MITOCHONDRIAL GENETIC EPIDEMIOLOGY OF LIFESPAN IN A POPULATION ISOLATE

By

Joseph Donald Bonner

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Comparative Medicine and Integrated Biology – Doctor of Philosophy

2013

ABSTRACT

MITOCHONDRIAL GENETIC EPIDEMIOLOGY OF LIFESPAN IN A POPULATION ISOLATE

By

Joseph Donald Bonner

Mitochondria play a crucial role in cellular processes. They are where oxidative phosphorylation (OXPHOS) occurs. OXPHOS produces ATP which powers most cellular functions. Mitochondria are also regulators of other cellular processes such as apoptosis, cell cycle regulation, calcium signaling and neuronal synaptic transmission. They contain the only extra-nuclear DNA in eukaryotes. Mitochondrial DNA (mtDNA) encodes crucial components required for OXPHOS. Reactive oxygen species (ROS) are a natural byproduct of OXPHOS. In healthy states, ROS production facilitates mitochondrial cell signaling. ROS levels are kept in homeostasis by naturally occurring anti-oxidants. When dysregulated, ROS can cause mutations in mtDNA ultimately thereby decreasing ATP production capacity.

Aging in humans manifests as a gradual decline in cellular functioning, leading to loss of tissue function and eventual death. The rate at which humans age and eventually die, is a complex interplay between genes and environment. Mitochondria are closely related to cellular function and their decline mirrors the aging process. These observations suggest that mitochondrial dysfunction might play a role in the aging process.

Human mtDNA is exclusively transmitted from mother to child with no possibility for paternal recombination. Certain patterns of mtDNA haplotypes cluster in geographic

regions. Observing these geographic differences has allowed researchers to put groups of mtDNA haplotypes into phylogenic trees that mirror human migration patterns. Epidemiologic studies have sought to associate disease patterns with mtDNA haplotype groups. Some studies of longevity have found that carriers of certain mtDNA haplotype groups are more common among aged cases than younger controls. However, these studies leave many questions unsettled. Prior studies do not account for nuclear genetic factors, have not achieved sufficient statistical power, have been unable to account for changes in age-at-death that have occurred over time and have not use time-to-event analysis.

This dissertation examines that hypothesis that mtDNA Haplogroup J carriers live longer than carriers of other haplogroups. This work focuses on a newly-identified population isolate in Mid-Michigan with a well-documented pedigree. It is within this pedigree that we propose to test the Haplogroup J and longevity or lifespan association.

.

Within hours of meeting my bride-to-be in 1989, I apologized for just having a bachelor's degree. With this dissertation, I fix that short-coming. Soon after we met, we started genealogy as a hobby. She encouraged me to pursue this degree and the master's degree before it. She stood by me during emotional angst with deadlines and contradictory feedback from reviewers, tricky financial times and several major job changes.

My father worked many years at a factory. He had several friends and a brother-in-law relocate with their families for ministerial studies. When I was young, he studied Theology via correspondence. I watched him prepare and review flashcards to study for exams and stay up late to do homework.

During these past five years, my precious children watched me prepare and review flashcards to study for exams and rising early to do homework. They observed me peeking around mountains of photocopied journal articles to catch glimpses of basketball practices. They watched me typing on a laptop analyzing data and writing these chapters; sometimes while in the car at soccer or football practice; sometime in school cafeterias or hallways or libraries; sometimes trying to ignore the television just to be in the same room with them. They watched me pipette, weigh, measure, prepare gels and clean up in the laboratory; sometimes helping. They reviewed colorful chromatographs of their own mitochondrial DNA. I hope they learned study habits and life-planning lessons from me who learned them from my father. I hope they share these skills with their children. I hope they plan their academics differently than my father and I.

To my bride and fellow genealogist Tracie, my father Thomas Henry II, my observant children Thomas Robert (Toby) and Roza, I dedicate this work.

ACKNOWLEDGEMENTS

I began this journey in September 2009, in October of that year my mother Joan Bonner passed away. She was my first scientific coach and cheerleader. Thanks Mom.

For my gracious and patient advisory committee Drs Brian Schutte, Susan Ewart, Shelagh Ferguson-Miller, John Fyfe and Rachel Fisher, I am thankful. I am doubly thankful to Dr. Fisher. This work is her brain-child and the genealogic database upon which I build, is her labor of love.

I recognize and acknowledge my cousins, Jenny and Michelle and Mrs. Kanoeznot. Our childhood play as scientists made lasting cerebral connections. It is intriguing how our shared early interests in scientific discovery guided me to this point. While this work is not a cure for cancer as we played, I am hopeful that others can stand on this work and my shoulders to achieve our childish lofty goals.

My Ph.D. journey has been balanced with a full-time job. To my employers Kyle Johnson of CHE/Trinity Health Information Technology Services and Phil Reed of Michigan State University Biomedical Research Informatics Core, I formally express my gratitude for your patience and flexibility.

To the communities of the CoSAGE cooperative, I am obliged for their choices and graciousness.

Could I have done this work without Dr. Yusbazian-Gurkan, the head of the Comparative Medicine and Integrated Biology Program? The answer is short – NO! Dr. Yusbazian-Gurkan is a solid as a rock and as warm as a saint and as approachable as a parent. I shall never forget the lunch meeting where she presented to me that the CMIB program studies “molecules in populations”. She followed that by saying that the program specializes in making an academic home for non-traditional students to pursue an advance scientific degree. I was speechless and breathless for the few halcyon seconds pondering the possibilities of studying molecular epidemiology while working full time with a family. Throughout the my pursuits, she served as a sounding board, a reality checkpoint, a source of inspiration, a creator of opportunities, a magnificent organizer, a tireless cheerleader, trusted mentor, equitable arbitrator, seasoned stateswoman, accomplished scholar, ready listener with a ready box of Kleenex and cup tea or very strong coffee. Thanks to you … more than my words can convey.

TABLE OF CONTENTS

.

LIST OF TABLES

LIST OF FIGURES

KEY TO ABBREVIATIONS

ADP – adenosine di-phosphate

ATP – Adenosine Tri-Phosphate

CoQ - Co-enzyme Q

COSAGE – Cooperative to study – across generations

DNA – Deoxyrobonucleic Acid

ETC – Electron Transport Chain

FAD/FADH – flavin adenine dinucleotide (oxidized)/ flavin adenine dinucleotide reduced

FMN – flavin mono-nucleotide/oxidized flavin

FMNH2 – reduced flavin

HVS1, HVS2 – Hyper-Variable Segments one and two

MIM – mitochondrial inner-membrane

MIS - mitochondrial inter-membrane space

mtDNA – Mitochondrial DNA

mtRNA – mitochondrial ribonucleic acid

mtTFH – mitochondrial DNA heavy strand promoter

mtTHL – mitochondrial DNA light strand promoter

mtTHx - mitochondrial DNA-TFMA binding site X

mtTHy – mitochondrial DNA-TFMA binding site Y

NAD/NADH - nicotinamide adenine dinucleotide (oxidized)/nicotinamide adenine dinucleotide reduced

OXPHOS – Oxidative Phosphorylation

P - Phosphate

rCRS – revised Cambridge Reference Sequence

RNA – ribonucleic acid

ROS – Reactive Oxygen Species

SNP – Single nucleotide polymorphism

TFAM – mitochondrial transcription factor A

CHAPTER 1: INTRODUCTION

1.1 Heritability of Longevity

In 1891, Oliver Wendell Holmes suggested that to live a long life one should choose one's parents well [1]. He followed that "Especially let (your) mother come from a race in which octogenarians and nonagenarians are very common." In 1899 and 1901, Beeton et al presented genealogical evidence to bolster Holmes' observation [2, 3]. Using genealogical databases, they showed that parent's ages at death correlated positively with children's age at death. Jalavisto (also using genealogic data) suggested that the mother-offspring correlates were significantly higher than the father-offspring. [4]. Table 1-1 presents the genealogic evidence to quantify Holmes' observation. Heritability estimates for age at death differ considerably, however mother-offspring estimates trend higher than father-offspring estimates. [2-4]

Table 1-1. Heritability of longevity

Citation	Father-Son	Father-Daughter	Mother-Son	Mother-Daughter	Population
Jalavisto [4]	0.12	0.04	0.16	0.20	Finnish and Swedish Nobility
Korpelianen [5]	0.48	0.44	0.84	0.66	Japanese Imperial Genealogy
Mayer [6]	0.1-0.3	0.12-0.21	0.1-0.32	0.17-0.33	North American/New England

These studies support a stronger maternal heritability of longevity than paternal. Among the many biologic contributions from parents to children, mothers give more. Fathers contribute only half a genome; but mothers give half a genome and the

cytoplasm within the ovum. The ovum cytoplasm contains the initial pool of mitochondria that will eventually populate every cell and power cellular processes throughout the life of the child. The background of maternal heritability of longevity in the setting of mitochondrial inheritance offers the context of the work herein presented.

1.2 Mitochondria

Mitochondria are cellular organelles that contain the energy producing machinery needed for tissue function. Thus they play a key role in energy production by synthesizing ATP along the electron transport chain (ETC). They are also the foci of apoptosis initiation. They buffer calcium and they regulate the cell cycle. [7-10]

1.3 Electron Transport Chain

The electron transport chain (ETC) is a series of enzymatic reactions that transform the potential energy of consumed calories into chemical energy for cellular work. ETC occurs within the mitochondrial inner membrane (MIM). Electrons move from a high to low reduction potential state through electron carriers in the chain. Their movement enables ETC complexes to translocate protons from the matrix into the inter-mitochondrial space thereby driving production of a proton gradient. The chemiosmotic pressure of the gradient drives the production of adenosine triphosphate (ATP), the chemical energy form used for cellular biologic processes. [11]

In humans most consumed calories are digested into glucose, amino acids and fatty acids. These molecules circulate through the blood stream and are imported into most cell types. Inside the cytosol, glycolysis culminates with two pyruvate molecules formed from each glucose molecule. Through dedicated shuttle systems, pyruvate,

amino acids and fatty acids are imported into mitochondrial via the inter-membrane space and into the matrix. Pyruvate dehydrogenase and the Krebs cycle, occurring inside the matrix, oxidize pyruvate to produce NADH. One step of the Krebs cycle oxidizes succinate into fumarate producing FADH2. Also inside the matrix, fatty acids are oxidized into Acetyl-Coenzyme A (Acetyl-CoA) which joins the Krebs cycle. The products of the Krebs cycle (NADH and FADH2) are the high-energy final products of consumed calories, carriers of the reducing equivalents (electrons and hydrogen atoms). The ETC transforms the energy from NADH, FADH2 into ATP for cellular utilization. [11]

The ETC consists of electron donors (NADH, FADH2), electron acceptors (Complex I, Complex II, Complex III, and Complex IV), electron carriers (Ubiquinone and Cytochrome C) and an ATP production complex (Complex V). The ETC enzymatically transforms the potential energy of consumed calories into chemical energy, heat, water and sometimes reactive oxygen species. [11]

1.3.1 Reactive Oxygen Species

When oxygen is partially reduced, a superoxide anion is formed. Anti-oxidant molecules and enzymes like manganese superoxide dismutase catalyze conversion of superoxide anions into hydrogen peroxide. Hydrogen peroxide can be further reduced to water or into the hydroxyl radical. The hydroxyl radical reacts with nitric oxide to form peroxynitrite. Both the hydroxyl radical and peroxynitrite readily and indiscriminately react with nucleic acids, lipids, or proteins. [11] [12]

Along the electron transport chain, electrons that build up on certain redox centers—particularly ubiquinol—can react with molecular oxygen, forming superoxide anions. This reaction is most likely to occur in periods of high electron-donor molecule availability, low activity of Complex IV, or high oxygen pressure. The mitochondrial proteins, the membranes, and the nucleic acid of the mitochondrial nucleoid can be targets of the reactive oxygen species (ROS) damage. [11] [12]

The full molecular integrity of the ETC complexes requires transcription and translation of mitochondrial DNA (mtDNA) located inside the matrix. Since reactive oxygen species have an affinity for nucleic acids, mtDNA is vulnerable to oxidative damage. When mtDNA is damaged, ETC complex subunits cannot be produced with fidelity or at all. When the ETC complexes are compromised, reactive oxygen species production increases, damage accelerates, mitochondrial bioenergy production capacity is decreased. [11]

1.4 Mitochondrial DNA

In 1962, Nass and Nass discovered that mitochondria present the only extra-nuclear source of DNA inside eukaryote cells.[13] mtDNA is a circular, double-stranded molecule of 16,569 nucleotide positions. [14] It can be visualized as a clock with nucleotide zero at noon and the bases numbered clockwise. mtDNA encodes 13 of the 86 polypeptides of the ETC complexes and some RNA molecules .[8] Positions 16024-16569 and 1-576, collectively known as the D-loop, are the only non-coding regions of the mitochondrial genome [15]. Positions 16024-16524 are known as hyper-variable segment one (HVS1). Positions 1-576 are known as hyper-variable segment 2 (HVS2). These regions contain expression and replication regulatory elements. Without active

coding, the D-loop region is thought to be neutral to selection pressures.[16-19] In the early 1980s, Anderson et al. published the reference sequence of human mtDNA [14]. The same group published a revision in 1999 known as the revised Cambridge Reference Sequence (rCRS). [20]

Collectively the five ETC complexes are made up of 86 polypeptides. Thirteen of these are encoded in mtDNA. Complex I has the most significant mtDNA presence with 7 of the 45 polypeptides encoded in the mitochondrial genome. Complex II has no sub-units encoded in mtDNA. (Table 1-2) [11]

Table 1-2. ETC Complex subunits and genome encoding

ETC Complex	Total Sub-units	Mitochondrial Sub-units	Nuclear sub-units
Complex I	45	7	28
Complex II	4	0	4
Complex III	10	1	9
Complex IV	13	3	10
Complex V	13	2	11

1.4.1 Mitochondrial DNA Expression and Replication

Within the D-loop are key sites that initiate, promote, and enhance mtDNA expression. Expression begins at transcription initiation sites. Promoters lie upstream from each initiation site. Farther upstream from each promoter lie strand-specific enhancer elements. The enhancer elements are the binding sites for mitochondrial transcription factor A (TFAM). [21] TFAM binds to the respective promoters and recruits mtRNA polymerase to initiate the mtDNA expression apparatus. [22-25]

The rate-limiting steps of mtDNA replication are currently unknown. However, it is known that mtDNA replication is independent of the cell cycle [26], and is correlated

with levels of the nuclear-encoded TFAM [27]. Mouse TFAM heterozygous knock-out models show mtDNA depletion phenotype and cardiac-respiratory deficiency.[28] Homozygous TFAM knock-out mice are embryonically lethal, demonstrating the critical function for TFAM in mtDNA replication [28-30]

The TFAM protein plays a pivotal role in the organization, replication, and expression of mtDNA. [22] It binds to the mtDNA molecule at four positions within the D-loop (heavy-strand promoter [mt-TFH], light-strand promoter [mt-TFL], TFAM binding site X [mt-TFX, positions 233-260], TFAM binding site Y [mt-TFY positions 276-303]).[23],

1.4.2 Nucleoid Organization

mtDNA is organized into nucleoids containing 5-7 mtDNA copies of the molecule, the minimal replisome components, TFAM, and other proteins [31]. Nucleoids have a layered structure with inner and outer zones. The inner zone is the locus of mtDNA replication [26]. The outer zone is the locus of mtDNA expression [26]. The mitochondrial nucleoid is considered to be the unit of inheritance, rather than a single mtDNA molecule [32]. Within the nucleoid, the ratio of TFAM to mtDNA is 1000:1 with TFAM blanketing mtDNA for some measure of histone-like protection. [22, 23, 33]

1.4.3 Mitochondrial Haplogroups

mtDNA is maternally inherited and thus does not recombine [34]. mtDNA mutates ten times faster than nuclear DNA.[35] These two properties make the molecule well suited to phylogenic analysis of mtDNA sequences between and within populations. In 1992, Torroni et al. found four distinct groups of mtDNA haplotypes from

a population of Native Americans, which they christened Haplogroups A, B, C and D [36]. Since then, other haplogroups have been discovered. It is now commonly accepted that mtDNA from human population groups can be classified into haplogroups using an alphabetical naming convention from A –Z [36, 37] in order of publication [38].

mtDNA haplogroups are determined by nucleotide sequence differences in the coding region [20]. Haplogroups of European populations can be deduced by sequencing seven SNPs of the coding region (mt11719A/G, mt7028T/C, mt12308A/G, mt10398A/G, mt15607A/G, mg12612A/G, mg1715C/T) [39]. Most haplotype grouping SNPs in the coding region are in disequlibrium with SNPs in HVS-I or HVS-2. Therefore, to infer the haplogroups when ancestry is unknown, one can sequence the full region of HSV-I and HSV-II, compare to a reference sequence and score based on common algorithms [18]. There are several haplogroup-scoring algorithms available. One relies on genotypes of coding region SNPs (MTDNAManager) [15], and another uses SNPs from coding and non-coding regions to produce a probability (HaploGrep) [40]. Haplogrouping efforts choose a deductive method when a subject's or population's ancestry is known, or an inductive/discovery method when ancestry is unknown.

1.4.5 Haplogroups in Genealogy

Archaeological evidence suggests that humans migrated out of Africa within the past 200k years [10]. Phylogenic trees of mtDNA haplogroups retrace human migration across time and geographic space. The African Haplogroup L is the trunk from which all other lines radiate. Haplogroup L gave rise to the branching macro Haplogroups M

and N.[38] European populations stem from Haplogroup N and Asian populations separate from Haplogroup M.[10, 38]

1.4.6 Mitochondrial Haplogroups and Biologic Variation

A growing body of evidence suggests that haplogroups associate with biologic variation and health states. Total energy expenditure is the daily requirement for dietary calories to produce physiologic energy. It measures the efficiency with which mitochondria convert calories to energy needs for respiration, brain/neural activity, and body heat. Tranah et al. reported in 2011 that total energy expenditures differ by macro haplogroups. They reported that elderly individuals with African Haplogroup L require 158 (7%) fewer daily calories than elderly individuals with European Haplogroup N [41].

In 2009, Beckstead et al. reported that the single nucleotide polymorphism 14766T>C is the principal difference between Haplogroup H and Haplogroup U. This single change substitutes a threonine for an isoleucine at Position 3 in of Complex III subunit 8. This changes make complex III within Haplogroup U more hydrophobic than the same complex within Haplogroup H. Beckstead reported that in times of caloric deprivation, the hydrophilic Haplogroup H confers a survival benefit over Haplogroup U. However, outside of times of caloric deprivation, the difference is unappreciated [42]. This group used a genealogical database like by Korpelianen [5], Jalavisto, [4]and Mayer [6]. Using genealogical records allows a test of survival benefit over time with the potential of large sample sizes.

Five of the six European haplogroup defining SNPs present synonymous mutations. The sixth is a non-synonymous mutation that lacks evolutionary conservation.[39] Since the variation is synonymous and lacks evolutionary

conservation, none of the differences have biologic consequence as measured in the coding region. Any biologic consequence to European haplogroup variation would need to come from non-coding region variation.

In 2009, Suissa et al. reported that the functional genetic difference between European Haplogroups H and J is the 295C>T SNP. This D-loop SNP is within the mt-TFY binding site. The change from a cytosine in Haplogroup H to the thymine in Haplogroup J strengthens TFAM binding in the nucleoid. To test the biological significance of this SNP, they injected human mitochondria from Haplogroups J and H into rho-zero (mitochondria-free) HeLa. The cultured cells showed significant differences in both TFAM binding affinity and mtDNA copy number. Suissa et al. reported that Haplogroup J cells have a greater mtDNA copy number than those of Haplogroup H. [43]. This greater copy number might delay the onset of mitochondrial dysregulation and sustain cellular functioning attributed to ROS damage and thereby postpone the aging process [43].

1.4.7 Mitochondrial Haplogroups and Longevity

In 1998, Tanaka et al reported mtDNA SNP 150C>T to be more frequent in a group of Japanese centenarian cases than in a group of younger controls [44]. mtDNA 150T is a signature SNP of Asian Haplogroup D. Extending the SNP evidence into haplogroups, De Benedictis et al. studied haplogroup frequency differences in Italian centenarians and younger controls. They showed that among males from Northern Italy, Haplogroup J is more frequent among centenarian cases than in younger male controls evidenced by a stratum specific odds ratio of 15 ($CI_{95\%}$1.7-132). The association was absent in males from southern Italy, or in females from either region. [45]. This high

odds ratio fueled other investigations of haplogroup frequency differences in older individuals vs. younger controls most of which did not replicate the initial findings. [39, 46-52].

1.5 Summary

The maternal heritability of longevity is well documented in genealogies in the nineteenth and twentieth century using pedigree data. It is commonly accepted that if a mother lives long, her children will live long. Since mothers alone pass down mitochondrial DNA, their genetic material is identical in all maternal relatives. Mitochondrial DNA clusters into haplogroups, giving a trace of ancient genealogy to contemporary individuals. *In-vivo* evidence suggests that the mitochondrial DNA of Haplogroup J might have greater protective properties in its increased TFAM binding within the nucleoid. TFAM is involved in both mtDNA expression and replication. Since the mtDNA D-Loop contains TFAM binding sites and haplogroup-defining SNPs, I hypothesize that SNP variation in this region can affect longevity through a pathway involving a protective effect of TFAM. The strengthened TFAM binding motif of Haplogroup J lends biologic plausibility to the hypothesis that haplogroup J confers a benefit of longevity. Several studies have investigated the Haplogroup J/longevity association in various populations. However no study has placed the longevity question in the context of pedigree data or “pedigree space” like the original studies of maternal heritability of longevity. Placing the association in “pedigree-space” 1) creates minimal control of nuclear genetic factors shared with most relatives 2) has the potential to generate a larger sample size 3) study changes over a time 4) use age at death as a metric for time-to-event analysis. I hypothesize that deceased relatives of current-living

individuals who carried Haplogroup J, lived longer than carriers of other haplogroups. Placing the question in pedigree space is the singular emphasis of this dissertation.

