging, this response can become impaired resulting in a decline in cell function, decreased longevity, and protein aggregate diseases such as AD, Parkinson's disease and cataracts [52, 59]. Specifically, chaperone activity is diminished which affects the folding and stability of proteins, and there is a decline in the protein quality control proteolytic systems [52] which may promote the development of age-related metabolic disorders [60].

Deregulated nutrient-sensing can have contradictory effects with regards to aging. Nutrient sensing and anabolism protect against nutrient scarcity, but may accelerate aging, whereas decreased nutrient signalling may extend longevity, which has been observed in dietary restriction experiments across multiple species [52, 61]. Dietary restriction is thought to affect the insulin and IGF-1 signaling (IIS) pathway, which plays a role in glucose-sensing. Downstream of this pathway are the FOXO transcription factors and mTOR complexes, which have also been implicated in aging [52]. Decreased IIS may be a self-protective response to cellular damage that aims to lessen cell growth and metabolism, and may be employed as a strategy to extend lifespan in aged organisms. Conversely, when IIS signals are too low, it contributes to aging [52]. Additionally,

interrelated nutrient-sensing systems include mTOR which senses high amino acid concentrations, AMPK which senses low energy states through AMP levels, and sirtuins which sense low energy states through NAD+ levels [52, 62]. These nutrient-sensing pathways are the focus of aging interventions. For instance, rapamycin [63, 64], metformin [65, 66], and NMN [67] administration may be beneficial to aging through mTOR inhibition, upregulation of AMPK, and upregulation of sirtuin pathways, respectively.

Mitochondrial dysfunction occurs through a combination of mtDNA mutations, destabilization of the electron transport chain, altered mitochondrial dynamics, defective mitophagy, and decreased mitochondrial biogenesis [52, 68]. Mitochondrial dysfunction can increase inflammatory factors, impact cellular signalling, and generate ROS which can be beneficial and mediate cell signalling and survival to restore homeostasis, but can also produce cellular damage that contributes to aging at high levels [52]. ROS has a negative reputation for causing oxidative stress, however antioxidants have shown inconsistent results in extending lifespan, and is some cases have shortened lifespan in model organisms [69-72].

Cellular senescence has the beneficial effect of preventing proliferation of damaged cells which is protective against tumorigenesis, however as the amount of senescent cells accumulate with aging due to increased production as well as reduced clearance, they have a number of deleterious effects such as decreased tissue function, increased inflammation, and stem cell exhaustion that contribute to aging [52, 73]. Senescent cells also affect the secretome, increasing pro-inflammatory factors [73]. This increase in inflammation, also termed inflammaging, may contribute to the development of disease, but may also be a necessary adaptation to changes in the immune system that occur during aging, and therefore its role in aging remains unclear [74]. Possibly, the beneficial effects of inflammation in the immune response occur in early and mid life, but by late life it leads to an increased susceptibility to disease and further plays a role in age-related diseases such as atherosclerosis, AD, and diabetes [75].

Stem cell exhaustion results in the diminution of the regenerative potential of tissues. Exhaustion of hematopoietic stem cells can lead to a decrease in the production of adaptive immune cells (immunosenescence), anemia, and myelodysplasia, exhaustion of mesenchymal stem cells can lead to osteoporosis and decreased fracture repair,

exhaustion of satellite cells can result in a decreased repair of muscle fibres, and exhaustion of intestinal epithelial stem cells can result in decreased intestinal function [52]. The deterioration of stem cell function with age is an area of focus for regenerative therapies including parabiosis, telomere lengthening, and the reprogramming of aged cells towards induced pluripotent stem cells (iPSCs) [76, 77].

Intercellular communication, for example renin-angiotensin and insulin-IGF1 signaling, can become deregulated as inflammaging increases and immunosurveillance decreases in aging [52, 78]. Proinflammatory cytokines can alter downstream pathways as well as play a role in the development of age-related diseases, and as the adaptive immune system declines there is a decreased ability to clear infectious agents, malignant cells, and senescent cells. Bystander effects also occur, for example, senescent cells can induce senescence in neighbouring cells through cell-cell contact [52]. As well, parabiosis experiments suggest that by altering the systemic environment, you can activate signaling pathways that lead to increased tissue regeneration [79].

1.4. Model organism genetics

Abnormal protein folding can be detrimental to the cell, so it is crucial to maintain proteostasis in the event of heat and other stressors [58]. A key mechanism for this maintenance is through molecular chaperones, which tend to be ubiquitous and act on a wide range of substrates [80]. In *Drosophila melanogaster,* heat shock protein 90 (Hsp90) is a well-known genetic capacitor [34]. Hsp90 is a molecular chaperone that acts on a diverse variety of proteins assisting in proper protein folding as well as refolding after stress and allows mutations to be present without affecting phenotype [34]. This buffering property has been seen in studies with mutations in Hsp90 where affected flies exhibit morphological defects [33, 81]. Hsp90 is a highly conserved HSP, and is a vital chaperone, as well as crucial element in the protective heat shock response across organisms from bacteria to mammals [58]. Hsp90 interacts with numerous co-chaperones, and its mechanism is highly dynamic is not fully understood. As well, in *Caenorhabditis elegans*, another HSP, Hsp-16, acts as a stress response protein and chaperone, and increases lifespan when an extra copy of the gene is introduced [82].

In humans, it is possible that long-lived individuals may also be carrying genetic capacitors that are buffering the phenotypic effects of deleterious variants. Hsp90 may be broadly protective against the effects of missense variants, and be protective against Fanconi anemia by binding and stabilizing mutant *FANCA* proteins [83]. Another HSP is Hsp70 which plays a role in substrate folding, disaggregation, refolding, and degradation [80], and can work in cooperation with Hsp90 [84]. It has been shown in a population of hospitalized elderly patients that serum concentration of Hsp70 was significantly higher in patients with higher degrees of inflammation, categorized by levels of C-reactive protein and fibrinogen as an indicator of inflammation status [85].

In *Caenorhabditis elegans,* lifespan is regulated by the insulin/IGF-1 signaling pathway [86]; specifically by *daf*-16, a homolog of FOXO transcription factors in humans, which regulates the transcription of numerous genes and is involved in oxidative stress, heat stress and starvation [87]. In humans, FoxO proteins have been implicated in AD due to their involvement in insulin resistance and oxidative stress pathways [88], and *FOXO3* has been associated with longevity [25].

1.5. Next-generation sequencing

Genetics research has evolved with advancements in sequencing technology and the decreasing prices of next-generation sequencing (NGS). My thesis project utilizes Illumina NGS, which offers short read, high-throughput sequencing [89]. Illumina NGS starts with preparation of sequencing libraries; a DNA or cDNA sample is randomly fragmented, adapters are ligated to both ends of the fragments, and PCR is used to amplify adapter-ligated fragments. The library is then loaded onto a flow cell for cluster generation, where fragments bind to surface-bound oligos that are complementary to the adapters in the library, and amplified into clonal clusters using bridge amplification. Sequencing then occurs through detection of fluorescently labeled nucleotides as they are incorporated [89].

RNA-seq uses NGS to sequence and quantify RNA transcripts [90]. RNA-seq libraries in this thesis were prepared using the Illumina TruSeq stranded mRNA kit (Illumina, Inc. USA). As per the TruSeq Stranded mRNA Sample Preparation Guide, this procedure involves purifying the RNAs with poly-A tails using magnetic beads attached to poly-T

oligos. The RNA is then fragmented and primed with random hexamers. The first stand of cDNA is synthesized from the cleaved RNA fragments using reverse transcriptase and random primers. The RNA template is then removed and replaced by a second strand of cDNA that incorporates dUTP instead of dTTP. The 3' ends of the blunt fragments are adenylated by adding an 'A' nucleotide to prevent fragments from ligating to each other, and to provide an overhang for ligating to the adapter. Indexing adapters are then ligated to the double stranded cDNA and the DNA fragments are enriched using PCR for fragments that have adapter molecules on both ends. During amplification, the PCR polymerase will not amplify the second strand containing the dUTP. Before flow cell loading there is a denaturation step so single stranded DNA goes onto the flow cell. Cluster generation and sequencing occurs as described above.

1.6. Genetic studies of aging and longevity

There are several studies focusing on the genetics of healthy aging and longevity that utilize a variety of study designs including family-based associated studies, longitudinal cohort studies, and case-control studies [9] (Table 1-1). Notably, different studies utilize different definitions for their long-lived groups, with studies varying from 80 to 100 years as the minimum age of their participants, as well as having different health inclusion criteria that range from none to stringent.

The Longevity Gene Study [91] consists of Ashkenazi Jews who are 95 years and older, and their offspring. Offspring of centenarians were found to have a favourable lipid profile compared to their spouse controls [91] and both long lived individuals and their offspring had significantly larger HDL and LDL particle sizes [92]. They also found that centenarians and their offspring better maintained their telomere length [93] and that exceptional longevity may be influenced by gender, with maternal inheritance affecting females, and both maternal and paternal inheritance affecting males [94]. The authors also proposed the idea that longevity genes have a buffering effect that allows for the accumulation of deleterious variants, and demonstrated this effect with *CETP-VV* as a favourable genotype that attenuates the effect of the deleterious *LPA* [32].

The Leiden Longevity Study [95] , which includes sib pairs aged 90 years and over, found evidence of genetic enrichment for exceptional survival when looking at standardized

mortality ratios in sibs, parents, and offspring, compared to the general Dutch population. They also examined 30 disease-associated SNPs in nonagenarians vs. younger controls and did not observe a difference in the number of disease risk alleles [96], and found that the major locus determining familial longevity in their GWAS was rs2075650, a SNP located in *TOMM40* that is close to and is in moderate linkage disequilibrium with *APOE* [97]. In a genome-wide gene expression study, they proposed an transcriptional aging-signature of 21 genes, with the most notable results being decreased expression of *ASF1A*, a histone chaperone involved in chromatin remodelling and *IL7R*, a receptor that regulates histone acetylation and is required for immune system maintenance [98].

The New England Centenarian Study [99] is a longitudinal study of centenarians living in the Boston area, their siblings and offspring, and a control group of the spouses of offspring, as well as families lacking longevity. They found that even among centenarians, at more extreme ages (100-104 years, versus 105-109, years versus 110+ years), there is evidence for progressive compression of disability and morbidity [99]. As well, in an informed GWAS that utilized the results of large studies on age-related diseases to increase statistical power, they were able to identify eight new extreme longevity loci, four of which they were able to replicate [30]. These included rs2075650/rs4420638 in the *APOE/TOMM40* region, rs4977756 in *CDKN2B/ANRIL*, rs514659 in *ABO*, and rs3184504 in the *SH2B3/ATXN2* region, which have associations with coronary artery disease and Alzheimer's disease.

The Long Life Family Study [100] consists of long-lived probands, their siblings, their offspring, and spouse controls. They found that their probands and offspring were less likely to have diabetes, chronic pulmonary disease and peripheral artery disease than similar aged cohorts, lower pulse pressure and triglycerides, higher high density lipid levels, better gait speed, and an overall later onset of decline with age [100]. Study participants also had a lower hazard ratio for cancer, CVD, severe dementia, diabetes, hypertension, osteoporosis and stroke [101]. Probands did not, however, have fewer risk alleles in SNPs associated with AD, CVD and stroke, type 2 diabetes or cancer, when compared to their offspring or spouses [102].

The Wellderly Study [103] comprises healthy elderly aged 80 and older with no chronic diseases and are not taking any chronic medications, however they do not have a control

group from the same population. They performed whole genome sequencing on 511 individuals and did not find evidence for enrichment of longevity-associated variants when compared to a younger control group further suggesting that healthy aging and longevity are distinct phenotypes [103, 104]. As well, while the Wellderly did not have a decreased genetic risk score for cancer, stroke, or type 2 diabetes, they did have a lower risk score for coronary artery disease and AD, the latter of which was strongly driven by an *APOE ε4* marker [103].

The Chinese Longitudinal Healthy Longevity Surveys [105] performed a GWAS of 2,578 centenarians and 2,387 geographically matched middle-age controls aged 40-59 [31]. They identified two novel loci: rs2069837 in *IL6* and rs2440012 in *ANKRD20A9P*, and in a gene set enrichment analysis of their GWAS results, they identified the starch, sucrose and xenobiotic metabolism pathway as a potential longevity pathway.

Okinawa Centenarian Study [106] is a population-based study of centenarians in Okinawa, Japan. Okinawans have a distinct traditional diet that is low in calories, nutrient-dense, antioxidant rich, low in glycemic load, features root vegetables, green and yellow vegetables, soybean-based foods, and medicinal plants, and may reduce the risk of chronic age-related diseases and promote healthy aging and longevity [107, 108]. They may also receive a life extension benefit from mild caloric restriction due to their cultural mindset regarding eating habits [109]. Female siblings of centenarians were 2.58 times more likely to survive to 90 years, and male siblings of centenarians were 5.43 times more likely [110], and participants aged 110+ delayed clinical major chronic disease and disability [111].

The Hawaii Lifespan Study [25] is a subset of the Honolulu Heart Program, which is a population-based, prospective study of Hawaiian men. When they looked at a long-lived group who survived to ≥ 95 years, and controls who died between 73-81 years of age, the *FOXO3* variant rs2802292 was found to be associated with longevity, as well as with plasma insulin levels, coronary artery disease, cancer, and type 2 diabetes prevalence [25]. The *G* allele of the rs2802292 SNP was also found to confer a 10% risk reduction for all-cause mortality in carriers, specifically due to decreased coronary heart disease mortality [112]. In a study where they followed men aged an average of 78 years, for 8

years, they found that participants with the rare G>A mutation at rs5742907 in *CETP* were more likely to be healthy beyond 90 years of age [113].

1.7. The Super-Senior Study

From 2004-2007, the Brooks-Wilson lab recruited participants for the Super-Seniors Study. Super-Seniors, born between 1901 and 1922, are individuals 85 years and older who have self-reported as never being diagnosed with cancer, cardiovascular disease (CVD), diabetes, dementia, or major pulmonary disease. The controls, born between 1952 and 1964, are aged 41-54 years and were not selected for health. Super-Seniors were interviewed at home and controls were interviewed by phone about personal and family medical history. Super-Seniors were also asked to perform standard geriatric tests for cognition and mobility. Each subject also gave a non-fasting blood sample that was extracted for DNA. My thesis work uses subsets of the Super-Senior Study population samples and data.

This thesis aims to explore the genetic contributions to healthy aging. While studying individual diseases provides important information about specific populations, healthy aging research provides a holistic view that spans disease phenotypes to form a more comprehensive model for human health. Such research will likely require comprehensive physiological and molecular theories in order to build an integrative model of avoiding chronic disease.

Healthy aging and longevity studies to date have found an overlap of healthy aging loci with disease loci and pathways implicated in disease [30, 103]. As well, there is evidence that successful aging may be a result of aging biologically slower than chronological age, for instance: offspring who have both parents surviving to over 90 years have a decreased risk of CVD [114], offspring of centenarians have a lipoprotein profile favourable to control groups [91], and female centenarians are more likely to have children in their 40s suggesting they are able to remain fertile longer [115].

Compared to other studies of healthy aging and longevity, the Super-Seniors had to pass stringent health inclusion criteria and have a well-characterized healthy aging phenotype in addition to surviving to 85 years. I hypothesize that Super-Seniors have an overall

decreased genetic susceptibility and increased resilience to age- related diseases that allows them to resist the effects of disease-risk variants that they carry. To investigate this, I will explore the Super-Senior phenotype as well as their common disease variants and potential protective variants compared to mid-life controls.

Table 1-1. Genetic studies of healthy aging and/or longevity.

Study name	Ages of long-lived group	Phenotype of long-lived group	Control and other groups	Ethnicity
Super-Senior Study	≥ 85	Never diagnosed with cancer, CVD, diabetes, AD, or major pulmonary disease	Controls aged 45-54	Mostly European
Leiden Longevity Study	≥ 89 men ≥ 91 women	Subjects must have a living sibling who also meets the age requirement	Offspring, spouses of offspring	Dutch and Caucasian
Long Life Family Study	72-110 probands	Families with multiple long-lived individuals	Offspring, spouses of offspring	White
Wellderly	≥ 80	Free of major diseases and long-term medications		Mostly European
Longevity Genes Project	95-107	Living independently at 95 years as a measure for health, and have an offspring willing to participate	Spouses of offspring	Ashkenazi Jewish
Okinawa Centenarian Study	≥ 100	Population-based	Family members	Okinawan (Japanese)
New England Centenarian Study	≥ 100	Population-based	Siblings, offspring, spouses of offspring	Living in the Boston area
Chinese Longitudinal Healthy Longevity Surveys (centenarian subset)	≥ 100	The CLHLS covered ~85% of the population of China	Geographically matched controls aged 40-59	Han Chinese
Hawaiian Lifespan Study and Honolulu Heart Program	Long-lived (age varies) participants of the HHP	Men; subset of the population-based, prospective Honolulu Heart Program	HHP participants with known younger ages of death	Japanese

Chapter 2.

The Super-Seniors Study: Phenotypic characterization of a healthy 85+ population[1]

Background: To understand why some people live to advanced age in good health and others do not, it is important to study not only disease, but also long-term good health. The Super-Seniors Study aims to identify genetic factors associated with healthy aging.

Methods: 480 healthy oldest-old 'Super-Seniors' aged 85 to 105 years and never diagnosed with cancer, cardiovascular disease, diabetes, dementia, or major pulmonary disease, were compared to 545 random population-based mid-life controls aged 41-54, who represent a group that is unselected for survival from late-life diseases. Health and lifestyle information, personal and family medical history, and blood samples were collected from all participants. Super-Seniors also underwent four geriatric tests.

Results: Super-Seniors showed high cognitive (Mini-Mental State Exam mean = 28.3) and functional capacity (Instrumental Activities of Daily Living Scale mean = 21.4), as well as high physical function (Timed Up and Go mean = 12.3 seconds) and low levels of depression (Geriatric Depression Scale mean = 1.5). Super-Seniors were less likely to be current smokers than controls, but the frequency of drinking alcohol was the same in both groups. Super-Seniors were more likely to have 4 or more offspring; controls were more likely to have no children. Female Super-Seniors had a mean age of last fertility 1.9 years older than controls, and were 2.3 times more likely to have had a child at ≥ 40 years. The parents of Super-Seniors had mean ages of deaths of 79.3 years for mothers, and 74.5 years for fathers, each exceeding the life expectancy for their era by a decade.

Conclusions: Super-Seniors are cognitively and physically high functioning individuals who have evaded major age-related chronic diseases into old age, representing the approximately top 1% for healthspan. The familiality of long lifespan of the parents of

[1] This chapter has been published in near-identical form. Halaschek-Wiener J*, Tindale LC*, Collins JA, Leach S, McManus B, Madden K, Meneilly G, Le ND, Connors JM, Brooks-Wilson AR. The Super-Seniors Study: Phenotypic characterization of a healthy 85+ population. PLOS One. 2018; 13(5): e0197578.

PLOS applies the Creative Commons Attribution (CC BY) license to works they publish.

Super-Seniors supports the hypothesis that heritable factors contribute to this desirable phenotype.

2.1. Introduction

Healthy aging and extreme longevity are phenotypes that many hope to achieve. For many, however, old age is accompanied by poor health. Longevity refers to the length of time an individual lives, their lifespan, whereas healthy aging refers to a person's 'health span'. The majority of longevity can be attributed to environmental and lifestyle factors; however, 15-30% of adult lifespan is heritable [6]. While the heritability of longevity is minimal before age 60, it increases at more advanced ages [7]. Some longevity genes have been identified in model organisms; however, few findings have been replicable in humans with the exceptions of *APOE* and *FOXO3*, reviewed in [9].

According to the US Centre for Disease Control, the leading causes of death over the age of 65, in descending order, are: diseases of the heart, malignant neoplasms, chronic lower respiratory diseases, cerebrovascular disease, Alzheimer disease (AD), and diabetes mellitus [4]. We describe here the ascertainment and characterization of 480 oldest old who have never been diagnosed with any of these diseases.

Several research studies have collected long-lived individuals [30, 95, 99, 100, 116, 117], most with the goal of studying longevity or exceptional longevity. Recent insights have also been gained from large cross-sectional population studies [118, 119]. The Super-Seniors Study examines individuals over 85 with a well-characterized health phenotype free of five specific major age-related diseases, in order to study healthy aging. This focus is shared by the Wellderly study [103], which researches healthy elderly aged 80 and over.

We have collected a group of oldest-old "Super-Seniors", aged 85 years and older who have never been diagnosed with cancer, cardiovascular disease (CVD), diabetes, dementia or major pulmonary disease; as well as a group of mid-life controls recruited randomly with respect to health. The Super-Seniors represent a group of individuals who have not only survived to at least 85, but have done so free of the major chronic diseases that lead to decreased quality of life and early death. The Super-Senior phenotype is also economically relevant – the diseases they have evaded are among the most common and

therefore expensive to provide care for within a healthcare system. Because most individuals in developed countries live to at least age 50, the mid-life group represents a set of individuals who are as yet unselected for the age-related diseases that the Super-Seniors have avoided.

To some, 85 years may now seem too young to be considered long-lived [120]. While individuals born more recently have tended to live longer, for those born in the years that the Super-Seniors were, reaching 85 years was quite a feat (Appendix A.1).

Future work will compare the Super-Seniors to population-based mid-life controls as a strategy to identify genetic factors that may contribute to their long-term good health. Here we characterize the health, family history and lifestyle of the Super-Seniors.

2.2. Design and Methods

2.2.1. Recruitment

This study was approved by the University of British Columbia (BC)-BC Cancer Agency Research Ethics Board and the Research Ethics Board of Simon Fraser University. All participants gave written informed consent.

Eligibility criteria for being a Super-Senior included: being 85 years or older and self-reporting as never having had cancer (except non-melanoma skin cancer), CVD, diabetes, dementia, or major pulmonary disease (except asthma). Controls were aged 41-54 years and not selected for health.

Using lists from the BC Ministry of Health Medical Services Plan (MSP), which includes 98% of BC residents, we contacted individuals living in Metro Vancouver, BC, Canada 85 or older, or 40-50 years old. Additional invitations were sent to potential controls accompanied by an offer of a $50 honorarium. Super-Seniors were also identified by BC Stats, which had confidential access to the Insurance Corporation of BC drivers license database and could allow contact of currently licensed drivers, or were those who volunteered following press coverage. In an attempt to collect additional Super-Seniors of

Asian ancestry, advertisements were placed in Chinese-language newspapers. Finally, to collect a greater number of European controls, later mail-outs included only midlife individuals with non-Asian surnames.

After initial contact by mail, potential Super-Seniors were screened by phone to determine eligibility. After receiving written informed consent, a home or telephone interview was arranged for Super-Seniors and controls, respectively. Ascertainment and sample collection took place between 2004 and 2007, resulting in the shifting age range from the intended 40-50 to 41-54.

2.2.2. Data collection

Super-Seniors were visited at home by an interviewer and asked for personal and family medical history, to show all prescription and non-prescription medications, and to perform geriatric tests. Test scores did not affect eligibility.

Controls were asked the same personal/family medical history questions, but were not selected for health or disease status.

Ethnicity of the four grandparents was collected, and a composite ethnicity was determined for each participant. A participant's ethnicity was categorized as unknown if they did not know the ethnicity of all four grandparents. If participants were unsure about the age of death of their parents, an approximation was made; for example, 'mid-sixties' was approximated as 65 years.

30mL of non-fasting blood was drawn from each participant. Super-Seniors were visited by a phlebotomist; controls visited a commercial clinical laboratory. DNA was extracted using the PureGene DNA isolation kit (Gentra Systems, MN).

2.2.3. Phenotypic review

Health data of the Super-Seniors were reviewed and potential cases were excluded based on presence of disease but not based on intermediate phenotypes such as high blood pressure. Medications were reviewed and potential Super-Seniors excluded if they were

taking a drug used exclusively to treat cancer, CVD, dementia, pulmonary disease or diabetes. Participants with borderline health status (generally those with an asymptomatic arrhythmia or chronic bronchitis) were excluded from subsequent analyses.

2.2.4. Statistical analysis

Statistical tests were run in JMP® Version 11 [121].

2.3. Results

2.3.1. Recruitment

Recruitment is summarized in Figure 2-1 and Appendix A.2. An initial mailing was sent to BC MSP subscribers: 8415 individuals aged 85 or more, and 3920 aged 40-50. Of potential Super-Seniors, 4261 (50.6%) had incorrect contact information, were unable to be contacted, or did not speak English; 628 (7.5%) were deceased; 1059 (12.6%) refused without determining eligibility; 2161 (25.7%) were not eligible; and 306 (3.6%) participated. Of the potential controls, 2884 (73.6%) had incorrect contact information, were unable to be contacted, or did not speak English; 491 (12.5%) refused, and 545 (13.9%) participated. 12.4% of potential Super-Seniors who were interested were eligible. The consent rate for controls was 52.6% (Appendix A.2).

In addition to those identified through BC MSP, 160 Super-Seniors volunteered and 94 were identified by BC Stats. After review of interview data and medications, 63 individuals were excluded and 17 were borderline, resulting in 480 Super-Seniors and 545 controls.

2.3.2. Descriptive statistics

There were 325 female and 155 male (32.3%) Super-Seniors and 336 female and 209 male (38.3%) controls (Table 2-1 and Appendix A.3) Super-Seniors were aged 85-105 (mean 88.5 years); controls were aged 41-54 (mean 46.7 years). 92.5% of Super-Seniors and 76.3% of controls were of European ancestry (Figure 2-1 and Appendix A.4). Super-

Seniors had a mean BMI of 25.7 (SD 4.7), while controls had a mean BMI of 24.5 (SD 3.9).

2.3.3. Functional tests

Super-Seniors scored high on the Mini-Mental State Exam (MMSE) [122], mean = 28.3 (SD 1.7) and Instrumental Activities of Daily Living Scale (IADL) [123], mean = 21.4 (SD 3.5), and low on the Timed Up and Go test (TUG) [124], mean = 12.3 seconds (SD 4.3) and Geriatric Depression Scale (GDS) [125], mean = 1.5 (SD 1.8) (Figure 2-2). Inability to complete the MMSE as a result of vision deficits resulted in 35 scores being excluded. The TUG was not administered if the participant was not able to ambulate independently. Seven MMSE, 15 TUG, and 6 GSD scores were missing. No differences were detected between the scores for men and women for any test.

2.3.4. Lifestyle

Smoking status was divided into current, never, and quit (Table 2-1). Controls were 8.0 fold more likely to be current smokers than cases (95%CI = 3.36-15.8, $p<0.0001$) (Appendix A.5). There was a difference in proportion between male and female never smokers and quitters in Super-Seniors ($X^2 = 24.2$, $p<0.0001$): female Super-Seniors were more likely to be never smokers (58.7%) than men (34.4%), and men were more likely to be quitters (65.6%) than women (41.3%). No difference in smoking status was detected between sexes in controls.

Among quitters, Super-Seniors smoked a mean of 29.4 years (SD = 17.8) and 19.3 pack years (SD = 24.6) (pack year = packs smoked/day*years smoked); whereas controls who quit smoked a mean of 13.8 years (SD = 9.2) and 9.6 pack years (SD = 9.9). Although there was no difference in the mean number of years that male and female Super-Seniors smoked, male Super-Seniors smoked more heavily, for a mean of 9.32 pack years more than females (SE = 3.4, 95%CI = 2.6– 16.0, p = 0.007). Super-Senior quitters started smoking at a mean age of 20.5 years (SD = 8.1) and quit at a mean age of 50.0 years (SD = 18.2); control quitters started smoking at a mean age of 16.6 years (SD = 3.3) and quit at a mean age of 30.5 years (SD = 8.8).

49% of Super-Seniors and 59% of controls reported exercising (X^2 = 10.44, p<0.0012), with controls more likely to engage in exercise other than walking. No significant difference was observed in the proportion of Super-Seniors and controls who drank alcohol.

2.3.5. Number of offspring and age of fertility

There was a difference in the proportion of number of offspring between Super-Seniors and controls (X^2 = 107.0, p<0.0001) (Appendix A.6). Super-Seniors had a mean number of offspring (2.6, SD = 1.7) that was higher than controls (mean = 1.6, SD = 1.2);

284 Super-Senior and 248 control females gave birth. Super-Senior women had a mean age of last fertility 1.9-years older than control women (SE = 0.5, 95%CI = -2.8– -1.9, p<0.0001). There was no evidence of a difference in mean age of first fertility. Super-Senior men also had their last child a mean of 1.7 years older than control men (SE = 07, 95%CI = -0.2– -3.1, p = 0.011).

47 (16.5%) Super-Seniors and 20 (8.1%) controls who reproduced gave birth at ≥40 years, at a mean age of 42.6 (SD = 2.2) and 41.3 (SD = 1.6) years, respectively. 128 (45.1%) Super-Seniors and 75 (30.2%) controls gave birth ≥35, at a mean age of 38.8 (SD = 3.4) and 37.9 (SD = 2.4). Among women who gave birth, Super-Seniors were 2.3 times more likely to have had a child at ≥40 years (95%CI = 1.3– 3.9, p = 0.004), and 1.9 times more likely to have had a child at ≥35 years (95%CI = 1.3-2.7).

2.3.6. The parents of Super-Seniors lived longer than their contemporaries

The Super-Seniors reported their parents' age at death to be between 21-110 years for their mothers (mean = 79.3, SD = 15.6), and 26-102 years for their fathers (mean = 74.5, SD = 16.0). The control parents' ages of death ranged from 26 years to still alive for mothers, and 29 to still alive for fathers.

The parents of the Super-Seniors were born between ~1880-1905. The earliest survival statistics for North America are Americans born in 1900 [4]. We compared age at death of the parents of Super-Seniors (who we know lived to reproductive age) to individuals born

in 1900 who survived to age 21. 50% of 21 year olds born in 1900 lived to 67 years (66 for men, 68 for women). The mothers and fathers of the Super-Seniors therefore lived 11.3 and 8.5 years (average of 9.9 years) longer than the 1900 birth cohort.

2.4. Discussion

We have established a collection of healthy oldest-old and a mid-life control group recruited from population-based lists. Here, we describe the characteristics and cognitive and physical function of the Super-Seniors. We also document major lifestyle factors such as smoking and alcohol consumption to allow adjustment for these factors in future genetic analyses.

Hidden differences in ethnicities in case-control studies can lead to false positive genetic findings. Early in recruitment, we noted a difference in ethnicity between the Super-Seniors and controls. Over time, Metro Vancouver has seen increasing immigration by non-European groups. We attempted to equalize the composition of the two groups by identifying Asian-ancestry Super-Seniors, with little success, so instead over-collected controls of European ancestry.

We define the Super-Senior phenotype as oldest-old (≥85 years) who have never been diagnosed with any of five major diseases that are the leading causes of death over the age of 65 [4]. 12.4% of seniors over age 85 who were contactable and interested were eligible. Given that 28.5% of Canadians age 85 or older have dementia [126], the eligibility rate of living individuals is therefore closer to 8.9%. Furthermore, only 9.0% of individuals born in 1916 lived to be 85 [127]. The proportion of the 1916 birth cohort who went on to become Super-Seniors is therefore approximately 0.80%, making Super-Senior status the ‘top 1%’ elite health and survival phenotype.