CHAPTER 2: PEDIGREE STRUCTURE AND KINSHIP MEASUREMENTS IN A MID-MICHIGAN COMMUNITY: A NEW NORTH AMERICAN POPULATION ISOLATE IDENTIFIED

IN PRESS with Human Biology:

JD Bonner, RA Fisher, J Klein^, Q Lu, E Wilch, KH Friderici, JL Elfenbein^, DL Schutte, BC Schutte

Contributions:

JDB: lead author, conceptualized community/pedigree properties within qualities of population isolate, programmed PedHunter, Oracle and SAS to produce statistics for description of isolation.

RAF: Caretaker of pedigree database.

JK, JLE deceased

QL: initial programming to describe family allocation of main pedigree database

EW, KHF: consulted with concept of the genetic value of population isolates, contributing scientist of initial description of genetics of hearing loss within the community

DLS: co-principal investigator for COSAGE

BCS: senior author/scientist and co-principal investigator for COSAGE

2.1 Abstract

Previous studies have identified a cluster of individuals with an autosomal recessive form of deafness residing in a small region of mid-Michigan. We hypothesized that affected members from this community descend from a defined founder population. Using public records and personal interviews, we constructed a genealogical database that includes the affected individuals and their extended families, and identified them as descendants of 461 settlers most of whom emigrated from the Eifel region of Germany between 1836 and 1875. The genealogical database represents a 13-generation pedigree that includes 27,747 descendants of these settlers. Among these descendants, 13,784 are presumed living. Many of the extant descendants reside in a 90-square-mile area and 52% were born to parents who share at least one common ancestor. Among those born to related parents, the median kinship coefficient is $3.7*10^{-3}$. While the pedigree contains 2,510 founders, 344 of the 461 original settlers accounted for 67% of the genome in the extant population. These data suggest that we identified a new population isolate in North America and, as demonstrated for congenital hearing loss, this rural mid-Michigan community is a new resource to discover heritable factors that contribute to common health-related conditions.

2.2 Introduction

Population isolates are not strictly defined. Rather, they share a number of general characteristics such as a small founder population, rapid outgrowth, a history of isolation, an endogamous genealogy, environmental and phenotypic homogeneity, and the likelihood of genetically recessive health conditions [53]. Endogamy is the practice

of marriage between individuals of a common ethnic, social, or religious group. Within a few generations, most individuals of the endogamous group can be related to each other, often through multiple genealogical lines [53]. Repeated generations of endogamy increase the likelihood of homozygosity of recessive alleles [54]. This homozygosity might increase the risk of health-related conditions [55] and common diseases such as cardiac conditions [56-58] and some cancers [59, 60]. While relatedness of parents is noted as a risk factor for some common complex diseases, individuals within population isolates are at no greater risk for childhood mortality [61] and may also share positive health attributes such as longevity [62] and increased fertility [63].

The basic characteristics of population isolates make them important resources for biomedical research [53]. For example, isolated populations were instrumental in the discovery of genetic factors for early-onset Alzheimer's disease [60, 64], Hirschsprung disease [65], hereditary non-polyposis colorectal cancer[66], schizophrenia [66, 67], Type II diabetes, [68] and hypertension [69]. These gene discoveries provided an immediate molecular diagnostic and screening benefit for the isolated population, and also benefited the broader population by advancing understanding of pathophysiological mechanisms and gene functions, and by accounting directly for some of the heritability of common diseases [70].

Over the last two decades, researchers have successfully identified genetic factors that cause some rare familial diseases. They have also identified common genetic factors that contribute risk for common diseases in broader populations. However, despite these successes, most of the heritability in many common disorders

remains elusive [71]. One possible source of the missing heritability in common diseases is rare DNA variants of moderate-to-strong effect. Because they are rare, such DNA variants are difficult to identify using strategies that rely on linkage disequilibrium or deep sequencing. Approaches based on linkage disequilibrium have very low power to detect rare DNA variants, while approaches based on deep sequencing require *a priori* knowledge of pathophysiological mechanisms and gene functions [71].

In a population isolate, the frequency of a rare pathology-inducing DNA variant may be higher than in the general population, thereby increasing the likelihood of finding a significant association with a common disease. For example, in the Icelandic population, two DNA variants—one in *TREM2* and one in *APP*—were found to be associated with late-onset Alzheimer's disease [70, 72]. Both DNA variants were found at a lower frequency in North American and other European populations. In fact, the frequency of the *APP* variant in North American and other European populations was too low to be tested for an association. Thus, these discoveries exemplify how a research partnership with a population isolate can be used to identify low-frequency etiologic DNA variants. This example also highlights the potential general impact of gene discoveries in population isolates. While both of these discoveries were highly significant because they provided insights into the etiology of Alzheimer's disease, the *TREM2* discovery had an additional immediate impact because the risk allele was found at a low, but significant, frequency in other populations worldwide. One way to maximize the impact of genetic discoveries from population isolates is to establish a research partnership with a population whose isolation was relatively recent. Because the

isolation of such a population was recent, the etiologic variant would more likely be found in the broader population and account for disease risk there, too.

North America is a challenging place to identify population isolates. While such isolates exist, including the Hutterite communities [73], Old Order Amish communities [74], and residents of Tangier Island [75], the relatively short time since initial immigration, the high mobility of citizens, and the paucity of distant historical and genealogical records make the identification of new population isolates difficult. In this paper, we present the history and genealogy of a community in mid-Michigan which demonstrates the features of a population isolate. This community came to our attention because of numerous cases of inherited deafness within a small geographical region. Genealogical analysis revealed a community of highly related individuals with a defined set of founders. Our data suggest that we have identified a new North American population isolate of European descent.

2.3 Materials and Methods

2.3.1 IRB and informed consent

All genealogical data for the community were collected as part of a study on congenital deafness [76-78]. That study was approved by the Michigan State University (MSU) Institutional Review Board (IRB) and two community-based committees. The two committees (Research Advisory Committee and Research Ethics Committee) each have six to nine community members and approve all research protocols and scholarly products from within the community, including this manuscript.

2.3.2 Data sources and management

The genealogical data were ascertained from multiple sources: private family genealogical documents; a community history publication; genealogical websites that included the International Genealogy Index (IGI) from the Church of Jesus Christ of Latter Day Saints (LDS), the searchmichigan.net website, individual family websites; cemetery searches in the community; verbal family histories; local historians; microfiche of early Michigan records at the local LDS Family History Library; and marriage banns, birth, death, and anniversary announcements in the local newspaper. Ancestry.com© served as a secondary source to verify data from other sources or to find ancillary dates of birth, immigration, or death. Collection of genealogical data for a given individual ended when they married someone who is not descended from the settlers.

The genealogical data were kept in a large spreadsheet file of multiple worksheets. For the present analysis, the multiple worksheets were concatenated into a large text file of one row per unique individual. The concatenated file was used as input for extended family allocation and relatedness calculations. Each row for an individual contained an identification number, sex, birth date, death date, and the names of parents and spouse(s).

2.3.3 Pedigree structure and relatedness measurements

We used PedHunter2.0 to allocate individuals into extended families, identify the presumed living (extant) population, and calculate measures of relatedness [79]. To allocate individuals to extended families, we wrote a Perl script to iteratively use individual identification numbers (ordered by birth year) as input to the “all_relatives”

query in PedHunter2.0. This query returns a list of all individuals connected to the input individual through parents, spouses, or offspring. The Perl script accepted the output list of relatives from “all_relatives” and assigned the input individual identification number as the extended family identification number.

Founders are individuals who start the genealogic lines because they lack parental names in a pedigree database. Some historic community settlers are pedigree founders. However, the genealogical database contains ancestors of some settlers, thereby creating a distinction between settlers and founders. Founders who contributed only one child to the pedigree were not relevant to this study and were trimmed from the list of founders. We used the “trim_pedigree” utility of PedHunter2.0 to purge all non-relevant founders. Pedigree trimming logic was performed iteratively until no further genetically irrelevant individuals were identified.

Kinship coefficients represent the probability of an allele being shared by an ancestor common to both parents. PedHunter2.0 uses the algorithm of Weir [80, 81] to calculate f-values as a kinship coefficient. We queried the database to produce a list of all mother-father combinations. We used this list as input to the “kinship” complex query. We used the output from “kinship” to assign to each child the kinship coefficient (f-value) of their parents.

Relative Founder Representation (RFR) or “r-value” is 2 times the kinship coefficient (f-value), and represents the proportion of a genome that is contributed by a specific ancestor. For example, an individual has an f-value of 0.125 with their grandparents. Conversely each grandparent contributes 25% of the genomes of their

grandchildren; therefore the r-value between grandparent and grandchild is 0.25. We generated the r-statistics for each relevant founder and their living descendants using the “calculate_r” and “average_r” complex queries of PedHunter2.0.

2.3.4 Plat map

We obtained plat maps of property ownership in 1999 from the respective county governments. We scaled the maps to a common unit, and highlighted the plots owned by individuals who have a surname in common with the settlers.

2.4 Results

2.4.1 Historical context

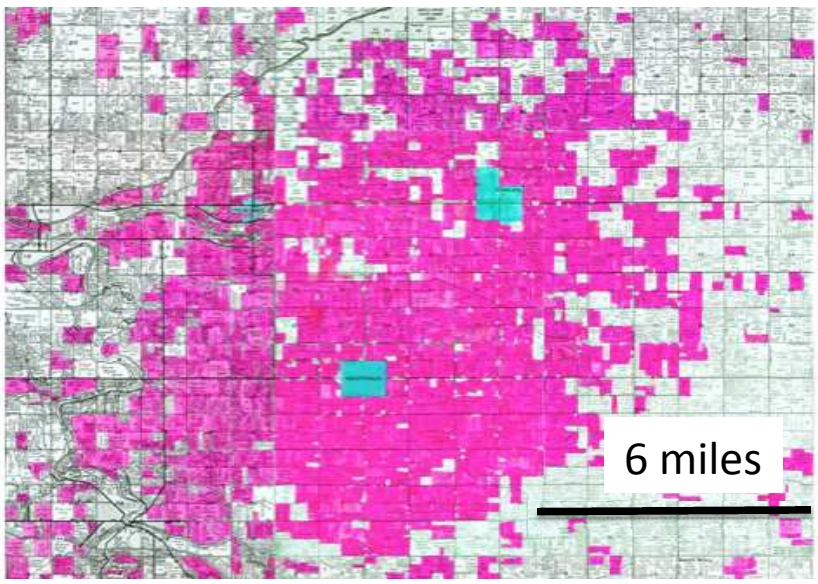
In 1997, we were consulted about the case of a three-year-old child with congenital hearing loss. The team met with the extended family and received a copy of the child’s pedigree going back six generations. The audiologist observed that the surname of an ancestor was the same as the surname of a child who was being counseled for post-cochlear implant speech therapy. With parental permission, this case joined the research study. A genealogist from the community took exhaustive family histories and established a genealogical link between the two families. In total, 15 cases of autosomal recessive deafness were identified in the community and linked within the pedigree. Genetic analyses of this extended pedigree showed that 11 cases were homozygous for the *35delG* mutation of *GJB2*, and four cases were compound heterozygous for *35delG* and a novel deletion located upstream of *GJB2* [76-78].

Further genealogical analyses revealed that a preponderance of the community descended from a group of settlers whose initial arrival antedates Michigan statehood in

1837. In 1836, six individuals, including a Roman Catholic priest, emigrated from the Eifel region of Germany to mid-Michigan. These settlers cleared land, built homesteads, and established a parish church and school. Between 1836 and 1875, 456 settlers followed. Most settlers were other family members and acquaintances from Germany and were in their third or fourth decade of life. The demographics of the settlers were 46% female and 54% male and 350 were married to other settlers at the time of immigration. There were 288 distinct surnames among the 461 settlers. The most common surname was shared by 28 (6%) settlers.

Many of the living descendants of the settlers reside within a 90-square-mile region. This region contains three villages. Much of the land surrounding the villages is owned by individuals whose surnames are the same as those of the settlers (Figure 2-1).

Figure 2-1. Plat map of partner community. Blue squares indicate the location of the three villages and magenta squares indicate property owned by an individual whose surname is the same as one of the original settlers. For interpretation of the references to color in this and all other figures, the reader is referred to the electronic version of this dissertation.



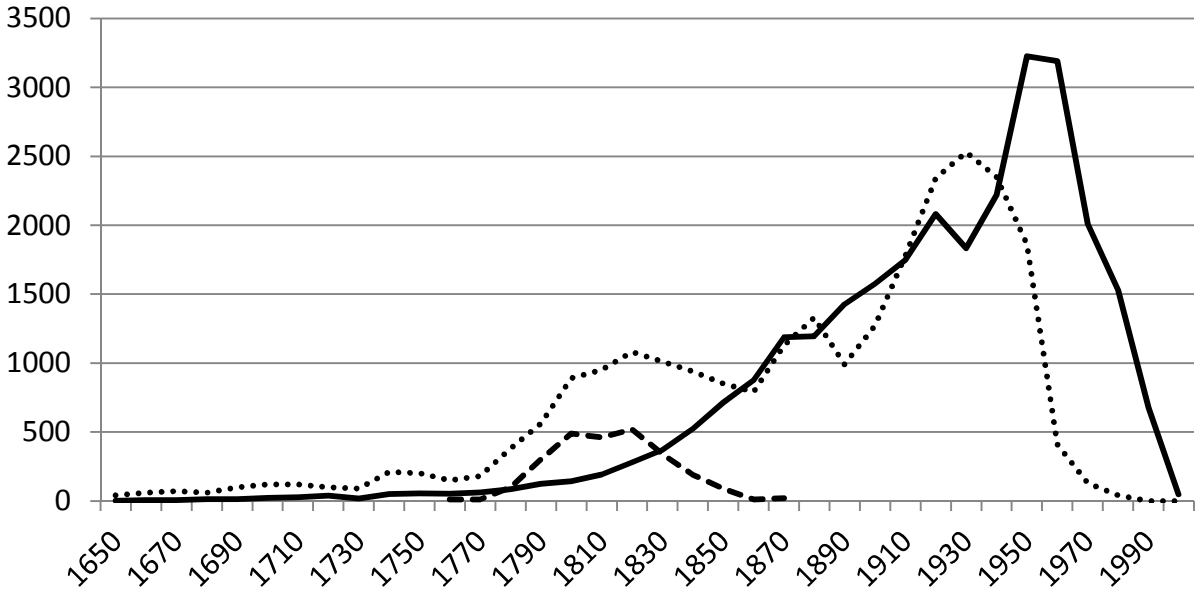
Although the community is not isolated geographically, the German Catholic settlers remained isolated by language and religion from their predominantly English-speaking, non-Catholic neighbors. For example, classes were taught in both German and English in the early years of the settlement. In addition, the school in the primary settlement was exclusively parochial until 1956, and the Catholic churches in each village remain a central social institution supporting the close-knit nature of the population [82] (Goris and Schutte, submitted manuscript).

2.4.2 Genealogical database

The primary sources of the genealogical database were family history publications, accounting for about 40% of the primary records. Internet searches and newspaper publications contributed another 30% and 12% of the records, respectively. The remaining sources (local history publications, cemetery searches, verbal family histories, local historians, early Michigan records, and parish marriage banns) each contributed 5% or less of the records. As of October 1, 2009, the genealogical database contained 28,256 ancestors or descendants of the original settlers. The birth decades of these individuals ranged from 1650 to 2000 (Figure 2-2). The number of births in the genealogical database rises rapidly in the early 1800s and corresponds with the birth decades of the original settlers (Figure 2-2). During the 20th century, the birth rates in the genealogical database precisely parallel the changes in the broader US population, with a sharp decrease in 1930s, a sharp increase in the 1950s and a sharp decrease in the 1960s. These data suggest that this community is strongly

integrated into the broader society both economically and culturally. The precipitous drop in the number of births after 1970 reflects missing data.

Figure 2-2. Numbers of community members, settlers and founders by birth decade. The Settler community experienced rapid growth from the time of the original settlement in 1836 to present (solid line). The apparent decline reflects the end of data collection. The original 254 settlers were born between 1780 and 1860 (dashed line). As expected, the birth decade of the founders (dotted line) tracks about two decades behind the birth decade for the total number of members in the community.



2.4.3 Pedigree analysis

The family allocation logic assembled all individuals from this genealogical database into 310 extended families. One of the extended families included 27,747 (98%) individuals, spanning 13 generations. The extended family within the database contained 2,754 founders. We defined founders as individuals who lacked parental information in the genealogical database. However, we trimmed 244 (9%) non-significant founders who only contributed one descendent to the pedigree, leaving 2,510 founders. While the birth decade of founders ranged from 1650 to 1982, there was a

rapid rise in the number of founders at the beginning of the 19th century (Figure 2-2). As expected, the birth decade of these early founders coincided with the birth decade for the settlers. In fact, the 461 settlers accounted for 344 of the founders in the pedigree. From 1850 to 1910, the rate of new founders declined or was steady (Figure 2-2). However, since the number of births increased during this same period, these data suggest a period of high endogamous growth. The rate of new founders increased after 1920, suggesting an increased rate of exogamous marriages. Like the decline in the total number of births, the decline in the number of founders after the 1970s reflects incompleteness of the genealogical data.

There are 14,756 (53%) individuals in the extended family who were born after 1930. Of these, 972 (3%) have a recorded death date. Thus, the genealogic database contains 13,784 living descendants of the 461 settlers, with equal numbers of males and females. Of the 461 settlers, 389 (84%) have living descendants among the presumed living population. One settler couple has more than 8,000 living descendants.

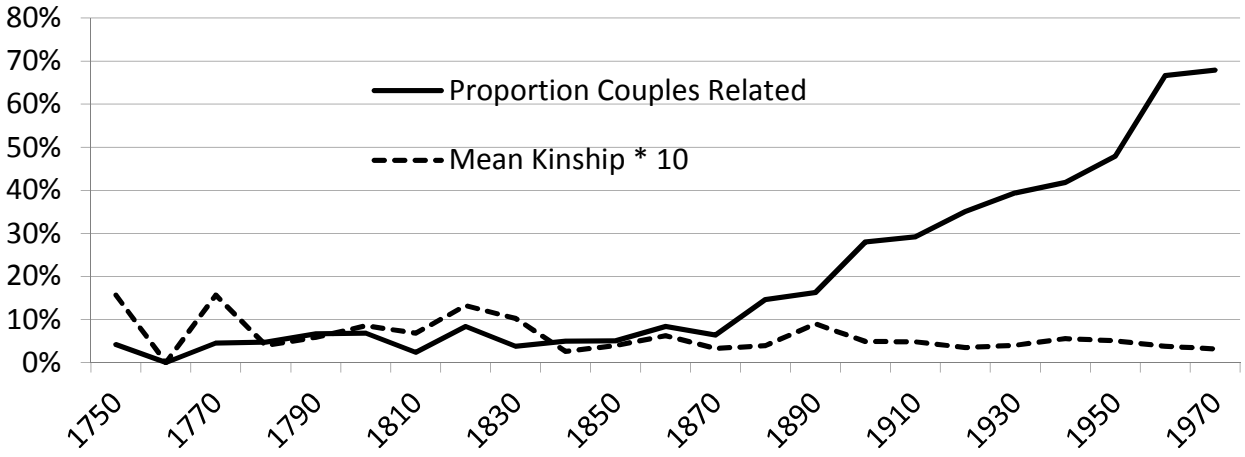
The extant descendants carry 963 surnames, 116 of which are shared with the settlers. The shared surnames are used by 70% (9,531/13,784) of the extant descendants. One surname is shared by 8% (1,047/13,784) of the living descendants.

2.4.4 Measures of relatedness

We measured relatedness in two ways: 1) the frequency of children born to parents who share at least one common ancestor, and 2) the median f-value (kinship coefficient) between all children born to related parents. For a closed population, both

of these values should increase with time, and following completion of the immigration phase, all children should be born to related parents. In this population we observed that the frequency of related parents peaked at 68% in 1970 and, after an initial modest rise, the kinship coefficient decreased (Figure 2-3). Again, these data suggest that while the community was isolated, it was not closed. However, the median f-value (kinship coefficient) was $3.7*10^{-3}$ (range $2.0*10^{-6}$ – $3.4*10^{-2}$), demonstrating the presence of measurable amounts of homozygosity within the extant population.

Figure 2-3. Relatedness of Settler community by birth decade. Proportion of couples who have a common ancestor (solid line) and average F value (dashed line) by birth decade of the female in each couple.



The extant genome represents the genetic contributions of the full set of founders. To estimate a settler's contribution to the extant genome, we calculated the relative founder representation (RFR) for all founders. One founder, born in 1811, contributed 1% of the extant genome. The settlers collectively contributed 67% of the extant genome (Figure 2-4). In comparison, the Old Order Amish population (a closed community) [79] has fewer founders and thus each has a greater contribution to the extant genome. (Figure 2-5).

Figure 2-4. Cumulative mean relative founder representation (RFR) of 13,874 living descended from 2500 founders (461 settlers) in the genealogical database. The RFR curve is organized from the founder with the highest to lowest representation in living members. The indicated points are compared with the cumulative RFR between the Settler (closed circles) and the Old Order Amish databases (open circles; data from Lee et al., 2010).