Super-Seniors had a mean TUG of 12.3 seconds, indicating that the majority are able to ambulate independently [124]. The Newcastle 85+ Study observed a baseline TUG of 18.6 ± 14.7 (n = 735) and a 5-year follow-up TUG of 20.7 ± 12.0 (n = 271) [128]. A study of community-dwelling Taiwanese individuals 65+ years of age found that physical fitness indicators, including the TUG, are associated with successful aging [129].

The median MMSE score of the Super-Seniors was 29, with a single participant scoring <19. In the Leiden Longevity Study, [95] (men 89+, women 91+) the median MMSE score was 25 and 14% scored <19. In the Taiwanese study only 73.5% had an MMSE ≥24 [129]. By selecting for a disease-free state that excluded dementia, the Super-Seniors are a cognitively high functioning group, even when compared to other long-lived groups.

A GDS score ≥ 5 has been found to be a sensitive and specific cutoff for depression [130]. The Super-Seniors had a mean GDS score of 1.5. The IADL assesses performance of daily tasks [123]. Although the mean IADL score of Super-Seniors was high, there were a few low scores, with the three lowest belonging to individuals who used wheelchairs. Because a minority of participants lived with family or in an assisted living community, lower IADL scores in some instances reflected their responsibilities rather than their abilities.

Several differences between the Super-Seniors and controls are expected and reflect population trends over time, including smoking habits and family size. Super-Senior women were less likely to be smokers than Super-Senior men, but there was no difference between male and female smoking rates in the controls. This is consistent with Canadian smoking trends [131]. Controls were more likely to be current smokers; one reason for this is that some smokers in the Super-Senior age range would likely have developed smoking-related diseases that would make them ineligible. There was no evidence of a difference in the proportion of Super-Seniors and controls who drank alcohol.

There were higher proportions of controls having no offspring and Super-Seniors having 4+ offspring. This likely reflects changes in family size over time; however, some controls in their early 40s could go on to have additional offspring.

Super-Senior women had an older mean age of last fertility and were 2.3 times more likely to give birth at ≥40 years, compared to controls. Associations have been found between late childbirth (40+) and increased survival in women [132-134]. Female centenarians were four times more likely to have had children in their forties than a group from the same birth year who died at age 73 [115]. Late fertility, and more specifically the ability to bear a child after age 40, may be a sign of slower biological aging [115]. On average, Super-Senior fathers had their last child at an older age than control fathers, suggesting that in

some cases the advanced maternal age of the Super-Senior women may have been because they had to wait until after the war to start their families.

Super-Senior parents lived substantially longer than their contemporaries, suggesting a familial tendency towards long life. The true difference in lifespan is probably greater, as many of the parents of the Super-Seniors were likely born before 1900, when life expectancy was even lower.

The mid-life control group is intended for use only for genetic comparisons with the Super-Seniors, not for epigenetics or formal comparison of lifestyle or other non-genetic factors. The latter quantities cannot be compared between these groups because they have lived in different eras and would be expected to show potentially confounding cohort effects.

The ideal control group for the Super-Seniors would be individuals from the same birth cohort who did not successfully achieve the Super-Senior phenotype. Clearly it is not feasible to obtain DNA from such a control group. Instead, we use the strategy of comparing the elite Super-Seniors to a group that has not been selected for survival in later life. From the survival curves in Appendix A.1B, in 1958 (the mean birth year of the controls), the curve is nearly flat until approximately age 50 because relatively few people born in 1958 died before that age; they were therefore largely unselected for survival up to that point.

If we were to compare Super-Seniors to age-matched individuals, we would be comparing healthy oldest old to unhealthy oldest old. Such individuals would have in common the fact that they all survived to at least age 85, and the phenotypic (and presumably genotypic) differences between them would be more subtle, and harder to detect, than those that could differ between the highly selected Super-Seniors vs. the largely unselected mid-life individuals. Importantly, an age-matched control group does not allow us to study genetic factors that affect survival to the oldest-old age category, or those genetic factors that might influence both survival and health late in life.

A limitation of our study design is that the midlife group is expected to contain a small minority of people (approximately 1%) who would have been destined to become Super-Seniors, had they been born circa 1916 (we note that being born in 1916 and surviving to 85, or being born in 1958 and surviving to 85, are not the same survival phenotype

because the threats to survival and health differ between the two eras). The presence of such rare individuals amongst the midlife controls slightly reduces the statistical power of the study. Another limitation is that small sample size limits the statistical power of the current sample set. To overcome this limitation a Phase 2 recruitment is in progress; we also intend to combine the data from the Super-Seniors Study with aging consortia for meta-analyses once we have completed initial genetic analyses. We plan a future recruitment of age-matched less healthy elderly to use as another comparison population.

We have established and characterized an initial cohort in which to study the genetics of healthy aging [135-138], with Super-Seniors representing an elite group in terms of healthspan. The study that is most comparable to the Super-Seniors is the Wellderly Study of healthy elderly aged 80 and older [103]. While healthy aging is defined similarly between these two studies, the Super-Seniors are on average 4.3 years older (average 88.5 years vs. 84.2 years) and have a higher proportion of women (67.7% vs. 60.7%). Geriatric test scores demonstrate that Super-Seniors are a cognitively and physically high functioning group in addition to being healthy oldest old. The long lifespan of the parents of Super-seniors supports the hypothesis that heritable factors contribute to this desirable phenotype.

Table 2-1. Characteristics of the study participants.

		Super-Seniors			Controls		
		Male	Female	Total	Male	Female	Total
Descriptive statistics	N	155	325	480	209	336	545
	Age - mean (SD) Range (years)	88.7 (2.9) 85-100	88.5 (2.9) 85-105	88.5 (2.9) 85-105	46.8 (3.2) 41-53	46.6 (3.4) 41-54	46.7 (3.3) 41-54
	Birth year - mean	1916	1916	1916	1958	1958	1958
	BMI - mean (SD) Range (kg/m^2)	25.0 (3.4) 19.0-42.5	24.3 (4.1) 15.1-42.1	24.5 (3.9) 15.1-42.5	26.7 (4.0) 18.4-46.8	25.0 (5.0) 16.8-48.4	25.7 (4.7) 16.8-48.4
Smoking	Smoker - current (%)	4 (2.6)	3 (0.9)	7 (1.5)	25 (12.0)	33 (9.8)	58 (10.6)
	Smoker - never (%)	52 (33.5)	189 (58.2)	241 (50.2)	107 (51.2)	161 (47.9)	268 (49.2)
	Smoker - quit (%)	99 (63.9)	133 (40.9)	232 (48.3)	77 (36.8)	142(42.3)	219 (40.2)
	Years smoked (among quitters) - mean (SD)	31.9 (17.4)	27.8 (17.8)	29.4 (17.8)	14.5 (10.2)	13.5 (8.6)	13.8 (9.2)
	Pack years smoked (among quitters) - mean (SD)	24.7 (28.2)	15.4 (20.8)	19.3 (24.6)	10.3 (10.9)	9.1 (9.4)	9.6 (9.9)
Activity	Activity - none (%)	25 (16)	71 (22)	96 (20)	51 (24)	81 (24)	132 (24)
	Activity - walking (%)	54 (35)	95 (29)	149 (31)	24 (11)	66 (20)	90 (17)
	Activity - exercise (%)	75 (49)	158 (49)	233 (49)	134 (64)	186 (56)	320 (59)
Alcohol	Alcohol - beer* (%)	28 (6)	24 (5)	52 (11)	130 (24)	75 (14)	205 (38)
	Alcohol - spirits* (%)	54 (11)	80 (17)	134 (28)	60 (11)	76 (14)	136 (25)
	Alcohol - wine* (%)	76 (16)	180 (38)	256 (53)	125 (23)	225 (41)	350 (64)
	Alcohol - none (%)	32 (7)	91 (19)	123 (26)	29 (5)	59 (11)	88(16)
Fertility	Number of offspring Range	2.6 (1.6) 0-12	2.5 (1.7) 0-8	2.6 (1.7) 0-12	1.5 (1.2) 0-5	1.7 (1.2) 0-5	1.6 (1.2) 0-5
	Had offspring	131	284	415	139	248	387
	Age of last fertility - mean (SD) Range (years)	35.8 (6.3) 23-54	33.7 (5.8) 19-47	- -	34.1 (5.6) 20-47	31.8 (5.1) 17-45	- -
	Had offspring 40+ years	34	47	-	25	20	-
	Age of 40+ parents - mean (SD)	44.1 (3.4)	42.6 (2.2)	-	42.4 (1.9)	41.3 (1.6)	-
	Had offspring 35+ years	68	128	-	63	75	-
	Age of 35+ parents - mean (SD)	40.7 (4.3)	38.8 (3.4)	-	39.2 (3.0)	37.9 (2.4)	-
Parents	Maternal age of death	79.6 (15.6)	79.1 (15.6)	79.3 (15.6)	-	-	-
	Paternal age of death	75.0 (15.9)	74.2 (16.1)	74.5 (16.0)	-	-	-
Geriatric tests	TUG - mean (SD)	12.2 (4.0)	12.3 (4.5)	12.3 (4.3)	-	-	-
	MMSE - mean (SD)	28.1 (1.7)	28.4 (1.7)	28.3 (1.7)	-	-	-
	GDS - mean (SD)	1.5 (1.8)	1.6 (1.8)	1.5 (1.8)	-	-	-
	IADL - mean (SD)	21.7 (3.0)	21.2 (3.7)	21.4 (3.5)	-	-	-

*Categories are not mutually exclusive

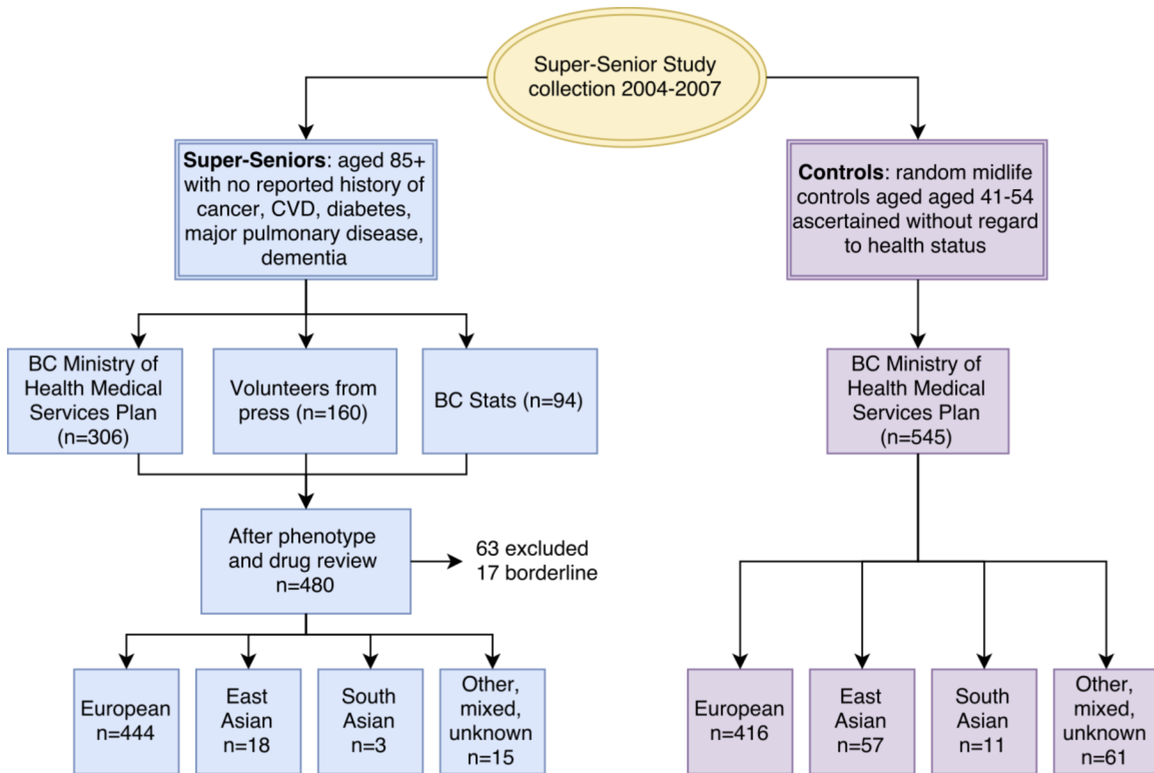


Figure 2-1. Recruitment of Super-Seniors and controls.

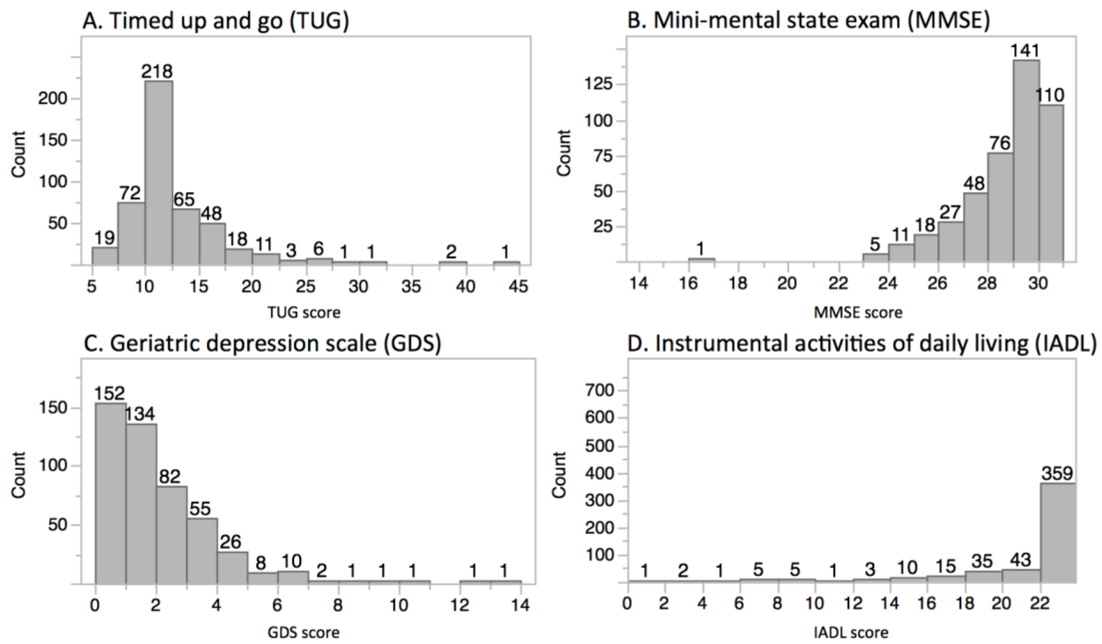


Figure 2-2. Distribution of functional test scores of the Super-Seniors.

Chapter 3.

10-year follow-up of the Super-Seniors Study: Compression of morbidity and genetic factors in survivors

Background: Super-Seniors are healthy, long-lived individuals who were recruited at age 85 years or older with no history of cancer, cardiovascular disease, diabetes, dementia, or major pulmonary disease. In a 10-year follow-up, we aimed to determine whether surviving Super-Seniors showed compression of morbidity, and to test whether the allele frequencies of longevity-associated variants in *APOE* and *FOXO3* were more extreme in such long-term survivors.

Methods: Super-Seniors who survived and were contactable were re-interviewed 10 years after initial characterization. Health and lifestyle were characterized via questionnaire. Geriatric tests including the Timed Up and Go, Geriatric Depression Scale, Instrumental Activities of Daily Living and the Mini-Mental State Exam were administered, and data were compared to results from on average 10 years earlier. As well, genotype and allele frequencies for SNPs rs7412 and rs429358 in *APOE*, and rs2802292 in *FOXO3* were compared to the frequencies in the original collection of Super-Seniors and mid-life controls.

Results: Of the 480 Super-Seniors recruited from 2004-2007, 13 were alive, contactable, and consented to re-interview (mean age = 100.1 ± 3.3). Eight of these 13 participants (62%) still met Super-Senior health criteria. Diseases that occurred in late life were cardiovascular (5 of 13; 38%) and lung disease (1 of 13; 8%). The surviving group of centenarians had a higher frequency of *APOE* and *FOXO3* longevity-associated variants even when compared to the original long-lived Super-Senior cohort.

Conclusions: Although physical and mental decline occurred in the decade between interviews, the majority of Super-Seniors re-interviewed still met the original health criteria. These observations are consistent with reports of compression of morbidity at extreme ages, particularly in centenarians. The increased frequency of longevity- associated

variants in this small group of survivors is consistent with studies that reported a larger role for genetics at older ages.

3.1. Introduction

Maintaining health while aging is important both for individual quality of life as well as costs to health care systems. Compression of morbidity refers to a shorter time between onset of disability and death, and was originally postulated to be due to chronic conditions having a greater capacity to be delayed than survival has to be increased [139]. A correlation between survival age and decreased morbidity has been observed, where older age groups (100-104 years, 105-109 years, and 110-119 years) experienced progressively delayed onset of disease and physical or cognitive impairment [140].

While the overall incidence of chronic conditions has been increasing in recent decades [4, 141], compression of morbidity has been observed in long-lived individuals [142-145]. As studies of younger groups (e.g., 51-61 years, [141]) did not reveal compression of morbidity, the reduction in years of disease may be limited to the high end of the human life span.

The Super-Seniors were collected as a phenotypically healthy oldest-old group in which to study genetic factors associated with healthy aging [138, 146]. Super-Seniors are individuals aged 85 years and older who reported never being diagnosed with cancer, cardiovascular disease (CVD), diabetes, dementia, or major pulmonary disease. Study controls are a population-based comparison group of mid-life individuals who resemble a genetic proxy group for the birth cohort of the Super-Seniors [146]. The initial collection and characterization of 480 Super-Seniors took place from 2004-2007 (Phase 1), with an additional and ongoing collection initiated in 2014 (Phase 2).

Approximately 10 years after the Phase 1 collection of Super-Seniors, we attempted to re-contact and re-interview surviving Super-Seniors. Participants still living would be in their late nineties or 100 or more, and we hypothesized that such individuals would have retained much of their health due to compression of morbidity.

3.2. Design and Methods

Research ethics board approval was received from the joint Clinical Research Ethics Board (REB) of BC Cancer and the University of British Columbia and from the REB of Simon Fraser University. All subjects gave written informed consent.

Phase 1 collection of Super-Seniors in 2004-2007 [146] did not include plans to follow participants longitudinally, so in 2010, Super-Seniors were mailed a letter requesting permission to re-contact them for future research. In 2016-17, contactable and interested Super-Seniors were invited for re-interview (Appendix B.1).

Super-Seniors were visited in their homes where they were asked personal and family medical history questions, and asked to perform geriatric tests as per their first interview. The Mini-Mental State Exam (MMSE) [122], Instrumental Activities of Daily Living (IADL) [123], Geriatric Depression Scale (GDS) [125], and Timed Up and Go (TUG) [124] were measured and compared to data from Phase 1 Super-Seniors collected approximately 10 years previously. Differences in scores were analyzed in one-tailed paired t-tests with the expectation that MMSE and IADL score would decrease with age, and that GDS and TUG scores would increase. BMI, HR, and BP were compared using two-tailed tests.

Genotype and allele frequencies for SNPs rs7412 and rs429358 in *APOE*, and rs2802292 in *FOXO3* were compared to the frequencies in the original Phase 1 Super-Seniors and mid-life controls. Genotypes for these select variants were extracted from data previously described [147]. Statistical analysis was done using R 3.2.2 and JMP 13.

3.3. Results

In 2010, we mailed 480 Super-Seniors and an additional 17 borderline phenotype individuals from Phase 1, excluding 9 for whom the study had been notified of their death. 26 were reported deceased by relatives, 30 declined (5/30 indicated they were “too sick”), 92 were unable to be contacted (mail was returned to sender), 246 did not respond, and 94 agreed to be re-contacted (Appendix B.1).

In 2016, we searched for online obituaries and notices of death, and sent another letter requesting permission for re-contact to 139 previous non-responders. By 2016, an additional 107 Super-Seniors were determined to be deceased through online records or reported to the study by family members. Of the 138 Super-Seniors mailed in 2016, 8 were reported deceased, 7 declined (3/7 indicated they were “too sick”), 26 were unable to be contacted, 89 did not respond, and 9 agreed to be re-contacted.

28 Super-Seniors were invited for re-interview. Of those, 13 accepted and completed the re-interview; 2 declined, 10 could not be reached, 1 died, and 2 were reported by relatives to be too ill to be interviewed. The 13 Super-Seniors re-interviewed in 2017 included 10 women and 3 men aged 96-106 (mean = 100.1, SD = 3.3), who were all of European ancestry.

The two Super-Seniors who were too ill to be interviewed included a 100-year-old woman reported by a family member to be confused much of the time, not doing very well, and feeling sick and tired. The other was a 106-year-old woman reported by a family member to be bedridden and mostly non-responsive. Deterioration reportedly took place around age 100-101 when she began suffering episodes of dementia. Other than dementia, she had no health problems and was very mobile until age 103.

Re-interviews took place 9.3-12.1 years (mean = 10.9, SD = 0.9) after the initial interviews. Of the 13 Super-Seniors who were re-interviewed, 8 still met the health criteria for being a Super-Senior, 5 women and all 3 of the men. Of the 5 Super-Seniors who did not meet the criteria for enrollment at the time of their re-interview in 2017, all had developed CVD: two had strokes (at ages 97 and 100, respectively), two had heart conditions (mitral valve issue [age unknown] and pacemaker [at age 99]), and one who was interviewed at age 96 had a minor heart attack, minor strokes, and COPD.

Of the 13 Super-Seniors re-interviewed, 6 were never smokers and the remaining 7 had quit. The quitters smoked between 5 and 51 years (mean = 29.7, SD = 16.0). Of note, the individual who had COPD was a housewife who never smoked.

Descriptive statistics and geriatric test scores are shown in Table 3-1 and Figure 3-1. MMSE mean scores declined from 28.7 (SD = 1.4) to 23.8 (SD = 4.2) points out of a possible 30, a mean decline of 4.9 points ($t = -4.5$, $p = 0.00036$). IADL scores declined

from 22.3 (SD = 1.5) to 15.6 (SD = 6.1) out of a possible 23 points, a mean decline of 6.7 points ($t = -3.7$, $p = 0.0016$). GDS scores increased from 0.5 (SD = 1.0) to 2.2 (SD = 2.5), a mean difference of 1.7 points ($t = 2.6$, $p = 0.011$). TUG scores increased from 9.8 (SD = 2.1) to 32.0 (28.0) seconds, a mean difference of 22.2 seconds ($t = 2.9$, $p = 0.0070$). Among the 9 individuals who did not use a walker as an aid, the mean TUG time was 17.9s (SD = 5.3).

There was no significant difference in BMI, heart rate, or blood pressure (BP) between the two interviews. Eight participants were taking BP medication at the time of both interviews, 4 were not taking any BP medication at either interview, and one man had discontinued taking BP medication by the time of the second interview.

Genotyping in re-interviewed Super-Seniors was not available for one participant in *APOE* and two participants in *FOXO3*. Genotype counts and minor allele frequency (MAF) values are shown in Table 3-2. Genotyping of Phase 1 Super-Seniors and controls was used for comparison [147].

3.4. Discussion

We attempted to re-contact Super-Seniors 9-12 years after they were enrolled. As expected for an older group, even one as healthy as the Super-Seniors, most participants had passed away within this time frame. We could confirm that at least 17 of the original 480 Super-Seniors were still living.

Eight of the 13 individuals re-interviewed in 2017 (54%) still met the criteria for being a Super-Senior; among the individuals who no longer met the Super-Senior criteria, most had developed their health problems in the previous three years. The most common diseases in the group were cardiovascular disease (5 of 13; 38%), and lung disease (1 of 13; 8%). According to the US National Center for Health Statistics from 2012-2013, 29.8% of Americans aged 65 years and older reported having heart disease, and 18.4% reported having cancer [4].

It is possible that compression of morbidity is occurring in these individuals. Compression of morbidity has been suggested in other studies of long-lived people including the New

England Centenarian Study, where they found that between nonagenarians, centenarians, semicentenarians (105-109 years), and supercentenarians (110-119 years), there was a later onset of major age-related diseases as the age group increased [140]. As well, among 15 Okinawa supercentenarians (age at death 110-112) it was found that they had delayed clinically apparent diseases until very late in life, with 83% not reporting clinically apparent disease until 105 years or later [111].

A German health insurance study also found that there was a lower prevalence of comorbidities among those who died as centenarians than those who died in their 80s [145]. Likewise, while 8 Super-Seniors were still disease free, 5 had one disease, and only one individual had co-morbidities.

We compared the current geriatric test scores of the Super-Seniors to their initial interview scores. MMSE and IADL scores declined in the decade between interviews, while GDS and TUG scores increased. For all four tests, the standard deviation increased from the first to second interviews indicating that there was a wider range of physical and cognitive function as the surviving participants aged.

Mean MMSE scores decreased from 28.7 to 23.8 points. Although this is a test of cognitive function, at least a portion of the decreased score was due to visual and/or hearing impairment, which affected the participant's ability to answer some questions. At the time of the second interview, two individuals were legally blind, another was unable to see the images, and three were hard of hearing. Vision impairment for these individuals was not noted at the time of the first interview. It is worth noting however, that the one perfect score among the re-interviews was by a 102-year-old woman with macular degeneration. The traditional cut-off score for being "normal" and without cognitive impairment is 24, however a study of highly educated Caucasians found that a cut-off of 24 resulted in a moderate sensitivity (0.66) and very high specificity (0.99), whereas a cut-off score of 27 achieved a balance between sensitivity (0.89) and specificity (0.91) [148]. At the second interview, the Super-Seniors mean score almost meets the traditional 24-point cut-off. Individually however, the range of scores decreased from 25-30 points to 17-30 points indicating that at while some Super-Seniors have retained their cognitive function, others are experiencing increasing impairment.

All geriatric scores declined with age, with the second interview showing that the Super-Seniors were performing fewer daily tasks in the IADL, having a mean higher depression score and decreased mobility. Their GDS score, however, was still a mean of 2.2 points, below the value of ≥5 that has been indicated as an appropriate cut-off score for depression [130]. As well, although the TUG time increased from a mean of 9.8 to 32.0 seconds, 4 individuals used a walker as a mobility aid and one used a cane. The mean TUG time for individuals without a walker was 17.9 seconds. Although there is no specific standard for TUG time, it has been found that physical fitness indicators such as the TUG are associated with successful aging [129]. It has also been suggested that among adults 65 years and older with a similar disease burden, those who were more physically vigorous experienced compression of morbidity and lived longer [149].

The Okinawa centenarian study found that both BMI and BP decreased with age as individuals transitioned from centenarians to supercentenarians [111]. Longitudinally, when they followed individuals from age groups 99-103 years, to 104-107 years, to 108-111 years, their BMI decreased from 21.47, to 18.81, to 17.43, respectively. As well, their BP (systolic/diastolic) decreased from 142/74mmHg, to 128/70mmHg, to 119/64mmHg. Similarly, following the Super-Seniors from ages 85-94 years, to 96-106 years, we saw a non-significant decrease in BMI from 25.8 (SD = 4.0) to 23.4 (SD = 3.9). Systolic BP of the Super-Seniors decreased from 142 mmHg (SD = 16) to 120mmHg (SD = 35) and diastolic BP remained the same at 72mmHg (SD = 9) and 72mmHg (SD = 10). Interestingly, the re-interviewed Super-Seniors had lower values at their initial interviews for both their systolic and diastolic BP than the mean of the overall Phase 1 Super-Seniors cohort, 152/78mmHg (Table 3-1). Both the values and the decline exhibited in Super-Seniors are consistent with what was observed in the Okinawa study.

APOE has been found in genome-wide scans and cohort studies of longevity; the minor allele of rs7412 is sometimes associated with longevity, and the minor allele of rs429358 is reliably associated with increased mortality [15, 16, 150]. Together these two variants determine *APOE* ε2/3/4 haplotypes. At rs7412, the re-interviewed Super-Seniors had a MAF 0.125, compared to MAF 0.092 in all Phase 1 Super-Seniors, and MAF 0.076 in mid-life controls. At rs429358, the re-interviewed Super-Seniors had MAF 0.083, compared to 0.105 in all Super-Seniors, and 0.157 in controls. This demonstrates that among these elite survivors, Super-Seniors who survived another ~10 years, there appears to be a

slightly higher frequency of the favorable longevity allele of rs7412, and a lower frequency of the deleterious mortality and AD-related allele of rs429358, even when compared to the original group of Phase 1 Super-Seniors.

In *FOXO3,* the G allele of rs2802292 has been associated with longevity [25]. 81.2% re-interviewed Super-Seniors carried at least one G allele, compared to 63.4% in all Phase 1 Super-Seniors and 58.7% of controls. Previously we did not find an association of rs2802292 with the Super-Senior phenotype, compared to mid-life controls, but we do see an apparent enrichment among this small group of survivors. This possible association may be stronger at more advanced ages, or it could be an artifact of the very small sample size.

The 13 Super-Seniors who were re-interviewed represent a small but elite group who delayed disease onset until very late in life, and exhibit compression of morbidity. Partially contributing to this low number, however, is the fact that the majority of Super-Seniors were either not contactable or did not respond and, while most will have passed away, we do not know the exact number. As well, although the average age at recruitment was 88.5 years, some Super-Seniors were already in or near their 100s at the time of initial recruitment and would therefore be less likely to be alive 10 years later.

Although physical and mental decline occurred in the decade between interviews, the majority of Super-Seniors re-interviewed still met the original health criteria. These observations are consistent with reports of compression of morbidity at extreme ages, particularly in centenarians [142, 144, 145]. The increased frequency of longevity-associated variants in this small group of survivors is consistent with studies that reported a larger role for genetics at older ages [6, 151].

Table 3-1. Characteristics of Super-Seniors at two interviews approximately 10 years apart.

	Super-Seniors Study	Re-contacts		*p*-value 1st and 2nd interviews
		First Interview	Second Interview	
Year	2004-2007	2004-2007	2016-2017	
N	480	13		
Male	155	3		
Female	325	10		
Age mean (SD) years	88.5 (2.9)	89.3 (2.7)	100.1 (3.3)	
Age range	85-105	85-94	96-106	
Years between interviews (SD)	--	10.9 (0.9)		
BMI mean (SD) kg/m^2	24.5 (3.9)	25.8 (4.0)	23.4 (3.9)*	0.071
BP mean systolic (SD) mmHg	152 (21)*	142 (16)*	120 (35)	0.077
BP mean diastolic (SD) mmHg	78 (11)*	72 (9)*	72 (10)	0.98
Heart rate (SD) beats per min	70 (11)*	67 (9)*	74 (11)	0.12
MMSE mean (SD)	28.3 (1.7)	28.7 (1.4)	23.8 (4.2)	0.00036
IADL mean (SD)	21.4 (3.5)	22.3 (1.5)	15.6 (6.1)	0.0016
GDS mean (SD)	1.5 (1.8)	0.5 (1.0)	2.2 (2.5)	0.011
TUG mean (SD) seconds	12.3 (4.3)	9.8 (2.1)	32.0 (28.0)	0.0070

*Blood pressure and heart rate were not available for all initial interviews. Super-Seniors Study blood pressure n = 298, heart rate n = 290. Re-contacts first interview blood pressure n = 12, heart rate n = 12. Re-contacts second interview BMI n = 12.