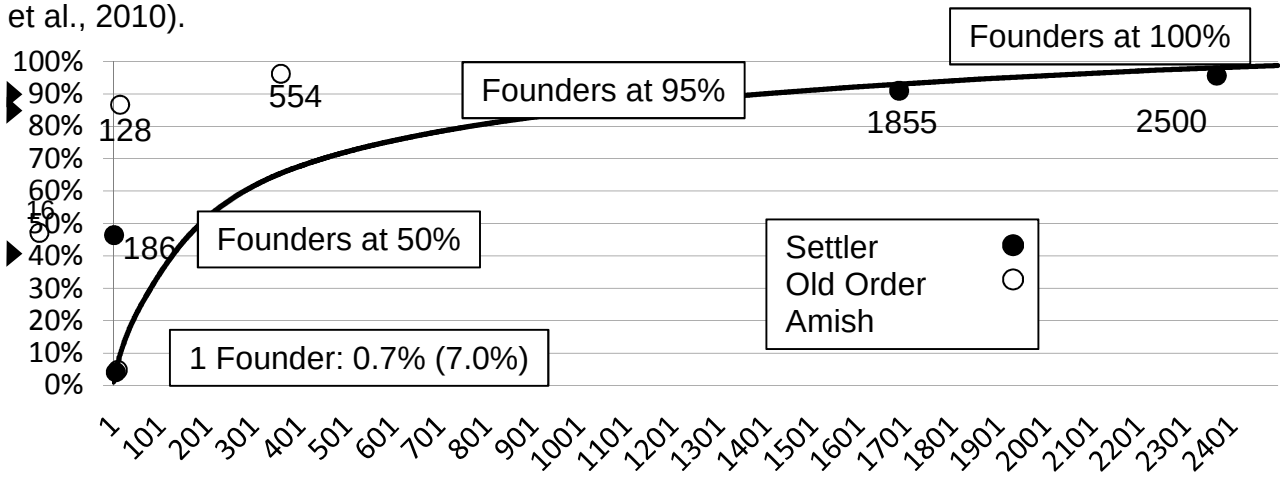
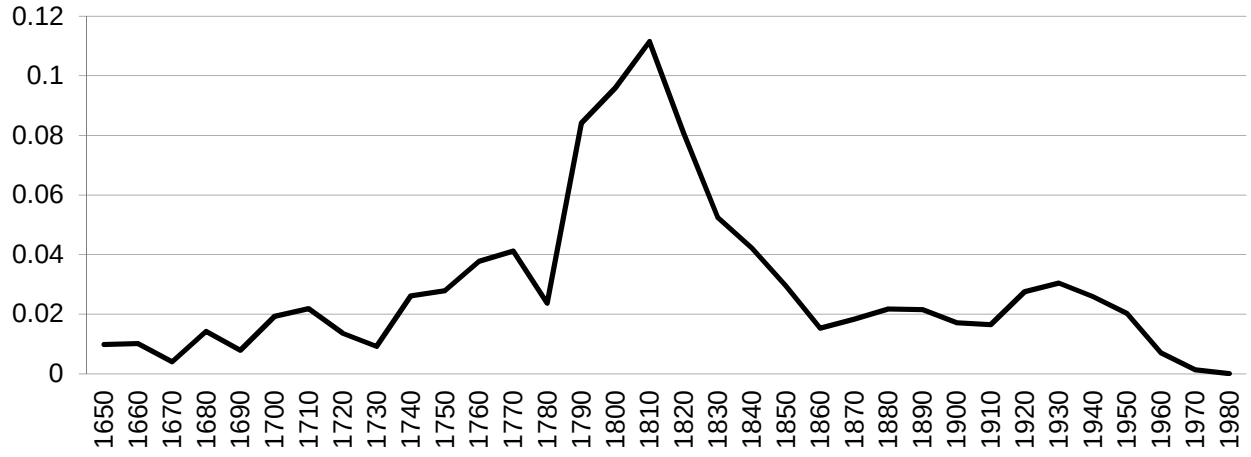


Figure 2-5. Relative founder representation (RFR) and number of founders by birth decade. The genomes of the 254 settlers born in the 1780s to the 1860s represent 62% of the genome of all living members of the community. The founders born since 1870 represent 17%; founders born before 1780 contribute 21%.



2.5 Discussion

The purpose of this study was to consider whether a community in mid-Michigan settled by a few immigrants in 1836 constitutes a new North American population

isolate. While there is no strict definition for a population isolate, we can compare the characteristics of this community with three known population isolates of European descent in North America: the Old Order Amish, the Hutterites of South Dakota, and Tangier Island (Table 2-1).

Table 2-1. Characteristics of population isolates of North America of European descent.

	German-Catholic	Old Order Amish	Hutterite (South Dakota)	Tangier Island
Date of origin	1836	1757	Early 1700s	1722
Founders	2510	554	64	104
Contemporary cohort	~14,000	295,095	722	~200
Isolation	Language/Religion	Religion	Religion	Geographic
Endogamy	42%			
Mean F-value	0.0037	0.015	0.034	0.018
Homogeneity	90 sq mile region		Communal lifestyle	
Recessive traits	Congenital hearing loss			Tangier disease
References		[83]	[73]	[75]

The origin of the isolated community was well-documented, beginning with 461 settlers who emigrated primarily from the Eifel region of Germany. In comparison, the Old Order Amish, Hutterites, and Tangier Island populations settled in North America a century earlier, indicating a more recent isolation in our subject community. The Old Order Amish and Hutterite communities are closed, while the present community is only isolated with 52% of the extant members born to parents that have a common ancestor. Additionally, the median kinship coefficient value for the community ($f = 3.7*10^{-3}$) indicates that the related individuals share a common ancestor within the past six to seven generations. This value is much lower than other North American isolates, yet

considerably higher than the broader population estimates of 99 generations [84] and is consistent with a population isolate.

The pedigree database at the core of this work was assembled by community members and the authors, with the goal of elucidating the biological relationships among the descendants of the settlers. Thus, the founders' descendants were tracked until an individual married someone who was not a descendant of the settlers. While this focus would tend to increase the proportion of individuals born to related parents, the focus does not inflate the degree of relatedness between parents.

Another characteristic of a population isolate is the occurrence of recessive disorders. We previously identified a number of cases of autosomal recessive deafness in the community [76] caused by homozygosity or compound heterozygosity for *GJB2* mutations. We discovered a novel mutant allele that bears a deletion of distant *GJB2*-regulatory sequence [25, 26]. This allele was found in compound heterozygosity with a common mutation of *GJB2* in four deaf individuals in separate sibships. These four individuals, and the other deletion carriers in our study population, share a set of four common ancestors born in Germany between 1702 and 1723, suggesting its presence among German descendants of these four individuals. However, this allele has not been found in other populations [78] and is likely to be rare. Although this rare mutation has not itself contributed to diagnostic yield in genetic testing for hearing loss outside of Michigan, it has provided strong evidence for the location of critical *GJB2*-regulatory elements.

Lastly, isolated populations share a relatively homogenous environment. Although this study did not directly study the environment of our partner community, the plat map provides a very broad overview. Both the rural setting and the close proximity in which descendants of the settlers currently reside are consistent with a shared and homogenous environment [85].

In summary, although the extent of isolation is somewhat lower than in other well-characterized North American isolates, the historical, genealogical, and geographical data are consistent with the conclusion that the community herein presented represents a newly identified North American population isolate.

2.6 Perspectives

We believe that we describe a newly identified population isolate. Importantly, the community in which the descendants reside has demonstrated an ongoing willingness to participate in biomedical research projects, which were focused initially on the genetic causes of congenital deafness. A community health-needs assessment and some early environmental descriptions are underway to guide an expanded research focus to examine the genetic and environmental causes of common complex diseases. The interests and priorities of the community will set the directives. We hypothesize that these processes, coupled with the unique community characteristics, provide a powerful tool to aid in searching for the missing heritability in common complex diseases.

CHAPTER 3: A META ANALYSIS OF MITOCHONDRIAL DNA HAPLOGROUP J AND LONGEVITY

3.1 Abstract

Epidemiologic studies of mitochondrial DNA haplogroups and longevity show mixed results. These studies require large sample sizes because they define population strata within a study cohort. To approach a large sample size, we performed a meta-analysis of the studies presenting mitochondrial DNA haplogroup frequency differences in longevious cases and younger controls. With the pooled measure, we can assess an overall effect size and estimate statistical power within the background haplogroup frequencies and understand the study-level differences contributing to the mixed results.

We followed the well-established meta-analysis steps of literature discovery, citation inclusion, data abstraction, study exclusion and meta-analysis, sensitivity analysis, inter-study heterogeneity assessment, and statistical power estimation.

We found nine studies that present Haplogroup J carrier frequency differences between longevious cases and younger controls. The independent studies have considerable inconsistency in definitions of case, control definitions, genotyping methods, and haplogroup scoring. Studies inconsistently present male:female ratios of their populations. Pooling the independent studies yields 1,224 cases and 1,734 controls. The Haplogroup J carrier rate was 128/1,224 (10.4%) among the longevious cases, and 146/1,734 (8.4%) among the younger controls for a pooled odds ratio of 1.18 (CI_{95} 0.8-1.76 p-value=0.408). The Cochran Q value of 14.424 (p-value = 0.044) indicates considerable inter-study effect size heterogeneity. The studies of definably

European populations show a pooled Haplogroup J carrier rate in longevious cases of 98/864 (11.3%) and 77/1,060 (7.3%) in the younger controls. The stratum-specific pooled odds ratio is 1.47 (CI_{95} 1.052-2.056). When stratified by European vs. Non-European cohorts, the Cochran Q value indicates considerably less inter-study heterogeneity among the European (p=0.267) vs. the Non-European cohorts (p=0.098).

In our pooled analysis of studies of proportionally represented European haplogroups, longevious cases are 1.47 times more likely to be Haplogroup J carriers than younger controls. Future studies in a population isolate with less environmental and genetic heterogeneity focusing on lifespan rather than longevity are required. To achieve sufficient statistical power to confirm or refute an effect of 1.47 in a population that contains the nine European haplogroups requires more than 3,000 cases and as many appropriately chosen controls.

3.2 Introduction

Mitochondria power most cellular processes by generating adenosine-triphosphate (ATP) through oxidative phosphorylation (OXPHOS) along the electron transport cycle (ETC). Mitochondria have their own DNA that is semi-independent from that of the nucleus and which encodes 13 of the 86 polypeptides required for OXPHOS. As a byproduct of the ETC, reactive oxygen species (ROS) or free radicals can be generated. ROS are mutagenic to DNA. The free-radical theory of aging suggests aging cells produce more ROS, thereby increasing the likelihood of mitochondrial DNA (mtDNA) mutations and leading to cellular dysfunction. Conversely, minimized ROS production or tighter mtDNA nucleoid binding might promote a longer lifespan. [9, 86]

Human mtDNA is a 16,569-base-pair circular molecule that has mostly coding nucleotides. Positions 16,024-16,569 and 1-576 are non-coding and hold control regions that house the binding sites for the initiation of transcription and replication. Since there is little selection pressure on mutations in the control regions, these areas are highly variable between individuals and by cell-type within an individual. [9]

Haplotypes of mtDNA polymorphisms can be clustered into haplogroups. Haplogroup names are given in alphabetical order based on their discovery in specific ancestral populations. Haplogroups correlate with the geographic origination of an individual's most distant maternal relative. European populations carry haplogroups H, I, J, K, T U, V, W, and X; while A, B, C, D, E, F, and G indicate Asian ancestry. Haplogroup L is the predominant African haplogroup. [9, 52, 87-91] mtDNA haplogroups can be organized into phylogenetic trees with the African Haplogroup L as a trunk, M and N as transitional branches from which the Asian and European groups stem, and sub-haplogroups as smaller branches. [38]

Europeans carry one of two possible mtDNA polymorphisms at coding position mt11719 (mt11719g or mt11719a). Haplotypes that include mt11719g form the basis of haplogroups H and V. Likewise, haplotypes of mt11917a form the basis of haplogroups I, J, K, U, W, X, and Z. All nine European haplogroups are haplotypes of coding region polymorphisms at mt11719, mt7028, mt12308, mt10393, mt15607, and mt1715. Each haplogroup is in disequilibrium with a haplogroup of non-coding region polymorphisms (e.g., mt73A is frequently presented with mt11719G). Sequencing the compact 1,069 base-pair non-coding region can inductively discover all possible haplogroups in a study cohort. Conversely, a limited set of polymorphisms deductively shows haplogroups

presented in a well-characterized cohort with limited discovery potential and diminished sub-group presentation capacity.

In 1998, Tanaka [44] observed that the coding-region polymorphisms mt5178A, mt8414T, and mt3010A were more frequent in Japanese centenarian cases than younger adult controls. Mt5178A is within Asian Haplogroup D [92]. Building on this evidence, De Benedictis et al. [45] reported finding more Haplogroup J carriers among male northern Italian centenarians than in younger adult controls. Feng et al. found that Haplogroup F is over-represented in Chinese healthy longevious cases compared with younger female controls [93]. It is possible that certain population-specific haplogroups associate with longevity.

Following the De Benedictis work, other groups have presented differing effect sizes on the Haplogroup J and longevity association. [94]. Despite the inconsistent results, the association was highlighted as a genetic variation favorable to longevity. [95]

Samuels showed that studying the effect of one haplogroup in the background of those in a population requires sample sizes larger than those presented in any single study. [94] They presented a universal formula to be used to estimate sample size and statistical power for haplogroup association studies. The formula offers a sample-size inflation factor to account for the multiple association testing that is inevitable when assessing one haplogroup against the background others present in a population.

In order to measure the effect of mtDNA Haplogroup J and longevity, we undertook this meta-analysis to assemble and analyze all the published evidence of the topic.

3.3 Methods

3.3.1 Four Stages

We followed the recommendations of the "Meta-Analysis of Observational Studies in Epidemiology (MOOSE)" [96] while executing the steps of Harrison [97]. To discover all available literature, we conducted a thorough literature search. Two of us (MF, JB) independently reviewed the abstracts of each discovered article while applying the inclusion criteria (Table 3-1).

Table 3-1**.** *A priori* Inclusion / Exclusion Criteria

Inclusion Criteria:

- Age of cases clearly described and presented
- Haplogroup frequency differences between cases and controls

Exclusion Criteria;

- Haplogroup J frequency not presented
- Controls are younger than cases

We then read the complete text of each included study to determine whether the methods and findings were consistent with our objectives. Studies that did not further this objective were dropped from further review. We then abstracted data into the data abstraction form (Table 3-2).

Table 3-2. *A priori* selected variables abstracted from each manuscript	
Population Ancestry Geography	From which geographic region was the population taken. Sub-population or minorities are favored over larger groups
Major Conclusion	What does the study cite as its major finding
Molecular Method	What areas of the mitochondrial genome are used to find haplogroup/SNPs coding region PCR fragments, HVS1, HVS2, Whole genome sequencing, SNP sequencing, RFLP)
Phylogeny Method	What method is used to ascribe haplogroups to the SNPs
Included in pooled analysis	Yes/No- if no, why
Case Definition	How does the study define its cases (e.g., Centenarians (100+))
Case Age Range	What is the age range of the cases
Median Case Age	What is a the median or mean age of cases
Control Age Range	What is the age range of controls
Median Control Age Range	What is the median or mean age of the controls
# Cases	How many controls
# Controls	How many cases
% Female Cases	What percentage of the cases were female
%Female Controls	What percentage of the controls were female
% J in Cases	What percentage of the cases were of Haplogroup J
% J in Controls	What percentage of the controls were of Haplogroup J
Haplogroup rank and % in Controls	List haplogroups in rank order of frequency among the controls
Point Estimate (Odds ratio and Confidence Interval)	Calculate the odds ratio of Haplogroup J in the cases (and 95% Confidence Interval)

Our literature search used PubMED and EMBASE to discover all potentially relevant articles on January 1, 2013. In PubMED, we searched using the MeSH terms and a keyword search. The selected MeSH search terms were: "DNA, Mitochondrial/genetics," "Aging/genetics," "humans," "Haplotype," “Haplogroup,” and “Longevity/genetics.” The keyword search terms used were: “mitochondrial DNA,” "Centenarian,*" "Octogenarian.*” and "Nonagenarian.*" We searched EMBASE using the following terms: “mitochondrial DNA,” “aging” “human,” “longevity,” “lifespan,” “haplotype,” and “haplogroup.”

We anticipated that most of the studies we discovered would present categorical dependent (longevious cases and younger controls) and ordinal (haplogroup frequencies) independent variables. However, it was also possible that studies present a continuous dependent variable (age of death). If a preponderance of included studies sustained the anticipated data types, we would use an odds-ratio statistic to measure the pooled effected. However, if the preponderance of studies presented with ordinal and continuous age at death, we would use a time-to-event statistical approach.

Using an *a-priori* established data-abstracted form, we reviewed established binary outcomes (case/control) and therefore the odds-ratio as the most appropriate effect-size estimating technique. A random-effect model was used to determine confidence intervals and significance levels. We used OpenMetaAnalyst© [7] as our database and analytic software for the pooled, heterogeneity, and sensitivity analyses.

To assess publication bias, we present the published studies in chronologic order to discover whether over time the point estimates migrate to unity.

3.4. Results

3.4.1 Literature Discovery, Study Inclusion, Data Abstraction and Study Exclusion

The PubMED search returned 334 citations. The EMBASE search returned 103 distinct citations. From a review of the abstract of these 437 discovered citations, and following the criteria in Table 3-2, we chose to include 85 unique citations for an in-depth review for possible selection. We dropped 67 studies for reasons listed in Table 3-3. We selected 17 citations for data abstraction and further consideration in the analysis stage. Two co-authors (JB and MF) independently abstracted data and jointly discerned which articles to exclude. Reasons for exclusion are presented in Table 3-4.

Nine studies of 17 were ultimately chosen for the analysis stage. All studies were in English with no references to unpublished datasets.

Table 3-3 Results of inclusion round			
#	Review Findings	Article Title	First Author – [Citation]
1	No Haplogroup J	Enrichment of longevity phenotype in mtDNA haplogroups D4b2b, D4a, and D5 in the Japanese population.	Alexe [92]
2	Not Longevity	Mutations for Leber hereditary optic neuropathy in patients with alcohol and tobacco optic neuropathy	Amaral-Fenandes [98]
3	Review	Aging Liver. A review.	Anantharaju [99]
4	Review	Biology and genetics of human longevity	Baggio 1998 [100]
5	No Haplogroup	Evolutionary pressure on mitochondrial cytochrome b is consistent with a role of CytbI7T affecting longevity during caloric restriction	Beckstead [42]
6	No Haplogroup	Mitochondrial DNA haplogroup D4a is a marker for extreme longevity in Japan.	Bilal [101]
7	Include		Benn [39]
8	Review	The genetic factors of longevity	Blanche [102]
9	Not mitochondrial haplogroups	Increase of homozygosity in centenarians revealed by a new inter-Alu PCR technique	Bonafe [103]
10	Not Mitochondria	An allele of HRAS1 3'variable number of tandem repeats is a frailty allele: implication for an evolutionarily-conserved pathway involved in longevity.	Bonafe [104]
11	Review	The genetics of human longevity	Capri 2008 [105]
12	Not Mitochondria	Human longevity within an evolutionary perspective: The peculiar paradigm of a post-reproductive genetics	Capri [106]
13	Include	Mitochondrial Polymorphisms Are Associated Both with Increased and Decreased Longevity	Castri [107]
14	Not mitochondrial haplogroups	Decreased reactive oxygen species production in cells with mitochondrial haplogroups associated with longevity.	Chen [108]
15	Not Mitochondria	A role for the CISD2 gene in lifespan control and human disease	Chen [109]

Table 3-3 Cont'd			
16	Not a primary study	MtDNA mutations in aging and apoptosis.	Chomyn [110]
17	Epistasis	Variation in mitochondrial genotype has substantial lifespan effects which may be modulated by nuclear background	Clancy [111]
18	Include	Data from complete mtDNA sequencing of Tunisian centenarians: testing haplogroup association and the golden mean" to longevity."	Costa [46]
19	Include	Mitochondrial Haplogroup X is associated with successful aging in the Amish	Courtenay [112]
20	Poster	Is mitochondrial genome chicken or egg of carcinogenesis?	Czarnecka [113]
21	Include	Association of the mitochondrial DNA haplogroup J with longevity is population specific	Dato [47]
22	Review	Does a retrograde response in human aging and longevity exist?	De Benedictis [114]
23	Review	Inherited variability of the mitochondrial genome and successful aging in humans.	De Benedictis [115]
24	Include	Mitochondrial DNA inherited variants are associated with successful aging and longevity in humans.	De Benedictis [45]
25	Not a primary study	Recent advances in human gene-longevity association studies	De Benedictis [116]
26	Not Longevity	Genetics and anthropology in studies on aging and Chagas disease	De Buenos Aries [117]
27	Not a primary study	Another case of Leber hereditary optic neuropathy in an octogenarian.	Decanini-Mancera [118]
28	Include	Association of mitochondrial haplogroup J and mtDNA oxidative damage in two different North Spain elderly populations	Dominguez-Garrido [48]
29	No Haplogroup J	Association of mtDNA haplogroup F with healthy longevity in the female Chuang population, China	Feng [93]
30	Not Mitochondria	Biomarkers of immunosenescence within an evolutionary perspective: The challenge of heterogeneity and the role of antigenic load	Franceschi [119]

Table 3-3 Cont'd			
31	Not Mitochondria	Genetics of healthy aging in Europe - The EU-integrated project GEHA (GEnetics of Healthy Aging)	Franceschi [120]
32	Not Mitochondria	Inflammaging and anti-inflammaging: A systemic perspective on aging and longevity emerged from studies in humans	Franceschi [121]
33	Not Mitochondria	Inflammaging as a major characteristic of old people: Can it be prevented or cured?	Franceschi [122]
34	Not mitochondrial haplogroups	Aging, resting metabolic rate, and oxidative damage; Results from the Louisiana healthy aging study	Frisard [123]
35	Not Longevity	Mitochondrial genotype frequent in centenarians predisposes resistance to adult-onset diseases	Gong [124]
36	Not Humans	Emerging area of aging research: long-lived animals with negligible senescence	Guerin [125]
37	Not mitochondrial haplogroups	Repeats, longevity and the sources of mtDNA deletions: evidence from 'deletional spectra'.	Guo [126]
38	Not mitochondrial haplogroups	Mitochondrial DNA, direct repeats, deletions, and centenarians: An unfinished puzzle	Guo [127]
39	Review	Inflammation and genetics: an insight in the centenarian model.	Iannitti [128]
40	Include	Mitochondrial genotype associated with French Caucasian centenarians.	Ivanova [129]
41	Not mitochondrial haplogroups	Aging-related occurrence in Ashkenazi Jews of leukocyte heteroplasmic mtDNA mutation adjacent to replication origin frequently remodeled in Italian centenarians.	Iwata [130]
42	Review	Current view on biology of aging	Kasabri [131]
43	Not Longevity	Cell-by-cell scanning of whole mitochondrial genomes in aged human heart reveals a significant fraction of myocytes with clonally expanded deletions.	Khrapko [132]
44	Not longevity	Mitochondrial genotype in vulvar carcinoma - cuckoo in the nest.	Klemba [133]

Table 3-3 Cont'd			
45	Not mitochondrial haplogroups	Longevity-associated mitochondrial DNA 5178 A/C polymorphism modulates effects of daily drinking and cigarette consumption on serum triglyceride levels in middle-aged Japanese men.	Kokaze [134]
46	Not Mitochondria	Gender-specific association between -1082 IL-10 promoter polymorphism and longevity.	Lio [135]
47	Include	Mitochondrial DNA in Polish centenarians.	Lorenc [136]
48	Not Mitochondria	TOMM40 rs10524523 polymorphism's role in late-onset Alzheimer's disease and in longevity	Maruszak [137]
49	Not Longevity	Defining the cytokine and immunoglobulin phenotype associated with mitochondrial J haplotype carriers in the BELFAST Elderly Longitudinal Free-Living Ageing Study (BELFAST)	Mcnerlan [138]
50	Not Longevity	Mitochondrial DNA 3644T-->C mutation associated with bipolar disorder.	Munakata [139]
51	Include	Mitochondrial DNA polymorphisms associated with longevity in a Finnish population	Niemi [52]
52	Include	A combination of three common inherited mitochondrial DNA polymorphisms promotes longevity in Finnish and Japanese subjects	Niemi [89]
53	Not mitochondrial haplogroups	Absence of association between mitochondrial DNA C150T polymorphism and longevity in a Han Chinese population	Pan [140]
54	Not Mitochondria	Y chromosome binary markers to study the high prevalence of males in Sardinian centenarians and the genetic structure of the Sardinian population	Passarino [141]
55	Include	Are mitochondrial haplogroups associated with extreme longevity? A study on a Spanish cohort.	Pinos [142]
56	Not a primary study	Association studies on human mitochondrial DNA: Methodological aspects and results in the most common age-related diseases	Raule [143]

Table 3-3 Cont'd			
57	Not Longevity	Mitochondrial J haplogroup is associated with lower blood pressure and anti-oxidant status: findings in octo/nonagenarians from the BELFAST Study.	Rea [144]
58	Include	MITOCHONDRIAL DNA HAPLOGROUPS IN A CHINESE UYGUR POPULATION AND THEIR POTENTIAL ASSOCIATION WITH LONGEVITY	Ren [145]
59	Epistasis	Somatic point mutations in mtDNA control region are influenced by genetic background and associated with healthy aging: a GEHA study.	Rose [146]
60	Not HPG	The mitochondrial DNA control region shows genetically correlated levels of heteroplasmy in leukocytes of centenarians and their offspring.	Rose [147]
61	Not a primary study	Paradoxes in longevity: sequence analysis of mtDNA haplogroup J in centenarians.	Rose [148]
62	Include	Mitochondrial DNA polymorphism: its role in longevity of the Irish population.	Ross [49]
63	Review	Genes, ageing and longevity in humans: problems, advantages and perspectives.	Salvioli [149]
64	Not a primary study	The impact of mitochondrial DNA on human lifespan: a view from studies on centenarians.	Salvioli [150]
65	Not a primary study	Mitochondrial DNA involvement in human longevity	Santoro [151]
66	Not mitochondrial haplogroups	PROSPECTS FOR THE GENETICS OF HUMAN LONGEVITY	Schachter [152]
67	Not Longevity	Genetic variability of HVRII mtDNA in cord blood and respiratory morbidity in children	Schmuczerova [153]
68	Include	Ashkenazi Jewish centenarians do not demonstrate enrichment in mitochondrial haplogroup J.	Shlush [50]
69	Not Longevity	Cardiovascular disease: Primary prevention, disease modulation and regenerative therapy	Sultan [154]

Table 3-3 Cont'd			
70	Not longevity	Association of a 5178C-->A (Leu237Met) polymorphism in the mitochondrial DNA with a low prevalence of myocardial infarction in Japanese individuals.	Takagi [155]
71	Not Longevity	Japanese Alzheimer's disease and other complex disorders diagnosis based on mitochondrial SNP haplogroups	Takasaki [156]
72	Not mitochondrial haplogroups	Mitochondrial SNPs associated with Japanese centenarians, Alzheimer's patients, and Parkinson's patients.	Takasaki[157]
73	Not Longevity	Mitochondrial haplogroups associated with Japanese centenarians, Alzheimer'spatients, Parkinson's patients, type 2 diabetic patients and healthy non-obese young males	Takasaki [158]
74	Review	Mitochondrial genotype associated with longevity and its inhibitory effect on mutagenesis.	Tanaka [159]
75	Not Longevity	Golden mean to longevity: rareness of mitochondrial cytochrome b variants in centenarians but not in patients with Parkinson's disease.	Tanaka [160]
76	Include	Mitochondrial genotypes and cytochrome b variants associated with longevity or Parkinson's disease.	Tanaka [161]
77	No Haplogroup	Mitochondrial genome single nucleotide polymorphisms and their phenotypes in the Japanese	Tanaka [162]
78	Not mitochondrial haplogroups	Common, rare, and individual variations of mitochondrial DNA associated with diseases or longevity	Tanaka [163]
79	Not mitochondrial haplogroups	Mitochondrial genotype associated with longevity and its inhibitory effect on mutagenesis	Tanaka [159]
80	Not mitochondrial haplogroups	Single nucleotide polymorphism spectra in newborns and centenarians: identification of genes coding for rise of mortal disease	Tomita-Mitchell [164]
81	Not mitochondrial haplogroups	Low level of the mtDNA(4977) deletion in blood of exceptionally old individuals.	von Wurmb-Schwark [165]
82	Not mitochondrial haplogroups	Genetic determinants of exceptional human longevity: insights from the Okinawa Centenarian Study	Willcox [166]

Table 3-3 Cont'd			
83	Include	Mitochondrial DNA polymorphisms are associated with the longevity in the Guangxi Bama population of China.	Yang [167]
84	Review	Genes and longevity: lessons from studies of centenarians.	Yashin [168]
85	Not mitochondrial haplogroups	Strikingly higher frequency in centenarians and twins of mtDNA mutation causing remodeling of replication origin in leukocytes.	Zhang [169]

Table 3-4**.** Abstracted data for meta-analysis from studies that passed the inclusion round						
First Author/ Year	Ancestry / Geography	Genotyp-ing source and method	Overall Odds Ratio CI_{95}	Case Definition Age Range Mid-point # (%Female)	Control Definition Age Range Mid-Point # (%Female)	Haplo-group Ranks by percent
1. DeBenedictis [45] 1999†	European Italian	Coding region RFLPs	2.4 (1.2-5.0)	Centenarian 100-109 104 N=212 (70%)	Younger Adults 20-75 47.5 N=275	H=40,U=17, T=11, Others=10, K=7, J=4, X=3, I=3, V=2 , W=1
2.Ross [49] 2001†	European Irish	Coding and Non-coding RFLPs	0.8 (0.4-1.8)s	80+ 80-97 N=129 (70%)	Younger Adults 18-45 31.5 N=100	*H=45,U=15**, J=14
3.Niemi[52] 2003†	European Finnish	Coding region RFLPs	1.4 (0.7-2.6)	90+ 90-97 94 N=225 (n/r**)	Younger Adults 18-65 41.5 N=400 Infants	H=51,U=21, J=6, K=6, T=6, V=4, I=3, X=2, W=1, M=1, OTHER=<1

Table 3-4 Cont'd						
4.Dato[47] 2004	European Southern Italy	Coding region RFLPs	Not available Exclude: Not a case/ control study.			Not stated
5.Lorenc[136] 2004 †	Polish	Coding region RFLPs	1.6 (ns)	Centenarians Range not specified	Anonymous controls	H=39,U=21,Others=13, T=13, J=8,K=6
6.Niemi[89] 2005	Finnish and Japanese	Coding region RFLPs	Exclude: No haplogroup frequency differences			
7. Benn[39] 2008	Danish	Selected coding region SNPs	Exclude: Not a case/control study			
8.Shlush[50] 2008 †	Ashkenazim	Sequencing of non-coding region	1.03 (0.6-1.8)	Centenarian 94-100 97 +/- 2.8 N=241 (79)	Spouses+ Israeli Controls 69 +/-8; 41.8 +/-16.2 N=683(n/r)	K=28,H=18, N=9, J=7, HV=5, U=5, T=1, V=3, R0=2, M=2 , L=1, W=1, X=<1, W=1, I=<1, A=<1

Table 3-4 Cont'd						
9.Ren[145] 2008†	Uygur/ Chinese	Full mtDNA sequence	0.4 (0.2-0.9) ††	Vitality 90+ 90-97 93.5 N=215 (n/r)	Healthy Uygur 45-55 50 N=117 (n/r)	H=38,J=19, U=9, T=8, K=6, V=6, I=4, OTHER=4, X=3, W=3
10.Bilal[101] 2008	Japanese	Full mtDNA sequence and PCA SNP selection	Exclude: HPG J not presented	>105 105-115	Several other groups 18-100	N/R
11.Costa[46] 2009 †	Tunisian	Full mtDNA Sequence	∞ p=0.2326	Centenarian Range not stated	Unrelated 19-45 32 N=22 (n/r)	L=55,U=22, R0=18, X=5
12.Castri[107] 2009	Costa Rican	HVS1 sequence/ Coding region RFLPs	Exclude: Not a case/control study, HPG J not presented			
13.Dominguez-Garrido[48] 2009 †	European Spanish	Coding region RFLP	2.2 (1.1-4.2).	Healthy 85+ n/r n/r N=138 (n/r)	Healthy blood donors 20-40 30 N=138 (45%)	H=44,U=23, J=11, R0=10, T=6, V=6
14.Pinos[142] 2011 †	European Spanish	Coding region RFLP	0.3 (0.06-1.2) Include cases with Dominguez controls[48]	Centenarian 100-109 104 N=65 (87%)	See Dominguez-Garrido	

Table 3-4 Cont'd						
15. Feng[93] 2011	Chinese	mtDNA Sequence and Coding RFLP	Exclude: HPG J not presented			
16. Courtenay [112] 2012	Amish	27 coding region SNP part of GWAS	Exclude: Controls same age as cases.	80+		
17. Yang [167] 2012	Chinese	mtDNA Sequence and Coding RFLP	Exclude: Haplogroup frequencies not presented in controls	Centenarian	10-79	Not presented
† Study selected for meta-analysis †† Inclusion is qualified based on the absence of Asian Haplogroups in a Uygur population. * Haplogroups H and U are estimated ** n/r= Not Reported						

3.4.2 Key Study Variables

Among the nine selected studies, five used 100 as the lower age limit for cases [46, 48, 50, 136, 142]; two studies included 90-year-olds [52, 145], one included 85-year-olds [48], and one included 80-year-olds [49]. The median age range of controls was 24 to 80 years, with no inter-study consistency. One study [136] did not present ages of the anonymous controls. Since the age of cases varies and age of controls is inconsistently presented, an age variable cannot be included in a meta-regression.

The genomic methods used by the nine selected studies varied considerably. One study used a full sequence of the mtDNA control region [50]. Five studies used restriction fragment length polymorphism (RFLP) of the coding region [45, 48, 52, 136,

142]. One study used RFLPs of both the coding and control regions [49]. Two studies used a full mtDNA sequence (coding and control regions) [46, 145]. Eight studies used established haplogroup scoring algorithms [45, 46, 48, 50, 52, 136, 142, 145], one study presented de-novo phylogeny and haplogroup nomenclature [49]. The frequencies of non-J and J sub-haplogroups, or sex ratios of cases and controls, were not consistently presented in the selected studies.

3.4.2 Critical Study Appraisal

In their discussion, Shlush et al. [50] used the Samuels [94] formula to estimate statistical power to find an odds ratio of 1.75. The Samuels [94] formula offers an inflation factor to accommodate the cryptic population stratification present between haplogroups when determining sample size and power for studies of haplogroup associations. The inflation factor uses the number of haplogroups within the study population. The Shlush study had 15 haplogroups; yet used 1 as the inflation factor. Rather, using an inflation factor of 15, the Samuels formula recommends 1,047 cases (400% more than were assembled). In the discussion section, the authors of Slush et al. state that 251 cases (4% more than they assembled) are necessary to find their point estimate (OR=1.75). Thus, the Samuels formula suggests that their study was under-powered.

Ren [145] presented the haplogroup frequencies among the Uygur minority in China. They present Haplogroups H, J, U, T, K, V, I, Others, X, and W (in rank order) among their controls. Surprisingly, the Asian haplogroups (A, B, D, F, and G) were absent in this study. In an earlier study of the Uygur minority [170], 17/47 (36%) individuals shared Asian haplogroups A, B, D, F, and G. The Ren study used a

sequence of the full mtDNA genome across several PCR fragments and aligned sequences to two reference sequences [14, 20]. The molecular method is comprehensive and inductive; yet, no Asian haplogroup carriers were found among the controls. We assigned this study population to be non-European, despite the European haplogroup distributions.

Dato et al. [47] analyzed Haplogroup J frequencies as a function of increasing age across categories within an Italian cohort. No age category could clearly be defined as a case group to be consistent with the other included studies. Extrapolating carrier frequencies from points on the chart would be too imprecise. A formal request for original data was declined (personal communication with S. Dato).

Niemi et al. [89], in a follow-up to their 2003 work, presented an association study of mt150T with longevity in Finnish and Japanese populations. They presented haplogroup frequencies among mt150T carriers, not between cases and controls.

Bilal et al. [101], Feng et al. [93], and Yang [167] presented haplogroup frequency among Asian cases and controls. These studies were excluded because, as expected, Haplogroup J was not observed in Asian populations.

Castri et al. [107] evaluated ancestor lifespan from within a pedigree of contemporary individuals. The group assembled living individuals who have a well-documented pedigree database. They assessed mtDNA haplogroups of the contemporary individuals and ascribed to ancestors in the pedigree. The study tested whether any ancestors carrying specific haplogroups lived longer than others using period-specific analysis of variance (ANOVA) models on age of death. The novel study

design does not present haplogroup frequency differences between longevious cases and younger controls, and thus cannot be included in a case-control meta-analysis.

Courtenay et al. [112] presented haplogroup frequency differences between healthy Amish 80+ year olds in comparison to less-healthy Amish 80+ year olds. They found that Haplogroup X is more frequent in their healthy 80+ year olds. This paper did not present haplogroup frequency differences between longevious cases and younger controls and was not suitable for the meta-analysis.

Benn et al. [39] presented a prospective study of 9,254 citizens of Denmark enrolled between 1976-1978 and followed for 25 years until 2004. Using personal interviews/examinations, electronic health records, and death registry, they have complete follow-up for all deaths and disease diagnoses. They used six coding region polymorphism to deduce the common European haplogroups. Correlating haplogroups with disease occurrence, cause of death and age at death, they found no positive association with any haplogroup. The study presented no longevious cases and younger controls, and thus cannot be included in the meta-analysis. Their equivocal results in a large sample followed prospectively might seem to offer a definitive answer. This study was eventually dropped from the selected group because the prospective design lacks case/control perspectives.

3.4.3 Synopsis of study findings

We reviewed nine studies that presented haplogroup frequency differences in longevious cases (80-100+) and younger controls. The range of ages of controls was inconsistently presented. All nine studies presented Haplogroup J frequency differences between cases and control, and were therefore included in a meta-analysis. Six studies

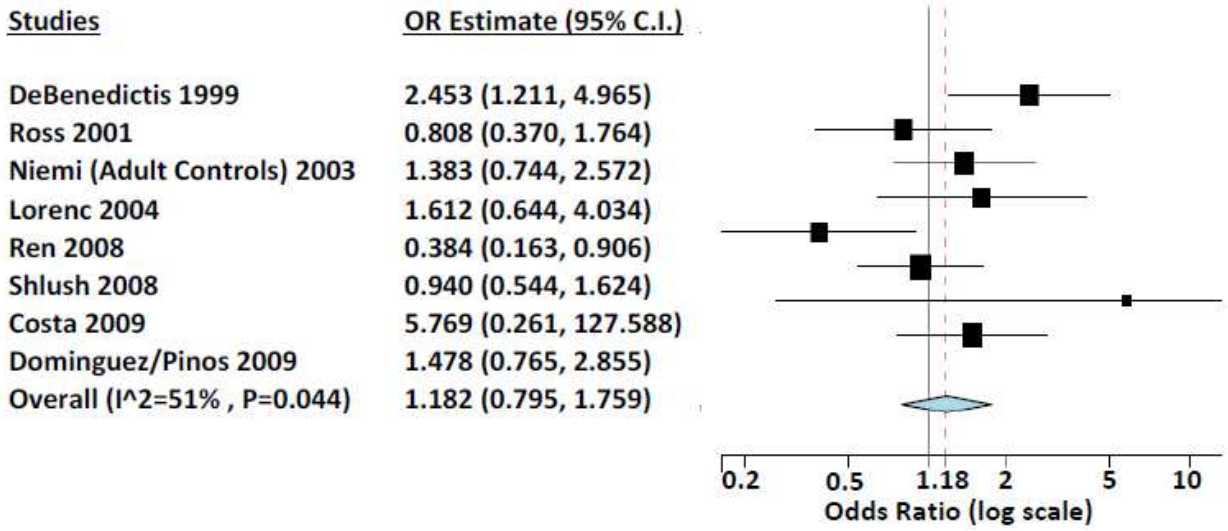
presented individuals from European ancestry [45, 48, 49, 52, 136, 142]. Three studies presented non-European populations (Ashkenazim, Tunisians, Uygur/Chinese) [46, 50, 145]. One study [142] repurposed controls from a prior study [48], and we included both as one entry in the meta-analysis. In total, we included seven independent studies and two joint studies in the meta-analysis, giving a sample size of eight. The abstracted data for the analysis are presented in Table 3-4.

3.4.4 Meta-analysis results

3.4.4.1 Main Results

Our pooled analysis of eight independent studies yields 1,224 longevious cases and 1,734 controls. The definition of longevious varies considerably across the studies, with the youngest cases being over 80 and the oldest 100 and over. Not all studies presented the ratio of males to females. The Haplogroup J carrier rate was 128/1,224 (10.4%) among the longevious cases, and 146/1,734 (8.4%) among the younger controls. The summary odds ratio is an equivocal 1.18 (CI_{95} 0.8-1.76; p-value = 0.408). The Cochran Q value was 14.424 (p-value = 0.044), indicating considerable inter-study heterogeneity of effect size [171, 172]. (Figure 3-1)

Figure 3-1: Forest Plot of Haplogroup J frequencies between cases and controls by year of publication. (Figure shows later studies not approaching unity of risk. Offers evidence of absence of a publication bias).



3.5 Discussion

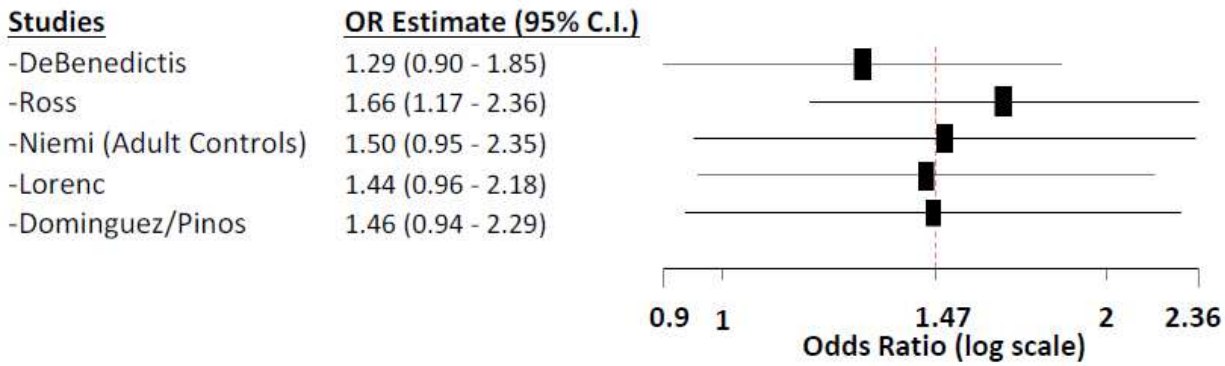
The studies of definably European populations show a pooled Haplogroup J carrier rate for a stratified pooled odds ratio of 1.47 (CI_{95} 1.052-2.056). When stratified by definable ancestry, the Cochran Q value indicates considerably less inter-study heterogeneity among the European (p = 0.267) vs. the Non-European cohorts (p = 0.098).

For a sub-group analysis we stratified studies by population ancestry. De Benedictis [45], Ross [49], Niemi [52], Dominguez-Garrido/Pinos [48, 142], and Lorenc [136] used explicitly European populations. The data from Ren [145], Costa [46], and Shlush [50] were grouped as “non-European” because the observed or expected haplogroup distributions were inconsistent with a European population.

For the European sub-group, Haplogroup J was more common in in longevious cases of 98/864 (11.3%) and 77/1,060 (7.3%) in the younger controls. The stratum-specific summary effect of OR=1.47 (CI_{95} 1.04-2.06). In the non-European strata, the effect was inconclusive (OR=0.749; CI_{95} 0.30-1.84). The stratified Cochran Q values indicate less inter-study heterogeneity in the sub-group analyses than with two groups combined. (p = 0.267 for European studies; p = 0.098 for Non-European studies).

To further analyze the pooled studies for potential stratification, we performed a leave-one-out sensitivity analysis. For the European sub-group, we observed a significant increase in the pooled odds ratio from 1.2 to 1.0 when the Ross [49] study data were removed. (Figure 3-2)

Figure 3-2: Leave-one-out sensitivity analysis of Haplogroup J studies among European cohorts. For interpretation of the references to color in this and all other figures, the reader is referred to the electronic version of this dissertation



We find the association between Haplogroup J and longevity is inconclusive. However, when we stratified by ancestry Haplogroup J is significantly over-represented in longevious cases than younger controls in the European sub-group (OR = 1.47; CI_{95} 1.04-2.06). Among Non-European populations, the association is equivocal (OR = 0.749; CI_{95} 0.30-1.84). Pooling all published literature fails to achieve sufficient power. Additionally, longevity has considerable time-period variation and appropriate controls

should be assessed contemporaneously with cases, making lifespan or age at death a more suitable statistical measure than longevity. Also, in humans, the semi-independent mitochondrial genome is in epistasis with the autosomal genome; knowledge of both genomes is required to fully appreciate the Haplogroup J and longevity (or lifespan) association.