Table 3-2. Genotype comparison between Super-Senior survivors, and the original Phase 1 collection of Super-Seniors and controls.

	Re-interviewed Super-Seniors (13)		**All Super-Seniors (466)**		**Controls (421)**	
	Genotypes	MAF	Genotypes	MAF	Genotypes	MAF
APOE rs7412	TT x 1 TC x 1 CC x 10	0.125	TT x 6 TC x 69 CC x 363	0.092	TT x 3 TC x 57 CC x 352	0.076
APOE rs429358	CC x 0 CT x 2 TT x 10	0.083	CC x 4 CT x 84 TT x 350	0.105	CC x 10 CT x 110 TT x 293	0.157
FOXO3 rs2802292	GG x 0 GT x 9 TT x 2	0.409	GG x 55 GT x 226 TT x 162	0.379	GG x 47 GT x 189 TT x 166	0.352

MAF = minor allele frequency

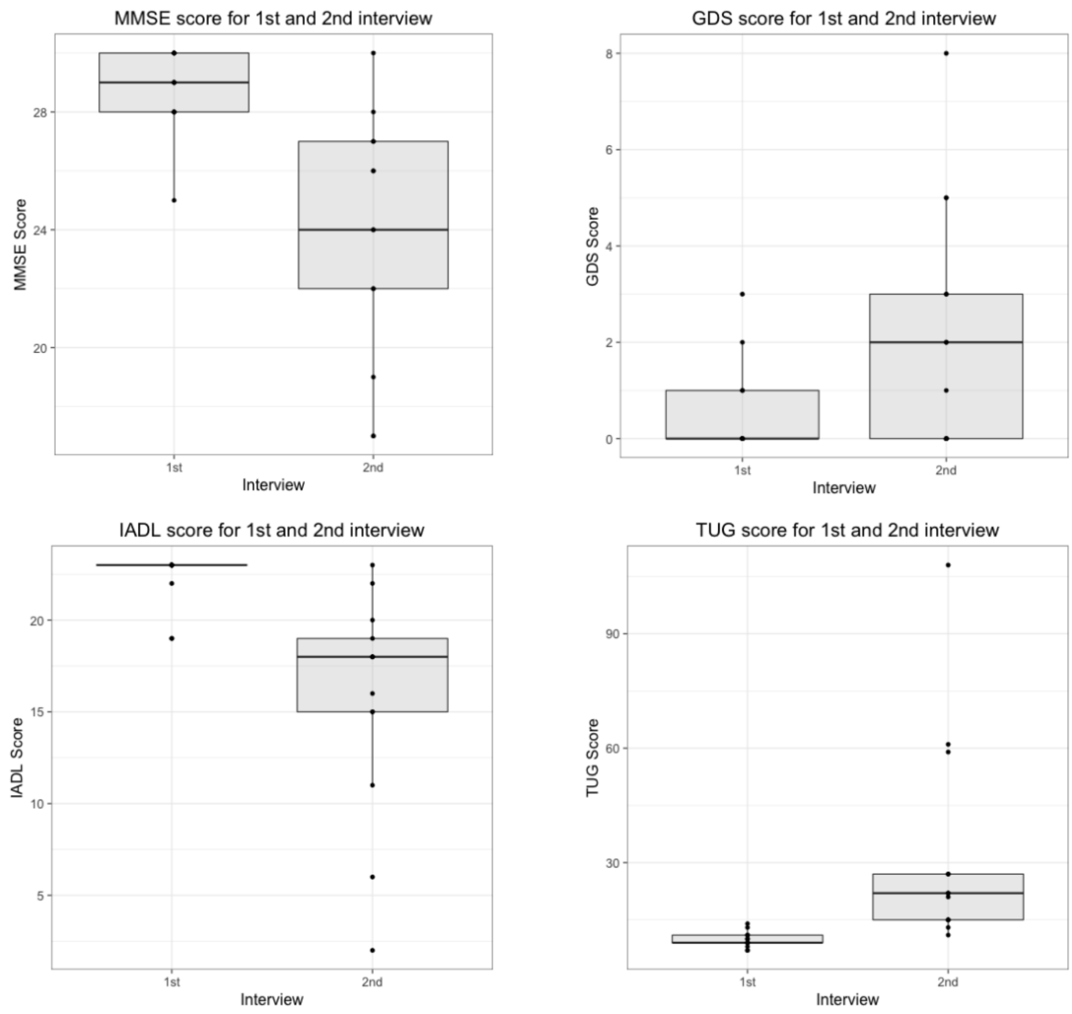


Figure 3-1. Super-Senior geriatric test scores at the first and second interviews.

Chapter 4.

Burden of common complex disease variants in the exomes of two healthy centenarian brothers[2]

Background: It is not understood whether long term good health is promoted by absence of disease risk variants, presence of protective variants, or both. We characterized the exomes of two exceptionally healthy centenarian brothers aged 106 and 109 who had never been diagnosed with cancer, cardiovascular disease, diabetes, Alzheimer disease or major pulmonary disease.

Objective: The aim of this study was to gain insight into whether exceptional health and longevity are a result of carrying fewer disease-associated variants than typical individuals.

Methods: We compared the number of disease-associated alleles, and the proportion of alleles predicted to be functionally damaging, between the centenarian brothers and published population data. Mitochondrial sequence reads were extracted from the exome data in order to analyze mitochondrial variants.

Results: The brothers carry a similar number of common disease-associated variants and predicted damaging variants compared to reference groups. They did not carry any high penetrance clinically actionable variants. They carry mitochondrial haplogroup T, and one brother has a single heteroplasmic variant.

Conclusions: Although our small sample size does not allow for definitive conclusions, a healthy aging and longevity phenotype is not necessarily due to a decreased burden of common disease-associated variants. Instead, it may be rare ‘positive’ variants that play a role in this desirable phenotype.

[2] This chapter has been published in near-identical form. Tindale LC, Zeng A, Bretherick KL, Leach S, Thiessen N and Brooks-Wilson AR. Burden of Common Complex Disease Variants in the Exomes of Two Healthy Centenarian Brothers. Gerontology. 2016; 62(1):58-62.

Written permission was received from S. Karger AG, Basel to republish this article in this thesis.

4.1. Introduction

Extreme longevity is a rare and complex phenotype. Family studies have established a role for genetics in longevity and healthy aging. Individuals have a four times greater chance of surviving to their early nineties if they are a sibling of a centenarian [152], indicating that there is a heritable component to longevity. Offspring of centenarians have a lower incidence of chronic health conditions including hypertension, diabetes mellitus, heart attacks and strokes, pointing to a genetic contribution to healthy aging [117]. To date, only variants in *APOE* and *FOXO3A* have been relatively consistently associated with longevity [9].

To investigate the genetic contributions to longevity and healthy aging, we resequenced the exomes of an unusual pair of brothers who not only lived into their 100s, but had also never been diagnosed with cardiovascular disease (CVD), diabetes, cancer, Alzheimer disease (AD), or major pulmonary disease. We hypothesized that the brothers have: fewer common complex disease associated genome-wide association study (GWAS) variants than other published exomes, fewer variants predicted to be functionally deleterious, and few or no clinically actionable high penetrance variants.

4.2. Design and Methods

4.2.1. Subjects

The younger brother (B1) is currently 106 and continues to ride a stationary bike every day, as well as participate in Tai Chi and yoga; and the older brother (B2) passed away a month before his 110th birthday. The brothers are of Ashkenazi Jewish decent.

In addition to the brothers' extraordinarily long life, they remained healthy and free of major disease to extreme ages. When they were first consented at the ages of 101 and 108, they had Mini Mental Status Exam [122] scores of 28 and 25 (out of 30), Timed Up and Go [124] scores of 11 seconds and 17 seconds, and Instrumental Activities of Daily Living Scale [123] scores of 22 and 18 (out of 23), respectively. This demonstrates that further to being healthy and long-lived, the brothers were also cognitively intact, independently mobile, and capable of performing day-to-day tasks.

The brothers also come from an exceptionally long-lived family. Five out of their six siblings lived to be over 80 years, their mother passed away at 100, their father at 85, and their maternal grandparents both lived into their eighties. Other notable family members include a maternal aunt who lived to 100.

Research ethics board approval was received from the joint Clinical Research Ethics Board of the BC Cancer Agency and the University of British Columbia. Subjects gave written informed consent. Subjects were interviewed as previously described [138].

4.2.2. GWAS variants

73 GWAS SNPs located in coding regions and associated with cancer, CVD, diabetes, AD, or major pulmonary disease in European ancestry populations were filtered from the National Human Genome Research Institute catalogue [153] (Appendix C.3). At the time of download (March 5, 2014) this included 12,899 SNPs from 1,827 publications. The functional impact of missense variants was predicted using SIFT [154].

Public data were used for comparison values. The NHLBI Exome Sequencing Project, Exome Variant Server, release ESP6500S1-V2 was accessed in May 2014 [155]. We used data from their European American cohort, with an average number of 4192 exomes represented for each of our selected GWAS SNPs. The HapMap CEU population included Utah residents with ancestry from northern and western Europe [156]. There were an average of 103 individuals represented for each of the GWAS SNPs.

4.2.3. High penetrance variants

The American College of Medical Genetics and Genomics (ACMG) published a list of 56 genes for which they consider discovery of a pathogenic variant to be an actionable incidental finding [157]. All SNPs in these 56 genes were extracted from sequence data of the 2 brothers and assessed for: previous phenotype associations, functional prediction by SIFT, and minor allele frequency (MAF).

4.2.4. Mitochondrial analysis

Mitochondrial reads were extracted from the whole exome sequencing data using MitoSeek [158]. Full details can be found in Appendix C.1-C.2.

4.3. Results

The brothers' exomes were sequenced with a mean coverage of 183x and 126x. B1 and B2 had 7629 and 7566 missense variants, 8811 and 8719 synonymous variants, and 55 and 48 nonsense variants, respectively.

4.3.1. GWAS variants

The 73 filtered GWAS SNPs included 46 synonymous, 17 missense, 3 nonsense, and 7 UTR variants. Of the 146 alleles at the 73 SNPs, B1 and B2 carried the disease risk variant at 59 (40%) and 56 (38%) alleles (Figure 4-1 A, B). The HapMap CEU population [156] carried risk alleles a mean of 39% of the time (5th percentile = 31%; 95th percentile = 46%) and the NHLBI Exome Sequencing Project (ESP) European Americans [155] carried the risk allele a mean of 38% of the time.

We compared our list of 73 GWAS SNPs to a published list of Ashkenazi Jewish disease alleles [159]. Two SNPs matched this list, rs1799945 and rs1800562.

7 of the 73 SNPs were predicted to be damaging by SIFT [154]. B1 and B2 were homozygous for the risk genotype 0/7 and 2/7, heterozygous 4/7 and 1/7, and homozygous for the non-risk genotype 3/7 and 4/7, respectively (Figure 4-1 C, D). Combined, the brothers had 32% risk alleles in the damaging SNPs; the HapMap CEU had 32% (5th percentile = 14%; 95th percentile = 50%), and the NHLBI ESP had 30%.

4.3.2. High penetrance variants

The brothers did not carry any known pathogenic variants in the 56 'ACMG' genes, however, B2 did have 6 heterozygous variants predicted damaging by SIFT [154], three of which were shared by B1 (Table 4-1).

4.3.3. Mitochondrial analysis

B1 and B2 had mean mitochondrial sequence depths of 29x and 99x, respectively. B1 had 34 homoplasmic variants relative to the Revised Cambridge Reference Sequence (rCRS, NC_012920). B2 shared all 34, with the addition of m.7521G>A, which was not called in B1 due to insufficient depth (Appendix C.4).

B1 and B2 were both assigned to haplogroup T2b25 with respective assignment qualities of 98.3% and 99.5%. Two local variants remained after filtering out haplogroup-specific variants: m.16519T>C (global MAF 61.9%) and m.16129G>A (global MAF 12.2%), present in the control region. One heteroplasmy at 16.7% heteroplasmic frequency (HF) was detected in B2 at position 887, located in the *MT-RNR1* gene. No positions with HF ≥ 10% were reported in B1.

4.4. Discussion

The brothers carried numbers of missense, synonymous, and nonsense variants comparable to those reported in other exome studies [160]. For both the 73 disease-associated GWAS SNPs and the 7 predicted damaging GWAS SNPs, the brothers had similar numbers of risk alleles as the HapMap CEU and NHLBI ESP European American comparison populations. The brothers therefore do not appear to carry a lower burden of common disease variants than reference samples do.

The two Ashkenazi Jewish disease alleles [159] that were included on the list of 73 GWAS SNPs, rs1799945 and rs1800562, are both located in *HFE*, which is associated with hereditary hemochromatosis. Both of these SNPs were considered damaging by SIFT and therefore were also on the list of 7 damaging variants. The brothers were both heterozygous at rs1799945, and were homozygous for the non-risk allele at rs1800562.

Our results are consistent with those of the Leiden Longevity Study, which examined 30 disease-associated SNPs in nonagenarians vs. younger controls and did not observe a difference in the number of disease risk alleles [96]. They are also consistent with those of the Long Life Family Study, where participants did not have fewer risk alleles in SNPs associated with AD, CVD and stroke, type 2 diabetes or cancer, when compared to their

offspring and spouses [102]. As well, Sebastiani and colleagues, did not find lower levels of disease-related alleles in whole genome sequences of a man and a women aged over 114 [161], and that while exceptional longevity is influenced by the combined effects of many SNPs in a genetic signature, the overall abundance of risk alleles was not markedly different between centenarians and controls [162]. Importantly, this new emerging evidence from multiple laboratories points away from lower numbers of disease alleles as a basis for healthy aging.

The brothers were found to carry variants in 'ACMG' genes, but not ones that were known to be pathogenic. Under the recommendations of the ACMG, none of these variants would be reportable to the subjects. The brothers shared rs79011683 in *CACNA1S*, which had unknown MAF and was not found in the literature. They also shared 2 common SNPs in *APOB*, rs676210 and rs679899, neither of which were associated with familial hypercholesterolemia. Members of long-lived families were previously found to have rare functional variants in *APOB* [163] defined as having MAF ≤1% and altering the open reading frame, however neither brother carried such variants, suggesting that genetic variation in *APOB* is probably not contributing to their longevity.

B2 had three additional 'ACMG gene' SNPs including a novel SNP at chr17:757826 in *TP53*, and rs12139527 in *CACNA1S* that has no reported phenotype associations. Also, rs766173 in *BRCA2*, while not significantly associated with breast cancer in Polish women [164], was one of 25 SNPs that showed a trend for rising breast cancer risk in individuals carrying increased numbers of the 25 SNPs [165]. Overall, the brothers did not have known pathogenic variants in ACMG genes, but did show some common and rare genetic variation in *CACNA1S, APOB, TP53* and *BRCA2* that appear unlikely to predispose to disease. Centenarians have also been shown to carry disease variants annotated in the ClinVar database [166] and pathogenic variants in ACMG genes, showing that carrying disease variants does not necessarily prevent exceptional longevity [167].

In the mitochondrial genome, the observation that B2 carried one heteroplasmic variant is consistent with reported population frequencies using 1000 Genomes data [168], in which 25-65% of individuals carried at least one heteroplasmy with HF ≥ 9%. The brothers mitochondrial DNA was of haplogroup T, which has been associated with coronary artery

disease [169] and age-related macular degeneration [170], but for which there has been no reported correlation with longevity.

The brothers shared the same genotype 82.2% of the time, and were concordant for 89.7% of alleles at the 73 SNPs. This is consistent with a genetic contribution to their shared phenotype. It is difficult, however, to distinguish whether this high degree of allele sharing is a result of their kinship, their shared healthy centenarian phenotype, or their founder population ancestry.

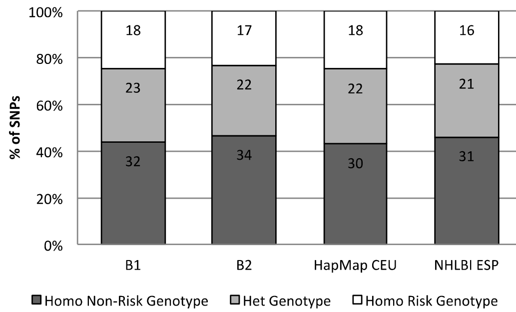
Our results alone are not conclusive due to small sample size. These two brothers, however, are exceptional in a number of ways; they are not only centenarians but are exceptionally healthy and free of common complex diseases, and they have a strong family history of longevity. These data add to a growing and important body of evidence from multiple studies [96, 102, 161, 162, 167] that the healthy longevity phenotype is not due to a lower burden of disease alleles. Instead, rare protective factors may underlie the genetic component of this desirable phenotype.

Table 4-1. Predicted damaging SNPs found in the brothers in ACMG genes.

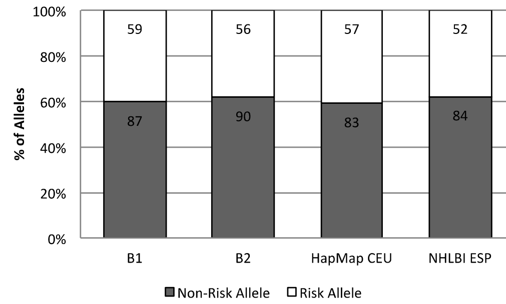
SNP	Alleles*	B1	B2	MAF#	Gene	AA
rs12139527	A/G	AA	**AG**	0.131	*CACNA1S*	Leu1800Ser
rs79011683	A/C	**AC**	**AC**	NA	*CACNA1S*	Val1449Gly
rs676210	G/A	**GA**	**GA**	0.217	*APOB*	Pro2739Leu
rs679899	G/A	**GA**	**GA**	0.462	*APOB*	Ala618Val
rs766173	A/C	AA	**AC**	0.037	*BRCA2*	Asn289His
17:7578265	A/G	AA	**AG**	NA	*TP53*	Ile195Thr

*major/minor, #MAF from the European American cohort of the NHLBI Exome Sequencing Project, MAF = minor allele frequency, B1 = younger brother, B2 = older brother, AA = amino acid, NA = not available.

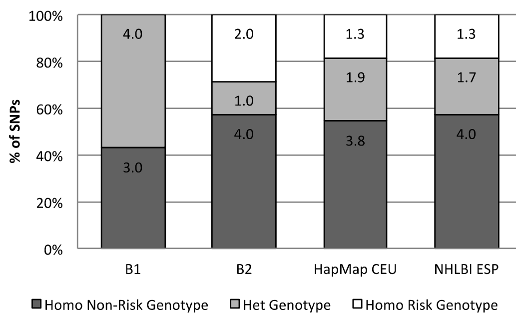
(A) Average number of genotypes for the 73 GWAS SNPs.



(B) Average number of alleles for the 73 GWAS SNPs.



(C) Average number of genotypes for the 7 damaging SNPs.



(D) Average number of alleles for the 7 damaging SNPs.

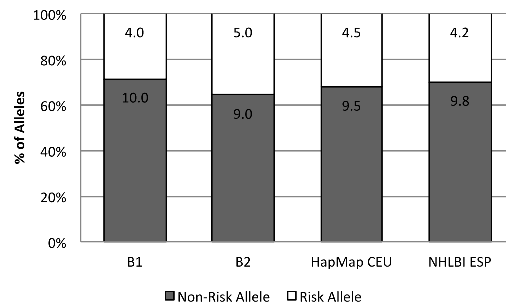


Figure 4-1. Proportion of risk genotypes and alleles in the brothers and other exome data sets.

The difference in the number of GWAS SNPs scored is due to SNPs not being included in the HapMap and NHLBI databases. ‘Damaging’ SNPs include the 7 GWAS SNPs predicted damaging by SIFT. GWAS = genome-wide association study, ESP = exome sequencing project.

Chapter 5.

The search for protective buffering variants: Lipid and Alzheimer's Disease genes associated with healthy aging and longevity in healthy oldest-old[3]

Several studies have found that long-lived individuals do not appear to carry lower numbers of common disease-associated variants than ordinary people; it has been hypothesized that they may instead carry protective variants. An intriguing type of protective variant is buffering variants that protect against variants that have deleterious effects. We genotyped 18 variants in 15 genes related to longevity or healthy aging that had been previously reported as having a gene-gene interaction or buffering effect. We compared a group of 446 healthy oldest-old 'Super-Seniors' (individuals 85 or older who have never been diagnosed with cancer, cardiovascular disease, dementia, diabetes or major pulmonary disease) to 421 random population-based midlife controls. Cases and controls were of European ancestry. Association tests of individual SNPs showed that Super-Seniors were less likely than controls to carry an *APOE ε4* allele or a haptoglobin *HP2* allele. Interactions between *APOE/FOXO3*, *APOE/CRYL1*, and *LPA/CRYL1* did not remain significant after multiple testing correction. In a network analysis of the candidate genes, lipid and cholesterol metabolism was a common theme. *APOE*, *HP*, and *CRYL1* have all been associated with AD, the pathology of which involves lipid and cholesterol pathways. Age-related changes in lipid and cholesterol maintenance, particularly in the brain, may be central to healthy aging and longevity.

5.1. Introduction

Healthy aging is the ability to age successfully without succumbing to disease, with an emphasis on healthspan over lifespan [9]. The genetics of healthy aging and longevity is complex, with few genetic associations replicating between studies. *APOE* (apolipoprotein

[3] This chapter has been published in near-identical form. Tindale LC et al. Lipid and Alzheimer's Disease genes associated with healthy aging and longevity in healthy oldest-old. Oncotarget. 2017; 8(13):pp: 20612-20621.

Oncotarget applies the Creative Commons Attribution 3.0 License to all works they publish.

E) is an exception; genetic variation in this gene has been associated with longevity in multiple genome-wide association studies (GWAS) and candidate gene studies [9, 171]. The *APOE ε4* allele is associated with increased mortality, and is also the major genetic risk factor for late onset Alzheimer's disease (AD). While the *ε4* allele is neither necessary nor sufficient for developing the disease, it increases risk in a dose-dependent manner [18].

Long-lived individuals have been found to carry a burden of disease-associated variants comparable to that observed in typical individuals [96, 102, 137]. One possible explanation for their ability to remain in good health to advanced ages, and still carry deleterious variants, is the concept of genetic buffering. Genetic buffering is a type of epistatic interaction in which a favourable genotype attenuates the effect of one or more deleterious variants. In this model, long-lived individuals may carry harmful (buffered) variants without developing disease, as a result of also carrying protective (buffering) variants. In a paper first suggesting the application of buffering to human longevity, Bergman and colleagues used changes in allele frequencies with age to show buffering of a deleterious *LPA* heterozygote by a buffering *CETP VV* genotype [32] in participants in the Longevity Genes Project [91].

We have assembled a list of genetic variants previously reported as having possible epistatic or buffering/buffered effects related to longevity in human studies. We examined these variants in individuals aged 85 years or older who had never been diagnosed with cancer, cardiovascular disease (CVD), diabetes, dementia, or major pulmonary disease; we call them the 'Super-Seniors' [138]. These healthy oldest-old were compared to random population-based middle-aged controls. We hypothesize that epistatic interactions, in which longevity-promoting buffering variants protect against the effects of deleterious buffered variants, contribute to the Super-Seniors' health and longevity.

5.2. Design and Methods

5.2.1. Subjects

The current analysis included 466 Super-Seniors (female = 312, male = 154; mean = 88.6 years, SD = 3.0, range = 85-108 years), and 421 mid-life controls (female = 253, male =

168; mean = 46.8 years, SD = 3.3, range = 40-54 years) [135, 138]. The Super-Senior group included 140 subjects 90 years and older, 4 of whom were centenarians. Both groups of unrelated individuals were of European ancestry and lived in Metro Vancouver, British Columbia (BC), Canada. Controls were random and population-based, and recruited randomly from BC Medical Services Plan lists. Research ethics board approval was received from the joint Clinical Research Ethics Board of the BC Cancer Agency and the University of British Columbia and Simon Fraser University. All subjects gave written informed consent.

5.2.2. Literature search

A literature search for protective buffering and deleterious buffered variants, as well as other epistatic effects associated with longevity was performed in PubMed. PubMed was chosen because of its biomedical and clinical focus. Search terms included combinations of: buffering, epistasis, aging, longevity, human, and genetics. Only variants found in human studies were considered. Variants located in the same gene were verified not to be in linkage disequilibrium at a threshold of $r^2>0.8$.

5.2.3. Genotyping

Sixteen SNPs were genotyped using Sequenom (San Diego, USA) iPLEX Gold technology at the McGill University and Genome Quebec Innovation Center. Two markers with a call rate below 95% were re-genotyped by the same method. 11 samples with a call rate <90% across all markers were excluded. Three markers that could not be genotyped by the Sequenom method were either replaced by another marker in linkage disequilibrium (rs2542052 in *LPA*) or genotyped by TaqMan® (rs56354395 in *ADIPOQ*) or PCR (rs72294371 in *HP*). Custom TaqMan® probes were designed using the Thermo Fisher Scientific (Waltham, USA) online tool (www.thermofisher.com).

A 1724 bp insertion in *HP* was genotyped by PCR using a two primer design as described by Koch *et al.* [172]; products were sized on an agarose gel. The first primer set: 5'-AGCCCACCCCTCCACCTATGTGCC-3' and 5'-GCTTAAGATCCCAGTCGCATACC-3' [173], yielded a 3221 or 4945bp product, corresponding to the *HP1* allele and *HP2* allele, respectively. Because the larger *HP2* product was not always clearly visible when the gel

was imaged, a second set of primers was used to detect this allele. The second primer set: 5’-CCCAGCCTCTTCTGCTCTTA-3’ and 5’-TGCACATCAATCTCCTTCCA-3’ yielded a 248bp product only when the *HP*2 allele was present.

5.2.4. Association tests of individual variants

Analyses were performed in R 3.2.2. Individual variants were tested using logistic regression to estimate odds ratios and 95% confidence intervals for associations between Super-Senior status and variants. Super-Seniors and controls were coded as 1 and 0. Models were adjusted for sex. Dominant and additive models were tested. In the dominant model, the major allele homozygote was coded as 0, and the heterozygote and minor allele homozygote were both coded as 1. Exceptions to this were *APOE* and *HP,* which were coded for the presence of carrying the risk-associated *APOE ε4* allele and *HP*2 allele, respectively. In the additive model, genotypes were coded as 0, 1, 2. All *p* values were determined using the likelihood ratio test. The false discovery rate (FDR threshold = 0.05) was used to adjust for multiple comparisons.

5.2.5. Gene-gene interaction analysis

Gene-gene interactions were tested using an additive-additive model. Logistic regression analysis was conducted as follows: *y ~ variant1 + variant2 + variant1 x variant2 + sex* (function = glm, family = binomial, link = logit). Super-Senior/control status was the outcome variable.

First, the 7 epistatic pairs from the literature were independently tested to see if they were observed in our population. Then all combinations of putative protective and deleterious variants were compared. FDR was used to adjust for multiple comparisons.

5.2.6. Network analysis

Pathway analysis was conducted using QIAGEN Ingenuity® Pathway Analysis (IPA®, QIAGEN Redwood City, www.qiagen.com/ingenuity) to characterize the types of genes identified during the literature search. IPA® uses the curated Ingenuity® Knowledge Base constructed from peer-reviewed journals and biomedical databases to construct networks

of connections between genes and molecules. The 15 genes from the literature search were entered into IPA® to produce a network that was then “grown” to include additional related molecules. IPA® was also used to identify functions and diseases that were most significantly represented in the network.

5.3. Results

5.3.1. Candidate variants

A search in PubMed of the combinations “epistasis AND aging”, “epistasis AND longevity”, “buffering AND aging”, “buffering AND longevity”, “human”, and “genetics” produced a list of 111 papers of interest. Manual review of the papers and, in some cases, references cited within them, identified 18 variants in 15 genes suspected as having an interaction related to aging or longevity (Table 5-1). This included 15 SNPs, a 1bp deletion, a 1724bp deletion, and the well-characterized *APOE* haplotype.

5.3.2. Genotypes and Quality Control

After excluding 11 samples with a call rate <90%, there were 459 (152 male, 307 female) Super-Seniors and 417 (166 male, 251 female) controls. The haptoglobin (*HP)* variant genotyped by PCR had a call rate of 93%. SNP call rates all exceeded 95%. *LPA* SNP rs3798220 had a minor allele frequency (MAF) < 5% in our study population so was excluded from analysis. There were no significant deviations from Hardy-Weinberg Equilibrium in controls when corrected using false discovery rate.

5.3.3. Association tests of individual variants

There was a greater proportion of female Super-Seniors [odds ratio (OR) 1.33, 95% confidence interval (CI) = 1.01-1.76], so sex was included in all models. Genotype frequencies for all variants are shown in Table 5-2. When the 17 variants were tested for association with healthy aging, under dominant and additive models, only the *HP* and *APOE* variants showed significant associations (Table 5-3 and Appendix D.1).

Super-Seniors were less likely than controls to carry the known disease risk alleles *HP2* or *APOE ε4*. Carriers of the *HP2* allele had decreased odds of being a Super-Senior, OR 0.63 (95% CI = 0.44-0.90, p = 0.010), as did *APOE ε4* allele carriers, OR 0.59 (95% CI = 0.43-0.81, p = 0.0010). The significance of the association with *HP* did not hold under application of the false discovery rate (FDR) (threshold = 0.05 for 17 comparisons), but *APOE* remained significant after FDR, p = 0.017.

Using an additive model, *HP* genotype was associated with healthy longevity with a per allele OR of 0.83 (95% CI = 0.68-1.00, p = 0.056). Compared to *HP*1 homozygotes, heterozygotes had an OR of 0.62 (95% CI = 0.43-0.90) and *HP2* homozygotes had an OR of 0.64 (95% CI = 0.42-0.96). Super-Senior status also differed significantly by *APOE* haplotype using an overall model with a per allele OR of 0.76 (95% CI = 0.67-0.87, p = 0.00017). Compared to *APOE ε3/3*, *APOE ε2/2* was associated with increased odds for healthy aging, OR = 5.33 (95% CI = 1.55-18.34), and *APOE ε3/4* and *APOE ε2/4* did not reach significance against healthy aging with odds ratios of 0.71 (95% CI = 0.50-1.01) and 0.40 (95% CI = 0.16-1.00), respectively.

5.3.4. Gene-gene interaction analysis

Among 7 previously reported gene-gene interactions, using an additive-additive model we did not observed any significant interactions (Appendix D.2). The interaction term between rs6455128 in *KHDRBS2* (KH domain containing, RNA binding, signal transduction associated 2) and rs7989332 in *CRYL1* (crystallin lambda 1), however, was p = 0.077. Because the original rs6455128/rs7989332 interaction was found in a genome-wide association interaction analysis for AD, we adjusted for *APOE ε4* carrier status and found that the p-value for the interaction decreased slightly (p = 0.061). Odds ratios for individual genotypes are shown in Appendix D.3. Per genotype, it appears that there may be an interaction between *CRYL1* GG and *KHDRBS2* AC/AA.