The association among Europeans was consistent across most published studies, with only one study [49] presenting an inverse effect. This single study with inconsistent results presented two sub-groups of Haplogroup J. One sub-group was in positive association with longevity; the other had a negative association. The sub-group in negative association was more frequent in this Irish population. The greater frequency of the negative sub-group in this study equivocated the overall Haplogroup J finding. The study used a *de-novo* nomenclature and a deductive RFLP-based genotyping; thus the sub-groups names were inconsistent with other studies.

If the true effect size of Haplogroup J and longevity is OR=1.47 among the nine European haplogroups, the Samuels universal statistical power formula estimates that more than 6,000 cases and an equal number of controls are needed to find the modest effect. This comprehensive meta-analysis could only assemble 864 European longevious cases and was thus considerably underpowered.

A recent review of the genetics of longevity presented Haplogroup J as a factor favorable to increased lifespan or longevity. [95] Among Asian populations, Haplogroup D appears more frequently in older individuals. Dato et al. suggested that Haplogroup J and longevity might be population specific. Our results support their suggestion by

showing that, among populations that follow European distributions of haplogroups H, I, J, K, T U, V, W, and X (in rank order), the cases are more likely than not to be in Haplogroup J.

One study [49] presented the inverse association. This study required a nuanced interpretation. In their work, Ross et al. [49] presented two sub-groups of Haplogroup J. However, their RFLP approach did not include diagnostic SNPs for the common Haplogroup J sub-group nomenclature. Therefore, it was possible that Haplogroup J1 might have a different effect than Haplogroup J2. It was also possible that other studies have a preponderance of the sub -haplogroup favoring longevity while the Ross study had an inverse proportion. Future studies should incorporate a standard set of SNPs and scoring methods to include sub-group assessment.

The usual suspects of meta-analysis limitations were unavoidable in this work [96]. The present work was particularly limited in three ways: 1) lack of uniformity in case definition (age range 80-100) and control definitions (youngest age range from 18-20 and oldest age range from 45-70), 2) genotyping methods, and 3) scoring methods. The case/control study design of the Haplogroup J and longevity association was sub-optimal. Cases that navigated environmental exposures and achieved an old age were different on many other factors, not just their inherited haplogroup. A more appropriate study design was one in which a time-to-event analytic strategy could be employed [173]. The modified kin-cohort method presented by Castri or Beckstead [42, 107] represents innovative method of achieving these sample sizes and analytic method. The most appropriate approach to studying the association in further studies might involve an isolated population such as the Amish using a kin-cohort method. If kin-

cohort methods were to be used, the impact of control age differences would be mitigated between studies.

CHAPTER 4: MITOCHONDRIAL GENETIC EPIDEMIOLOGY OF LIFESPAN IN A POPULATION ISOLATE

4.1 Abstract

Previous studies suggest that mtDNA Haplogroup J is more prevalent among older people than younger people, especially in European populations. However these studies could not establish whether Haplogroup J carriers actually live longer than other carriers because those studies use a sub-optimal study design in which older people lack contemporaneity with the younger people being studied. Sub-optimal study design of prior work also prevented time-to-event analysis and could not control for autosomal genetics or environmental factors. We sought to test the Haplogroup J association in the context of genealogic records of a population isolate. Within this context our work should allow for a more suitable analytic strategy, offer contemporaneity between study groups and lend some control for other genetic and environmental factors.

This work is presented in the context of a broader cooperative of studies across generations (CoSAGE). CoSAGE is a community setting within which a large extended family resides. Most members of the community are descendants of the original settlers of the area. The extended family has a pedigree database that serves as a sampling frame within which to study historical age-at-death phenotype patterns. Community members who participate in CoSAGE serve as the source of genetic records of their distant ancestors documented in the pedigree database. We determined the mtDNA haplogroup of study participants and imputed the haplogroups for maternal ancestors and relatives in the pedigree database. We used age at death and proportion of octogenarians to assess extended lifespan.

We recruited 118 participants for the study. However, due to the inter-relatedness of the community members only 32 distinct maternal lines were represented among the participants. We found eight of the nine common European haplogroups in the 31 samples that were successfully genotyped. One maternal line was a Haplogroup J carrier. This single line had 80 mitochondrial relatives in the database, 71 of whom were born in the study time frame. None of the measures tested supported our hypothesis of extended lifespan in the haplogroup J lines. However we observed that male carriers and a mean age at death of 72 whereas female carriers died at 69 on average (p-value=0.254).

4.2 Introduction

Mitochondria power cellular function through the synthesis of ATP from consumed calories. They possess their own non-nuclear DNA outside the nucleus. Byproducts of ATP synthesis are highly mutagenic to mitochondrial DNA (mtDNA). Modern theories of aging suggest that as cells age the mitochondria suffer more mutagenic events and thereby a decline in ATP synthesis and thus cellular function capacity.

Some studies of European mtDNA haplogroups and longevity show more Haplogroup J carriers among older people than younger people. A recent meta-analysis of the association, showed that the association is present only in studies of European populations. This meta-analysis finds among European populations older people are 1.47 (CI_{95} 1.04-2.06) times more likely to be Haplogroup J carriers than younger people. The meta-analysis included a study among Irish centenarians that demonstrated that the positive benefit to longevity might be isolated to one sup-group of

the population. These finding beg the question of whether Haplogroup J carriers have increased lifespan. In this work, we examine the Haplogroup J and longevity association in a time-to-event or lifespan context, building on prior cross-sectional studies of longevity.

4.3 Background

4.3.1.The community and population setting

We have presented evidence supporting the identification of a population isolate in Michigan (Chapter 2). A large extended family forms the nucleus of a community that spreads across 90 square mile, encompassing three municipalities. The family and communities are partners in the Community-Based Cooperative for Studies Across Generations (CoSAGE). The present study of Mitochondrial DNA Patterns and Lifespan (MDPL) builds on the CoSAGE community cooperative. Our prior work demonstrates that this community and extended family possess the qualities of a population isolate. A preponderance of the individuals in the region are descendants from a small group of who settled the land in the mid-1830s. The group experienced a population growth rate not dissimilar from the United States and is now estimated to be near 14,000. The group came to the attention of MSU researchers because of the frequency of hearing loss. The genetic causes of hearing loss in the extended is linked to two well-characters mutations. [77, 78]. One unique feature (among many) of this community is an extended pedigree database. The database enumerates over 30,000 individuals most of whom descend from the 1830s settlers/founders. The present work used living individuals of the CoSAGE cooperative and the sampling space created by the pedigree

database (pedigree space) to study whether Haplogroup J carriers live longer than other carriers of other haplogroups.

4.4 Methods

4.4.1 Pedigree Data

Genealogical data was collected as described previously [174]. To increase the genealogical data, including death dates, we searched www.familysearch.org™ for new death dates. To minimize potential bias of completeness and validity, we established a range of continuous birth decades during which the individuals born have a high chance of being dead and with a high proportion of death dates are recorded. Using a chi-squared likelihood ratio test in SAS™ we analyzed whether the proportion of individuals missing death dates was significantly different between males and females and across mtDNA haplogroups. We also decided that individuals who died after the age of 100 were not eligible to be included in the indirect cohort due to blunt death-date invalidity. Using these criteria, we established a birth-cohort-decade series that was minimally biased by invalid or incomplete death dates from genealogic sources.

We wrote a SAS™ script to find records with potentially invalid parental ties to other individuals in the pedigree database. The script checks for records with wrong-sex parents, parents born after children, and for children born more than one year after a parent's death date. Individuals with questionable parental ties and their parents were ineligible for this analysis.

4.4.2 Recruitment

Participants in the umbrella CoSAGE registry were recruited through word-of-mouth referrals, informational sessions, community newspaper advertisements, community bulletin boards, church bulletins, health fairs, community events, an internet web page, Facebook.com, and routine “walk-in” information sessions at the project office. CoSAGE sponsors health fairs and community events that allowed interested individuals an opportunity to complete a contact request card. A member of the study team scheduled a meeting to inform the potential participant about the project and review and discuss the informed consent process. The informed consent sessions occur in the Project Office or the participant’s home. Consenting participants completed self-report questionnaires and a CoSAGE team member collected an anthropometric assessment, a whole blood sample (by peripheral phlebotomy) or cheek cell (by saliva collection kit). Consenting participants are asked to supply the names of their parents and grandparents to merge them into the pedigree database.

CoSAGE participants were mailed an MDPL study packet that included an introductory letter, a consent document approved by the Michigan State University Institutional Review Board, a Genetic Information Non-Disclosure Act (GINA) pamphlet, and a postage-paid return envelope. The consent form sought permission to use the participant’s banked DNA sample and genealogy data for this study. CoSAGE participants who consented to the MDPL comprised a direct cohort for this analysis. Deceased relatives of direct cohort identified in the genealogy database form a larger, indirect cohort sampled from pedigree space. The lifespan phenotype of the indirect cohort was calculated from the birth year and death year as recorded and validated in

the database. Lifespan phenotype of the indirect cohort is the principal variable for this analysis.

4.4.4 DNA isolation and genotyping

CoSAGE participants donated venous blood or saliva samples. DNA was isolated from saliva samples using Oragene™ DNA Self-Collection Kit as recommended (DNA Genotek Inc., Ottawa ON, Canada). DNA was isolated from venous blood samples using the Quick-Gene -810 as recommended (AutoGen, Holliston, MA, USA).

Mitochondrial haplogroups were based on DNA sequence analysis of the non-coding region of mtDNA15998-mt326. Aliquots of banked DNA samples from MDPL volunteers were amplified with a forward primer (mt15596 5'-AAG TCT TTA ACT CCA CCA TTA GCA-3') and a reverse primer (mt326 5'- GGC TGG TGT TAG GGT TCT TTG- 3'). The PCR amplification condition was an initial temperature elevation to 94^0 held for 3 minutes, followed by 25 cycles of 94^0 for 30 seconds, 60^0 for 3 minutes and 72^0 for 1 minute, and a final cycle at 72^0 for 1 minute and held at 15^0.

Sequencing aliquots of each PCR was prepared with the primers listed above and an internal primer (mt16523 5'-ACT TCA GGG CCA TAA AGC CT-3'). PCR products were sequenced using an ABI 3730 Genetic Analyzer (Applied Biosysems, Foster City, CA, USA).

4.4.5 Bioinformatics Methods

The files of sequence data were imported into Sequencher V2.1.6 (Gene Codes, Ann Arbor, MI, USA). DNA sequences were trimmed and ambiguous calls were reviewed by two investigators independently. Trimmed sequences were aligned with the

revised Cambridge Reference Sequence (rCRS). [20] DNA sequence variants from rCRS were exported and processed into standard SNP nomenclature using a script written in SAS v9.1.3 ™ (SAS Corp., Cary, NC, USA) . The SNP variants were processed into a file for import into HaploGrep™. [40] Haplogrep™-scored haplogroup values were exported and merged with the pedigree data for statistical analysis using SAS™.

Two scripts were programmed using Oracle™ PL/SQL (Oracle Corp., Redwood, CA, USA). One to recursively query the pedigree database to discover the most distant maternal ancestor (or “Eve”) for each member of the direct cohort. A second script was written to discover all the descendants of each “Eve” who share her mitochondrial DNA (Mitochondrial Relatives). The mitochondrial relatives are potential members of the indirect cohort. The haplogroup assigned to each member of the direct cohort member was ascribed to all their mitochondrial relatives.

4.4.6 Analytic Methods

Ate at death (or lifespan) was calculated as death year minus birth year. Age at death was analyzed as a quantitative trait (mean, median, and time to death) using a Student’s t-test and proportional hazards model. Secondarily age at death was analyzed as a categorical trait (before or after 80 years) using chi statistics from a contingency table and an odds ratio. Each statistical test contrasted the ascribed Haplogroup J with all other haplogroups combined then stratified by sex and across time periods. All statistical analyses were conducted in SAS™.

4.5 Results

4.5.1 Pedigree Data

This work presents the pedigree database with changes and additions to the database up to April 1, 2013. The database contained 37,315 individuals spanning birth years from 1511 to 2013. We discovered 9,055 “Eves” in the pedigree database. Of these female pedigree founders, 7,685/9,055 (86%) have shallow contributions to the pedigree database consisting of themselves and only one additional generation of descendants. The remaining 1,750/9,055 (14%) female founders contributed more than two generations and accounted for 29,850/37,315 (80%) of the individuals in the pedigree.

The majority of individuals in the pedigree database were born after 1930 (Figure 4-1). There is little difference in the ratio of males to females in any birth decade (Figure 4-1). Individuals born between 1810 and 1909 had the highest proportion of recorded death dates (Figure 4-2). In the 100 years between 1810 and 1909, 11,989 individuals were born and 9,013 (76%) have a recorded death date. It is highly unlikely that any individual born in this epoch still lives. Using this 100 year window we maximize the number of records to calculate age at death (or lifespan) while minimizing potential bias due to missing death dates.

Figure 4-1 Pedigree Individuals by Birth Decade and Sex (potential participants of indirect cohort).

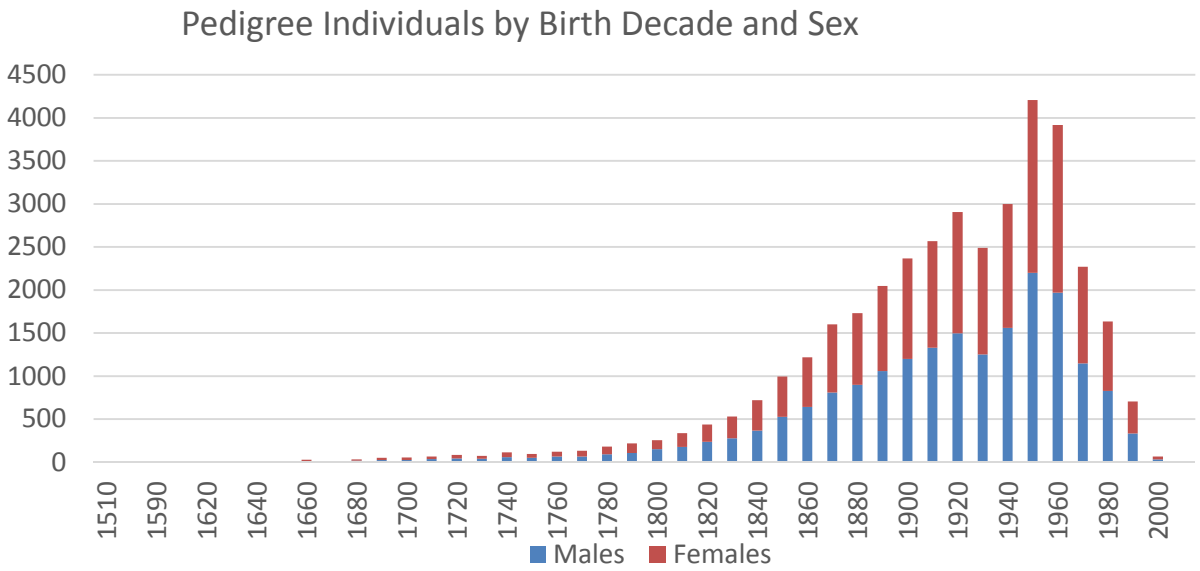
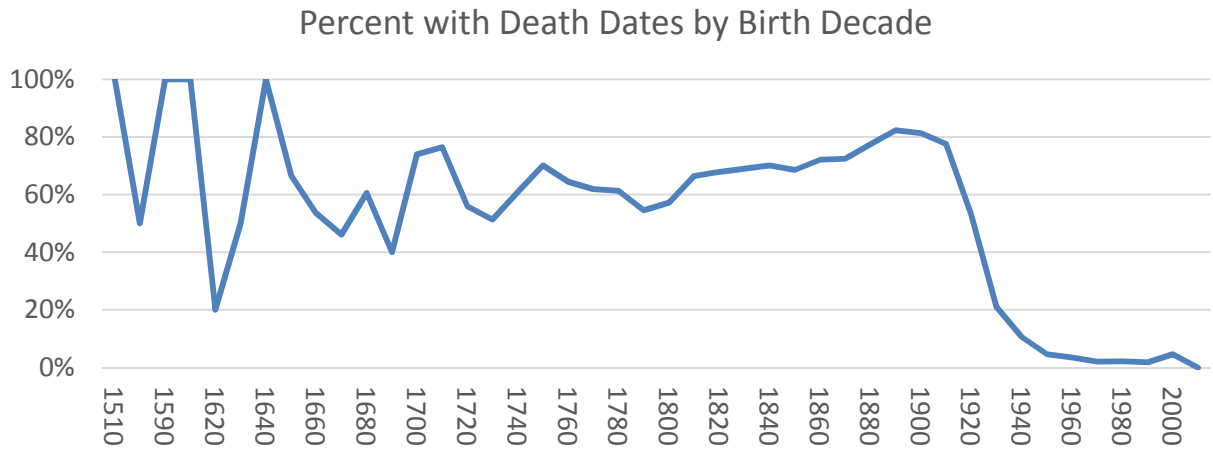


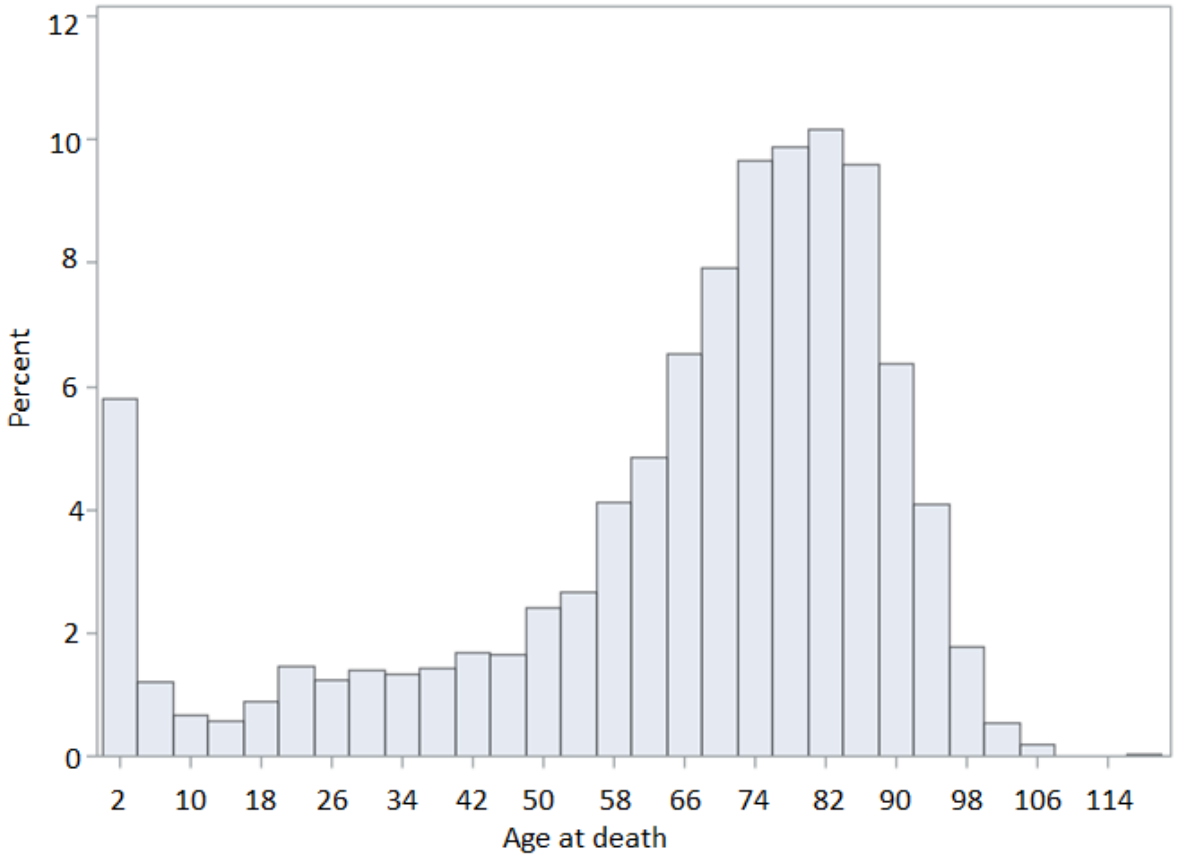
Figure 4-2: Percent of individuals with death date present by birth decade.



Of the 11,989 individuals born in the study time frame, 5,792 (49%) were female, 6,116 (51%) were male, and 80 were of unknown sex nor was there a record of them as a parent. Individuals without a recorded sex were ineligible for the indirect cohort. Significantly more males 4,741/6,116 (78%) have a death date recorded than females 4,272/5,792 (74%) (p-value < 0.001).

Of the 9,013 individuals born between 1810 and 1909 who have a recorded death date, the median age at death was 74 years. There are 831 deaths before age 20; and 68 deaths after 100 (Figure 4-3). We removed the childhood deaths (before 20) from this study of adulthood lifespan. We also removed the deaths over 120 because of poor data-quality suspicions. The potential pool of individuals from pedigree space eligible for the indirect cohort was 8,172.

Figure 4-3: Distribution of age at death among pedigree members born between 1810 and 1909 who have a death date documented.



4.5.2 MDPL Recruitment

As of February 28, 2013, CoSAGE had 250 registry participants, each of whom was sent an MDPL study packet. By June 1, 2013, 118 participants returned signed consent forms. Of these participants, 101/118 were members of the extended pedigree, and 98 had available banked DNA samples.

Among the 98 eligible participants, there were 46 distinct maternal lineages (Eves). Of these lines, 32 were deeper than two generations. There were 18 deep maternal lines represented by more than one descendant. These lines could be used to validate

genotyping data. When multiple descendants of the same “Eve” were available, we chose the oldest participant to be the haplogroup-generating individual and the youngest to be the replicate sample. The direct cohort was 32 individuals (19 women, 13 men, median age = 66.5, age range 34-85).