Since most variants did not have a known interaction partner, we then tested for interactions between all combinations of the nine protective and eight deleterious variants (72 interaction tests) using an additive-additive model. Sex was included in all models. Three additional variant pairs showed evidence of interactions. *APOE* haplotype and rs9486902 in *FOXO3* (forkhead box O3) had a significant interaction (p = 0.035), as did

rs10455872 in *LPA* (lipoprotein(a)) and rs7989332 *CRYL1* (p = 0.041). *APOE* haplotype also showed evidence of an interaction with rs7989332 *CRYL1* (p = 0.049). No interactions withstood FDR correction. Odds ratios for individual genotypes in interacting pairs are shown in Appendix D.3; due to low frequencies only *APOE ε4* carrier vs. non-carrier status, *CRYL1*/*LPA* and *CRYL1*/*KHDRBS2* dominant models are presented. There is some evidence that the *APOE ε4* allele interacts with the rs9486902 *FOXO3* CC genotype, as well as the rs7989332 *CRYL1* GG and GT genotypes.

There were also two pairs of SNPs with possible interactions. Rs1853021 in *LPA* and rs2802292 in *FOXO3* (p = 0.052) and rs1800562 in *HFE* (hemochromatosis) and rs56354395 in *ADIPOQ* (adiponectin, C1Q and collagen domain containing) (p = 0.059).

5.3.5. Network analysis

Network analysis was done using QIAGEN's Ingenuity® Pathway Analysis (IPA®, QIAGEN Redwood City, www.qiagen.com/ingenuity) to characterize the types of genes identified during the literature search. Using the 15 genes from the literature search produced a network that connected 13 of the 15 query genes. The "grow" feature of IPA® was used to expand the network to include additional molecules (Figure 5-1). When growing the network, priority is given to molecules that have the most overlap with the parts of the existing network that are the least connected. Of note, *KHDRBS2* and *CRYL1* did not have a known network connection with each other; however, *CRYL1* was connected to *APOE* by one node. *APOE* was also connected by one node to *FOXO3,* which was connected by one node to *LPA. HFE* and *ADIPOQ* were connected by a single edge.

The top functional and disease category represented in this network was metabolic disease, followed by hematological disease, lipid metabolism, and molecular transport. Multiple pathways are related to the genes in this network; the top 20 functions and diseases of those pathways are listed in Table 5-4; all of the top terms relate to lipids or cholesterol.

5.4. Discussion

APOE has three major alleles: *ε2* has been associated with decreased mortality [15, 16], *ε3* can be considered neutral, and *ε4* is associated with increased risk of AD and mortality [18]. Super-Seniors were less likely to carry an *APOE ε4* allele, a finding that we previously published [135]. The *APOE ε2/2* diplotype was protective. A larger sample size would be needed to more confidently determine the effects of other diplotypes.

The *HP*2 allele contains a duplication of exons 3 and 4 of the haptoglobin gene [172], making the *HP1* and *HP2* alleles functionally different. Although the inverse association of *HP2* carrier status with healthy aging was not significant after FDR correction, it is consistent with the idea that the *HP*1/1 genotype is associated with longevity [174].

Several gene-gene interaction tests gave results that approached but did not achieve statistical significance. While we cannot reject the null hypothesis of no interaction in these cases, they represent candidate pairs with potentially intriguing biological roles that would be worth testing in other studies. One such pair is rs6455128 in *KHDRBS2* and rs7989332 in *CRYL1,* previously associated with AD [175]. Gusareva *et al*. hypothesized that the *CRYL1* encoded crystallin protein may act as a stress-protective heat-shock protein that could have a functional interaction with KHDRBS2, which also has a potential role in response to stress [175]. Gusareva *et al*. postulated that this interaction may occur within the TOR pathway [176], which influences β-amyloid plaques (Aβ) and AD-like deficits in a mouse model [177] and life span in model organisms [178, 179].

Some interactions between all combinations of variants were significant prior to multiple testing correction and may therefore be candidates for future replication analyses. *APOE* haplotype and rs9486902 in *FOXO3* showed an interaction effect. Per genotype, there may be an interaction between the *APOE ε4* allele and the *FOXO3 CC* genotype; *FOXO3 CC* could be a buffering genotype for the deleterious *APOE ε4*. Pathway analysis in IPA® showed that one mechanism of interaction could be through amyloid beta precursor protein, APP. *FOXO3* is part of the insulin/insulin-like growth factor 1 signal pathway and has been associated with longevity [25], and FoxO proteins have been implicated in AD [88].

APOE haplotype and *LPA* rs10455872 had significant interaction effects with rs7989332 in *CRYL1*. The interaction between *APOE* and *CRYL1* may originate from an interaction between the *APOE ε4* allele and the *CRYL1* GG and GT genotypes. Interestingly, the *CRYL1* GG genotype also showed evidence of an interaction with the *KHDRBS2* AC/AA genotypes.

Another example is rs1853021 in *LPA,* which showed p = 0.052 for interaction with rs2802292 in *FOXO3*. Rs1853021 has been associated with elevated Lp(a) lipoprotein level, which is a risk factor for coronary disease, carotid atherosclerosis, and stroke [180]. Rs2802292 has been associated with longevity [25] and all-cause mortality [112].

Rs1800562 in *HFE* and rs56354395 in *ADIPOQ* (p = 0.059) were connected by a single edge in IPA®. The minor allele in *HFE* rs1800562 has been associated with risk of death, but has been seen to increase in frequency at older ages [181]. Increased serum adiponectin levels have been associated with longevity [182, 183]. Two variants in *ADIPOQ*, including rs56354395, have been associated with increased adiponectin levels and the del/del genotype had a higher prevalence in long-lived men [184].

When looking at the overall network, metabolic disease, hematological disease, lipid metabolism, and molecular transport were the most represented functional and disease categories. Despite the fact that many of the individual genes did not show significant differences in our population, it is interesting that lipid and cholesterol functions were significantly over-represented in the network. As well, a review of GWAS-identified risk genes for AD found that the associated genes clustered into three pathways: cholesterol and lipid metabolism, immune system and inflammatory response, and endosome vesicle cycling [185]. The idea that longevity is associated with a favourable lipid profile is not new. It has been found that individuals with exceptional longevity and their offspring have HDL and LDL particle sizes that are significantly larger than controls [92], that offspring of centenarians have favourable lipid profiles compared to their spouse controls [91], and that favourable HDL phenotypes and genotypes may contribute to a lower incidence of age-related diseases such as CVD and decreased mortality [186]. These results are all consistent with lipid and cholesterol maintenance being a key mediator in healthy aging and longevity.

Many of the candidate genes in our study were chosen in the literature reports by the original investigators due to their potential function in longevity. As a result, the selection of genes is biased; however, it is still valuable to examine themes, especially among the genes that were also significant in our study population, which represents long-term good health more than extreme longevity.

CVD and AD are age-related chronic diseases that decrease quality of life and increase risk of mortality. *APOE ε4* confers a dose dependent increased risk for developing AD [18], and it was found in a meta-analysis that while the global frequency of the *ε4* allele is 13.7%, the allele frequency in AD patients is 36.7% [187]. *APOE ε4* is also associated with hyperlipidemia and hypercholesterolemia, and causes neuroinflammation resulting in neurovascular dysfunction [188].

The two main neuropathological features seen in the brains of patients with AD are Aβ and neurofibrillary tangles [18]. ApoE is thought to help to remove Aβ from the brain by transporting it across the blood brain barrier; however, ApoE4 lipoproteins have a decreased binding affinity for Aβ compared to ApoE3 lipoproteins and may therefore be less efficient. ApoE also mediates delivery of cholesterol to neurons in the CNS, which is less efficient by ApoE4 than ApoE3 [188]. The CNS contains about 25% of total body cholesterol, which plays a key role in synaptic plasticity [189]. With age, there are system-wide changes in cholesterol metabolism, and this altered metabolism in the brain may relate to AD development [189]. There is also a decreased amount of cholesterol in the hippocampus and cortical areas in AD patients compared to age-matched controls [188].

Cardiovascular and neurovascular health share common risk factors including diabetes mellitus and hypertension [190]. Cognitively normal individuals with controlled hypertension have less Aβ accumulation than those with unmedicated hypertension. As well, the combination of carrying an *APOE ε4* allele and having unmedicated hypertension increased the risk for Aβ accumulation [190].

Hp is an extracellular chaperone that acts as an antioxidant and anti-inflammatory by binding free hemoglobin, which it then transports to the liver [191]. Hp is produced in the brain in response to stress stimuli; it is increased in the cerebral spinal fluid of patients with AD and other neurodegenerative disorders [192]. Patients with AD consistently show signs of inflammation in their brains and oxidative stress is strongly implicated in AD

etiology [191]. Hp has been found to be more oxidized in AD patients, and *in vitro*, oxidized Hp is less able to perform its chaperone function and inhibit Aβ aggregates [192, 193]. Aβ also competes with hemoglobin for binding to Hp, thus impairing its antioxidant function [192]. There is strong support that Aβ is central in AD pathogenesis and it is thought to trigger oxidative stress-mediated damage that leads to neuronal death [193].

*HP*1 and *HP*2 alleles form structurally different proteins that differ in hemoglobin binding and antioxidant capacity, and may be related to autoimmune and inflammatory disorders [194]. Despite the association of *HP*1/1 with longevity, there are conflicting results from studies looking at *HP* in relation to coronary heart disease (CHD) [195-197].

Our findings provide further evidence that *APOE* and genes in associated pathways are key players in healthy aging. This is consistent with a recent informed GWAS that utilized knowledge about age-related diseases to identify new extreme longevity loci that overlap with those associated with coronary artery disease and AD [30]. As well, in a whole genome sequencing study in a healthy aging cohort aged over 80 years, the Topol and Torkamani group found that healthy aging is associated with reduced genetic susceptibility to AD and coronary artery disease, but not cancer or diabetes [103]. In addition to *APOE* being the most replicable signal in GWAS of longevity, the search for more complex longevity haplotypes and interactions points towards mechanisms related to *APOE,* AD, and lipids.

Our results highlight pathways related to AD and reinforce the importance of lipids and cholesterol in healthy aging and longevity. Due to the exploratory nature of finding epistatic effects, it is unsurprising that the observed effects do not remain significant after multiple testing correction. However, these results are noteworthy as they represent additional candidates for buffering pairs that may be tested in other studies. The study of epistatic interactions, particularly buffering/ buffered pairs, is important as the identification of such pairs may help identify therapeutic drug targets for use in aiding individuals who do not carry health-protective longevity variants.

Table 5-1. Candidate genes and candidate epistatic variants.

Gene	ID	Effect	Proposed Interaction	Reference
APOA1	rs670	Deleterious	Buffered	Garasto et al., 2003 [198]
APOE	*APOE*ε4	Deleterious	*APOE*ε4 buffered by *HP1/1*	Napolioni et al., 2011 [173]
HFE	rs1800562	Deleterious	*HFE T* allele buffered	Tan et al., 2003 [181]
KL	rs9536314	Deleterious	*KL het* buffered	Bergman et al., 2007 [32]
LPA(1)	rs1853021	Deleterious	*LPA het* buffered by *CETP VV*	Bergman et al., 2007 [32]
LPA(2)	rs3798220	Deleterious	Risk for coronary disease	Clarke et al., 2009 [180]
LPA(3)	rs10455872	Deleterious	Risk for coronary disease	Clarke et al., 2009 [180]
MTTP	rs2866164	Deleterious	*MTTP CC* buffered by *APOC3 CC*, *CETP VV*, *ADIPOQ del/del*	Huffman et al., 2012 [35]
PON1	rs662	Deleterious	*PON1 het* buffered	Bonafè et al., 2002 [199]
ADIPOQ	rs56354395	Protective	*ADIPOQ del/del* buffers *MTTP CC*	Atzmon et al., 2008 [184]
APOC3	rs595049 (LD with rs2542052)	Protective	*APOC3 CC* buffers *MTTP CC*	Atzmon et al., 2006 [200]
CETP	rs5882	Protective	*CETP VV* buffers *MTTP CC*, *LPA*(1) *het*	Barzilai et al., 2003 [92]
CRYL1	rs7989332	Protective	AD-associated with *KHDRBS2*	Gusareva et al., 2014 [175]
FOXO1	rs2701858	Protective	Joint effect with *FOXO3*(1) for longevity	Tan et al., 2013 [201]
FOXO3(1)	rs9486902	Protective	Joint effect with *FOXO1* for longevity	Tan et al., 2013 [201]
FOXO3(2)	rs2802292	Protective	*FOXO3 GG* buffering	Willcox et al., 2008 [25]
HP	rs72294371	Protective	*HP1/1* buffers *APOE*ε4	Napolioni et al., 2011 [173]
KHDRBS2	rs6455128	Protective	AD-associated with *CRYL1*	Gusareva et al., 2014 [175]

Effect indicates whether the variant was considered be deleterious or protective in the original literature report. Het = heterozygous, AD = Alzheimer's disease.

Table 5-2. Genotype counts and frequencies in Super-Seniors and controls.

				Super-Seniors			Controls		
Gene	**ID**	**Alleles***	**MAF in study**	**Homo major allele**	**Het**	**Homo minor allele**	**Homo major allele**	**Het**	**Homo minor allele**
ADIPOQ	rs56354395	A>del	0.37	182	212	54	159	192	63
APOA1	rs670	C>T	0.158	308	123	8	283	110	8
APOC3	rs595049	T>G	0.345	204	196	59	176	191	50
APOE	*APOE*ε4	ε2/ ε3>ε4	0.128	365	84	4	293	109	10
CETP	rs5882	T>C	0.304	209	190	44	198	171	32
CRYL1	rs7989332	G>T	0.261	249	179	30	224	169	24
FOXO1	rs2701858	G>A	0.065	388	63	2	371	41	2
FOXO3	rs9486902	C>T	0.142	341	100	13	305	97	12
FOXO3	rs2802292	A>C	0.366	162	226	55	166	189	47
HFE	rs1800562	C>T	0.067	394	64	1	367	49	1
HP	rs72294371	*HP*2>*HP*1	0.448	126	199	99	123	202	63
KHDRBS2	rs6455128	C>A	0.178	321	117	21	279	122	16
KL	rs9536314	T>G	0.163	334	114	11	283	116	17
LPA	rs1853021	C>T	0.152	324	119	9	303	96	15
LPA	rs3798220	T>C	0.017	445	14	0	402	15	0
LPA	rs10455872	T>C	0.07	403	54	2	357	56	4
MTTP	rs2866164	C>G	0.256	234	168	30	229	137	29
PON1	rs662	A>G	0.284	228	177	38	212	151	37

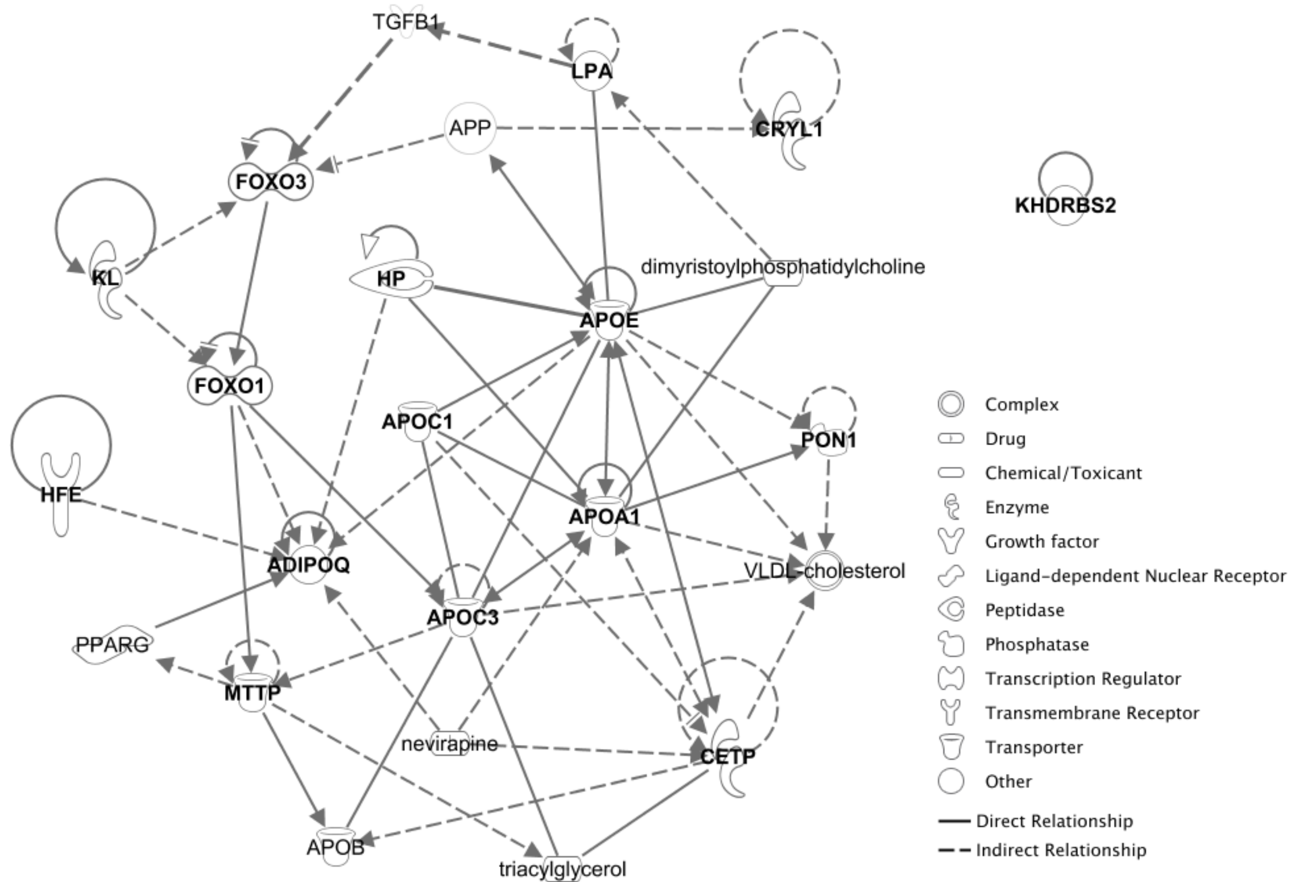
*Major allele > minor allele. Minor allele frequency (MAF) was calculated from the entire study population.

Table 5-3. Odds ratios and 95% confidence intervals for the association between variants in *APOE* and *HP* and healthy aging.

Variant	Model	Super-Seniors	Controls	Genotype	Odds ratio (95% CI)	*p* value
HP rs72294371	Dominant	99	63	1/1	1	0.010 (df=1)
		325	325	1/2 or 2/2	0.63 (0.44-0.90)	
	Additive	99	63	1/1	1	0.056 (df=1)
		199	202	1/2	0.62 (0.43-0.90)	
		126	123	2/2	0.64 (0.42-0.96)	
APOE haplotype	ε4 Dominant	365	293	Non- ε4 carrier	1	0.0010 (df=1)
		88	119	ε4 carrier	0.59 (0.43-0.81)	
	Overall	283	248	ε3/ε3	1	0.00017 (df=5)
		18	3	ε2/ε2	5.33 (1.55-18.34)	
		64	42	ε2/ε3	1.32 (0.86-2.02)	
		77	94	ε3/ε4	0.71 (0.50-1.01)	
		7	15	ε2/ε4	0.40 (0.16-1.00)	
		4	10	ε4/ε4	0.35 (0.11-1.12)	

Table 5-4. The top 20 functions and diseases represented in a candidate gene network in IPA®.

Rank	Diseases and Functions
1	Disorder of lipid metabolism
2	Dyslipidemia
3	Concentration of sterol
4	Quantity of steroid
5	Concentration of triacylglycerol
6	Concentration of lipid
7	Atherosclerosis
8	Metabolism of triacylglycerol
9	Concentration of cholesterol
10	Hyperlipoproteinemia
11	Hyertriglyceridemia
12	Area of atherosclerotic lesion
13	Accumulation of lipid
14	Size of atherosclerotic lesion
15	Efflux of cholesterol
16	Homeostasis of lipid
17	Concentration of cholesterol ester
18	Hyperlipidemia
19	Dementia
20	Transport of lipid



© 2000-2016 QIAGEN. All rights reserved.

Figure 5-1. A network including 15 candidate epistatic longevity genes.

This diagram was created using QIAGENS's Ingenuity® Pathway Analysis software.

Chapter 6.

Allele-specific expression: A pilot study in healthy centenarians

Background: The genetic basis of healthy aging and longevity remains largely unexplained. One hypothesis as to why long-lived individuals do not appear to have a decreased number of common-complex disease risk alleles, is that despite carrying these risk alleles, they express the disease-linked variants at a lower level than the wild-type allele. Heterozygous SNPs can act as indicators to determine if one copy of a gene is more highly expressed. The difference in expression between the two haplotypes of a diploid individual is referred to as allele-specific expression (ASE). We hypothesized that centenarians will have an overall pattern of decreased expression of disease-associated alleles compared to mid-life controls. This would effectively allow them to favour expression of the non-risk allele, despite also carrying the disease-causing variant. In the present study, whole blood is an appropriate tissue in which to achieve our goal of assessing overall disease risk and patterns of ASE.

Methods: We sequenced the transcriptomes of four healthy centenarians with no reported history of cancer, cardiovascular disease, diabetes, dementia, or major pulmonary disease, and four mid-life controls using whole blood. CIBERSORT was used to estimate cell fractions.

Results: Neutrophils were the most abundant source of RNA, followed by CD8+ T cells, resting NK cells, and monocytes. ASE variants were more common in non-coding regions, and centenarians and controls had a comparable distribution of ASE variants by predicted effect. We did not observe an overall bias in expression towards the major or minor allele. Immune pathways were most highly represented among the gene set that showed ASE.

Conclusions: Although we found evidence of ASE occurring in some disease-associated genes, we did not observe any differences in the pattern of expression between centenarians and controls in this small pilot study.

Introduction

Long-lived individuals do not appear to carry a substantially decreased number of common-complex disease risk alleles compared to typical individuals [102, 137]. It may be that despite carrying risk alleles, they express the disease-related variants at a lower level. The difference in expression between the two haplotypes of a diploid individual is referred to as allele-specific expression (ASE). ASE is distributed across the genome in different biological processes and across multiple tissues [39, 202]. ASE can be caused by allele-specific alternative splicing, variation in transcriptional start or stop sites, *cis*-acting regulatory variants, epigenetic differences such as imprinting, DNA methylation, or chromatin state, and differences in mRNA stability [39]. ASE has also been shown to be under genetic control in a study using lymphoblastoid cell lines, where monozygotic twins were found to have a more similar degree of ASE than unrelated controls [46].

Although different tissues have been shown to have important differences in expression [203, 204], whole blood (WB) analysis has advantages. Peripheral blood is a cost-effective as well as clinically relevant tissue because its collection is minimally invasive and it provides a useful overall view of the body since it comes into contact with almost all organs and tissues [205]. WB is also useful when looking for biomarkers for diseases where the tissue of interest is not readily available; for example, the brain or the heart [206]. One study was able to identify and validate a gene expression signature in PAXgene collected blood that distinguished between AD patients and cognitively healthy controls, showing that AD could be detected away from the primary site of the disease [207]. As well, it has been shown that tissue-restricted mRNAs are able to be detected in WB from PAXgene tubes as seen in a study where mutations in cardiac-restricted genes for Long QT Syndrome, Marfan Syndrome, and hypertrophic cardiomyopathy were identified [208]. In the present study, WB is an appropriate tissue in which to achieve our goal of assessing the overall disease risk and patterns of ASE, which may contribute to disease.

In a pilot study of four healthy centenarians and four mid-life controls, we hypothesized that centenarians exhibit an overall pattern of decreased expression of disease-associated alleles. This would effectively allow them to favour expression of the non-risk allele, despite also carrying the disease-causing variant. Heterozygous SNPs in the transcribed region of a gene act as indicators to determine if one copy of the DNA is more

highly expressed. We tested whether preferential expression is observed, and whether it is skewed in the centenarians versus controls by looking at ASE in disease-associated genes and variants.

6.1. Design and Methods

6.1.1. Subjects

Subjects were selected from the Super-Senior Study [146]. Super-Seniors are aged 85 years and older with no reported history of cancer, cardiovascular disease (CVD), diabetes, major pulmonary disease, or dementia. Four Super-Senior centenarians aged 100-104 years, and four controls aged 50-56 years who were not selected for heath, were asked to participate in a transcriptome pilot study. Research ethics board approval was received from the joint Clinical Research Ethics Board of the BC Cancer Agency and the University of British Columbia and Simon Fraser University. All subjects gave written informed consent.

6.1.2. Sample preparation

Whole blood (WB) was collected in PAXgene tubes, which have been designed to preserve the *in vivo* transcription levels of the sample. RNA was extracted using the PAXgene blood miRNA kit (QIAGEN) and globin depleted using the GLOBINclear™kit (Thermo Fisher Scientific, Massachusetts, USA). Transcriptome and exome libraries were constructed using the Illumina TruSeq stranded mRNA kit and Agilent Technologies SureSelect v4+UTR kit, respectively.

6.1.3. Cell fraction prediction

CIBERSORT [209], an *in silico* flow cytometry tool, was used to estimate immune cell type abundances from gene expression data. The signature gene file of 22 distinct immune cell types provided by CIBERSORT was used as a reference of gene expression signatures.

6.1.4. Variant filtering

Exome sequence calling was used to determine heterozygous SNPs, which were then matched to the transcriptome sequences. Autosomal variants were filtered for a read depth (RD) ≥ 30 and an alternate allele frequency > 0.7 or < 0.3.

6.1.5. Analysis of ASE

Variants with ASE were compared by SNPEff predicted effect [210] and by major or minor allele bias as determined from minor allele frequencies reported in dbSNP [211]. To look for disease associations we compared genes showing ASE in our dataset to lists of publically available disease-associations. We looked for evidence of ASE in any of the 60 ACMG recommended genes for reporting exonic incidental findings as clinically relevant [157], as well as variants with SNP-trait associations that reached genome-wide significance listed in the NHGRI-EBI GWAS catalog v1.0.1. Gene lists were entered into QIAGEN Ingenuity® Pathway Analysis to determine enriched pathways and physiological functions.

6.2. Results

6.2.1. Cell fraction prediction

CIBERSORT estimated the mean major cell type in our samples to be neutrophils, followed by CD8+ T cells, monocytes, and resting natural killer (NK) cells (Appendix E.1 and E.3).

6.2.2. Variant filtering

After transcriptome alignment there were ~240-320 million mapped reads per sample. Centenarians had a mean of 344,167 variants per transcriptome (SD = 14,369) and 130,068 heterozygous exome variants (SD = 10,188), while controls had a mean of 355,22 variants per transcriptome (SD = 136,358) and 128,775 (SD = 3,029) heterozygous exome variants. This resulted in a mean of 30,335 (SD = 1,480) matches between the transcriptome and heterozygous exome variants in centenarians, and 28,670 (SD = 2,822)

matches in controls. After filtering for RD ≥ 30, there were 16,679 (SD = 1,108) matches in centenarians, and 15,000 (SD = 1,967) matches in controls. After filtering for an alternate allele frequency > 0.7 or < 0.3, the final list of ASE variants contained 1,145 (SD = 92) variants in centenarians and 1,048 (SD = 136) variants in controls.

6.2.3. Distribution and function of variants with ASE

By predicted effect, variants with ASE were most frequently non-coding: 3 prime UTR variants were the most frequent, followed by intronic variants, downstream gene variants, and then coding missense and synonymous variants (Appendix E.2).

There was no distinguishable difference in the proportion of variants with major allele bias versus minor allele bias when comparing centenarians and controls (Figure 6-1A). Among ASE variants with an allelic ratio ≥ 0.7:0.3, centenarians had 49.1% of variants skewed to the minor allele, compared to 59.9% of variants in controls. At the more extreme ASE variants with an allelic ratio ≥ 0.9:0.1, centenarians had 53.3% of variants skewed to the minor allele, compared to 51.2% among controls. There was no difference in the proportion of the number of variants skewed to the minor allele between centenarians and controls when filtering for an allelic ratio ≥ 0.7:0.3 (X^2=3.4; p=0.065) or for an allelic ratio ≥ 0.9:0.1 (X^2=3.33; p=0.068). There was also no difference in the proportion in % skewed to the minor allele between the allelic ratio ≥ 0.7:0.3 and the allelic ratio ≥ 0.9:0.1 sets (X^2 = 0.06, p=0.806). Figure 6-1B shows the proportion of major and minor allele bias in coding region variants by predicted effect in centenarians and controls. A similar pattern was observed when filtering for the more extreme ASE with an alternate allele frequency > 0.9 or < 0.1 (Figure 6-1C).

6.2.4. Disease association of genes and variants with ASE

82 SNPs with ASE overlapped with SNPs in the NHGRI-EBI GWAS catalog. Trait associations from the catalog were filtered to keep only disease-associations that were excluded for in the Super-Seniors criteria. 21 disease-associated SNPs remained (Table 6-1). Of these, 9 were skewed to the minor allele and 12 were skewed to the major allele. 7 were skewed towards the SNP associated with the disease trait, 9 were skewed away from the SNP associated with the disease trait, and 5 had an unclear disease-associated

allele. 7 SNPs had ASE in multiple samples, all of which were skewed in the same direction.

There were 3373 genes where ASE was observed in either centenarians or controls. When the gene list was entered into QIAGEN Ingenuity® Pathway Analysis to determine enriched pathways and physiological functions, the top canonical pathways were: the antigen presentation pathway; natural killer cell signaling; CD28 signaling in T helper cells; type 1 diabetes mellitus signaling; and the Th1 and Th2 activation pathway (Table 6-2).

Top centenarian genes with ASE were defined as genes for which there was evidence of ASE in 4 centenarians and only 1 or no controls, or in 3 centenarians and no controls; and vice versa for top control genes with ASE. There were 35 top centenarian genes with ASE and 23 top control genes with ASE (Appendix E.4). In QIAGEN Ingenuity® Pathway Analysis, the top canonical pathways for the top centenarian genes with ASE were, in order: citrulline-nitric oxide cycle; LXR/RXR activation; IL-12 signaling and production in macrophages; superpathway of citrulline metabolism; and production of nitric oxide; and ROS in macrophages. The top canonical pathways for the top control genes with ASE were: LXR/RXR activation; role of osteoblast, osteoclasts and chondrocytes in rheumatoid arthritis; IL-10 signaling; CCR5 signalling in macrophages; and Fc-gamma Receptor-mediated phagocytosis in macrophages and monocytes.