4.5.3 Indirect cohort

To assemble the indirect cohort, we used the Eve script to discover the maternal lines of the direct cohort. We then discovered all mitochondrial relatives of the “Eves” represented in the direct cohort. With this approach, the indirect cohort consisted of the 1,490 descendants born between 1810 and 1909 and for whom a death date and sex were recorded. The indirect cohort was equally split between males and females. The sex of 16 individuals was unknown. The average age at death was 72 (SD +/- 17) years old, with significant differences between men and women (70 +/- 16 versus 73 +/- 18; p-value <0.001). It is well-established that lifespan has increased over time. This change was evident in our indirect cohort. Individuals born in the first quarter-century died at an average age of 67 (+/-19) while those born in the last quarter century died at 73 (+/-16). (Figure 4-4) The time-to-death curve showed that past the age of 55 there are more women alive than men (Figure 4-5) between 1810 and 1909.

Figure 4-4 Age at death of the Indirect cohort presented in various categories.

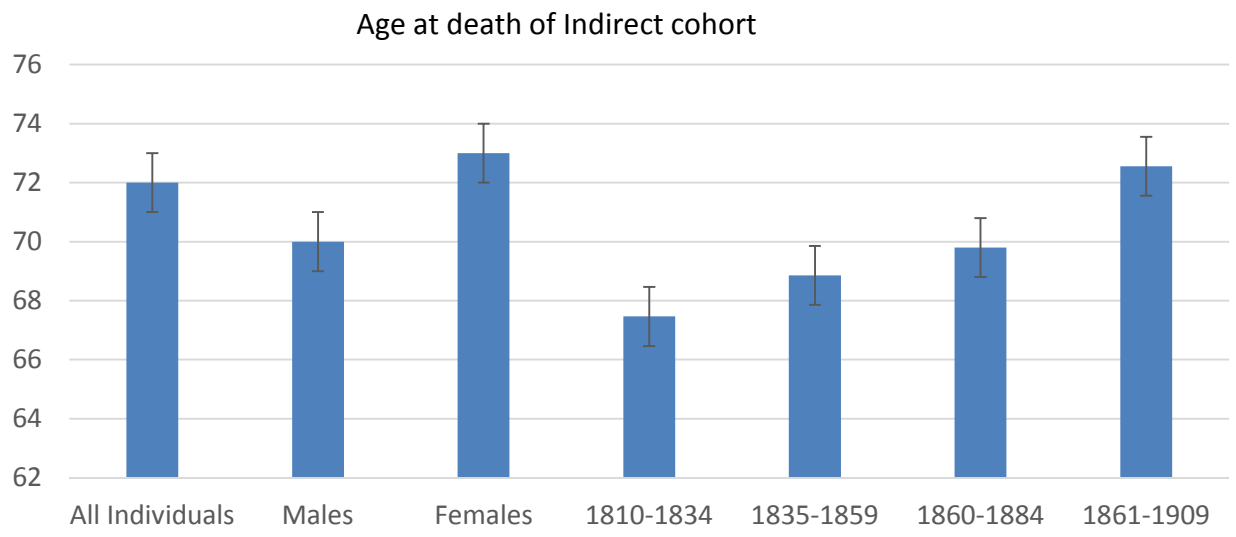
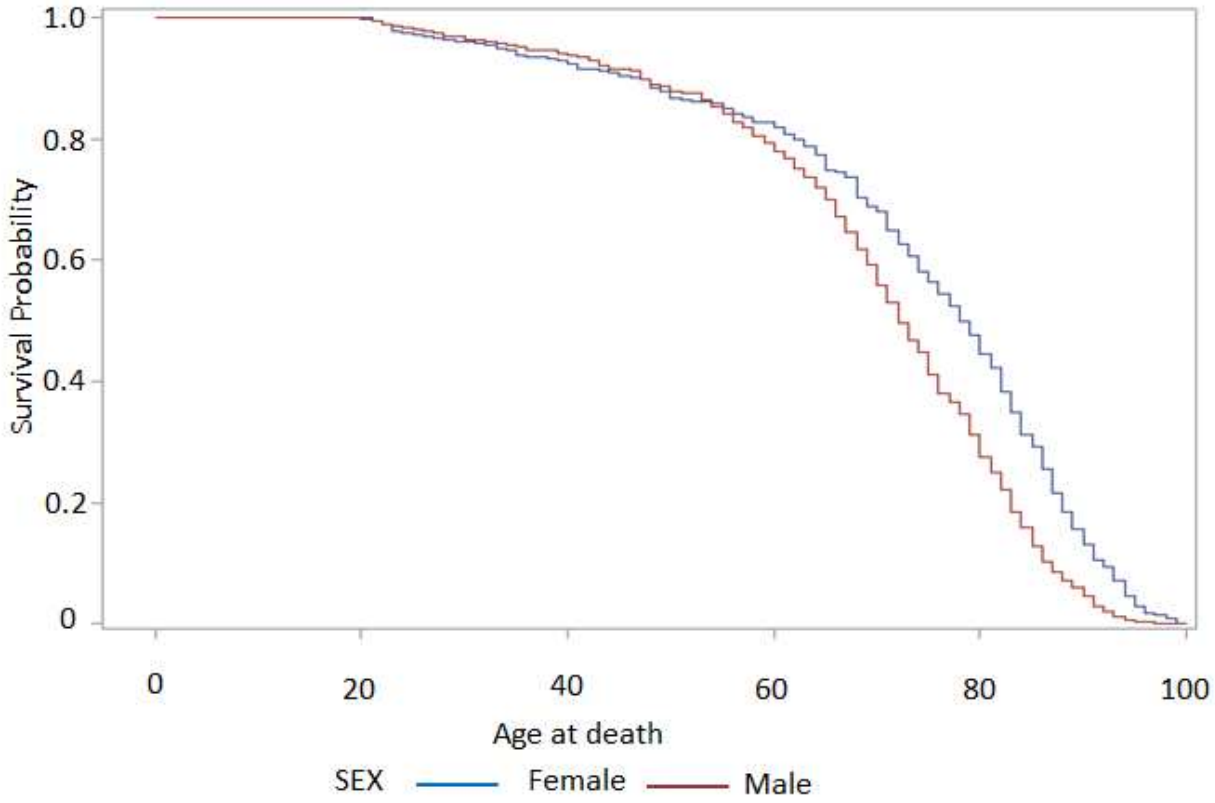


Figure 4-5: Time-to-death plot in men vs. women in indirect cohort (p-value <0.001) Hazard ratio = 1.517. For interpretation of the references to color in this and all other figures, the reader is referred to the electronic version of this dissertation



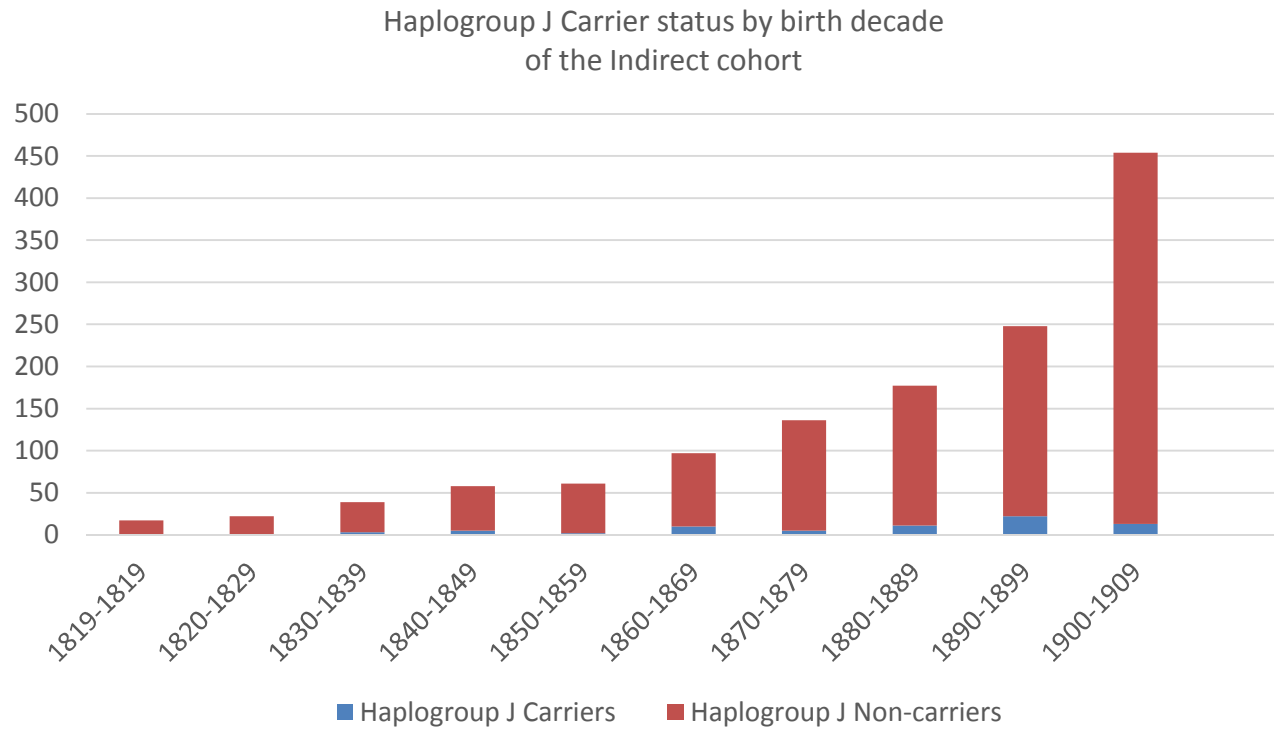
4.5.4 Haplogroups

We successfully genotyped 31/32 samples of the direct cohort and 14 of the validation samples.(Table 4-1). All replicate samples matched the haplogroup assessments of the primary samples. The distribution of the eight haplogroups followed an expected pattern for a European population [175]. Two samples were mt16189C within poly-c microsatellite. These samples required a reverse sequence to fully ascribe their haplogroup. However, since our statistical contrasts are Haplogroup J versus all others and this haplogroup is mt16189T, the reverse sequencing was not done.

The direct cohort had only one Haplogroup J thus one “Eve”. This single Eve had 71 mitochondrial relatives born in the study century. The Haplogroup J carriers in the indirect cohort were not born at regular intervals across the study century. (Figure 4-6)

Table 4-3: Distribution of Haplogroups among the 31 genotyped samples		
Haplo-sub-group	**Number in Direct cohort**	**Macro Group**
Mt16189C	2	16189C
H1a1	2	H
H1e1a1	1	
H1e1a4	1	
H2a2a	8	
H4a1a+195	1	
H6	2	
I	1	I
J1c2	1	J
K	1	K
K2b2	1	
T1a	1	T
T1a1'3	1	
T2	2	
T2a1b	1	
T2b	1	
U4a2	1	U
U5a1	1	
U7	2	
Total	31	

Figure 4-6: Haplogroup J carriers status by birth decade (indirect cohort). For interpretation of the references to color in this and all other figures, the reader is referred to the electronic version of this dissertation



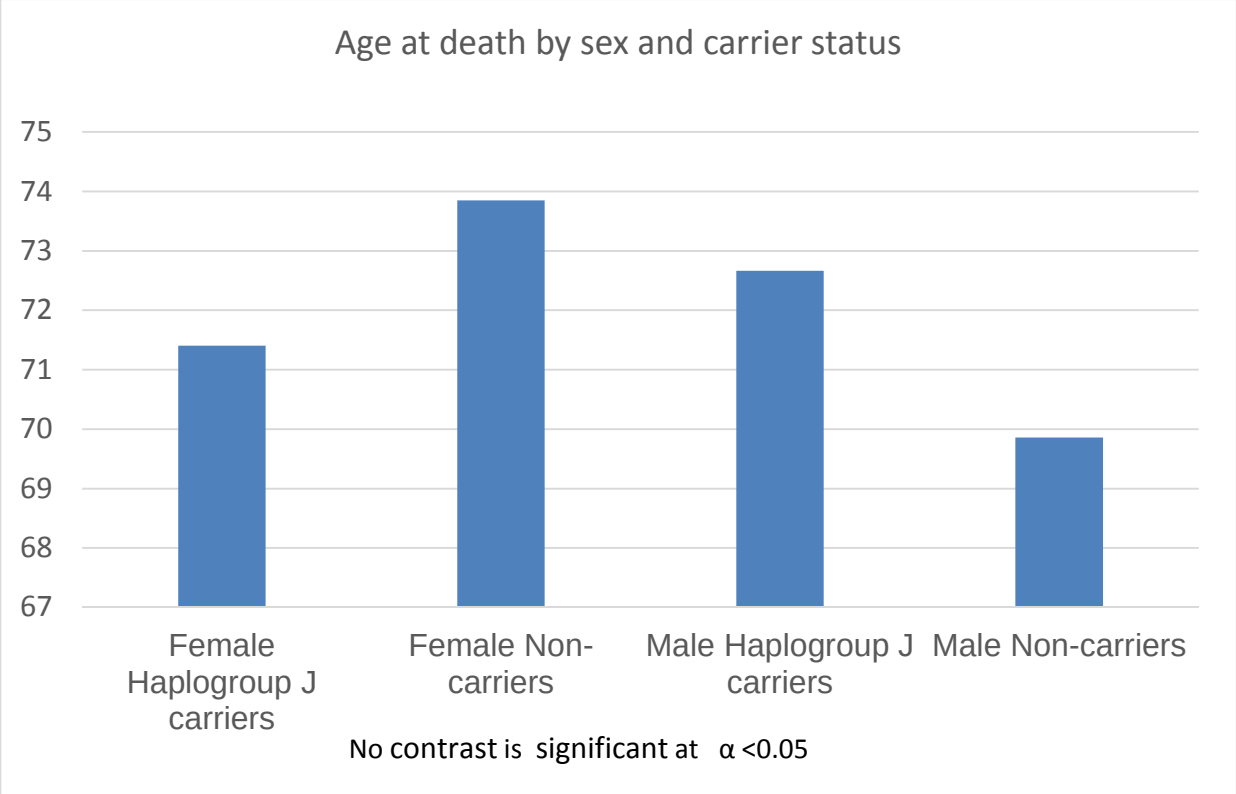
4.5.5 Lifespan of Haplogroup J carriers in the indirect cohort

In the indirect cohort, we observed no significant difference in the proportion of individuals who are missing death date (7/87) 8% of Haplogroup J and 217/1,627 (13%) of all other haplogroups (p-value=0.1537). This result indicated the absence of potential differential bias of missing death dates with respect to Haplogroup J carrier status.

The mean age at death among the Haplogroup J carriers was 72.1 (+/- 17) and among the non-J carriers it was 71.8 (+/- 17) (p-value=0.875). The male Haplogroup J carriers had an average age at death of 73 (+/- 13) versus 69 (+/- 16) among the non-carriers. In stark contrast, the women carriers had a younger age at death (71.4 +/- 21) than non-carriers (73.8 +/- 18); however neither difference was significant. (Figure 4.7) The uneven distribution of Haplogroup J carriers across the study century prevented an age at death analysis by carrier status across the quarter centuries

.

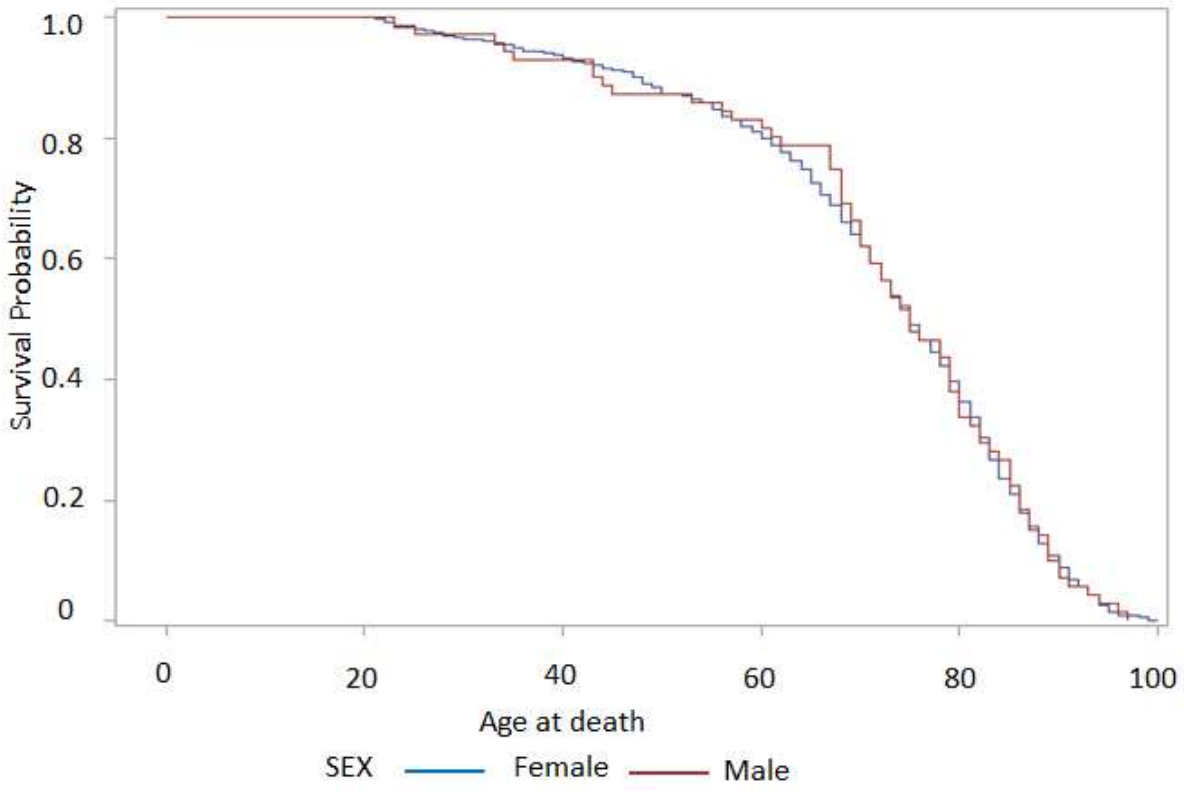
Figure 4-7 Age at death by carrier status and sex.



Time-to-death curve showed that at no age between 20 and 100 was a greater proportion of Haplogroup J carriers alive than non-J carriers. (Figure 4-8) (Hazard ratio = 1.004; p-value=0.954). However, when stratified by sex, there was a non-significant difference in age at death. As foreshadowed by the differences in mean age at death, the time-to-death curves for men showed that for most ages between 20 and 100 there were more or just as many carriers alive until after 85. (Figure 4-9). This curve is the inverse among the women, where at most ages, there were fewer carriers alive than non-carriers until after 83. However, these differences are non-significant. When displayed directly among carriers only, the time-to-death curve showed that before the age of 65 more of the haplogroup J carrier men were alive than women. The inverse

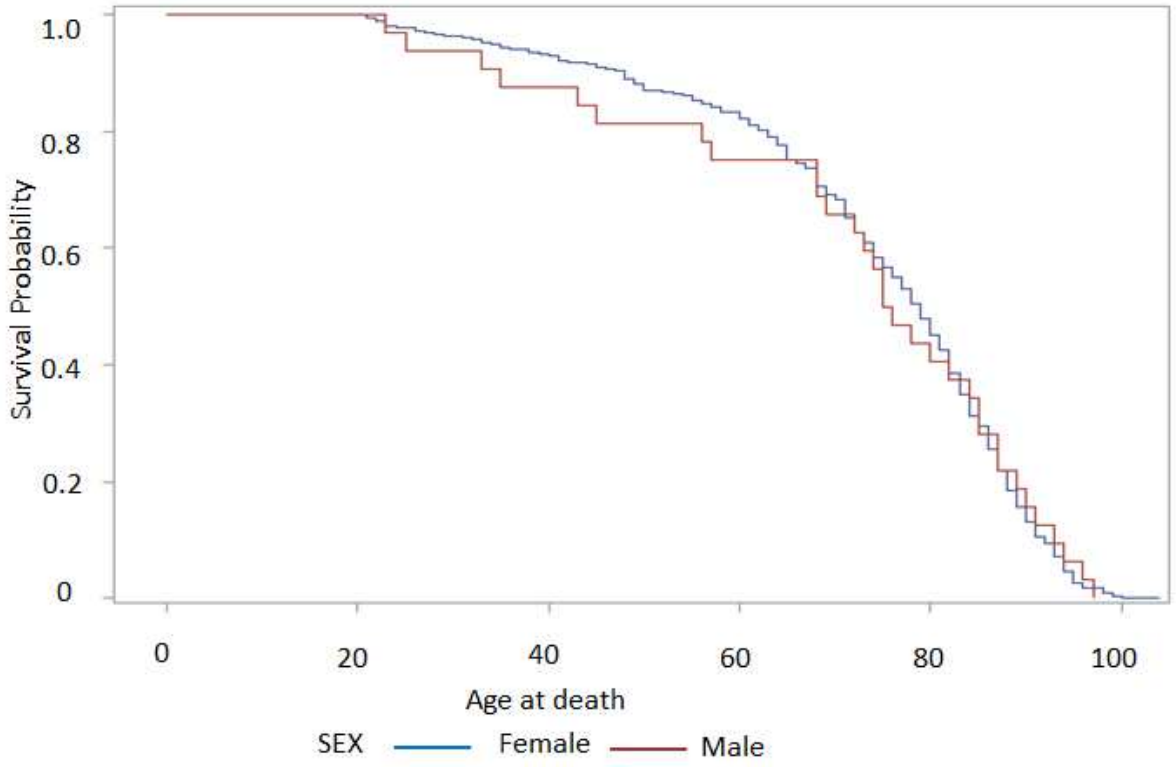
was true after the age of 65 when more women than men were alive (among the carriers); however the difference is not significant.

Figure 4-8: Time-to-death curve of Haplogroup J carriers vs. non-j carriers. Hazard ratio 1.005).



.

Figure 4-9 Time-to-death curve contrasting Haplogroup J carriers and non-carriers among women (n=31).



Among the Haplogroup J carriers 27/71 (38%) and 493/1,244 (40%) of the non-carrier individuals died as octogenarians. Conversely, the proportion of carriers who died before 20 was 16/87 (18%) and among non-carriers 383/1,627 (24%). There was no significant difference in the proportion of the indirect cohort who died as children (p-value=0.2682) or as octogenarians (p-value=0.7883)

4.6 Discussion

Using pedigree data and three metrics of lifespan (mean age at death, time to death and proportion of octogenarians), we found no difference in lifespan between Haplogroup J carriers and non-J carriers born between 1810 and 1909. Although 24%

of the possible indirect cohort was missing a death date, we found no differential bias of death date reporting between Haplogroup J carriers and others.

Although not significant, we found that Haplogroup J carrier men diee older than non-carrier men and Haplogroup J carrier women died younger than non-carrier women. Additionally we found that among carriers before the age 65 more men were alive than women. This trend reversed after 65.

These trends were similar to De Benedictis in 1998 [45] who presented more Haplogroup J carriers among the male centenarians from Northern Italy and no appreciable differences in rates in females or among southern Italians. We showed in Chapter 3 that among definably European populations that older cases were more likely to be Haplogroup J carriers than younger controls. However inconsistent presentations of demographic data precluded pooled assessment of the De Benedictis male-female dimorphism as hinted in our data. These time to death findings offer new insight into how more Haplogroup J carriers were present in groups of older cases and probably more pronounced among males.

For argument sake, please consider that if women live longer than men, then a case-control study of old people as cases would include more women than men. If the study were to use younger controls say younger than 65 there might be fewer haplogroup J females available as controls. Figure 4-10 demonstrates that beyond the age of 65, there were more women Haplogroup J carriers alive than men; however younger than 65 the opposite was true and around 65 the rates were similar. I postulate that the choice of age of controls, in the light of the evidence of Figure 4-11, created

artificially high or inconclusive odds-ratio estimates in cross-sectional studies. If a study chooses controls considerably younger than 65 the odds-ratio would present carrier status as a risk factor. Likewise, if the study chooses controls around 65, the odds ratio would be equivocal because the time-to-death curves cross each other at that age. Biology and genetics, aside, if Figure 4.10 is the true phenomena, then the choice of age of controls in cross-sectional studies creates an artifact in the true effect measure.

Figure 4-10 Time-to-death curve contrasting Haplogroup J carriers and non-carriers among men (n=39).

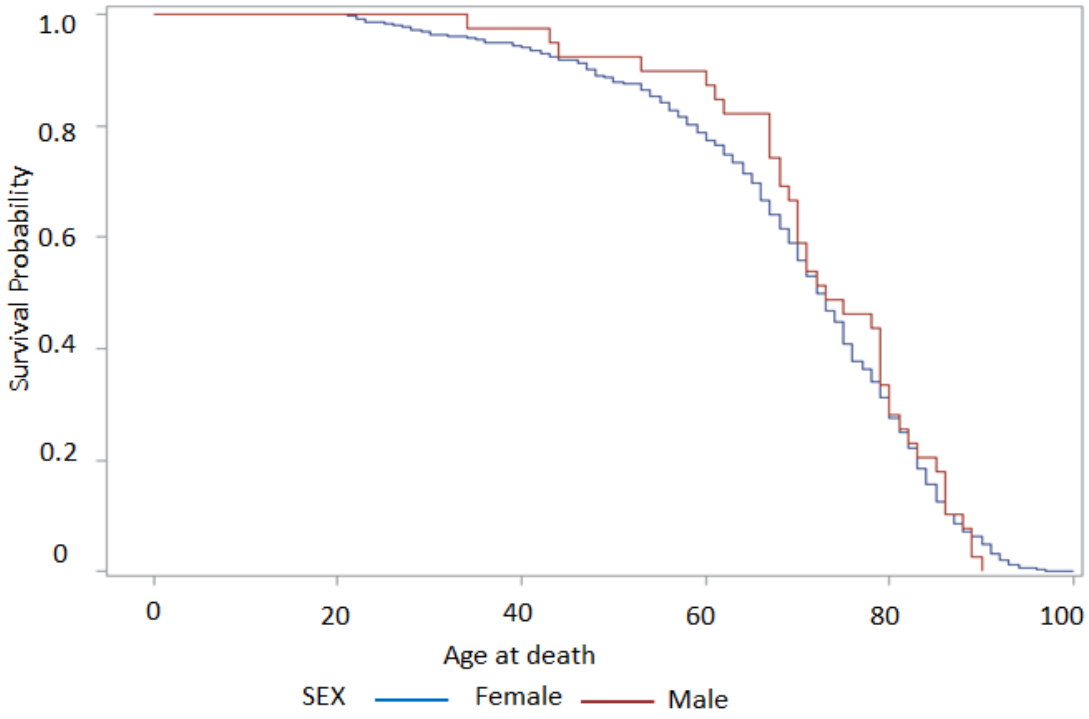
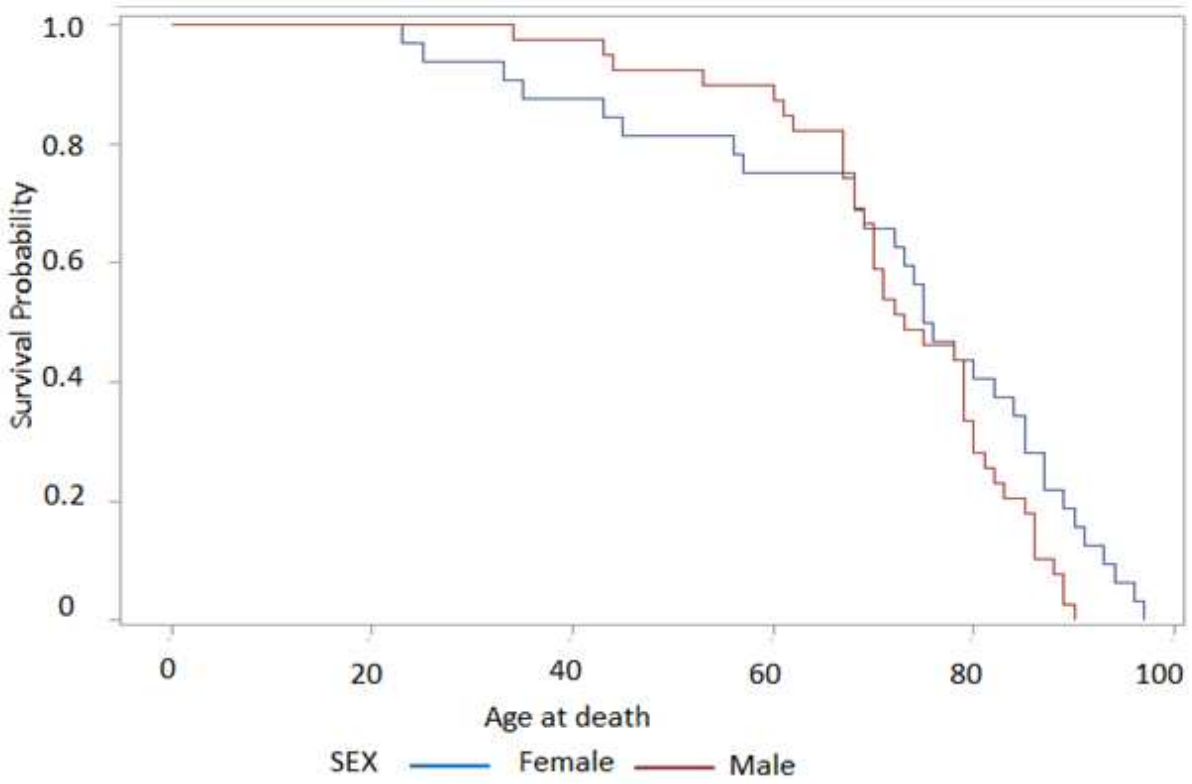


Figure 4-11: Time-to-death curve of Haplogroup J carrier men (n=39) versus women (n=31).



This male-female dimorphism has overtones in the epidemiology of other mitochondrial diseases [176] like Leber's hereditary optic neuropathy (LHON), caused by one of several heritable mtDNA mutations. There are fewer female LHON patients than male. Nearly 85% of LHON cases are male. Yet there is no evidence that the causative mtDNA mutation rates differ by sex [177]. This observed male-female dimorphism of the LHON pathologic phenotype is counter-intuitive to the sex-specific benefit to Haplogroup J and the longevity benefit. If one carries a LHON mutation it is harmful to be male; whereas our results hint at a benefit of being male if you are a Haplogroup J carrier.

One plausible biologic explanation of the dimorphism pivots on the fact that mtDNA is maternally inherited and encodes a small proportion (13 of 86) of the ETC polypeptide sub-units. This leaves 75 of 86 polypeptides encoded in nuclear genome. The 75 nuclear encoded ETC polypeptides traverse a complex path to find their way into a mitochondrion. Most of the path is regulated by other nuclear encoded proteins. Epistasis between paternally inherited alleles of mitochondrial proteins might play a role in the male-female dimorphism. [10] [178]

Despite the temptation, Occam's razor, the lack of statistical significance and the solitary maternal Haplogroup J line in the MDPL caution against over interpretation.

CHAPTER 5: SUMMARY AND FUTURE DEVELOPMENT

There is ample evidence that mitochondria play a role in cellular function and that loss of mitochondrial function is a hallmark of the aging phenotype. We demonstrated that in cross-sectional studies among definably European populations older people are more likely by half to be carriers of Haplogroup J than other haplogroups. Historically, studies of maternal contributions to longevity or lifespan have used genealogical databases. In this work we presented a study of the purported Haplogroup J association and lifespan in a genealogical context, but the association was not found. This chapter presents questions that future investigators might consider.

The principal weakness of this study was incompleteness. Despite our efforts at covering the full pedigree within CoSAGE, the MDPL recruitment effort was only able to cover 31 of an estimated 1,099 deep maternal lines in the pedigree. Due to the MSU Institutional Review Board's position, we were not able to directly approach members of the community and CoSAGE participants who would have given better coverage of the pedigree. Further, given the high degree of relatedness among the direct participants it is unlikely that all possible maternal lines are represented. In fact, in MDPL, 67% of the consented participants were unusable because they were duplicates within a maternal line. Future studies of the Haplogroup J/Lifespan association should wait until the CoSAGE recruitment has a more complete documentable coverage of the pedigree. A future study might be successful were it not saddled with a requirement for direct consent for an indirect cohort. An improved approach would be to use the entire CoSAGE registry plus other studies of this population and direct approach to individuals who map to maternal lines covering more of the database. However, even with the

improved approach, the observation that 1 of 32 samples in the direct cohort was Haplogroup J suggests that the CoSAGE population might be too isolated to contain sufficient Haplogroup J Eves to study the association.

Samuels et al.[94] offers evidence that the cross-sectional approach of haplogroup studies must apply a sample size inflation factor when testing one group in the background of population. The inflation factor increases the sample size required based on the odds ratio effect size estimate and the number of haplogroups in the background population. The literature is silent on whether the time-to-event studies would use exactly the same inflation. We postulate that a different formula would be required. Future studies of the association in the chosen populations should investigate this important methodical venue and make necessary adjustments.

Another future effort might be to assemble the original data behind the studies presented in Chapter 2 to assess whether the male-female dimorphism exists in all studies and was just over-looked. Reassembling these source data might also assess whether the pooled odds ratio is affected by the age or sex of the controls as posited in Chapter 4.

It seems clear that studying this association is best in an isolated population. Posing the question in the Amish or the population of Iceland would be a logical next step. However, these populations might also be too isolated to contain sufficient Haplogroup J Eves and mitochondrial relatives to satisfy Samuel's inflation factor .

If the dataset were assembled of an isolated population with multiple maternal lines of Haplogroup J carriers, it would be prudent to build a model in which the f-value

or degree of relatedness of an individual's parents could also be tested. Testing the significance and effect of relatedness would help tease apart the nuclear components from the mitochondrial and lend evidence towards possible epistasis.

LITERATURE CITED

LITERATURE CITED

1. Holmes, O.W., *Over the teacups*. 1891.

2. Beeton M, P.K., *Data for the problem of evolution of man. III. A first study of the inheritance of longevity and the selective death-rate in man.* Proc. Roy. Soc. London, 1899. **65**: p. 290-302.

3. Beeton M, P.K., *On the inheritance of the Duratinon of Life, and on the Intensity of Natural Selection in Man.* Biometrika, 1901. **1**(1): p. 50-89.

4. Jalavisto, E., *Inheritance of longevity according to Finnish and Swedish Geneaolgies.* Annales medicinae internae Fenniae, 1951. **40**(263-274): p. 263.

5. Korpelainen, H., *Genetic maternal effects on human life span through the inheritance of mitochondrial DNA.* Hum Hered, 1999. **49**(4): p. 183-5.

6. Mayer, P.J., *Inheritance of longevity evinces no secular trend among members of six New England families born 1650–1874.* American Journal of Human Biology, 1991. **3**(1): p. 49.

7. Wallace, B.C., C.H. Schmid, J. Lau, and T.A. Trikalinos, *Meta-Analyst: software for meta-analysis of binary, continuous and diagnostic data.* BMC Med Res Methodol, 2009. **9**: p. 80.

8. Wallace, D.C., *A mitochondrial paradigm of metabolic and degenerative diseases, aging, and cancer: a dawn for evolutionary medicine.* Annu Rev Genet, 2005. **39**: p. 359-407.

9. Wallace, D.C., M.D. Brown, and M.T. Lott, *Mitochondrial DNA variation in human evolution and disease.* Gene, 1999. **238**(1): p. 211-30.

10. Wallace, D.C., E. Ruiz-Pesini, and D. Mishmar, *mtDNA variation, climatic adaptation, degenerative diseases, and longevity.* Cold Spring Harb Symp Quant Biol, 2003. **68**: p. 479-86.

11. Bhagavan, N.V., *Medical Biochemistry*. 4th ed. 2001, San Diego, CA: Harcourt Academic Press.

12. Turrens, J.F., *Mitochondrial formation of reactive oxygen species.* J Physiol, 2003. **552**(Pt 2): p. 335-44.

13. Nass, S. and M.M. Nass, *Intramitochondrial Fibers with DNA Characteristics. Ii. Enzymatic and Other Hydrolytic Treatments.* J Cell Biol, 1963. **19**: p. 613-29.

14. Anderson, S., A.T. Bankier, B.G. Barrell, M.H. de Bruijn, A.R. Coulson, J. Drouin, I.C. Eperon, D.P. Nierlich, B.A. Roe, F. Sanger, P.H. Schreier, A.J. Smith, R. Staden, and I.G. Young, *Sequence and organization of the human mitochondrial genome.* Nature, 1981. **290**(5806): p. 457-65.

15. Lee, H.Y., I. Song, E. Ha, S.B. Cho, W.I. Yang, and K.J. Shin, *mtDNAmanager: a Web-based tool for the management and quality analysis of mitochondrial DNA control-region sequences.* BMC Bioinformatics, 2008. **9**: p. 483.

16. Cann, R.L., W.M. Brown, and A.C. Wilson, *Polymorphic sites and the mechanism of evolution in human mitochondrial DNA.* Genetics, 1984. **106**(3): p. 479-99.

17. Cann, R.L., M. Stoneking, and A.C. Wilson, *Mitochondrial DNA and human evolution.* Nature, 1987. **325**(6099): p. 31-6.

18. Macaulay, V., M. Richards, E. Hickey, E. Vega, F. Cruciani, V. Guida, R. Scozzari, B. Bonne-Tamir, B. Sykes, and A. Torroni, *The emerging tree of West Eurasian mtDNAs: a synthesis of control-region sequences and RFLPs.* Am J Hum Genet, 1999. **64**(1): p. 232-49.

19. Torroni, A., M. Richards, V. Macaulay, P. Forster, R. Villems, S. Norby, M.L. Savontaus, K. Huoponen, R. Scozzari, and H.J. Bandelt, *mtDNA haplogroups and frequency patterns in Europe.* Am J Hum Genet, 2000. **66**(3): p. 1173-7.

20. Andrews, R.M., I. Kubacka, P.F. Chinnery, R.N. Lightowlers, D.M. Turnbull, and N. Howell, *Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA.* Nat Genet, 1999. **23**(2): p. 147.

21. Taanman, J.W., *The mitochondrial genome: structure, transcription, translation and replication.* Biochim Biophys Acta, 1999. **1410**(2): p. 103-23.

22. Kang, D. and N. Hamasaki, *Mitochondrial transcription factor A in the maintenance of mitochondrial DNA: overview of its multiple roles.* Ann N Y Acad Sci, 2005. **1042**: p. 101-8.

23. Kang, D., S.H. Kim, and N. Hamasaki, *Mitochondrial transcription factor A (TFAM): roles in maintenance of mtDNA and cellular functions.* Mitochondrion, 2007. **7**(1-2): p. 39-44.

24. Falkenberg, M., M. Gaspari, A. Rantanen, A. Trifunovic, N.G. Larsson, and C.M. Gustafsson, *Mitochondrial transcription factors B1 and B2 activate transcription of human mtDNA.* Nat Genet, 2002. **31**(3): p. 289-94.

25. Falkenberg, M., N.G. Larsson, and C.M. Gustafsson, *DNA replication and transcription in mammalian mitochondria.* Annu Rev Biochem, 2007. **76**: p. 679-99.

26. Bogenhagen, D.F., D. Rousseau, and S. Burke, *The layered structure of human mitochondrial DNA nucleoids.* J Biol Chem, 2008. **283**(6): p. 3665-75.

27. Bonawitz, N.D., D.A. Clayton, and G.S. Shadel, *Initiation and beyond: multiple functions of the human mitochondrial transcription machinery.* Mol Cell, 2006. **24**(6): p. 813-25.

28. Larsson, N.G., G.S. Barsh, and D.A. Clayton, *Structure and chromosomal localization of the mouse mitochondrial transcription factor A gene (Tfam).* Mamm Genome, 1997. **8**(2): p. 139-40.

29. Moraes, C.T., *What regulates mitochondrial DNA copy number in animal cells?* Trends Genet, 2001. **17**(4): p. 199-205.

30. Clay Montier, L.L., J.J. Deng, and Y. Bai, *Number matters: control of mammalian mitochondrial DNA copy number.* J Genet Genomics, 2009. **36**(3): p. 125-31.

31. Holt, I.J., J. He, C.C. Mao, J.D. Boyd-Kirkup, P. Martinsson, H. Sembongi, A. Reyes, and J.N. Spelbrink, *Mammalian mitochondrial nucleoids: organizing an independently minded genome.* Mitochondrion, 2007. **7**(5): p. 311-21.

32. Jacobs, H.T., S.K. Lehtinen, and J.N. Spelbrink, *No sex please, we're mitochondria: a hypothesis on the somatic unit of inheritance of mammalian mtDNA.* Bioessays, 2000. **22**(6): p. 564-72.

33. Takamatsu, C., S. Umeda, T. Ohsato, T. Ohno, Y. Abe, A. Fukuoh, H. Shinagawa, N. Hamasaki, and D. Kang, *Regulation of mitochondrial D-loops by transcription factor A and single-stranded DNA-binding protein.* EMBO Rep, 2002. **3**(5): p. 451-6.

34. Giles, R.E., H. Blanc, H.M. Cann, and D.C. Wallace, *Maternal inheritance of human mitochondrial DNA.* Proc Natl Acad Sci U S A, 1980. **77**(11): p. 6715-9.

35. Brown, W.M., M. George, Jr., and A.C. Wilson, *Rapid evolution of animal mitochondrial DNA.* Proc Natl Acad Sci U S A, 1979. **76**(4): p. 1967-71.

36. Torroni, A., T.G. Schurr, C.C. Yang, E.J. Szathmary, R.C. Williams, M.S. Schanfield, G.A. Troup, W.C. Knowler, D.N. Lawrence, K.M. Weiss, and et al., *Native American mitochondrial DNA analysis indicates that the Amerind and the Nadene populations were founded by two independent migrations.* Genetics, 1992. **130**(1): p. 153-62.

37. Kong, Q.P., H.J. Bandelt, C. Sun, Y.G. Yao, A. Salas, A. Achilli, C.Y. Wang, L. Zhong, C.L. Zhu, S.F. Wu, A. Torroni, and Y.P. Zhang, *Updating the East Asian mtDNA phylogeny: a prerequisite for the identification of pathogenic mutations.* Hum Mol Genet, 2006. **15**(13): p. 2076-86.

38. van Oven, M. and M. Kayser, *Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation.* Hum Mutat, 2009. **30**(2): p. E386-94.

39. Benn, M., M. Schwartz, B.G. Nordestgaard, and A. Tybjaerg-Hansen, *Mitochondrial haplogroups: ischemic cardiovascular disease, other diseases, mortality, and longevity in the general population.* Circulation, 2008. **117**(19): p. 2492-501.

40. Kloss-Brandstatter, A., D. Pacher, S. Schonherr, H. Weissensteiner, R. Binna, G. Specht, and F. Kronenberg, *HaploGrep: a fast and reliable ntealgorithm for automatic classification of mitochondrial DNA haplogroups.* Hum Mutat, 2011. **32**(1): p. 25-32.

41. Tranah, G.J., T.M. Manini, K.K. Lohman, M.A. Nalls, S. Kritchevsky, A.B. Newman, T.B. Harris, I. Miljkovic, A. Biffi, S.R. Cummings, and Y. Liu, *Mitochondrial DNA variation in human metabolic rate and energy expenditure.* Mitochondrion, 2011.

42. Beckstead, W.A., M.T. Ebbert, M.J. Rowe, and D.A. McClellan, *Evolutionary pressure on mitochondrial cytochrome b is consistent with a role of CytbI7T affecting longevity during caloric restriction.* PLoS One, 2009. **4**(6): p. e5836.

43. Suissa, S., Z. Wang, J. Poole, S. Wittkopp, J. Feder, T.E. Shutt, D.C. Wallace, G.S. Shadel, and D. Mishmar, *Ancient mtDNA genetic variants modulate mtDNA transcription and replication.* PLoS Genet, 2009. **5**(5): p. e1000474.

44. Tanaka, M., J.S. Gong, J. Zhang, M. Yoneda, and K. Yagi, *Mitochondrial genotype associated with longevity.* Lancet, 1998. **351**(9097): p. 185-6.