Among the 60 ACMG genes, we found that centenarians showed ASE in 8 genes, and controls showed ASE in 11 genes (Table 6-3). Of note, in *LDLR*, which is associated with familial hypercholesterolemia, ASE was observed in 1 centenarian and all 4 controls; and in *TMEM43*, which is associated with arrhythmogenic right ventricular cardiomyopathy, type 5, ASE was observed in 3 centenarians and no controls. ASE variants in the ACMG genes were unique with the exception of two SNPs in *PRKAG2,* which were each present in two subjects. One variant had a high predicted functional impact: a TTCT>TTCTCT indel at chr18: 31068033 in *DSC2*.

6.3. Discussion

Greater read depth (RD) increases the power to detect subtler ASE. Li et al. [212] computed that at a RD of 30 per variant, there is ~60% statistical power to detect an allelic

ratio of 0.7:0.3. RD = 30 provides ~90% power to detect an allelic ratio of 0.8:0.2, and nearly 100% power to detect an allelic ratio of 0.9:0.1. They further calculated that to achieve ~90% of variants in exons having a RD = 30 would require ~190 million reads, whereas to achieve ~95% of variants in exons having a RD = 30 would require ~240 million reads. Since this is an exploratory study, to detect allelic ratios with a difference ≥ 0.7:0.3, samples were sequenced to 330-400 million raw reads per sample.

The main components of WB are red blood cells (RBC), platelets, and white blood cells. RBCs contribute globin transcripts which have been found to make up on average 60% of the mRNA transcripts of WB; however, after globin reduction this decreases to 0.1-0.4% [205]. White blood cells are comprised of numerous cell types including neutrophils, eosinophils, basophils, lymphocytes, and monocytes. Some studies utilize lymphoblastoid cell lines (LCL) that are immortalized by Epstein-Barr virus (EBV), and gene expression has been shown to vary greatly between LCL and WB samples [213]. While LCLs are simpler to interpret because they represent a single cell type, they may be less likely to reflect *in vivo* expression levels due to expression changes during the EBV transformation process [214] and degradation between collection and preparation [215].

Overall, we found that centenarians and controls had a comparable distribution of ASE variants with no difference in proportion by predicted effect. ASE variants were also more common in non-coding regions. We did not observe an overall bias in expression towards the major or minor alleles in centenarians or controls.

When looking at ACMG genes with potential clinical therapeutic value, we found that there is evidence of ASE occurring in some genes that are related to disease etiology. Although it may be an artifact of small sample size, *LDLR*, for which ASE was observed in 1 centenarian and all 4 controls, and *TMEM43*, for which ASE was observed in 3 centenarians and no controls, are potentially interesting to look at in a larger study. *LDLR* is associated with familial hypercholesterolemia. *TMEM43* is associated with arrhythmogenic right ventricular cardiomyopathy, type 5. According to ACMG guidelines [157], pathogenic exonic incidental findings may be considered clinically relevant. It is possible that among patients with disorders associated with these genes, even when a pathogenic variant is not identified, ASE may be playing a role in disease pathogenicity if a *cis*-acting factor is affecting the gene transcript.

Immune pathways were most highly represented among the gene set that showed ASE. Immunosenescence is seen in the adaptive as well as innate immune systems [216]. The occurrence of immune dysfunction with aging is well established, including increased susceptibility to infection, increased frequency of neoplasia, increased inflammation, and autoimmune response [217]. A common factor underlying the pathogenesis of many age-related chronic diseases including CVD, cancer, neurodegenerative diseases, and diabetes is inflammation [78, 216]. It has been previously suggested that centenarians may have genetic factors that allow them to better maintain their immune function as they age [218-220].

As this was a pilot study, the sample size was very small, however, we did not find evidence for a decreased expression of disease-linked variants in centenarians compared to mid-life controls. The presence of ASE in some ACMG genes, which are strongly implicated in clinically relevant disorders, suggest that in patients with disorders that are highly associated with a certain gene, but no known causal variant, looking at ASE may be of interest. As well, the representation of immune pathways among the ASE gene set may indicate an underlying mechanism for decreased immune function with aging.

Table 6-1. NHGRI-EBI GWAS catalog SNPs with evidence of allele-specific expression in centenarians of controls.

Gene	rs ID	Associated disease/trait	Skewed to MA?	Skewed to associated allele from GWAS?*	Cents	Controls
MTX1	rs1057941	Multiple cancers	NO	NO	1	0
TRIM66	rs11042023	Obesity	NO	NO	0	1
GSDMB	rs11078927	Asthma	NO	YES	2	1
HIP1	rs1167827	Body mass index	NO	n/a	1	0
SREBF1	rs11868035	Parkinson's disease	YES	NO	0	1
LYSMD4	rs12185079	Post bronchodilator FEV1/FVC ratio in COPD	YES	YES	1	0
SLC25A44	rs2072499	Testicular germ cell tumor	NO	NO	0	1
OR52K2-OR52K1	rs2278170	Amyotrophic lateral sclerosis	NO	n/a	2	0
PRC1	rs2290203	Breast cancer	NO	YES	1	0
GSDMB	rs2290400	Type 1 diabetes, Bronchial hyper-responsiveness in asthma	NO	NO	3	1
GSDMB	rs2305480	Asthma, Ulcerative colitis	NO	n/a	2	1
TNS1	rs2571445	Pulmonary function	YES	YES	1	1
CASP8	rs3769823	Basal cell carcinoma	NO	YES	1	0
TRIM8	rs3850699	Prostate cancer	YES	NO	0	1
SFTPD	rs721917	Chronic obstructive pulmonary disease	NO	NO	1	0
ASAH1	rs7508	Atrial fibrillation	YES	YES	1	1
CD226	rs763361	Type 1 diabetes	YES	n/a	0	1
C9orf72	rs774359	Amyotrophic lateral sclerosis	YES	n/a	0	1
RCCD1	rs79548680	Type 2 diabetes	YES	YES	2	1
RNF207	rs846111	QT interval	NO	NO	1	0
SDCCAG8	rs953492	Diastolic blood pressure	YES	NO	0	1

*n/a indicates that the allele associated with the trait was not clear.
MA = minor allele

Table 6-2. Top canonical pathways and top disease categories among genes with allele-specific expression as determined by QIAGEN Ingenuity® Pathway Analysis.

Top canonical pathways	*p*-value	# genes in pathway
1 Antigen presentation pathway	1.33E-06	18/38
2 Natural killer cell signaling	5.44E-06	37/122
3 CD28 signaling in T helper cells	1.56E-05	38/132
4 Type 1 diabetes mellitus signalling	2.69E-05	33/111
5 Th1 and Th2 activation pathway	1.23E-04	46/185

Table 6-3. ACMG genes with evidence of allele-specific expression in centenarians of controls.

Gene	Disease	Centenarians	Controls
APC	Adenomatous polyposis coli		1
BRCA1	Breast-ovarian cancer, familial 1		1
DSC2	Arrhythmogenic right ventricular cardiomyopathy, type 11	1	
KCNH2	Long QT syndrome 2	1	
LDLR	Familial hypercholesterolemia	1	4
MYH11	Aortic aneurysm, familial thoracic 4	1	
NF2	Neurofibromatosis, type 2		1
PRKAG2	Familial hypertrophic cardiomyopathy 6	2	1
SDHAF2	Paragangliomas 2		1
SMAD3	Loeys-Dietz syndrome type 3		1
SMAD4	Juvenile polyposis syndrome,	1	
TGFBR1	Loeys-Dietz syndrome type 1A Loeys-Dietz syndrome type 2A Marfan's syndrome		1
TMEM43	Arrhythmogenic right ventricular cardiomyopathy, type 5	3	
TNNI3	Familial hypertrophic cardiomyopathy 7		1
TPM1	Familial hypertrophic cardiomyopathy 3		1
TSC1	Tuberous sclerosis 1		1
VHL	Von Hippel-Lindau syndrome	1	
	TOTAL	**11**	**14**

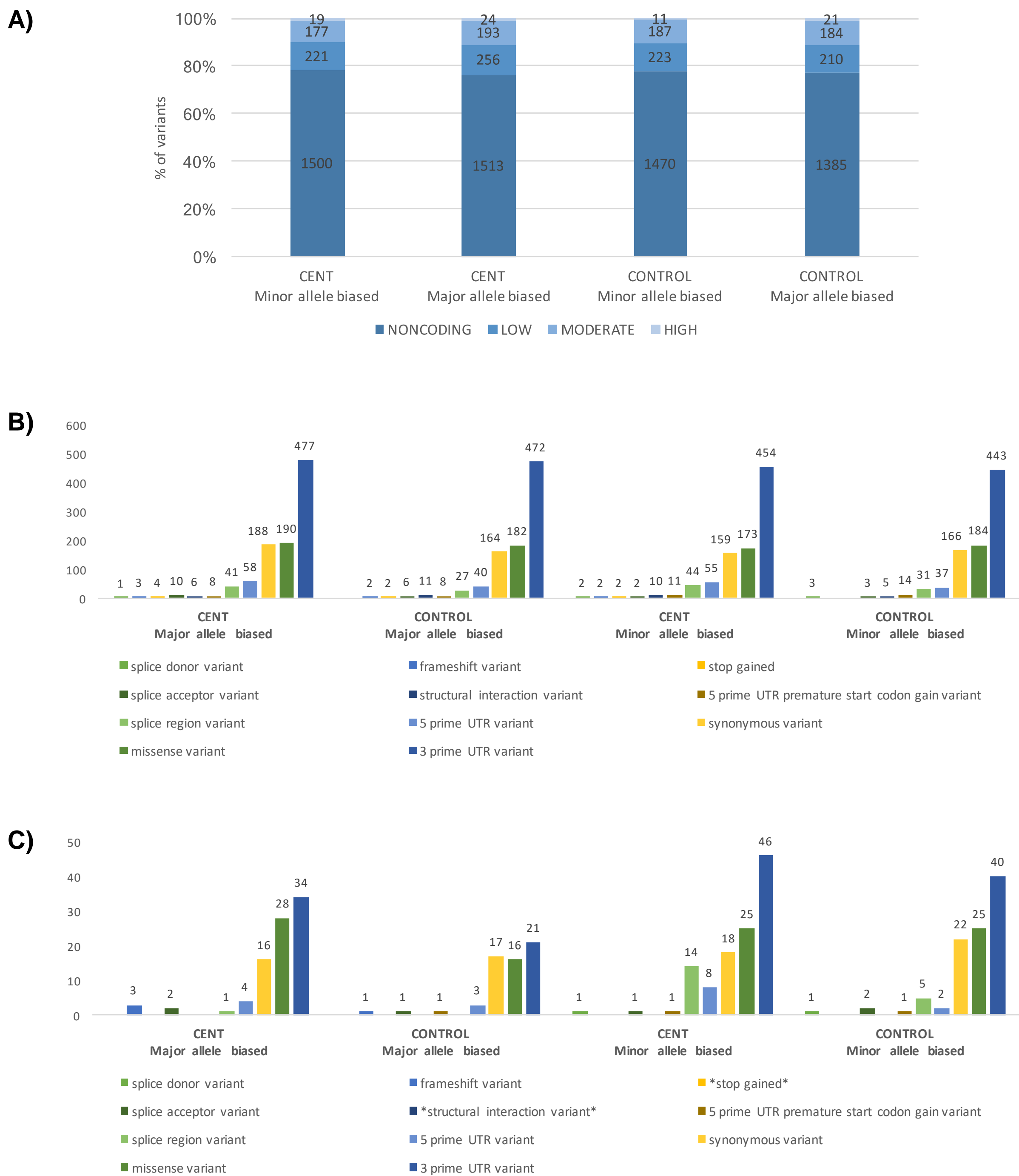


Figure 6-1. A) Proportion of major and minor allele bias by group. B) Coding region variants by major and minor allele bias (n = 3700 coding variants with ASE ≥ 0.7:0.3). C) Coding region variants by major and minor allele bias with an alternate allele frequency > 0.9 or < 0.1 (n = 360 coding variants with ASE ≥ 0.9:0.1).

Chapter 7.

Discussion and Conclusions

7.1. The contribution of this thesis to the field of healthy aging

The vast majority of individuals will develop at least one chronic aging-related disease in their lifetime, resulting in a decreased quality of life and an increased burden on the health care system [4, 5]. We hypothesized that the reason why some individuals are able to age more successfully than others is because they have genetic factors that decrease disease susceptibility and increase disease resilience. Phenotypically, the Super-Seniors are cognitively and physically superior to other aged groups in the literature, and since healthy aging and longevity appear to have a heritable component, it is likely that the Super-Seniors' genetics are contributing to their long and healthy life. By studying an exceptionally healthy long-lived group of individuals in the Super-Senior Study, we hoped to gain a better understanding of the underlying genetics and molecular mechanisms contributing to successful aging and resistance to age-related diseases, which may one day be used to develop therapies to improve the quality of human health.

Other healthy aging and longevity studies include: the Longevity Gene Study which consists of Ashkenazi Jews who are 95 years and older, and their offspring [91], The Leiden Longevity Study which includes sib pairs aged 90 years and over [95], The New England Centenarian Study which is a longitudinal study of centenarians living in the Boston area, their siblings and offspring, and a control group of the spouses of offspring, as well as families lacking longevity [99], The Long Life Family Study which consists of long-lived probands, their siblings, their offspring, and spouse controls [100], the Wellderly cohort who are >80 years old with no chronic diseases and are not taking an chronic medications, however they do not have a control group from the same population [103], the Chinese Longitudinal Healthy Longevity Surveys which includes over 2500 centenarians along with middle-aged controls of Han Chinese decent [31], and the Okinawa Centenarian Study which is a population-based study of centenarians in Okinawa, Japan [106]. The main findings from each of these studies have been highlighted in Chapter 1.

Compared to the other studies of healthy aging and longevity, the major strength of the Super-Senior Study is the thorough characterization of the Super-Seniors' health in addition to their exceptional longevity. We maintained a stringent health inclusion criterion, ensuring that the Super-Seniors truly represent a healthy aging phenotype. Remaining cognitively intact, physically able, and free of disease is the true sign of successful aging, and this is what makes the Super-Seniors so unique.

How elite is the Super-Senior phenotype? From Appendix A.2, 12.4% of seniors over age 85 who were contactable and interested were eligible for the Super-Seniors Study. Taking into account that 28.5% of Canadians aged 85 or older have dementia [126] and that such individuals are unlikely to be contactable and able to understand a phone interviewer, the eligibility rate of living individuals would be closer to 12.4% x (1-0.285) = 8.9%. Furthermore, only 9.0% of individuals born in 1916 lived to be 85 [127]; interpolated from data for 1900 and 1950 birth cohorts. Therefore, the proportion of the 1916 birth cohort who went on to become Super-Seniors is approximately 8.9% x 9.0% = 0.80%.

This thesis explored different aspects of the Super-Seniors' successful aging including their healthy phenotype, their burden of disease variants, the presence of protective buffering variants, and evidence of allele specific expression (ASE). First, it was important to characterize the phenotype of the Super-Seniors in order to more clearly define what "healthy aging" refers to in our population (Chapters 2 and 3). Chapters 4-6 explored three different mechanisms that may prevent disease and premature aging. A decreased burden of chronic disease variants (Chapter 4) could theoretically allow the Super-Seniors to avoid disease by having a decreased overall disease risk. Protective buffering variants (Chapter 5) could provide an explanation for how the Super-Seniors are able to effectively tolerate deleterious variants and attenuate their effects, thereby increasing their resistance to developing disease. Similar to buffering variants, ASE (Chapter 6) could decrease the effectiveness of disease-risk variants by preferentially expressing transcripts with the non-disease associated allele.

A description of the 2004-2007 collection of the Super-Senior Study had not yet been published. The Super-Senior Study provides an important contribution to the field of healthy aging research because while there are other studies of long-lived individuals, the Super-Seniors have a uniquely well-characterized healthy aging phenotype that is not

present in the other studies. We applied strict health criteria during collection to ensure that the Super-Seniors represented an even more exclusive population that have aged successfully, free of disease. In Chapter 2 we showed that Super-Seniors are cognitively and physically high functioning individuals who have evaded major age-related chronic diseases into old age, representing the approximately top 1% for healthspan. The familiality of long lifespan of the parents of Super-Seniors supports the hypothesis that heritable factors contribute to this desirable phenotype.

In Chapter 3 we found evidence for compression of morbidity in the elite subset of Super-Seniors who were re-interviewed 10 years after their initial collection. Although physical and mental decline occurred in the decade between interviews, the majority of Super-Seniors re-interviewed still met the original health criteria. These observations were consistent with reports of compression of morbidity at extreme ages, particularly in centenarians. The increased frequency of longevity- associated variants in this small group of survivors was also consistent with studies that reported a larger role for genetics at older ages.

Many common-complex disease variants that are identified by GWAS have low effect sizes with each variant contributing to a small percentage of the heritability of a disease. Therefore, one hypothesis for how some individuals can escape developing age-related chronic diseases is that they carry an overall burden of fewer disease-associated variants. In Chapter 4 we found that our centenarian brothers did not appear to have decreased disease susceptibility due to carrying fewer chronic disease associated variants. Although our small sample size did not allow for definitive conclusions, a healthy aging and longevity phenotype is not necessarily due to a decreased burden of common disease-associated variants. Overall, the brothers showed a high degree of concordance, sharing 89.7% of their alleles at the 73 GWAS SNPs. This suggests that they may have a shared genetic component contributing to their shared healthy longevity phenotype. While no definitive conclusions can be made due to the small sample size, our results supported a growing body of literature that suggests that long-lived individuals do not carry fewer disease-associated variants than you would expect to see in a typical individual [96, 102], and that the search for the heritability of longevity should focus on the discovery of protective variants.

Protective buffering variants may explain why long-lived individuals are able to carry as many disease-associated variants as the general population and still remain disease-free. We predicted that deleterious/protective buffering pairs are contributing to the successful aging of the Super-Seniors and will show evidence of epistatic interaction effects in association with healthy aging. In Chapter 5 we showed that there is evidence that epistatic interactions may play a role in healthy aging, possibly through buffering variants increasing resistance to developing disease. Searching for epistatic interactions has large multiple comparison issues and therefore it was expected that the majority of our findings would not withstand multiple testing correction. Nonetheless, our results highlighted pathways related to AD, reinforced the importance of lipids and cholesterol in healthy aging and longevity, and provided additional candidates for buffering pairs that may be tested in future studies.

In Chapter 6, although we found evidence of ASE occurring in some disease-associated genes, we did not find decreased expression of disease linked variants in centenarians compared to mid-life controls, nor did we observe any differences in the pattern of ASE between centenarians and controls in this small pilot study. We did, however, see an enrichment of ASE in immune pathways, which may suggest an underlying mechanism for decreased immune function with aging.

I originally hypothesized that Super-Seniors had an overall decreased genetic susceptibility and increased resilience to age- related diseases that allowed them to resist the effects of disease-risk variants that they carry. I found corroborating evidence that the Super-Seniors' health and longevity is at least partially due to their genetic make-up, however whether this contribution is from gene-environmental interactions, gene-gene interactions, carrying protective variants, the presence of ASE in key pathways, or other genetic factors, remains to be elucidated.

7.2. Future directions

7.2.1. Sample size and power

Lack of replication in genes associated with healthy aging and longevity could be due to small sample sizes resulting in false positives, or different replication populations having

different environmental or genetic backgrounds that are interacting with the variant of interest [32]. An undetected underlying gene-environment or gene-gene interactions could lead to an unaccounted for confounder that masks the effect seen in the initial study population.

The main challenge with studying a complex trait such as healthy aging, is obtaining a large enough sample size to detect the likely small effect sizes that we are looking for. Polygenic traits can have large numbers of associated variants that each have a small but important contribution to the trait [10-12]. Adding in the search for gene-gene interactions exponentially increases the number of tests we are conducting, further increasing the need for larger sample sizes.

The only way to truly understand the mystery of aging is through collaboration. The Brooks-Wilson lab in conducting a GWAS of the Super-Seniors Study, and once the initial analysis is complete, we plan to combine our data with an aging consortia for meta-analysis. It is important to follow-up on variants that have trended towards significance in our study as well as others, in a well-powered meta-analysis.

The Cohorts For Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium has conducted meta-analyses with the aim of finding new associations and replicating previous longevity associations [221, 222]. Their cases are aged 90 years and older, and their controls died between 55 and 80 years of age. In their GWAS meta-analysis that included 6,036 cases and 3,757 controls, they were able to replicate the associations of *APOE* and *FOXO3* with longevity [21]. During their discovery phase, none of their associations were genome-wide significant, however they did identify new potential longevity candidates that trended towards significant. This further demonstrates the need for even larger sample sizes, especially when employing a GWAS strategy.

The International Centenarian Consortium – Dementia (ICC-Dementia) is another consortia focusing on individuals aged 95 years and older; they include studies that have at least 80 eligible participants with cognitive function data [223]. The goal of the consortia is to answer epidemiological questions regarding the prevalence, dynamics, and risks of dementia.

Further compounding the need for larger sample sizes, is the need for more extreme age cut-offs for longevity. 80 years was previously considered to be “oldest-old”, but as this thesis as well as other studies have found, the genetic component to healthy aging may be more substantial at even older age groups [99, 219, 224], and there is a need to collaborate and build a cohort of centenarians. Centenarians are rare, but are becoming more common, and it has been reported that while in 1994 the estimated prevalence of centenarians in the United States was 1 in 10,000, by 2012 that estimate had increased to 1 in 5,000 [99]. It is worth noting, however, that meta-analyses can introduce new issues and Sebastiani *et al.* caution researchers to clearly define the definition of longevity across studies, and to pay attention to the effects of sex and ethnicity [225].

7.2.2. Sick elderly as a control group

The Super-Seniors Study also plans to collect a comparison group of age-matched sick elderly. This control group, who will be 85 years and older and will NOT meet the health criteria for being a Super-Senior, will be recruited from assisted living homes and long-term residential care homes. Compared to the mid-life controls, this group of age-matched sick elderly will allow for the comparison of metabolomic and epigenetic factors that may have era effects. While sick elderly will still have lived to a fairly old-age, the question remains as to how the Super-Seniors are able to live to the same age while staying free of disease. The age-matched group will reduce the number of confounding variables in the Super-Seniors studying including the environmental and societal conditions that the participants grew up in.

7.2.3. Gut microbiome of the Super-Seniors

The Super-Seniors Study has begun to collect fecal samples to examine the gut microbiome (GM) profile of our healthy, long-lived group. The microbiome refers to the collection of microbes, their genetic information and the environment in which they interact [226]. The GM is perhaps the most adaptable genetic component in humans, and can adapt to many factors including lifestyle, diet, and aging; as a result, profiling the GM may give clues into the comprehensive mechanisms of aging. The GM has also been associated with many chronic diseases [226], and examining what a healthy aging GM

profile looks like may help us better understand human health and how to maintain a life free of disease.

It has been suggested that nuclear DNA, mitochondrial DNA, and the GM interact and form a more holistic view of human genetics [227]. GM studies that have been conducted in elderly populations to date have found that older populations have a decreased diversity of GM and an increase in pro-inflammatory bacteria [227], as well as a GM profile associated with inflammaging [228]. A Chinese study also found that the gut microbiota of centenarians (100-108 years) was more diverse than younger elderly (80-99 years), and there were differences in the abundances of certain bacterial phyla [229].

In addition to comparing the GM between Super-Seniors and controls, we can also examine the proposed interaction between the GM and host-genetics [227]. The Super-Senior GWAS genotyping results can be used in an integrated analysis with the GM profile to look for interactions between a healthy GM signature and GWAS SNPs. Understanding a healthy aging GM signature may lead to therapeutic interventions to shift the microbiota of patients to one that is favourable to health and longevity.

7.2.4. Epigenetics

Epigenetics, specifically DNA methylation, is a key area for future research in the field of healthy aging. The epigenetic clock uses DNA methylation as a biomarker of biological age that more accurately predicts the functional capability of a tissue or person over chronological age by using a mathematical algorithm to estimate age based on methylation states of 353 CpGs [56]. Epigenetic age, or DNA methylation (DNAm) age, strongly correlates with the chronological age of sorted cells, tissues, and organs [57]. In cell culture, it has also been shown that cellular aging is independent from senescence, telomere length, or DNA damage response, and that the epigenetic clock may be a distinct and intrinsic property of cells [230].

Among neurodegenerative diseases, increased DNAm age was found in the blood independently of blood cell counts, and in the immune related components of blood in patients with Parkinson's Disease [231]. Also, increased DNAm age in the dorsolateral prefrontal cortex was correlated with AD neuropathic markers which included diffuse

plaques, neuritic plaques, and amyloid load, as well as decline in global cognitive functioning, episodic memory, and working memory in patients with AD [232].

DNAm has been found to be a predictor of all-cause mortality in a meta-analysis [233]. In a study that looked at the difference between DNAm age and chronological age, they found that individuals with a 5-year higher difference was associated with a 16% increased mortality risk even after adjusting for health status, lifestyle factors and *APOE e4* status [234]. DNAm age has also been shown to be a potential biomarker for predicting lung cancer susceptibility [235] and as a predictor of mortality among patients with acute ischemic stroke [236]. Further, DNAm age acceleration in the liver has been correlated with increased BMI [237] and decreased grip strength [238].

In a study of Italian long-lived families, it was found that their semi-centenarians (n = 82, mean age = 105.6 years ± 1.6) had a DNAm age (measured in their peripheral blood mononuclear cells) that was 8.6 years younger than their chronological age, and that their offspring had a lower DNAm age that was 5.1 years less than age-matched controls [239]. In female monozygotic twins, it was found that twins had high within-pair correlations for DNAm age, age acceleration, and leukocyte telomere length [238]. These findings suggest that in addition to the strong morbidity and mortality associations of DNAm, there is also has a heritable component.

7.2.5. Translation of genetics to the clinic

Once a healthy aging- associated variant, gene, or pathway is identified, there is the possibility of translating this discovery to the clinic. An example of basic aging research that has become a potential longevity intervention is nicotinamide mononucleotide (NMN), which is currently being tested for safety and bioavailability in human phase I clinical trials for nutraceutical development [240]. In mice, NMN intervention has been shown to mitigate the age-associated physiological decline of normal aging [241], as well as improve impairments in a variety of disease mouse models [242-247]. As reviewed by Imai and Guarente [248] and rooted in the mechanism behind calorie restriction as an aging intervention, nicotinamide adenine dinucleotide (NAD+) regulates sirtuin activity. Sirtuins are a family of proteins that regulate many biological processes including aging and longevity, and interact with other aging-related pathways including FOXO and mTOR

pathways [67]. A promising way to increase NAD+ appears to be through the administration of its precursors, NMN and nicotinamide riboside (NR). NR is also currently in human phase clinical trials as a safe and effective way to increase NAD+ levels [249].

Another anti-aging drug currently under study is metformin. Metformin is already a well-established FDA approved drug for the treatment of type 2 diabetes, but there is evidence in the clinic as well as in model organisms that is may have a wider effect on aging [250]. An initial phase 4 clinical trial running from 2014-2017 aimed to examine whether metformin treatment can restore the gene expression profile of older adults with impaired glucose tolerance back to a younger healthier profile in muscle and adipose tissue (ClinicalTrials.gov Identifier: NCT02432287). The Targeting Aging with Metformin (TAME) trial, which plans to start in 2018, will include patients without diabetes, and look at the time to development of age-related diseases such as cancer, CVD, type 2 diabetes, and dementia [251]. TAME is a randomized, controlled clinical trial that will aim to demonstrate that in addition to metformin's effects on diabetes, it also modulates aging and age-related diseases [252]. Metformin effects glucose metabolism and may also influence inflammation, oxidative damage, diminished autophagy, cell senescence, and apoptosis (ClinicalTrials.gov Identifier: NCT02432287). If TAME, which will take 5-7 years and aims to enrol 3000 participants aged 70-80 years, succeeds, researchers hope that metformin will be the first FDA approved drug against aging [253].

With the discovery of additional aging and longevity pathways using basic research, or the validation of current prospective pathways, there is the opportunity for the translation of more therapies into the clinic.

References.

1. *Human Mortality Database*. University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). Available at http://www.mortality.org or http://www.humanmortality.de (data downloaded on Feb 12, 2015).

2. Cardona C and Bishai D. The slowing pace of life expectancy gains since 1950. BMC Public Health. 2018; 18(1):151.

3. Crimmins EM, Zhang Y and Saito Y. Trends Over 4 Decades in Disability-Free Life Expectancy in the United States. Am J Public Health. 2016; 106(7):1287-1293.

4. National Center for Health Statistics. Health, United States, 2014: With Special Feature on Adults Aged 55–64. Hyattsville, MD. 2015.

5. Canadian Institute for Health Information. Patient Cost Estimator. Available at http://www.cihi.ca (data accessed on Feb 23, 2016).

6. Murabito JM, Yuan R and 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(5):470-479.

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

8. Flachsbart F, Franke A, Kleindorp R, Caliebe A, Blanche H, Schreiber S and Nebel A. Investigation of genetic susceptibility factors for human longevity - a targeted nonsynonymous SNP study. Mutat Res. 2010; 694(1-2):13-19.

9. Brooks-Wilson AR. Genetics of healthy aging and longevity. Hum Genet. 2013; 132(12):1323-1338.

10. Ali O. Genetics of type 2 diabetes. World J Diabetes. 2013; 4(4):114-123.

11. Perola M, Sammalisto S, Hiekkalinna T, Martin NG, Visscher PM, Montgomery GW, Benyamin B, Harris JR, Boomsma D, Willemsen G, Hottenga JJ, Christensen K, Kyvik KO, et al. Combined genome scans for body stature in 6,602 European twins: evidence for common Caucasian loci. PLoS Genet. 2007; 3(6):e97.

12. Yang J, Benyamin B, McEvoy BP, Gordon S, Henders AK, Nyholt DR, Madden PA, Heath AC, Martin NG, Montgomery GW, Goddard ME and Visscher PM. Common

SNPs explain a large proportion of the heritability for human height. Nat Genet. 2010; 42(7):565-569.
13. Deelen J, Beekman M, Uh HW, Helmer Q, Kuningas M, Christiansen L, Kremer D, van der Breggen R, Suchiman HE, Lakenberg N, van den Akker EB, Passtoors WM, Tiemeier H, et al. Genome-wide association study identifies a single major locus contributing to survival into old age; the APOE locus revisited. Aging Cell. 2001; 10(4):686-698.
14. Nebel A, Kleindorp R, Caliebe A, Nothnagel M, Blanche H, Junge O, Wittig M, Ellinghaus D, Flachsbart F, Wichmann HE, Meitinger T, Nikolaus S, Franke A, et al. A genome-wide association study confirms APOE as the major gene influencing survival in long-lived individuals. Mech Ageing Dev. 2011; 132(6-7):324-330.
15. Schupf N, Barral S, Perls T, Newman A, Christensen K, Thyagarajan B, Province M, Rossi WK and Mayeux R. Apolipoprotein E and familial longevity. Neurobiol Aging. 2013; 34(4):1287-1291.
16. Frisoni GB, Louhija J, Geroldi C and Trabucchi M. Longevity and the e2 Allele of Apolipoprotein E: The Finnish Centenarians Study. Journal of Gerontology Series A, Biological Sciences and Medical Sciences. 2001; 56(2):M75-78.
17. Cherbuin N, Leach LS, Christensen H and Anstey KJ. Neuroimaging and APOE genotype: a systematic qualitative review. Dement Geriatr Cogn Disord. 2007; 24(5):348-362.
18. Rocchi A, Pellegrini S, Siciliano G and Murri L. Causative and susceptibility genes for Alzheimer’s disease: a review. Brain Research Bulletin. 2003; 61(1):1-24.
19. Genin E, Hannequin D, Wallon D, Sleegers K, Hiltunen M, Combarros O, Bullido MJ, Engelborghs S, De Deyn P, Berr C, Pasquier F, Dubois B, Tognoni G, et al. APOE and Alzheimer disease: a major gene with semi-dominant inheritance. Mol Psychiatry. 2011; 16(9):903-907.
20. Shadyab AH and LaCroix AZ. Genetic factors associated with longevity: a review of recent findings. Ageing Res Rev. 2015; 19:1-7.
21. Broer L, Buchman AS, Deelen J, Evans DS, Faul JD, Lunetta KL, Sebastiani P, Smith JA, Smith AV, Tanaka T, Yu L, Arnold AM, Aspelund T, et al. GWAS of longevity in CHARGE consortium confirms APOE and FOXO3 candidacy. J Gerontol A Biol Sci Med Sci. 2015; 70(1):110-118.
22. Tissenbaum HA and Guarente L. Increased dosage of a sir-2 gene extends lifespan in Caenorhabditis elegans. Nature. 2001; 410:227-230.

23. van der Horst A and Burgering BM. Stressing the role of FoxO proteins in lifespan and disease. Nat Rev Mol Cell Biol. 2007; 8(6):440-450.

24. Bartke A. Impact of reduced insulin-like growth factor-1/insulin signaling on aging in mammals: novel findings. Aging Cell. 2008; 7(3):285-290.

25. Willcox BJ, Donlon TA, He Q, Chen R, Grove JS, Yano K, Masaki KH, Willcox DC, Rodriguez B and Curb JD. FOXO3A genotype is strongly associated with human longevity. Proc Natl Acad Sci U S A. 2008; 105(37):13987-13992.

26. Morris BJ, Donlon TA, He Q, Grove JS, Masaki KH, Elliott A, Willcox DC and Willcox BJ. Association analyses of insulin signaling pathway gene polymorphisms with healthy aging and longevity in Americans of Japanese ancestry. J Gerontol A Biol Sci Med Sci. 2014; 69(3):270-273.

27. Flachsbart F, Caliebe A, Kleindorp R, Blanche H, von Eller-Eberstein H, Nikolaus S, Schreiber S and Nebel A. Association of FOXO3A variation with human longevity confirmed in German centenarians. Proc Natl Acad Sci U S A. 2009; 106(8):2700-2705.

28. Flachsbart F, Dose J, Gentschew L, Geismann C, Caliebe A, Knecht C, Nygaard M, Badarinarayan N, ElSharawy A, May S, Luzius A, Torres GG, Jentzsch M, et al. Identification and characterization of two functional variants in the human longevity gene FOXO3. Nat Commun. 2017; 8(1):2063.

29. Nygaard HB, Erson-Omay EZ, Wu X, Kent BA, Bernales CQ, Evans DM, Farrer MJ, Vilarino-Guell C and Strittmatter SM. Whole Exome Sequencing of an Exceptional Longevity Cohort. J Gerontol A Biol Sci Med Sci. 2018.

30. Fortney K, Dobriban E, Garagnani P, Pirazzini C, Monti D, Mari D, Atzmon G, Barzilai N, Franceschi C, Owen AB and Kim SK. Genome-Wide Scan Informed by Age-Related Disease Identifies Loci for Exceptional Human Longevity. PLoS Genet. 2015; 11(12):e1005728.

31. Zeng Y, Nie C, Min J, Liu X, Li M, Chen H, Xu H, Wang M, Ni T, Li Y, Yan H, Zhang JP, Song C, et al. Novel loci and pathways significantly associated with longevity. Sci Rep. 2016; 6:21243.

32. Bergman A, Atzmon G, Ye K, MacCarthy T and Barzilai N. Buffering mechanisms in aging: a systems approach toward uncovering the genetic component of aging. PLoS Comput Biol. 2007; 3(8):e170.

33. Chen B and Wagner A. Hsp90 is important for fecundity, longevity, and buffering of cryptic deleterious variation in wild fly populations. BMC Evolutionary Biology. 2012; 12(25):12-25.

34. Rutherford S, Hirate Y and Swalla BJ. The Hsp90 capacitor, developmental remodeling, and evolution: the robustness of gene networks and the curious evolvability of metamorphosis. Crit Rev Biochem Mol Biol. 2007; 42(5):355-372.
35. Huffman DM, Deelen J, Ye K, Bergman A, Slagboom EP, Barzilai N and Atzmon G. Distinguishing between longevity and buffered-deleterious genotypes for exceptional human longevity: the case of the MTP gene. J Gerontol A Biol Sci Med Sci. 2012; 67(11):1153-1160.
36. Heap GA, Yang JH, Downes K, Healy BC, Hunt KA, Bockett N, Franke L, Dubois PC, Mein CA, Dobson RJ, Albert TJ, Rodesch MJ, Clayton DG, et al. Genome-wide analysis of allelic expression imbalance in human primary cells by high-throughput transcriptome resequencing. Hum Mol Genet. 2010; 19(1):122-134.
37. Gaur U, Li K, Mei S and Liu G. Research progress in allele-specific expression and its regulatory mechanisms. J Appl Genet. 2013; 54(3):271-283.
38. Yan H and Zhou W. Allelic variations in gene expression. Current Opinions in Oncology. 2004; 16(1):39-43.
39. Pastinen T. Genome-wide allele-specific analysis: insights into regulatory variation. Nat Rev Genet. 2010; 11(8):533-538.
40. Galupa R and Heard E. X-chromosome inactivation: new insights into cis and trans regulation. Curr Opin Genet Dev. 2015; 31:57-66.
41. Serre D, Gurd S, Ge B, Sladek R, Sinnett D, Harmsen E, Bibikova M, Chudin E, Barker DL, Dickinson T, Fan JB and Hudson TJ. Differential allelic expression in the human genome: a robust approach to identify genetic and epigenetic cis-acting mechanisms regulating gene expression. PLoS Genet. 2008; 4(2):e1000006.
42. Vidal DO, de Souza JE, Pires LC, Masotti C, Salim AC, Costa MC, Galante PA, de Souza SJ and Camargo AA. Analysis of allelic differential expression in the human genome using allele-specific serial analysis of gene expression tags. Genome. 2011; 54(2):120-127.
43. Blomquist TM, Crawford EL and Willey JC. Cis-acting genetic variation at an E2F1/YY1 response site and putative p53 site is associated with altered allele-specific expression of ERCC5 (XPG) transcript in normal human bronchial epithelium. Carcinogenesis. 2010; 31(7):1242-1250.
44. Aminoff A, Ledmyr H, Thulin P, Lundell K, Nunez L, Strandhagen E, Murphy C, Lidberg U, Westerbacka J, Franco-Cereceda A, Liska J, Nielsen LB, Gafvels M, et al.

Allele-specific regulation of MTTP expression influences the risk of ischemic heart disease. J Lipid Res. 2010; 51(1):103-111.
45. Grievink H and Stowell KM. Allele-specific differences in ryanodine receptor 1 mRNA expression levels may contribute to phenotypic variability in malignant hyperthermia. Orphanet J Rare Dis. 2010; 5:10.
46. Cheung VG, Bruzel A, Burdick JT, Morley M, Devlin JL and Spielman RS. Monozygotic twins reveal germline contribution to allelic expression differences. Am J Hum Genet. 2008; 82(6):1357-1360.
47. Birney E, Lieb JD, Furey TS, Crawford GE and Iyer VR. Allele-specific and heritable chromatin signatures in humans. Hum Mol Genet. 2010; 19(R2):R204-209.
48. McDaniell R, Lee BK, Song L, Liu Z, Boyle AP, Erdos MR, Scott LJ, Morken MA, Kucera KS, Battenhouse A, Keefe D, Collins FS, Willard HF, et al. Heritable individual-specific and allele-specific chromatin signatures in humans. Science. 2010; 328(5975):235-239.
49. Stepanow S, Reichwald K, Huse K, Gausmann U, Nebel A, Rosenstiel P, Wabitsch M, Fischer-Posovszky P and Platzer M. Allele-specific, age-dependent and BMI-associated DNA methylation of human MCHR1. PLoS One. 2011; 6(5):e17711.
50. Hutchinson JN, Raj T, Fagerness J, Stahl E, Viloria FT, Gimelbrant A, Seddon J, Daly M, Chess A and Plenge R. Allele-specific methylation occurs at genetic variants associated with complex disease. PLoS One. 2014; 9(6):e98464.
51. Trerotola M, Relli V, Simeone P and Alberti S. Epigenetic inheritance and the missing heritability. Hum Genomics. 2015; 9:17.
52. Lopez-Otin C, Blasco MA, Partridge L, Serrano M and Kroemer G. The hallmarks of aging. Cell. 2013; 153(6):1194-1217.
53. Hayflick L and Moorhead PS. The serial cultivation of human diploid cell strains. Experimental Cell Research. 1961; 25:585-621.
54. Hayflick L. A brief history of the mortality and immortality of cultured cells. The Keio Journal of Medicine. 1998; 47(3):174-182.
55. Mather KA, Jorm AF, Parslow RA and Christensen H. Is telomere length a biomarker of aging? A review. The Journals of Gerontology: Series A. 2011; 66(2):202-213.
56. Horvath S and Raj K. DNA methylation-based biomarkers and the epigenetic clock theory of ageing. Nat Rev Genet. 2018.

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

58. Taipale M, Jarosz DF and Lindquist S. HSP90 at the hub of protein homeostasis: emerging mechanistic insights. Nat Rev Mol Cell Biol. 2010; 11(7):515-528.

59. Calderwood SK, Murshid A and Prince T. The shock of aging: molecular chaperones and the heat shock response in longevity and aging--a mini-review. Gerontology. 2009; 55(5):550-558.

60. Tomaru U, Takahashi S, Ishizu A, Miyatake Y, Gohda A, Suzuki S, Ono A, Ohara J, Baba T, Murata S, Tanaka K and Kasahara M. Decreased proteasomal activity causes age-related phenotypes and promotes the development of metabolic abnormalities. Am J Pathol. 2012; 180(3):963-972.

61. Fontana L, Partridge L and Longo VD. Extending healthy life span--from yeast to humans. Science. 2010; 328(5976):321-326.

62. Houtkooper RH, Williams RW and Auwerx J. Metabolic networks of longevity. Cell. 2010; 142(1):9-14.

63. Harrison DE, Strong R, Sharp ZD, Nelson JF, Astle CM, Flurkey K, Nadon NL, Wilkinson JE, Frenkel K, Carter CS, Pahor M, Javors MA, Fernandez E, et al. Rapamycin fed late in life extends lifespan in genetically heterogeneous mice. Nature. 2009; 460(7253):392-395.

64. Ehninger D, Neff F and Xie K. Longevity, aging and rapamycin. Cell Mol Life Sci. 2014; 71(22):4325-4346.

65. Mair W, Morantte I, Rodrigues AP, Manning G, Montminy M, Shaw RJ and Dillin A. Lifespan extension induced by AMPK and calcineurin is mediated by CRTC-1 and CREB. Nature. 2011; 470(7334):404-408.

66. Onken B and Driscoll M. Metformin induces a dietary restriction-like state and the oxidative stress response to extend C. elegans Healthspan via AMPK, LKB1, and SKN-1. PLoS One. 2010; 5(1):e8758.

67. Satoh A, Imai SI and Guarente L. The brain, sirtuins, and ageing. Nat Rev Neurosci. 2017; 18(6):362-374.

68. Wang K and Klionsky DJ. Mitochondria removal by autophagy. Autophagy. 2014; 7(3):297-300.

69. Gladyshev VN. The free radical theory of aging is dead. Long live the damage theory! Antioxid Redox Signal. 2014; 20(4):727-731.

70. Mockett RJ, Sohal BH and Sohal RS. Expression of multiple copies of mitochondrially targeted catalase or genomic Mn superoxide dismutase transgenes does not extend the life span of Drosophila melanogaster. Free Radic Biol Med. 2010; 49(12):2028-2031.

71. Pérez VI, Van Remmen H, Bokov A, Epstein CJ, Vijg J and Richardson A. The overexpression of major antioxidant enzymes does not extend the lifespan of mice. Aging Cell. 2009; 8(1):73-75.

72. Bonawitz ND, Chatenay-Lapointe M, Pan Y and Shadel GS. Reduced TOR signaling extends chronological life span via increased respiration and upregulation of mitochondrial gene expression. Cell Metab. 2007; 5(4):265-277.

73. Rodier F and Campisi J. Four faces of cellular senescence. J Cell Biol. 2011; 192(4):547-556.

74. Fulop T, Larbi A, Dupuis G, Le Page A, Frost EH, Cohen AA, Witkowski JM and Franceschi C. Immunosenescence and Inflamm-Aging As Two Sides of the Same Coin: Friends or Foes? Front Immunol. 2017; 8:1960.

75. Franceschi C, Bonafè M, Valensin S, Olivieri F, De Luca M, Ottaviani E and De Benedictis G. Inflamm-aging. An evolutionary perspective on immunosenescence. Annals of the New York Academy of Sciences. 2000; 908:244-254.

76. Ahmed AS, Sheng MH, Wasnik S, Baylink DJ and Lau KW. Effect of aging on stem cells. World J Exp Med. 2017; 7(1):1-10.

77. Rando TA and Chang HY. Aging, rejuvenation, and epigenetic reprogramming: resetting the aging clock. Cell. 2012; 148(1-2):46-57.

78. Bektas A, Schurman SH, Sen R and Ferrucci L. Human T cell immunosenescence and inflammation in aging. J Leukoc Biol. 2017; 102(4):977-988.

79. Conese M, Carbone A, Beccia E and Angiolillo A. The Fountain of Youth: A Tale of Parabiosis, Stem Cells, and Rejuvenation. Open Med (Wars). 2017; 12:376-383.

80. Saibil H. Chaperone machines for protein folding, unfolding and disaggregation. Nat Rev Mol Cell Biol. 2013; 14(10):630-642.

81. Rutherford SL and Lindquist S. Hsp90 as a capacitor for morphological evolution. Nature. 1998; 396(6709):336-342.

82. Walker GA and Lithgow GJ. Lifespan extension in C. elegans by a molecular chaperone dependent upon insulin-like signals. Aging Cell. 2003; 2:pp131-139.

83. Karras GI, Yi S, Sahni N, Fischer M, Xie J, Vidal M, D'Andrea AD, Whitesell L and Lindquist S. HSP90 Shapes the Consequences of Human Genetic Variation. Cell. 2017; 168(5):856-866 e812.

84. Johnson JL. Evolution and function of diverse Hsp90 homologs and cochaperone proteins. Biochim Biophys Acta. 2012; 1823(3):607-613.

85. Njemini R, Demanet C and Mets T. Infammatory status as an important determinant of heat shock protein 70 serum concentrations during aging. Biogerontology. 2004; 5(1):31-38.

86. Hsin H and Kenyon C. Signals from the reproductive system regulate the lifespan of C. elegans. Nature. 1999; 399(6734):362-366.

87. Hesp K, Smant G and Kammenga JE. Caenorhabditis elegans DAF-16/FOXO transcription factor and its mammalian homologs associate with age-related disease. Exp Gerontol. 2015; 72:1-7.

88. Manolopoulos KN, Klotz LO, Korsten P, Bornstein SR and Barthel A. Linking Alzheimer's disease to insulin resistance: the FoxO response to oxidative stress. Mol Psychiatry. 2010; 15(11):1046-1052.

89. Mardis ER. Next-generation sequencing platforms. Annu Rev Anal Chem (Palo Alto Calif). 2013; 6:287-303.

90. Kumar R, Ichihashi Y, Kimura S, Chitwood DH, Headland LR, Peng J, Maloof JN and Sinha NR. A High-Throughput Method for Illumina RNA-Seq Library Preparation. Front Plant Sci. 2012; 3:202.

91. Barzilai N, Gabriely I, Gabriely M, Iankowitz N and Sorkin JD. Offspring of centenarians have a favorable lipid profile. Journal of the American Geriatrics Society. 2001; 49(1):76-79.

92. Barzilai N, Atzmon G, Schechter C, Schaefer EJ, Cupples AL, Lipton R, Cheng S and Shuldiner AR. Unique lipoprotein phenotype and genotype associated with exceptional longevity. The Journal of the American Medical Association. 2003; 290(15):2030-2040.

93. Atzmon G, Cho M, Cawthon RM, Budagov T, Katz M, Yang X, Siegel G, Bergman A, Huffman DM, Schechter CB, Wright WE, Shay JW, Barzilai N, et al. Evolution in health and medicine Sackler colloquium: Genetic variation in human telomerase is associated with telomere length in Ashkenazi centenarians. Proc Natl Acad Sci U S A. 2010; 107 Suppl 1:1710-1717.

94. Deluty JA, Atzmon G, Crandall J, Barzilai N and Milman S. The influence of gender on inheritance of exceptional longevity. Aging. 2015; 7(6):412-418.

95. Schoenmaker M, de Craen AJ, de Meijer PH, Beekman M, Blauw GJ, Slagboom PE and Westendorp RG. Evidence of genetic enrichment for exceptional survival using a family approach: the Leiden Longevity Study. Eur J Hum Genet. 2006; 14(1):79-84.

96. 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, et al. GWAS identified disease risk alleles do not compromise human longevity. Proc Natl Acad Sci U S A. 2010; 107(42):18046-18049.

97. Deelen J, Beekman M, Uh HW, Helmer Q, Kuningas M, Christiansen L, Kremer D, van der Breggen R, Suchiman H, Lakenberg N, van den Akker EB, Passtoors WM, Tiemeier H, et al. Genome-wide association study identifies a single major locus contributing to survival into old age; the APOE locus revisited. Aging Cell. 2011; 10(4):686-698.

98. Passtoors WM, Boer JM, Goeman JJ, Akker EB, Deelen J, Zwaan BJ, Scarborough A, Breggen R, Vossen RH, Houwing-Duistermaat JJ, Ommen GJ, Westendorp RG, van Heemst D, et al. Transcriptional profiling of human familial longevity indicates a role for ASF1A and IL7R. PLoS One. 2012; 7(1):e27759.

99. Sebastiani P and Perls TT. The genetics of extreme longevity: lessons from the new England centenarian study. Front Genet. 2012; 3:277.

100. Newman AB, Glynn NW, Taylor CA, Sebastiani P, T.T. P, Mayeux R, Christensen K, Zmuda JM, Barral S, Lee JH, Simonsick EM, Walston JD, Yashin AI, et al. Health and function of participants in the Long Life Family Study: A comparison with other cohorts. Aging. 2011; 3(1):63-76.

101. Sebastiani P, Sun FX, Andersen SL, Lee JH, Wojczynski MK, Sanders JL, Yashin A, Newman AB and Perls TT. Families Enriched for Exceptional Longevity also have Increased Health-Span: Findings from the Long Life Family Study. Front Public Health. 2013; 1:38.

102. Stevenson M, Bae H, Schupf N, Andersen S, Zhang Q, Perls TT and Sebastiani P. Burden of disease variants in participants of the long life family study. Aging. 2015; 7(2):123-132.

103. Erikson GA, Bodian DL, Rueda M, Molparia B, Scott ER, Scott-Van Zeeland AA, Topol SE, Wineinger NE, Niederhuber JE, Topol EJ and Torkamani A. Whole-Genome Sequencing of a Healthy Aging Cohort. Cell. 2016; 165(4):1002-1011.
104. Kauwe JSK and Goate A. Genes for a 'Wellderly' Life. Trends Mol Med. 2016; 22(8):637-639.
105. Zeng Y. Towards Deeper Research and Better Policy for Healthy Aging --Using the Unique Data of Chinese Longitudinal Healthy Longevity Survey. China Economic J. 2012; 5(2-3):131-149.
106. Willcox DC, Willcox BJ, Hsueh WC and Suzuki M. Genetic determinants of exceptional human longevity: insights from the Okinawa Centenarian Study. Age (Dordr). 2006; 28(4):313-332.
107. Willcox DC, Scapagnini G and Willcox BJ. Healthy aging diets other than the Mediterranean: a focus on the Okinawan diet. Mech Ageing Dev. 2014; 136-137:148-162.
108. Willcox DC, Willcox BJ, Todoriki H and Suzuki M. The Okinawan Diet: Health Implications of a Low-Calorie, Nutrient-Dense, Antioxidant-Rich Dietary Pattern Low in Glycemic Load. Journal of the American College of Nutrition. 2009; 28(sup4):500S-516S.
109. Willcox DC, Willcox BJ, Todoriki H, Curb JD and Suzuki M. Caloric restriction and human longevity: what can we learn from the Okinawans? Biogerontology. 2006; 7(3):173-177.
110. Willcox BJ, Willcox DC, He Q, Curb JD and Suzuki M. Siblings of Okinawan centenarians share lifelong mortality advantages. The journals of gerontology Series A, Biological sciences and medical sciences. 2006; 61(4):345-354.
111. Willcox DC, Willcox BJ, Wang N, He Q, Rosenbaum M and Suzuki M. Life at the Extreme Limit: Phenotypic Characteristics of Supercentenarians in Okinawa. J Gerontol A Biol Sci Med Sci. 2008; 63(11):1201-1208.
112. Willcox BJ, Tranah GJ, Chen R, Morris BJ, Masaki KH, He Q, Willcox DC, Allsopp RC, Moisyadi S, Poon LW, Rodriguez B, Newman AB, Harris TB, et al. The FoxO3 gene and cause-specific mortality. Aging Cell. 2016; 15(4).
113. Koropatnick TA, Kimbell J, Chen R, Grove JS, Donlon TA, Masaki KH, Rodriguez BL, Willcox BJ, Yano K and Curb JD. A prospective study of high-density lipoprotein cholesterol, cholesteryl ester transfer protein gene variants, and healthy aging in very old

Japanese-american men. The journals of gerontology Series A, Biological sciences and medical sciences. 2008; 63(11):1235-1240.

114. Tyrovolas S, Polychronopoulos E, Mariolis A, Piscopo S, Valacchi G, Makri K, Zeimbekis A, Tyrovola D, Bountziouka V, Gotsis E, Metallinos G, Tur JA, Matalas A, et al. Is Parental Longevity Associated With the Cardiovascular Risk and the Successful Aging of Their Offspring? Results From the Multinational MEDIS Study. Angiology. 2016.

115. Perls TT, Alpert L and Fretts RC. Middle-aged mothers live longer. Nature. 1997; 389(6647):133.

116. Malovini A, Illario M, Iaccarino G, Villa F, Ferrario A, Roncarati R, Anselmi CV, Novelli V, Cipolletta E, Leggiero E, Orro A, Rusciano MR, Milanesi L, et al. Association study on long-living individuals from Southern Italy identifies rs10491334 in the CAMKIV gene that regulates survival proteins. Rejuvenation Research. 2011; 14(3):283-291.

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

118. Joshi PK, Fischer K, Schraut KE, Campbell H, Esko T and Wilson JF. Variants near CHRNA3/5 and APOE have age- and sex-related effects on human lifespan. Nat Commun. 2016; 7:11174.

119. McDaid AF, Joshi PK, Porcu E, Komljenovic A, Li H, Sorrentino V, Litovchenko M, Bevers RPJ, Rueger S, Reymond A, Bochud M, Deplancke B, Williams RW, et al. Bayesian association scan reveals loci associated with human lifespan and linked biomarkers. Nat Commun. 2017; 8:15842.

120. Sebastiani P, Gurinovich A, Bae H, Andersen S, Malovini A, Atzmon G, Villa F, Kraja AT, Ben-Avraham D, Barzilai N, Puca A and Perls TT. Four Genome-Wide Association Studies Identify New Extreme Longevity Variants. J Gerontol A Biol Sci Med Sci. 2017; 72(11):1453-1464.

121. SAS Institute Inc. Cary, NC. 1989-2007.

122. Folstein MF, Folstein SE and McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res. 1975; 12(3):189-198.

123. Lawton MP and Brody EM. Assessment of older people: self-maintaining and instrumental activities of daily living. Gerontologist. 1969; 9(3):179-186.

124. Podsiadlo D and Richardson S. The timed "Up & Go": a test of basic functional mobility for frail elderly persons. J Am Geriatr Soc. 1991; 36(2):142-148.

125. Sheikh JI and Yesavage JA. Geriatric Depression Scale (GDS) Recent evidence and development of a shorter version. Clinical Gerontologist. 1986; 5:165–173.
126. Ebly EM, Parhad IM, Hogan DB and Fung TS. Prevalence and types of dementia in the very old: results from the Canadian Study of Health and Aging. Neurology. 1994; 44(9):1593-1600.
127. Bell FC and Miller ML. Life Tables for the United States Social Security Area 1900-2100. Social Security Administration. 2005; Actuarial Study No. 120(SSA Pub. No. 11-11536).
128. Granic A, Jagger C, Davies K, Adamson A, Kirkwood T, Hill TR, Siervo M, Mathers JC and Sayer AA. Effect of Dietary Patterns on Muscle Strength and Physical Performance in the Very Old: Findings from the Newcastle 85+ Study. PLoS One. 2016; 11(3):e0149699.
129. Lin PS, Hsieh CC, Cheng HS, Tseng TJ and Su SC. Association between Physical Fitness and Successful Aging in Taiwanese Older Adults. PLoS One. 2016; 11(3):e0150389.
130. Marc LG, Raue PJ and Bruce ML. Screening performance of the 15-item geriatric depression scale in a diverse elderly home care population. Am J Geriatr Psychiatry. 2008; 16(11):914-921.
131. Corsi DJ, Boyle MH, Lear SA, Chow CK, Teo KK and Subramanian SV. Trends in smoking in Canada from 1950 to 2011: progression of the tobacco epidemic according to socioeconomic status and geography. Cancer Causes Control. 2014; 25(1):45-57.
132. Ehrlich S. Effect of fertility and infertility on longevity. Fertil Steril. 2015; 103(5):1129-1135.
133. Grundy E and Tomassini C. Fertility history and health in later life: a record linkage study in England and Wales. Soc Sci Med. 2005; 61(1):217-228.
134. Doblhammer G. Reproductive history and mortality later in life: a comparative study of England and Wales and Austria. Popul Stud (Camb). 2000; 54(2):169-176.
135. Tindale LC, Leach S, Ushey K, Daley D and Brooks-Wilson AR. Rare and common variants in the Apolipoprotein E gene in healthy oldest old. Neurobiol Aging. 2014; 35(3):727 e721-723.
136. Halaschek-Wiener J, Vulto I, Fornika D, Collins J, Connors JM, Le ND, Lansdorp PM and Brooks-Wilson A. Reduced telomere length variation in healthy oldest old. Mech Ageing Dev. 2008; 129(11):638-641.

137. Tindale LC, Zeng A, Bretherick KL, Leach S, Thiessen N and Brooks-Wilson AR. Burden of Common Complex Disease Variants in the Exomes of Two Healthy Centenarian Brothers. Gerontology. 2016; 62(1):58-62.
138. Halaschek-Wiener J, Amirabbasi-Beik M, Monfared N, Pieczyk M, Sailer C, Kollar A, Thomas R, Agalaridis G, Yamada S, Oliveira L, Collins JA, Meneilly G, Marra MA, et al. Genetic variation in healthy oldest-old. PLoS One. 2009; 4(8):e6641.
139. Fries JF. Aging, Natural Death, and the Compression of Morbidity. The New England Journal of Medicine. 1980; 303(3):245-250.
140. Andersen SL, Sebastiani P, Dworkis DA, Feldman L and Perls TT. Health span approximates life span among many supercentenarians: compression of morbidity at the approximate limit of life span. J Gerontol A Biol Sci Med Sci. 2012; 67(4):395-405.
141. Beltran-Sanchez H, Jimenez MP and Subramanian SV. Assessing morbidity compression in two cohorts from the Health and Retirement Study. J Epidemiol Community Health. 2016; 70(10):1011-1016.
142. Ismail K, Nussbaum L, Sebastiani P, Andersen S, Perls T, Barzilai N and Milman S. Compression of Morbidity Is Observed Across Cohorts with Exceptional Longevity. J Am Geriatr Soc. 2016; 64(8):1583-1591.
143. Christensen K, McGue M, Petersen I, Jeune B and Vaupel JW. Exceptional longevity does not result in excessive levels of disability. Proc Natl Acad Sci U S A. 2008; 105(36):13274-13279.
144. Kheirbek RE, Fokar A, Shara N, Bell-Wilson LK, Moore HJ, Olsen E, Blackman MR and Llorente MD. Characteristics and Incidence of Chronic Illness in Community-Dwelling Predominantly Male U.S. Veteran Centenarians. J Am Geriatr Soc. 2017; 65(9):2100-2106.
145. Gellert P, von Berenberg P, Oedekoven M, Klemt M, Zwillich C, Horter S, Kuhlmey A and Drager D. Centenarians Differ in Their Comorbidity Trends During The 6 Years Before Death Compared to Individuals Who Died in Their 80s or 90s. J Gerontol A Biol Sci Med Sci. 2017.
146. Halaschek-Wiener J, Tindale LC, Collins JA, Leach S, McManus B, Madden K, Meneilly G, Le ND, Connors JM and Brooks-Wilson AR. The Super-Seniors Study: Phenotypic characterization of a healthy 85+ population. Plos One. 2018; 13(5):e0197578.

147. Tindale LC, Leach S, Spinelli JJ and Brooks-Wilson A. Lipid and Alzheimer's disease genes associated with healthy aging and longevity in healthy oldest-old. Oncotarget. 2017; 8(13):pp20612-20621.

148. O'Bryant SE, Humphreys JD, Smith GE, Ivnik RJ, Graff-Radford NR, Petersen RC and Lucas JA. Detecting dementia with the mini-mental state examination in highly educated individuals. Arch Neurol. 2008; 65(7):963-967.

149. Sanders JL, Arnold AM, Hirsch CH, Thielke SM, Kim D, Mukamal KJ, Kizer JR, Ix JH, Kaplan RC, Kritchevsky SB and Newman AB. Effects of Disease Burden and Functional Adaptation on Morbidity and Mortality on Older Adults. J Am Geriatr Soc. 2016; 64(6):1242-1249.

150. Ryu S, Atzmon G, Barzilai N, Raghavachari N and Suh Y. Genetic landscape of APOE in human longevity revealed by high-throughput sequencing. Mech Ageing Dev. 2016; 155:7-9.

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

152. Perls TT, Bubrick E, Wager CG, Vijg J and Kruglyak L. Siblings of centenarians live longer. The Lancet. 1998; 351(9115):1560.

153. Hindorff LA, MacArthur J, Morales J, Junkins HA, Hall PN, Klemm AK and Manolio TA. A Catalog of Published Genome-Wide Association Studies. Accessed Mar 2014:Available at: http://www.genome.gov/gwastudies.

154. Kumar P, Henikoff S and Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat Protoc. 2009; 4(7):1073-1081.

155. Exome Variant Server. NHLBI GO Exome Sequencing Project (ESP). Accessed May 2014:Seattle, WA (URL: http://evs.gs.washington.edu/EVS/).

156. International HapMap Consortium. A second generation human haplotype map of over 3.1 million SNPs. Nature. 2007; 449(7164):851-861.

157. Green RC, Berg JS, Grody WW, Kalia SS, Korf BR, Martin CL, McGuire AL, Nussbaum RL, O'Daniel JM, Ormond KE, Rehm HL, Watson MS, Williams MS, et al. ACMG recommendations for reporting of incidental findings in clinical exome and genome sequencing. Genet Med. 2013; 15(7):565-574.

158. Guo Y, Li J, Li CI, Shyr Y and Samuels DC. MitoSeek: extracting mitochondria information and performing high-throughput mitochondria sequencing analysis. Bioinformatics. 2013; 29(9):1210-1211.

159. Ostrer H and Skorecki K. The population genetics of the Jewish people. Hum Genet. 2013; 132(2):119-127.

160. Kiezun A, Garimella K, Do R, Stitziel NO, Neale BM, McLaren PJ, Gupta N, Sklar P, Sullivan PF, Moran JL, Hultman CM, Lichtenstein P, Magnusson P, et al. Exome sequencing and the genetic basis of complex traits. Nat Genet. 2012; 44(6):623-630.

161. Sebastiani P, Riva A, Montano M, Pham P, Torkamani A, Scherba E, Benson G, Milton JN, Baldwin CT, Andersen S, Schork NJ, Steinberg MH and Perls TT. Whole genome sequences of a male and female supercentenarian, ages greater than 114 years. Front Genet. 2011; 2:90.

162. 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, et al. Genetic signatures of exceptional longevity in humans. PLoS One. 2012; 7(1):e29848.

163. Cash TP, Pita G, Domínguez O, Alonso MR, Moreno LT, Borrás C, Rodríguez-Mañas L, Santiago C, Garatachea N, Lucia A, Avellana JA, Viña J, González-Neira A, et al. Exome sequencing of three cases of familial exceptional longevity. Aging Cell. 2014; Epub ahead of print.

164. Ledwoń JK, Hennig EE, Maryan N, Goryca K, Nowakowska D, Niwińska A and Ostrowski J. Common low-penetrance risk variants assoc with breast cancer in Polish women. BMC Cancer. 2013; 13(510).

165. Johnson N, Fletcher O, Palles C, Rudd M, Webb E, Sellick G, dos Santos Silva I, McCormack V, Gibson L, Fraser A, Leonard A, Gilham C, Tavtigian SV, et al. Counting potentially functional variants in BRCA1, BRCA2 and ATM predicts breast cancer susceptibility. Hum Mol Genet. 2007; 16(9):1051-1057.

166. Landrum MJ, Lee JM, Riley GR, Jang W, Rubinstein WS, Church DM and Maglott DR. ClinVar: public archive of relationships among sequence variation and human phenotype. Nucleic Acids Res. 2014; 42(Database issue):D980-985.

167. Freudenberg-Hua Y, Freudenberg J, Vacic V, Abhyankar A, Emde AK, Ben-Avraham D, Barzilai N, Oschwald D, Christen E, Koppel J, Greenwald B, Darnell RB, Germer S, et al. Disease variants in genomes of 44 centenarians. Mol Genet Genomic Med. 2014; 2(5):438-450.

168. Ye K, Lu J, Ma F, Keinan A and Gu Z. Extensive pathogenicity of mitochondrial heteroplasmy in healthy human individuals. Proc Natl Acad Sci U S A. 2014; 111(29):10654-10659.

169. Kofler B, Mueller EE, Eder W, Stanger O, Maier R, Weger M, Haas A, Winker R, Schmut O, Paulweber B, Iglseder B, Renner W, Wiesbauer M, et al. Mitochondrial DNA haplogroup T is associated with coronary artery disease and diabetic retinopathy: a case control study. BMC Med Genet. 2009; 10:35.

170. SanGiovanni JP, Arking DE, Iyengar SK, Elashoff M, Clemons TE, Reed GF, Henning AK, Sivakumaran TA, Xu X, DeWan A, Agron E, Rochtchina E, Sue CM, et al. Mitochondrial DNA variants of respiratory complex I that uniquely characterize haplogroup T2 are associated with increased risk of age-related macular degeneration. PLoS One. 2009; 4(5):e5508.

171. Beekman M, Blanche H, Perola M, Hervonen A, Bezrukov V, Sikora E, Flachsbart F, Christiansen L, De Craen AJ, Kirkwood TB, Rea IM, Poulain M, Robine JM, et al. Genome-wide linkage analysis for human longevity: Genetics of Healthy Aging Study. Aging Cell. 2013; 12(2):184-193.

172. Koch W, Latz W, Eichinger M, Roguin A, Levy AP, Schomig A and Kastrati A. Genotyping of the Common Haptoglobin Hp 1/2 Polymorphism Based on PCR. Clinical Chemistry. 2002; 48(9):1377–1382.

173. Napolioni V, Gianni P, Carpi FM, Predazzi IM and Lucarini N. APOE haplotypes are associated with human longevity in a Central Italy population: evidence for epistasis with HP 1/2 polymorphism. Clin Chim Acta. 2011; 412(19-20):1821-1824.

174. Napolioni V, Gianni P, Carpi FM, Concetti F and Lucarini N. Haptoglobin (HP) polymorphisms and human longevity: a cross-sectional association study in a Central Italy population. Clin Chim Acta. 2011; 412(7-8):574-577.

175. Gusareva ES, Carrasquillo MM, Bellenguez C, Cuyvers E, Colon S, Graff-Radford NR, Petersen RC, Dickson DW, Mahachie John JM, Bessonov K, Van Broeckhoven C, Consortium G, Harold D, et al. Genome-wide association interaction analysis for Alzheimer's disease. Neurobiol Aging. 2014; 35(11):2436-2443.

176. Berchtold D, Piccolis M, Chiaruttini N, Riezman I, Riezman H, Roux A, Walther TC and Loewith R. Plasma membrane stress induces relocalization of Slm proteins and activation of TORC2 to promote sphingolipid synthesis. Nat Cell Biol. 2012; 14(5):542-547.

177. Pierce A, Podlutskaya N, Halloran JJ, Hussong SA, Lin PY, Burbank R, Hart MJ and Galvan V. Over-expression of heat shock factor 1 phenocopies the effect of chronic inhibition of TOR by rapamycin and is sufficient to ameliorate Alzheimer's-like deficits in mice modeling the disease. J Neurochem. 2013; 124(6):880-893.

178. Powers RW, Kaeberlein M, Caldwell SD, Kennedy BK and Fields S. Extension of chronological life span in yeast by decreased TOR pathway signaling. Genes & Development. 2006; 20:174-184.

179. Kapahi P, Zid BM, Harper T, Koslover D, Sapin V and Benzer S. Regulation of lifespan in Drosophila by modulation of genes in the TOR signaling pathway. Curr Biol. 2004; 14(10):885-890.

180. Clarke R, Peden JF, Hopewell JC, Kyriakou T, Goel A, Heath SC, Parish S, Barlera S, Franzosi MG, Rust S, Bennett D, Silveira A, Malarstig A, et al. Genetic Variants Associated with Lp(a)
Lipoprotein Level and Coronary Disease. The New England Journal of Medicine. 2009; 361(26):2518-2528.

181. Tan Q, Bathum L, Christiansen L, Benedictis GD, Dahlgaard J, Frizner N, Vach W, Vaupel JW, Yashin AI, Christensen K and Kruse TA. Logistic Regression Models for Polymorphic and Antagonistic Pleiotropic Gene Action on Human Aging and Longevity. Annals of Human Genetics. 2003; 67(6):598-607.

182. Arai Y, Nakazawa S, Kojima T, Takayama M, Ebihara Y, Shimizu K, Yamamura K, Homma S, Osono Y, Gondo Y, Masui Y, Inagaki H, Kitagawa K, et al. High adiponectin concentration and its role for longevity in female centenarians. Geriatrics & Gerontology International. 2006; 6(32-39).

183. Miura Y, Hashii N, Tsumoto H, Takakura D, Ohta Y, Abe Y, Arai Y, Kawasaki N, Hirose N, Endo T and Sonic. Change in N-Glycosylation of Plasma Proteins in Japanese Semisupercentenarians. PLoS One. 2015; 10(11):e0142645.

184. Atzmon G, Pollin TI, Jill Crandall J, Tanner K, Schechter CB, Scherer PE, Rincon M, Siegel G, Katz M, Lipton RB, Shuldiner AR and Barzilai N. Adiponectin Levels and Genotype: A Potential Regulator of Life Span in Humans. The journals of gerontology Series A, Biological sciences. 2008; 63(5):447-453.

185. Van Cauwenberghe C, Van Broeckhoven C and Sleegers K. The genetic landscape of Alzheimer disease: clinical implications and perspectives. Genet Med. 2015.

186. Milman S, Atzmon G, Crandall J and Barzilai N. Phenotypes and genotypes of high density lipoprotein cholesterol in exceptional longevity. Current Vascular Pharmacology. 2014; 12(5):690-697.

187. Farrer LA, Cupples LA, Haines JL, Hyman B, Kukull WA, Mayeux R, H. Myers RH, Pericak-Vance MA, Risch N, van Duijn CM and Consortium ftAaADMA. Effects of

Age, Sex, and Ethnicity on the Association Between Apolipoprotein E Genotype and Alzheimer Disease A Meta-analysis. JAMA. 1997; 278:1349-1356.
188. Liu CC, Kanekiyo T, Xu H and Bu G. Apolipoprotein E and Alzheimer disease: risk, mechanisms and therapy. Nat Rev Neurol. 2013; 9(2):106-118.
189. Yanagisawa K. Cholesterol and pathological processes in Alzheimer's disease. J Neurosci Res. 2002; 70(3):361-366.
190. Rodrigue KM, Rieck JR, Kennedy KM, Devous MD, Sr., Diaz-Arrastia R and Park DC. Risk factors for beta-amyloid deposition in healthy aging: vascular and genetic effects. JAMA Neurol. 2013; 70(5):600-606.
191. Thambisety M. Do extracellular chaperone proteins in plasma have potential as Alzheimer's disease biomarkers? Biomarkers in medicine. 2010; 4(6):831–834.
192. Spagnuolo MS, Maresca B, La Marca V, Carrizzo A, Veronesi C, Cupidi C, Piccoli T, Maletta RG, Bruni AC, Abrescia P and Cigliano L. Haptoglobin interacts with apolipoprotein E and beta-amyloid and influences their crosstalk. ACS Chem Neurosci. 2014; 5(9):837-847.
193. Cocciolo A, Di Domenico F, Coccia R, Fiorini A, Cai J, Pierce WM, Mecocci P, Butterfield DA and Perluigi M. Decreased expression and increased oxidation of plasma haptoglobin in Alzheimer disease: Insights from redox proteomics. Free Radic Biol Med. 2012; 53(10):1868-1876.
194. Langlois MR and Delanghe JR. Biological and clinical significance of haptoglobin polymorphism in humans. Clinical Chemistry. 1996; 42(10):1589-1600.
195. De Bacquer D, De Backer D, Langlois M, Delanghe J, Kesteloot H and Kornitzer M. Haptoglobin polymorphism as a risk factor for coronary heart
disease mortality. Atherosclerosis. 2001; 157:161-166.
196. Levy AP, Larson MG, Corey D, Lotan R, Vita JA and Benjamin EJ. Haptoglobin phenotype and prevalent coronary heart disease in the Framingham offspring cohort. Atherosclerosis. 2004; 172(2):361-365.
197. Simpson M, Snell-Bergeon JK, Kinney GL, Lache O, Miller-Lotan R, Anbinder Y, Rewers MJ and Levy AP. Haptoglobin genotype predicts development of coronary artery calcification in a prospective cohort of patients with type 1 diabetes. Cardiovasc Diabetol. 2011; 10:99.
198. Garasto S, Rose G, Derango F, Berardelli M, Corsonello A, Feraco E, Mari V, Maletta R, Bruni A, Franceschi C, Carotenuto L and De Benedictis G. The Study of

APOA1, APOC3 and APOA4 Variability in Healthy Ageing People Reveals Another Paradox in the Oldest Old Subjects. Annals of Human Genetics. 2003; 67:54-62.
199. Bonafe M, Marchegiani F, Cardelli M, Olivieri F, Cavallone L, Giovagnetti S, Pieri C, Marra M, Antonicelli R, Troiano L, Gueresi P, Passeri G, Berardelli M, et al. Genetic analysis of Paraoxonase (PON1) locus reveals an increased frequency of Arg192 allele in centenarians. Eur J Hum Genet. 2002; 10(5):292-296.
200. Atzmon G, Rincon M, Schechter CB, Shuldiner AR, Lipton RB, Bergman A and Barzilai N. Lipoprotein genotype and conserved pathway for exceptional longevity in humans. PLoS Biol. 2006; 4(4):e113.
201. Tan Q, Soerensen M, Kruse TA, Christensen K and Christiansen L. A novel permutation test for case-only analysis identifes epistatic effects on human longevity in the FOXO gene family. Aging Cell. 2013; 12:690-694.
202. Palacios R, Gazave E, Goni J, Piedrafita G, Fernando O, Navarro A and Villoslada P. Allele-specific gene expression is widespread across the genome and biological processes. PLoS One. 2009; 4(1):e4150.
203. Fairfax BP, Makino S, Radhakrishnan J, Plant K, Leslie S, Dilthey A, Ellis P, Langford C, Vannberg FO and Knight JC. Genetics of gene expression in primary immune cells identifies cell type-specific master regulators and roles of HLA alleles. Nat Genet. 2012; 44(5):502-510.
204. Liu P, Barb J, Woodhouse K, Taylor JGt, Munson PJ and Raghavachari N. Transcriptome profiling and sequencing of differentiated human hematopoietic stem cells reveal lineage-specific expression and alternative splicing of genes. Physiol Genomics. 2011; 43(20):1117-1134.
205. Mastrokolias A, den Dunnen JT, van Ommen GB, t Hoen PA and van Roon-Mom WM. Increased sensitivity of next generation sequencing-based expression profiling after globin reduction in human blood RNA. BMC Genomics. 2012; 13:28.
206. Kennedy L, Vass JK, Haggart DR, Moore S, Burczynski ME, Crowther D and Miele G. Hematopoietic Lineage Transcriptome Stability and Representation in PAXgeneTM Collected Peripheral Blood Utilising SPIA Single-Stranded cDNA Probes for Microarray. Biomarker Insights. 2008; 3:403-417.
207. Booij BB, Lindahl T, Wetterberg P, Skaane NV, Saebo S, Feten G, Rye PD, Kristiansen LI, Hagen N, Jensen M, Bardsen K, Winblad B, Sharma P, et al. A gene expression pattern in blood for the early detection of Alzheimer's disease. J Alzheimers Dis. 2011; 23(1):109-119.

208. Miller TE, You L, Myerburg RJ, Benke PJ and Bishopric NH. Whole blood RNA offers a rapid, comprehensive approach to genetic diagnosis of cardiovascular diseases. Genetics in Medicine. 2007; 9(1):23-33.

209. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, Hoang CD, Diehn M and Alizadeh AA. Robust enumeration of cell subsets from tissue expression profiles. Nat Methods. 2015; 12(5):453-457.

210. Cingolani P, Platts A, Wang le L, Coon M, Nguyen T, Wang L, Land SJ, Lu X and Ruden DM. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly (Austin). 2012; 6(2):80-92.

211. Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM and Sirotkin K. dbSNP: the NCBI database of genetic variation. Nucleic acids research. 2001; 29(1):308-311.

212. Li G, Bahn JH, Lee JH, Peng G, Chen Z, Nelson SF and Xiao X. Identification of allele-specific alternative mRNA processing via transcriptome sequencing. Nucleic Acids Res. 2012; 40(13):e104.

213. Powell JE, Henders AK, McRae AF, Wright MJ, Martin NG, Dermitzakis ET, Montgomery GW and Visscher PM. Genetic control of gene expression in whole blood and lymphoblastoid cell lines is largely independent. Genome Res. 2012; 22(3):456-466.

214. Joehanes R, Johnson AD, Barb JJ, Raghavachari N, Liu P, Woodhouse KA, O'Donnell CJ, Munson PJ and Levy D. Gene expression analysis of whole blood, peripheral blood mononuclear cells, and lymphoblastoid cell lines from the Framingham Heart Study. Physiol Genomics. 2012; 44(1):59-75.

215. Rainen L, Oelmueller U, Jurgensen S, Wyrich R, Ballas C, Schram J, Herdman C, Bankaitis-Davis D, Nicholls N, Trollinger D and Tryon V. Stabilization of mRNA Expression in Whole Blood Samples. Clinical Chemistry. 2002; 48(11):1883-1890.

216. Fulop T, Witkowski JM, Le Page A, Fortin C, Pawelec G and Larbi A. Intracellular signalling pathways: targets to reverse immunosenescence. Clin Exp Immunol. 2017; 187(1):35-43.

217. Ponnappan S and Ponnappan U. Aging and immune function: molecular mechanisms to interventions. Antioxid Redox Signal. 2011; 14(8):1551-1585.

218. Ostan R, Bucci L, Capri M, Salvioli S, Scurti M, Pini E, Monti D and Franceschi C. Immunosenescence and immunogenetics of human longevity. Neuroimmunomodulation. 2008; 15(4-6):224-240.

219. Franceschi C and Bonafè M. Centenarians as a model for healthy aging. Biochemical Society transactions. 2003; 31(2):457-461.
220. Effros RB. Genetic alterations in the ageing immune system: impact on infection and cancer. Mechanisms of Ageing and Development. 2003; 124(1):71-77.
221. Psaty BM, O'Donnell CJ, Gudnason V, Lunetta KL, Folsom AR, Rotter JI, Uitterlinden AG, Harris TB, Witteman JC, Boerwinkle E and Consortium C. Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium: Design of prospective meta-analyses of genome-wide association studies from 5 cohorts. Circ Cardiovasc Genet. 2009; 2(1):73-80.
222. 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, et al. 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. 2010; 65(5):478-487.
223. Brodaty H, Woolf C, Andersen S, Barzilai N, Brayne C, Cheung KS, Corrada MM, Crawford JD, Daly C, Gondo Y, Hagberg B, Hirose N, Holstege H, et al. ICC-dementia (International Centenarian Consortium - dementia): an international consortium to determine the prevalence and incidence of dementia in centenarians across diverse ethnoracial and sociocultural groups. BMC Neurol. 2016; 16:52.
224. Wheeler HE and Kim SK. Genetics and genomics of human ageing. Philos Trans R Soc Lond B Biol Sci. 2011; 366(1561):43-50.
225. Sebastiani P, Bae H, Gurinovich A, Soerensen M, Puca A and Perls TT. Limitations and risks of meta-analyses of longevity studies. Mech Ageing Dev. 2017; 165(Pt B):139-146.
226. Cho I and Blaser MJ. The human microbiome: at the interface of health and disease. Nat Rev Genet. 2012; 13(4):260-270.
227. Garagnani P, Pirazzini C, Giuliani C, Candela M, Brigidi P, Sevini F, Luiselli D, Bacalini MG, Salvioli S, Capri M, Monti D, Mari D, Collino S, et al. The three genetics (nuclear DNA, mitochondrial DNA, and gut microbiome) of longevity in humans considered as metaorganisms. Biomed Res Int. 2014; 2014:560340.
228. Rampelli S, Candela M, Turroni S, Biagi E, Collino S, Franceschi C, O'Toole PW and Brigidi P. Functional metagenomic profiling of intestinal microbiome in extreme ageing. Aging. 2013; 5(12):902-912.

229. Wang F, Yu T, Huang G, Cai D, Liang X, Su H, Zhu Z, Li D, Yang Y, Shen P, Mao R, Yu L, Zhao M, et al. Gut Microbiota Community and Its Assembly Associated with Age and Diet in Chinese Centenarians. J Microbiol Biotechnol. 2015; 25(8):1195-1204.
230. Lowe D, Horvath S and Raj K. Epigenetic clock analyses of cellular senescence and ageing. Oncotarget. 2016; 7(8):8524-8531.
231. Horvath S and Ritz BR. Increased epigenetic age and granulocyte counts in the blood of Parkinson's disease patients. Aging. 2015; 7(12):1130-1142.
232. Levine ME, Lu AT, Bennett DA and Horvath S. Epigenetic age of the pre-frontal cortex is associated with neuritic plaques, amyloid load, and Alzheimer's disease related cognitive functioning. Aging. 2015; 7(12):1198-1211.
233. Chen BH, Marioni RE, Colicino E, Peters MJ, Ward-Caviness CK, Tsai PC, Roetker NS, Just AC, Demerath EW, Guan W, Bressler J, Fornage M, Studenski S, et al. DNA methylation-based measures of biological age: meta-analysis predicting time to death. Aging. 2016; 8(9):1844-1865.
234. Marioni RE, Shah S, McRae AF, Chen BH, Colicino E, Harris SE, Gibson J, Henders AK, Redmond P, Cox SR, Pattie A, Corley J, Murphy L, et al. DNA methylation age of blood predicts all-cause mortality in later life. Genome Biol. 2015; 16:25.
235. Levine ME, Hosgood HD, Chen B, Absher D, Assimes T and Horter S. DNA methylation age of blood predicts future onset of lung cancer in the women's health initiative. Aging. 2015; 7(9):690-700.
236. Soriano-Tarraga C, Giralt-Steinhauer E, Mola-Caminal M, Ois A, Rodriguez-Campello A, Cuadrado-Godia E, Fernandez-Cadenas I, Cullell N, Roquer J and Jimenez-Conde J. Biological Age is a predictor of mortality in Ischemic Stroke. Sci Rep. 2018; 8(1):4148.
237. Horvath S, Erhart W, Brosch M, Ammerpohl O, von Schonfels W, Ahrens M, Heits N, Bell JT, Tsai PC, Spector TD, Deloukas P, Siebert R, Sipos B, et al. Obesity accelerates epigenetic aging of human liver. Proc Natl Acad Sci U S A. 2014; 111(43):15538-15543.
238. Sillanpaa E, Laakkonen EK, Vaara E, Rantanen T, Kovanen V, Sipila S, Kaprio J and Ollikainen M. Biological clocks and physical functioning in monozygotic female twins. BMC Geriatr. 2018; 18(1):83.
239. Horvath S, Pirazzini C, Bacalini MG, Gentilini D, Di Blasio A, Delledonne M, Mari D, Arosio B, Monti D, Passarino G, De Rango F, D'Aquila P, Giuliani C, et al. Decreased

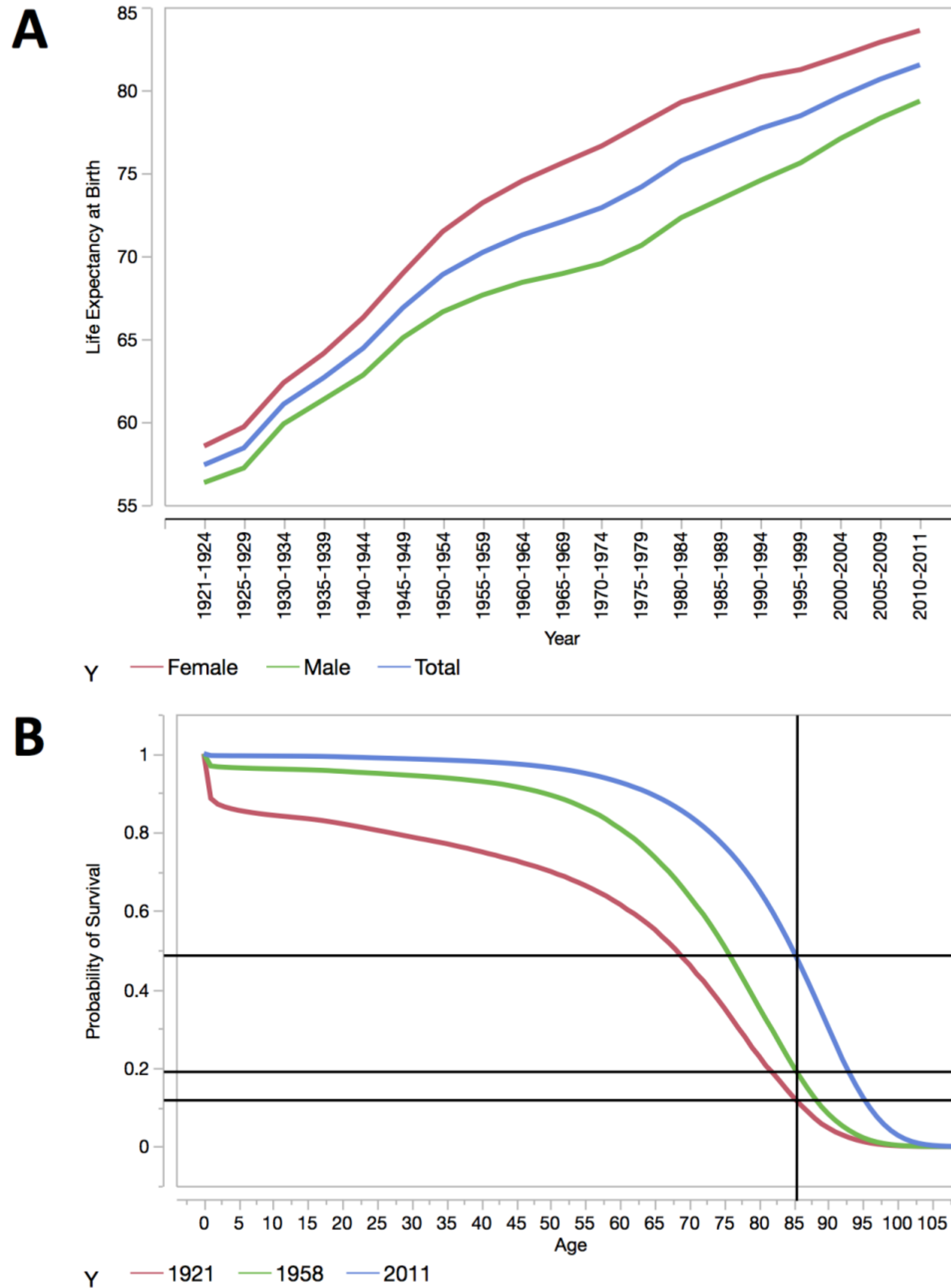
epigenetic age of PBMCs from Italian semi-supercentenarians and their offspring. Aging. 2015; 7(12):1159-1170.
240. Tsubota K. The first human clinical study for NMN has started in Japan. NPJ Aging Mech Dis. 2016; 2:16021.
241. Mills KF, Yoshida S, Stein LR, Grozio A, Kubota S, Sasaki Y, Redpath P, Migaud ME, Apte RS, Uchida K, Yoshino J and Imai SI. Long-Term Administration of Nicotinamide Mononucleotide Mitigates Age-Associated Physiological Decline in Mice. Cell Metab. 2016; 24(6):795-806.
242. Yoshino J, Mills KF, Yoon MJ and Imai S. Nicotinamide mononucleotide, a key NAD(+) intermediate, treats the pathophysiology of diet- and age-induced diabetes in mice. Cell Metab. 2011; 14(4):528-536.
243. Yamamoto T, Byun J, Zhai P, Ikeda Y, Oka S and Sadoshima J. Nicotinamide mononucleotide, an intermediate of NAD+ synthesis, protects the heart from ischemia and reperfusion. PLoS One. 2014; 9(6):e98972.
244. Wang X, Hu X, Yang Y, Takata T and Sakurai T. Nicotinamide mononucleotide protects against beta-amyloid oligomer-induced cognitive impairment and neuronal death. Brain Res. 2016; 1643:1-9.
245. Gong B, Pan Y, Vempati P, Zhao W, Knable L, Ho L, Wang J, Sastre M, Ono K, Sauve AA and Pasinetti GM. Nicotinamide riboside restores cognition through an upregulation of proliferator-activated receptor-gamma coactivator 1alpha regulated beta-secretase 1 degradation and mitochondrial gene expression in Alzheimer's mouse models. Neurobiol Aging. 2013; 34(6):1581-1588.
246. Long AN, Owens K, Schlappal AE, Kristian T, Fishman PS and Schuh RA. Effect of nicotinamide mononucleotide on brain mitochondrial respiratory deficits in an Alzheimer's disease-relevant murine model. BMC Neurol. 2015; 15:19.
247. Caton PW, Kieswich J, Yaqoob MM, Holness MJ and Sugden MC. Nicotinamide mononucleotide protects against pro-inflammatory cytokine-mediated impairment of mouse islet function. Diabetologia. 2011; 54(12):3083-3092.
248. Imai SI and Guarente L. It takes two to tango: NAD(+) and sirtuins in aging/longevity control. NPJ Aging Mech Dis. 2016; 2:16017.
249. Dellinger RW, Santos SR, Morris M, Evans M, Alminana D, Guarente L and Marcotulli E. Repeat dose NRPT (nicotinamide riboside and pterostilbene) increases NAD(+) levels in humans safely and sustainably: a randomized, double-blind, placebo-controlled study. NPJ Aging Mech Dis. 2017; 3:17.

250. Newman JC, Milman S, Hashmi SK, Austad SN, Kirkland JL, Halter JB and Barzilai N. Strategies and Challenges in Clinical Trials Targeting Human Aging. J Gerontol A Biol Sci Med Sci. 2016; 71(11):1424-1434.
251. Valencia WM, Palacio A, Tamariz L and Florez H. Metformin and ageing: improving ageing outcomes beyond glycaemic control. Diabetologia. 2017; 60(9):1630-1638.
252. Barzilai N, Crandall JP, Kritchevsky SB and Espeland MA. Metformin as a Tool to Target Aging. Cell Metab. 2016; 23(6):1060-1065.
253. Hayden EC. Anti-ageing pill pushed as bona fide drug. Nature News. 2015; 522:265-266.
254. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M and DePristo MA. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 2010; 20(9):1297-1303.
255. Dohm JC, Lottaz C, Borodina T and Himmelbauer H. Substantial biases in ultra-short read data sets from high-throughput DNA sequencing. Nucleic Acids Res. 2008; 36(16):e105.
256. Li M and Stoneking M. A new approach for detecting low-level mutations in next-generation sequence data. Genome Biol. 2012; 13(5):R34.
257. Kloss-Brandstatter A, Pacher D, Schonherr S, Weissensteiner H, Binna R, Specht G and Kronenberg F. HaploGrep: a fast and reliable algorithm for automatic classification of mitochondrial DNA haplogroups. Hum Mutat. 2011; 32(1):25-32.
258. van Oven M and Kayser M. Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. Hum Mutat. 2009; 30(2):E386-394.
259. Ruiz-Pesini E, Lott MT, Procaccio V, Poole JC, Brandon MC, Mishmar D, Yi C, Kreuziger J, Baldi P and Wallace DC. An enhanced MITOMAP with a global mtDNA mutational phylogeny. Nucleic Acids Res. 2007; 35(Database issue):D823-828.

Appendices.

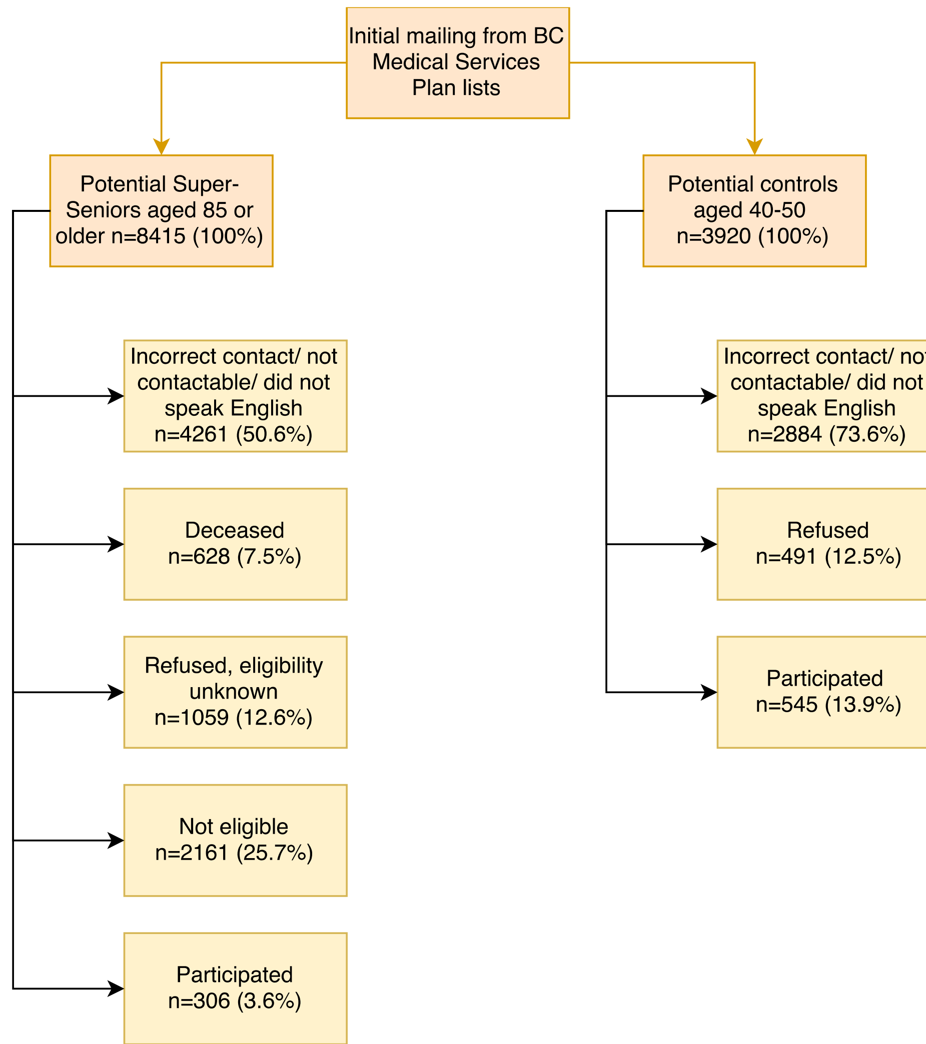
Appendix A.

A.1. Life expectancy in Canada



Data from the Human Mortality Database [1]. (A) Life expectancy at birth in Canada from 1921 to 2011. (B) Survival curve for Canada in 1921, 1958 and 2011. The Super-Seniors were born between 1901 and 1922 with a mean birth year of 1916; the controls were born between 1952 and 1964 with a mean birth year of 1958.

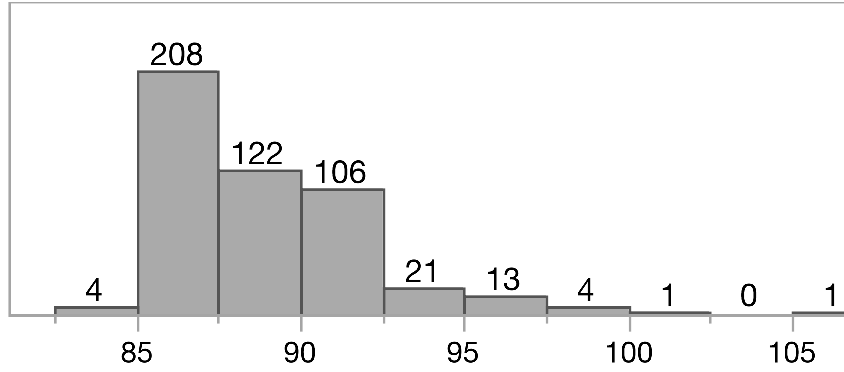
A.2. Collection of Super-Seniors and controls from BC Medical Services Plan lists



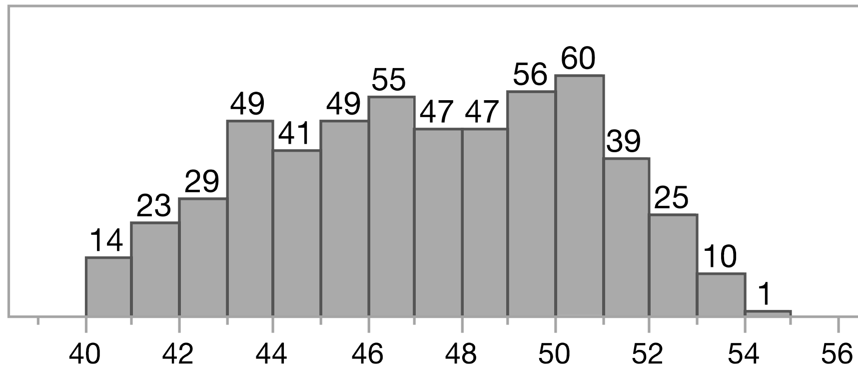
Consent rate for controls = 545 / (545+491) = 52.6%
Eligibility rate of contactable potential Super-Seniors = 306 / (306+2161) = 12.4%
How elite is the Super-Senior phenotype? 12.4% of seniors over age 85 who were contactable and interested were eligible. Taking into account that 28.5% of Canadians age 85 or older have dementia [126] and that such individuals are unlikely to be contactable and able to understand a phone interviewer, the eligibility rate of living individuals would be closer to 12.4% x (1-0.285) = 8.9%. Furthermore, only 9.0% of individuals born in 1916 lived to be 85 [127]; interpolated from data for 1900 and 1950 birth cohorts. Therefore, the proportion of the 1916 birth cohort who went on to become Super-Seniors is approximately 8.9% x 9.0% = 0.80%.

A.3. Age distribution of participants in the Super-Seniors Study

A. Age distribution of Super-Seniors in years.



B. Age distribution of mid-life controls in years.



A.4. Distribution of ethnicity in Super-Seniors and controls

	Super-Seniors			**Controls**
	Exclude	Borderline	Include	Include
First nations			1	
African				1
East Asian	2		18	57
European	56	16	444	416
Latin American			1	4
Middle Eastern	1		1	6
Mixed	1		3	24
South Asian	1		3	11
Unknown	2	1	9	26
Total	63	17	480	545

A.5. Contingency table of smoking status in Super-Seniors and controls

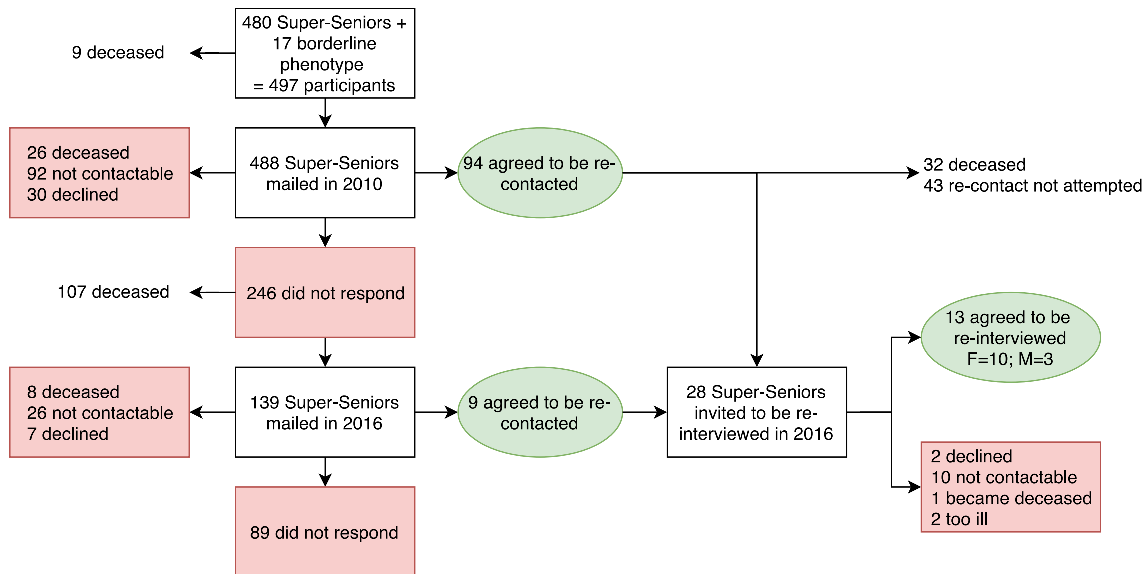
		Current	Never	Quit	
Super-Seniors	Count	7	241	232	480
	Cell X^2	18.0	0.03	2.0	
Control	Count	58	268	219	545
	Cell X^2	15.9	0.03	1.8	
Total		66	509	450	1025

A.6. Contingency table of the number of offspring in Super-Seniors and controls

Offspring		0	1	2	3	4+	
Super-Seniors	Count	49	63	140	107	121	480
	Cell X^2	19.3	0.2	2.5	3.4	31.5	
Control	Count	145	79	202	84	35	545
	Cell X^2	17.0	0.2	2.2	3.0	27.7	
Total		194	142	342	191	156	1025

Appendix B.

B.1. Follow-up of participants in the Super-Seniors Study



Appendix C.

C.1. Supplementary methods: Exome sequencing

Exome libraries were prepared using the Agilent Sure Select All Exon 50Mb kit (Agilent Technologies, CA, USA) from isolated DNA. Each library was run in a single lane on a HiSeq2000 (Illumina Inc. CA, USA) using 100bp PE sequencing, at Canada's Michael Smith Genome Sciences Centre.

Exomes were aligned to hg19 using BWA-MEM (Li, 2013). Variants were called using the Genome Analysis Toolkit [254], and filtered for Mapping Quality ≥ 40, Quality by Depth ≥ 2.0, and Fisher Strand test ≥ 200. Functional annotations for predicted pathogenicity were run using snpEff [210].

C.2. Supplementary methods: Mitochondrial sequencing analysis

In this analysis, the primary allele refers to the most common allele in the read counts, and the secondary allele refers to the second most common allele in the read counts. The reference allele refers to the allele of the rCRS reference sequence at a given position, while the alternate allele refers to the most common allele in the read counts that is not the reference allele. The total allele count refers to the sum of all alleles at a given position, excluding any '*n'* calls.

Mitochondrial reads were extracted from the whole exome sequencing data using MitoSeek [158], which was modified for analysis of both heteroplasmic and homoplasmic variants. An hg19 to rCRS conversion step was implemented to ensure that variant analysis would be performed relative to the rCRS reference sequence (NC_012920). Poor quality reads were filtered out on the criteria of Mapping Quality Score ≥ 20 and Base Quality Score ≥ 20. Positions with read depth ≤ 5 were considered unsuitable for variant analysis and excluded. A minimum secondary allele frequency of 10% and alternate allele count of 5 were set as the frequency cut-off values. When a potential heteroplasmy was called, the surrounding nucleotides were manually assessed with consideration to previously reported sequencing error trends [255, 256] to determine whether variants were located in an 'error hotspot'.

Heteroplasmic variant calls that were outside of an error hotspot and that passed the 10% secondary allele cutoff as well as the 5-read alternate allele cutoff were taken as true heteroplasmies. Positions that did not meet the 10% secondary allele frequency cut-off were then evaluated as possible homoplasmic variants by comparing the primary allele with the reference allele at that position. Heteroplasmic frequencies of such variants were reported as (Alternate Allele Count)/(Total Allele Count). Amino acid differences were also reported for non-synonymous variants, accounting for the differences in genetic code between nuclear and mitochondrial DNA.

HaploGrep [257, 258] was used to predict the subjects' haplogroup from the list of variants and to identify variants not accounted for by the assigned haplogroup. Global minor allele frequencies of reported variants were estimated using MITOMAP [259], which reported the frequency of each variant using 26850 mitochondrial GenBank sequences.

C.3. List of 73 filtered GWAS variants from the National Human Genome Research Institute catalogue

Gene	SNP
BABAM1	rs8170
MLPH	rs2292884
WFS1	rs1801214
IFIH1	rs1990760
BUD13	rs11820589
SLC39A8	rs13107325
HFE	rs1799945
IL6R	rs2229238
SLC17A4	rs11754288
CCDC170	rs3734805
MC1R	rs1805007
FAM208B	rs2797501
TFPI	rs7586970
FADS1	rs174546
PRKCQ	rs11258747
CHRNA3	rs1051730
IREB2	rs13180
HERC5	rs10516809
ATF1	rs17291650
OASL	rs3213545
APOB	rs693
CCHCR1	rs130067
FARP2	rs757978
PPP1R3B	rs3748140
PRRC2C	rs2421847
LOC390956	rs7245858
AGER	rs2070600
P2RX7	rs3751143
AGL	rs17121403
TNS1	rs2571445
LPA	rs3798220
HFE	rs1800562
HNF4A	rs1800961
BCHE	rs1803274
LPL	rs268
BRCA2	rs11571833
PCIF1	rs7679
SCGB1A1	rs3741240
MED24	rs2302777
IL13	rs20541
ANKLE1	rs2363956
SH2B3	rs3184504
KCNJ11	rs5215
IL7R	rs6897932
CD226	rs763361
KIAA1462	rs3739998
EDC4	rs8060686
CAPSL	rs1445898
PPARG	rs1801282
ZC3HC1	rs11556924
RFWD3	rs4888262
CLEC2D	rs3764021
ADAMTS7	rs3825807
MYNN	rs10936599
ADH7	rs971074
CHRNA3	rs8040868
CAMK2B	rs1127065
LDLR	rs2228671
ULK4	rs1052501
ADH1B	rs1229984
ATM	rs1801516
ULK4	rs2272007
SLC30A8	rs13266634
TYK2	rs2304256
THADA	rs7578597
ABCA7	rs3752246
GCKR	rs1260326
LPL	rs328
SLC30A8	rs3802177
MARCH10	rs2251393
KCNJ11	rs5219
APOE	rs429358
TREM2	rs75932628

C.4. Homoplasmic variants found in the mitochondrial genomes of the brothers

Name	Depth (B1)	Depth (B2)	MAF	Gene or Region	Function	AA diff
m.73A>G	11	36	0.73	Control Region	n/a	n/a
m.263A>G	14	44	0.93	Control Region	n/a	n/a
m.709G>A	37	125	0.13	*MT-RNR1*	n/a	n/a
m.750A>G	33	129	0.99	*MT-RNR1*	n/a	n/a
m.930G>A	22	63	0.02	*MT-RNR1*	n/a	n/a
m.1438A>G	21	97	0.94	*MT-RNR1*	n/a	n/a
m.1888G>A	25	72	0.06	*MT-RNR2*	n/a	n/a
m.2706A>G	25	83	0.76	*MT-RNR2*	n/a	n/a
m.4216T>C	20	61	0.10	*MT-ND1*	NS	Tyr->His
m.4769A>G	32	81	0.98	*MT-ND2*	S	n/a
m.4917A>G	26	61	0.05	*MT-ND2*	NS	Asn->Asp
m.5147G>A	7	17	0.00	*MT-ND2*	S	n/a
m.7028C>T	22	96	0.78	*MT-COX1*	S	n/a
m.7521G>A	n/a *	6	0.07	*MT-TRND*	n/a	n/a
m.8697G>A	38	145	0.05	*MT-ATP6*	S	n/a
m.8860A>G	35	121	0.99	*MT-ATP6*	NS	Thr->Ala
m.8934C>T	18	69	0.00	*MT-ATP6*	S	n/a
m.10463T>C	17	50	0.05	*MT-TRNR*	n/a	n/a
m.11251A>G	34	106	0.10	*MT-ND4*	S	n/a
m.11719G>A	31	110	0.74	*MT-ND4*	S	n/a
m.11812A>G	39	111	0.04	*MT-ND4*	S	n/a
m.13368G>A	20	83	0.05	*MT-ND5*	S	n/a
m.14233A>G	22	124	0.04	*MT-ND6*	S	n/a
m.14766C>T	33	93	0.74	*MT-CYTB*	NS	Thr->Ile
m.14905G>A	48	114	0.05	*MT-CYTB*	S	n/a
m.15326A>G	52	129	0.99	*MT-CYTB*	NS	Thr->Ala
m.15452C>A	31	109	0.10	*MT-CYTB*	NS	Leu->Ile

m.15607A>G	25	88	0.05	*MT-CYTB*	S	n/a
m.15928G>A	6	34	0.05	*MT-TRNT*	n/a	n/a
m.16126T>C	22	85	0.12	Control Region	n/a	n/a
m.16129G>A	20	87	0.12	Control Region	n/a	n/a
m.16294C>T	11	56	0.09	Control Region	n/a	n/a
m.16296C>T	7	25	0.02	Control Region	n/a	n/a
m.16304T>C	55	177	0.07	Control Region	n/a	n/a
m.16519T>C	18	60	0.62	Control Region	n/a	n/a

MAF = global minor allele frequency (as reported by MitoMap), AA = amino acid, n/a = not applicable, S = synonymous, NS = nonsynonymous.

* m.7521G>A was not reported in B1 because the data at that position did not meet the minimum depth requirement of 5 reads.

Appendix D.

D.1. *p*-values for main effects examined

Gene	ID	Dominant model *p* value (df)	Additive model *p* value (df)
ADIPOQ	rs56354395	0.50 (1)	0.24 (1)
APOA1	rs670	0.85 (1)	0.88 (1)
APOC3	rs595049	0.61 (1)	0.88
APOE	*APOE* haplotype	0.0010 (1)*	0.00017 (5)*
CETP	rs5882	0.54 (1)	0.36 (1)
CRYL1	rs7989332	0.88 (1)	0.96 (1)
FOXO1	rs2701858	0.089 (1)	0.11 (1)
FOXO3	rs9486902	0.57 (1)	0.60 (1)
FOXO3	rs2802292	0.16 (1)	0.24 (1)
HFE	rs1800562	0.32 (1)	0.34 (1)
HP	rs72294371	0.010 (1)*	0.056 (1)
KHDRBS2	rs6455128	0.29 (1)	0.48 (1)
KL	rs9536314	0.12 (1)	0.070 (1)
LPA	rs1853021	0.58 (1)	0.99 (1)
LPA	rs10455872	0.31 (1)	0.25 (1)
MTTP	rs2866164	0.34 (1)	0.53 (1)
PON1	rs662	0.72 (1)	0.93 (1)

D.2. *p*-values for interaction tests identified a priori

Variant 1	Variant 2	*p* value (df)
CRYL1 rs798933	*KHDRBS2* rs6455128	0.077 (1)
FOXO3 rs9486902	*FOXO1* rs2701858	0.38 (1)
LPA rs1853021	*CETP* rs5882	0.63 (1)
MTTP rs2866164	*CETP* rs5882	0.57 (1)
MTTP rs2866164	*APOC3* rs595049	0.42 (1)
MTTP rs2866164	*ADIPOQ* rs56354395	0.10 (1)
APOE haplotype	*HP* rs72294371	0.37 (1)

D.3. Interaction odds ratios and 95% confidence intervals between variants

		CRYL1 rs7989332	
KHDRBS2 rs6455128		GG	GT or TT
	CC	1.0	0.80 (0.58-1.11)
	AC or AA	0.64 (0.43-0.94)*	0.96 (0.64-1.46)

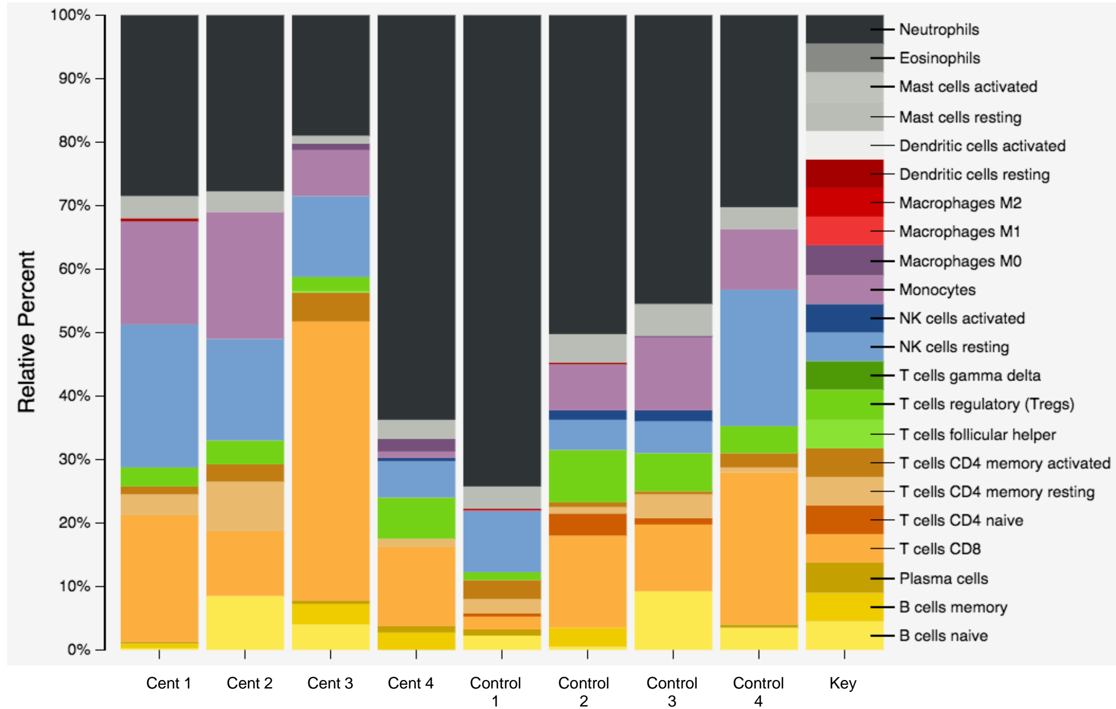
		APOE4 carrier	
FOXO3 rs948602		Non-carrier	Carrier
	CC	1.0	0.51 (0.35-0.73)*
	CT	0.83 (0.58-1.20)	0.70 (0.36-1.35)
	TT	0.61 (0.25-1.5)	2.71 (0.3-24.5)

		CRYL1 rs7989332	
LPA rs10455872		GG	GT or TT
	TT	1.0	1.07 (0.80-1.42)
	CT or CC	1.09 (0.66-1.81)	0.54 (0.28-1.03)

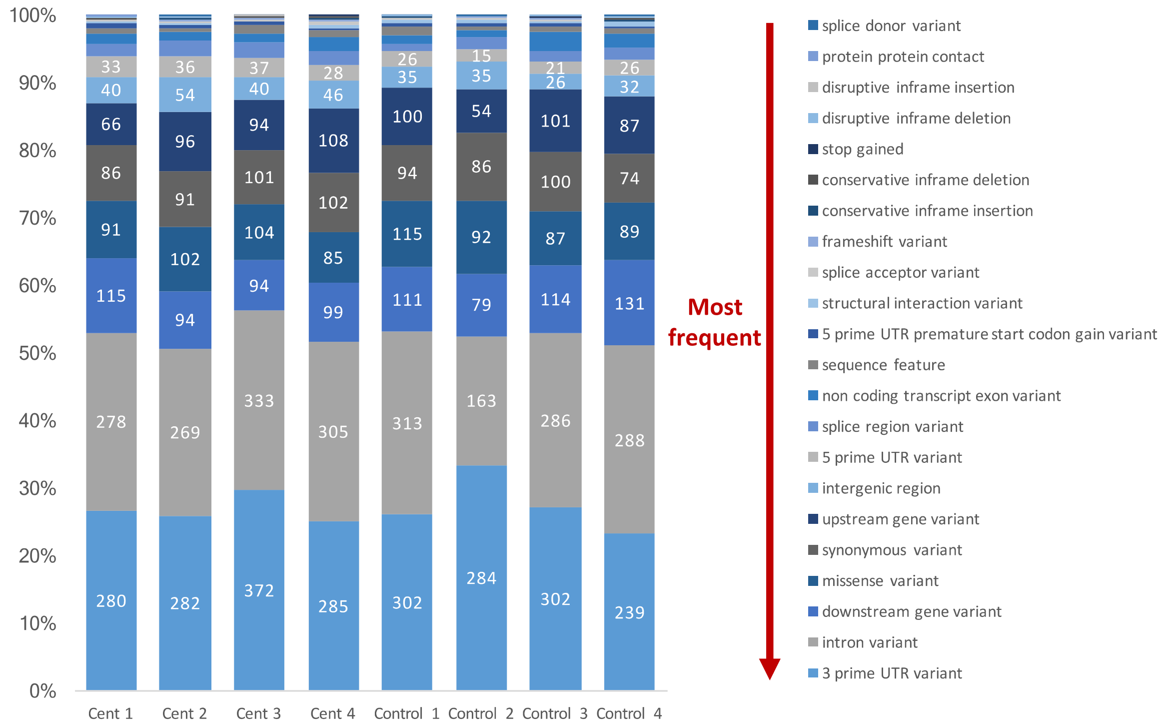
		APOE4 carrier	
CRYL1 rs7989332		Non-carrier	Carrier
	GG	1.0	0.50 (0.32-0.76)*
	GT	0.89 (0.64-1.22)	0.53 (0.32-0.90)*
	TT	0.76 (0.39-1.48)	1.53 (0.51-4.59)

Appendix E.

E.1. Proportion of immune cell types as estimated by CIBERSORT by sample



E.2. Proportion of variants showing allele-specific expression by predicted effect



E.3. Estimated average cell fractions of immune cell types from CIBERSORT. The signature gene file of 22 distinct immune cell types provided by CIBERSORT was used as a reference for gene expression signatures

	AVG RELATIVE %	AVG RELATIVE CELL FRACTION	SD	MIN	MAX
Neutrophils	42.37%	0.42	0.19	0.19	0.74
T cells CD8	17.21%	0.17	0.13	0.022	0.44
NK cells resting	12.30%	0.12	0.072	0.050	0.23
Monocytes	9.08%	0.091	0.068	0.0026	0.20
T cells regulatory (Tregs)	4.36%	0.044	0.023	0.013	0.081
B cells naive	3.54%	0.035	0.036	0	0.092
Mast cells resting	3.44%	0.034	0.011	0.013	0.048
T cells CD4 memory resting	2.51%	0.025	0.025	0	0.077
T cells CD4 memory activated	1.88%	0.019	0.015	0.0012	0.046
B cells memory	1.25%	0.012	0.015	0	0.034
T cells CD4 naive	0.63%	0.0063	0.012	0	0.036
Macrophages M0	0.46%	0.0046	0.0079	0	0.022
NK cells activated	0.43%	0.0043	0.0069	0	0.016
Plasma cells	0.38%	0.0038	0.0036	0	0.0097
Dendritic cells resting	0.11%	0.0011	0.0016	0	0.0037
T cells follicular helper	0.05%	0.00048	0.0014	0	0.0039
Dendritic cells activated	0.01%	0.000050	0.00014	0	0.00040
T cells gamma delta	0.00%	0	0	0	0
Macrophages M1	0.00%	0	0	0	0
Macrophages M2	0.00%	0	0	0	0
Mast cells activated	0.00%	0	0	0	0
Eosinophils	0.00%	0	0	0	0

E.4. List of 35 top biased centenarian and 23 top biased control genes. Top biased centenarian genes were defined as genes were there was evidence of ASE for a particular gene in 4 centenarians and only 1 or no controls, or in 3 centenarians and no controls; and vice versa for top biased control genes

Subjects with ASE	Gene
4 Cents - 0 Controls	LINC01060
4 Cents - 0 Controls	LINC01262
4 Cents - 0 Controls	NOS2
4 Cents - 0 Controls	TCF25
3 Cents - 0 Controls	ATF7IP
3 Cents - 0 Controls	C3orf58
3 Cents - 0 Controls	CAMK2N1
3 Cents - 0 Controls	CD151
3 Cents - 0 Controls	CEP295
3 Cents - 0 Controls	EIF4G2
3 Cents - 0 Controls	ELMSAN1
3 Cents - 0 Controls	FES
3 Cents - 0 Controls	JAML
3 Cents - 0 Controls	LILRA1
3 Cents - 0 Controls	LINC00226
3 Cents - 0 Controls	LINC00221
3 Cents - 0 Controls	LLGL2
3 Cents - 0 Controls	LPIN1
3 Cents - 0 Controls	MED16
3 Cents - 0 Controls	MEGF6
3 Cents - 0 Controls	NELFCD
3 Cents - 0 Controls	ORM1
3 Cents - 0 Controls	PDCD6IP
3 Cents - 0 Controls	RNF44
3 Cents - 0 Controls	RRN3P2
3 Cents - 0 Controls	SCRN1
3 Cents - 0 Controls	TMEM43
3 Cents - 0 Controls	TRIM39
3 Cents - 0 Controls	UNC13D
3 Cents - 0 Controls	WBP2

Subjects with ASE	Gene
3 Cents - 0 Controls	ZFP57
3 Cents - 0 Controls	ZNF718
4 Cents - 1 Controls	KRT72
4 Cents - 1 Controls	SNORA10
4 Cents - 1 Controls	TPTE2P5
4 Cents - 1 Controls	TRG-AS1
0 Cents - 4 Controls	WDR90
1 Cents - 4 Controls	DLGAP4
1 Cents - 4 Controls	IL18RAP
1 Cents - 4 Controls	LDLR
1 Cents - 4 Controls	LDOC1L
0 Cents - 3 Controls	ANXA5
0 Cents - 3 Controls	COX5BP7
0 Cents - 3 Controls	CARD8
0 Cents - 3 Controls	CD247
0 Cents - 3 Controls	CDC42EP1
0 Cents - 3 Controls	CPA5
0 Cents - 3 Controls	FAM118A
0 Cents - 3 Controls	PSPHP1
0 Cents - 3 Controls	FYB
0 Cents - 3 Controls	HSD17B1
0 Cents - 3 Controls	HVCN1
0 Cents - 3 Controls	IL1RN
0 Cents - 3 Controls	MFSD9
0 Cents - 3 Controls	NOC4L
0 Cents - 3 Controls	PTK2B
0 Cents - 3 Controls	SGSH
0 Cents - 3 Controls	SNX22
0 Cents - 3 Controls	UPK3A
0 Cents - 3 Controls	VNN1