45. De Benedictis, G., G. Rose, G. Carrieri, M. De Luca, E. Falcone, G. Passarino, M. Bonafe, D. Monti, G. Baggio, S. Bertolini, D. Mari, R. Mattace, and C. Franceschi, *Mitochondrial DNA inherited variants are associated with successful aging and longevity in humans.* FASEB J, 1999. **13**(12): p. 1532-6.

46. Costa, M.D., L. Cherni, V. Fernandes, F. Freitas, A.B. Ammar El Gaaied, and L. Pereira, *Data from complete mtDNA sequencing of Tunisian centenarians: testing haplogroup association and the "golden mean" to longevity.* Mech Ageing Dev, 2009. **130**(4): p. 222-6.

47. Dato, S., G. Passarino, G. Rose, K. Altomare, D. Bellizzi, V. Mari, E. Feraco, C. Franceschi, and G. De Benedictis, *Association of the mitochondrial DNA*

haplogroup J with longevity is population specific. Eur J Hum Genet, 2004. **12**(12): p. 1080-2.

48. Dominguez-Garrido, E., D. Martinez-Redondo, C. Martin-Ruiz, A. Gomez-Duran, E. Ruiz-Pesini, P. Madero, M. Tamparillas, J. Montoya, T. von Zglinicki, C. Diez-Sanchez, and M.J. Lopez-Perez, *Association of mitochondrial haplogroup J and mtDNA oxidative damage in two different North Spain elderly populations.* Biogerontology, 2009. **10**(4): p. 435-42.

49. Ross, O.A., R. McCormack, M.D. Curran, R.A. Duguid, Y.A. Barnett, I.M. Rea, and D. Middleton, *Mitochondrial DNA polymorphism: its role in longevity of the Irish population.* Exp Gerontol, 2001. **36**(7): p. 1161-78.

50. Shlush, L.I., G. Atzmon, R. Weisshof, D. Behar, G. Yudkovsky, N. Barzilai, and K. Skorecki, *Ashkenazi Jewish centenarians do not demonstrate enrichment in mitochondrial haplogroup J.* PLoS One, 2008. **3**(10): p. e3425.

51. Pinos, T., G. Nogales-Gadea, J.R. Ruiz, G. Rodriguez-Romo, C. Santiago-Dorrego, C. Fiuza-Luces, F. Gomez-Gallego, A. Cano-Nieto, N. Garatachea, M. Moran, M. Angel Martin, J. Arenas, A.L. Andreu, and A. Lucia, *Are mitochondrial haplogroups associated with extreme longevity? A study on a Spanish cohort.* Age (Dordr), 2011.

52. Niemi, A.K., A. Hervonen, M. Hurme, P.J. Karhunen, M. Jylha, and K. Majamaa, *Mitochondrial DNA polymorphisms associated with longevity in a Finnish population.* Hum Genet, 2003. **112**(1): p. 29-33.

53. Arcos-Burgos, M. and M. Muenke, *Genetics of population isolates.* Clin Genet, 2002. **61**(4): p. 233-47.

54. McQuillan, R., A.L. Leutenegger, R. Abdel-Rahman, C.S. Franklin, M. Pericic, L. Barac-Lauc, N. Smolej-Narancic, B. Janicijevic, O. Polasek, A. Tenesa, A.K. Macleod, S.M. Farrington, P. Rudan, C. Hayward, V. Vitart, I. Rudan, S.H. Wild, M.G. Dunlop, A.F. Wright, H. Campbell, and J.F. Wilson, *Runs of homozygosity in European populations.* Am J Hum Genet, 2008. **83**(3): p. 359-72.

55. Macgregor, S., C. Bellis, R.A. Lea, H. Cox, T. Dyer, J. Blangero, P.M. Visscher, and L.R. Griffiths, *Legacy of mutiny on the Bounty: founder effect and admixture on Norfolk Island.* Eur J Hum Genet, 2010. **18**(1): p. 67-72.

56. Youhanna, S., D.E. Platt, A. Rebeiz, M. Lauridsen, M.E. Deeb, A. Nasrallah, S. Alam, H. Puzantian, S. Kabbani, M. Ghoul, T.G. Zreik, H. El Bayeh, A. Abchee, and P. Zalloua, *Parental consanguinity and family history of coronary artery disease strongly predict early stenosis.* Atherosclerosis, 2010.

57. Ushasree, B., V. Shivani, A. Venkateshwari, R.K. Jain, C. Narsimhan, and P. Nallari, *Epidemiology and genetics of dilated cardiomyopathy in the Indian context.* Indian J Med Sci, 2009. **63**(7): p. 288-96.

58. Ismail, J., T.H. Jafar, F.H. Jafary, F. White, A.M. Faruqui, and N. Chaturvedi, *Risk factors for non-fatal myocardial infarction in young South Asian adults.* Heart, 2004. **90**(3): p. 259-63.

59. Denic, S., *Consanguinity as risk factor for cervical carcinoma.* Med Hypotheses, 2003. **60**(3): p. 321-4.

60. Levy-Lahad, E., E.M. Wijsman, E. Nemens, L. Anderson, K.A. Goddard, J.L. Weber, T.D. Bird, and G.D. Schellenberg, *A familial Alzheimer's disease locus on chromosome 1.* Science, 1995. **269**(5226): p. 970-3.

61. Bittles, A.H. and J.V. Neel, *The costs of human inbreeding and their implications for variations at the DNA level.* Nat Genet, 1994. **8**(2): p. 117-21.

62. Altukhov Yu, P. and V.A. Sheremet'eva, *Genomic heterozygosity and human longevity.* Dokl Biol Sci, 2000. **371**: p. 197-9.

63. Helgason, A., S. Palsson, D.F. Gudbjartsson, T. Kristjansson, and K. Stefansson, *An association between the kinship and fertility of human couples.* Science, 2008. **319**(5864): p. 813-6.

64. Rogaev, E.I., R. Sherrington, E.A. Rogaeva, G. Levesque, M. Ikeda, Y. Liang, H. Chi, C. Lin, K. Holman, T. Tsuda, and et al., *Familial Alzheimer's disease in*

kindreds with missense mutations in a gene on chromosome 1 related to the Alzheimer's disease type 3 gene. Nature, 1995. **376**(6543): p. 775-8.

65. Puffenberger, E.G., E.R. Kauffman, S. Bolk, T.C. Matise, S.S. Washington, M. Angrist, J. Weissenbach, K.L. Garver, M. Mascari, R. Ladda, and et al., *Identity-by-descent and association mapping of a recessive gene for Hirschsprung disease on human chromosome 13q22.* Hum Mol Genet, 1994. **3**(8): p. 1217-25.

66. Hovatta, I., J.D. Terwilliger, D. Lichtermann, T. Makikyro, J. Suvisaari, L. Peltonen, and J. Lonnqvist, *Schizophrenia in the genetic isolate of Finland.* Am J Med Genet, 1997. **74**(4): p. 353-60.

67. Leach, F.S., N.C. Nicolaides, N. Papadopoulos, B. Liu, J. Jen, R. Parsons, P. Peltomaki, P. Sistonen, L.A. Aaltonen, M. Nystrom-Lahti, and et al., *Mutations of a mutS homolog in hereditary nonpolyposis colorectal cancer.* Cell, 1993. **75**(6): p. 1215-25.

68. Hanson, R.L., M.G. Ehm, D.J. Pettitt, M. Prochazka, D.B. Thompson, D. Timberlake, T. Foroud, S. Kobes, L. Baier, D.K. Burns, L. Almasy, J. Blangero, W.T. Garvey, P.H. Bennett, and W.C. Knowler, *An autosomal genomic scan for loci linked to type II diabetes mellitus and body-mass index in Pima Indians.* Am J Hum Genet, 1998. **63**(4): p. 1130-8.

69. Hegele, R.A., J.H. Brunt, and P.W. Connelly, *Genetic and biochemical factors associated with variation in blood pressure in a genetic isolate.* Hypertension, 1996. **27**(2): p. 308-12.

70. Jonsson, T., H. Stefansson, D.S. Ph, I. Jonsdottir, P.V. Jonsson, J. Snaedal, S. Bjornsson, J. Huttenlocher, A.I. Levey, J.J. Lah, D. Rujescu, H. Hampel, I. Giegling, O.A. Andreassen, K. Engedal, I. Ulstein, S. Djurovic, C. Ibrahim-Verbaas, A. Hofman, M.A. Ikram, C.M. van Duijn, U. Thorsteinsdottir, A. Kong, and K. Stefansson, *Variant of TREM2 Associated with the Risk of Alzheimer's Disease.* N Engl J Med, 2012.

71. Manolio, T.A., F.S. Collins, N.J. Cox, D.B. Goldstein, L.A. Hindorff, D.J. Hunter, M.I. McCarthy, E.M. Ramos, L.R. Cardon, A. Chakravarti, J.H. Cho, A.E. Guttmacher, A. Kong, L. Kruglyak, E. Mardis, C.N. Rotimi, M. Slatkin, D. Valle, A.S. Whittemore, M. Boehnke, A.G. Clark, E.E. Eichler, G. Gibson, J.L. Haines,

T.F. Mackay, S.A. McCarroll, and P.M. Visscher, *Finding the missing heritability of complex diseases.* Nature, 2009. **461**(7265): p. 747-53.

72. Jonsson, T., J.K. Atwal, S. Steinberg, J. Snaedal, P.V. Jonsson, S. Bjornsson, H. Stefansson, P. Sulem, D. Gudbjartsson, and J. Maloney, *A mutation in APP protects against Alzheimer/'s disease and age-related cognitive decline.* Nature, 2012. **488**(7409): p. 96-99.

73. Steinberg AG, B.H., Kurczynski TW, Martin AO, Kurczynski EM *Genetic studies in an inbred human isolate*. in *Proceedings of the Third International Congress of Human Genetics.* 1967. Johns Hopkins University Press, Baltimore.

74. Agarwala, R., L.G. Biesecker, and A.A. Schaffer, *Anabaptist genealogy database.* Am J Med Genet C Semin Med Genet, 2003. **121C**(1): p. 32-7.

75. Mathias, R.A., C.A. Bickel, T.H. Beaty, G.M. Petersen, J.B. Hetmanski, K.Y. Liang, and K.C. Barnes, *A study of contemporary levels and temporal trends in inbreeding in the Tangier Island, Virginia, population using pedigree data and isonymy.* Am J Phys Anthropol, 2000. **112**(1): p. 29-38.

76. Rothrock, C.R., A. Murgia, E.L. Sartorato, E. Leonardi, S. Wei, S.L. Lebeis, L.E. Yu, J.L. Elfenbein, R.A. Fisher, and K.H. Friderici, *Connexin 26 35delG does not represent a mutational hotspot.* Hum Genet, 2003. **113**(1): p. 18-23.

77. Wilch, E., M. Zhu, K.B. Burkhart, M. Regier, J.L. Elfenbein, R.A. Fisher, and K.H. Friderici, *Expression of GJB2 and GJB6 is reduced in a novel DFNB1 allele.* Am J Hum Genet, 2006. **79**(1): p. 174-9.

78. Wilch, E., *A novel DFNB1 deletion allele supports the existence of a distant CIS-regulatory region that controls GJB2 and GJB6 expression.* Clinical Genetics, 2010.

79. Lee, W.J., T.I. Pollin, J.R. O'Connell, R. Agarwala, and A.A. Schaffer, *PedHunter 2.0 and its usage to characterize the founder structure of the Old Order Amish of Lancaster County.* BMC Med Genet, 2010. **11**(1): p. 68.

80. Weir, B.S., *Genetic Data Analysis II: Methods for Discrete Population Genetic Data*. 1996, Sunderland, MA: Sinauer.

81. Weir, B.S. and C.C. Cockerham, *Estimating F-statistics for the analysis of population structure.* Evolution, 1984. **1984**(38): p. 1358-1370.

82. Norris, T.L., *Acculturation in a German Catholic Community*, in *Department of Sociology and Anthropology*. 1950, Michigan State University: East Lansing, MI.

83. Agarwala, R., A.A. Schaffer, and J.F. Tomlin, *Towards a complete North American Anabaptist Genealogy II: analysis of inbreeding.* Hum Biol, 2001. **73**(4): p. 533-45.

84. Lachance, J., *Inbreeding, pedigree size, and the most recent common ancestor of humanity.* J Theor Biol, 2009. **261**(2): p. 238-47.

85. Igl, W., A. Johansson, and U. Gyllensten, *The Northern Swedish Population Health Study (NSPHS)--a paradigmatic study in a rural population combining community health and basic research.* Rural Remote Health, 2010. **10**(2): p. 1363.

86. Harman, D., *Aging: a theory based on free radical and radiation chemistry.* J Gerontol, 1956. **11**(3): p. 298-300.

87. Kong, Q.P., Y.G. Yao, C. Sun, H.J. Bandelt, C.L. Zhu, and Y.P. Zhang, *Phylogeny of east Asian mitochondrial DNA lineages inferred from complete sequences.* Am J Hum Genet, 2003. **73**(3): p. 671-6.

88. Herrnstadt, C., J.L. Elson, E. Fahy, G. Preston, D.M. Turnbull, C. Anderson, S.S. Ghosh, J.M. Olefsky, M.F. Beal, R.E. Davis, and N. Howell, *Reduced-median-network analysis of complete mitochondrial DNA coding-region sequences for the major African, Asian, and European haplogroups.* Am J Hum Genet, 2002. **70**(5): p. 1152-71.

89. Niemi, A.K., J.S. Moilanen, M. Tanaka, A. Hervonen, M. Hurme, T. Lehtimaki, Y. Arai, N. Hirose, and K. Majamaa, *A combination of three common inherited*

mitochondrial DNA polymorphisms promotes longevity in Finnish and Japanese subjects. Eur J Hum Genet, 2005. **13**(2): p. 166-70.

90. Tranah, G.J., *Mitochondrial-nuclear epistasis: implications for human aging and longevity.* Ageing Res Rev, 2011. **10**(2): p. 238-52.

91. Silva, W.A., Jr., S.L. Bonatto, A.J. Holanda, A.K. Ribeiro-Dos-Santos, B.M. Paixao, G.H. Goldman, K. Abe-Sandes, L. Rodriguez-Delfin, M. Barbosa, M.L. Paco-Larson, M.L. Petzl-Erler, V. Valente, S.E. Santos, and M.A. Zago, *Mitochondrial genome diversity of Native Americans supports a single early entry of founder populations into America.* Am J Hum Genet, 2002. **71**(1): p. 187-92.

92. Alexe, G., N. Fuku, E. Bilal, H. Ueno, Y. Nishigaki, Y. Fujita, M. Ito, Y. Arai, N. Hirose, G. Bhanot, and M. Tanaka, *Enrichment of longevity phenotype in mtDNA haplogroups D4b2b, D4a, and D5 in the Japanese population.* Hum Genet, 2007. **121**(3-4): p. 347-56.

93. Feng, J., J. Zhang, M. Liu, G. Wan, K. Qi, C. Zheng, Z. Lv, C. Hu, Y. Zeng, S.G. Gregory, and Z. Yang, *Association of mtDNA haplogroup F with healthy longevity in the female Chuang population, China.* Exp Gerontol, 2011. **46**(12): p. 987-93.

94. Samuels, D.C., A.D. Carothers, R. Horton, and P.F. Chinnery, *The power to detect disease associations with mitochondrial DNA haplogroups.* Am J Hum Genet, 2006. **78**(4): p. 713-20.

95. Fallin, M.D. and A. Matteini, *Genetic epidemiology in aging research.* J Gerontol A Biol Sci Med Sci, 2009. **64**(1): p. 47-60.

96. Stroup, D.F., J.A. Berlin, S.C. Morton, I. Olkin, G.D. Williamson, D. Rennie, D. Moher, B.J. Becker, T.A. Sipe, and S.B. Thacker, *Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group.* JAMA, 2000. **283**(15): p. 2008-12.

97. Harrison, F., *Getting started with meta-analysis.* Methods in Ecology and Evolution, 2011. **2**: p. 1-10.

98. Amaral-Fernandes, M.S., A.M. Marcondes, P.M.d.A.D. Miranda, A.T. Maciel-Guerra, and E.L. Sartorato, *Mutations for Leber hereditary optic neuropathy in patients with alcohol and tobacco optic neuropathy.* Molecular vision, 2011. **17**: p. 3175.

99. Anantharaju, A., A. Feller, and A. Chedid, *Aging liver.* Gerontology, 2002. **48**(6): p. 343-353.

100. Baggio, G.F., C; Mari, D; Herskind, AM; Andersen-Ranberg, K; Jeune, B, *Biology and Genetics of Human Longevity.* Australasian Journal on Ageing, 1998. **17**(s1): p. 8-10.

101. Bilal, E., R. Rabadan, G. Alexe, N. Fuku, H. Ueno, Y. Nishigaki, Y. Fujita, M. Ito, Y. Arai, N. Hirose, A. Ruckenstein, G. Bhanot, and M. Tanaka, *Mitochondrial DNA haplogroup D4a is a marker for extreme longevity in Japan.* PLoS One, 2008. **3**(6): p. e2421.

102. Blanche, H., *[Genetic factors in longevity].* Presse Med, 2003. **32**(8): p. 365-9.

103. Bonafè, M., M. Cardelli, F. Marchegiani, L. Cavallone, S. Giovagnetti, F. Olivieri, R. Lisa, C. Pieri, and C. Franceschi, *Increase of homozygosity in centenarians revealed by a new inter-Alu PCR technique.* Experimental Gerontology, 2001. **36**(7): p. 1063-1073.

104. Bonafe, M., C. Barbi, F. Olivieri, A. Yashin, K.F. Andreev, J.W. Vaupel, G. De Benedictis, G. Rose, G. Carrieri, and S.M. Jazwinski, *An allele of< i> HRAS1</i> 3' variable number of tandem repeats is a frailty allele: implication for an evolutionarily-conserved pathway involved in longevity.* Gene, 2002. **286**(1): p. 121-126.

105. Capri, M., S. Salvioli, D. Monti, C. Caruso, G. Candore, S. Vasto, F. Olivieri, F. Marchegiani, P. Sansoni, G. Baggio, D. Mari, G. Passarino, G. De Benedictis, and C. Franceschi, *Human longevity within an evolutionary perspective: the peculiar paradigm of a post-reproductive genetics.* Exp Gerontol, 2008. **43**(2): p. 53-60.

106. Capri, M., S. Salvioli, D. Monti, C. Caruso, G. Candore, S. Vasto, F. Olivieri, F. Marchegiani, P. Sansoni, and G. Baggio, *Human longevity within an evolutionary*

perspective: the peculiar paradigm of a post-reproductive genetics. Experimental gerontology, 2008. **43**(2): p. 53-60.

107. Castri, L., M. Melendez-Obando, R. Villegas-Palma, R. Barrantes, H. Raventos, R. Pereira, D. Luiselli, D. Pettener, and L. Madrigal, *Mitochondrial polymorphisms are associated both with increased and decreased longevity.* Hum Hered, 2009. **67**(3): p. 147-53.

108. Chen, A., N. Raule, A. Chomyn, and G. Attardi, *Decreased reactive oxygen species production in cells with mitochondrial haplogroups associated with longevity.* PLoS One, 2012. **7**(10): p. e46473.

109. Chen, Y.F., C.Y. Wu, R. Kirby, C.H. Kao, and T.F. Tsai, *A role for the CISD2 gene in lifespan control and human disease.* Annals of the New York Academy of Sciences, 2010. **1201**(1): p. 58-64.

110. Chomyn, A. and G. Attardi, *MtDNA mutations in aging and apoptosis.* Biochem Biophys Res Commun, 2003. **304**(3): p. 519-29.

111. Clancy, D.J., *Variation in mitochondrial genotype has substantial lifespan effects which may be modulated by nuclear background.* Aging Cell, 2008. **7**(6): p. 795-804.

112. Courtenay, M.D., J.R. Gilbert, L. Jiang, A.C. Cummings, P.J. Gallins, L. Caywood, L. Reinhart-Mercer, D. Fuzzell, C. Knebusch, R. Laux, J.L. McCauley, C.E. Jackson, M.A. Pericak-Vance, J.L. Haines, and W.K. Scott, *Mitochondrial haplogroup X is associated with successful aging in the Amish.* Hum Genet, 2012. **131**(2): p. 201-8.

113. Czarneckza, A.M., *Is mitochondrial genome chicken or egg of carcinogenesis?* FEBS Journal, 2010. **277**(Supplement 1): p. 219.

114. De Benedictis, G., G. Carrieri, S. Garasto, G. Rose, O. Varcasia, M. Bonafe, C. Franceschi, and S.M. Jazwinski, *Does a retrograde response in human aging and longevity exist?* Exp Gerontol, 2000. **35**(6-7): p. 795-801.

115. de Benedictis, G., G. Carrieri, O. Varcasia, M. Bonafe, and C. Franceschi, *Inherited variability of the mitochondrial genome and successful aging in humans.* Ann N Y Acad Sci, 2000. **908**: p. 208-18.

116. De Benedictis, G., Q. Tan, B. Jeune, K. Christensen, S.V. Ukraintseva, M. Bonafè, C. Franceschi, J.W. Vaupel, and A.I. Yashin, *Recent advances in human gene–longevity association studies.* Mechanisms of Ageing and Development, 2001. **122**(9): p. 909-920.

117. de Buenos Aires, A., *Genetics and anthropology in studies on aging and Chagas disease.* Genetics, 2010. **88**: p. 245-250.

118. Decanini-Mancera, A., A.R. Harrison, and M.S. Lee, *Another case of leber hereditary optic neuropathy in an octogenarian.* J Neuroophthalmol, 2009. **29**(2): p. 159-60.

119. Franceschi, C., S. Valensin, F. Fagnoni, C. Barbi, and M. Bonafè, *Biomarkers of immunosenescence within an evolutionary perspective: the challenge of heterogeneity and the role of antigenic load.* Experimental gerontology, 1999. **34**(8): p. 911-921.

120. Franceschi, C., V. Bezrukov, H. Blanché, L. Bolund, K. Christensen, G.D. BENEDICTIS, L. Deiana, E. Gonos, A. Hervonen, and H. Yang, *Genetics of healthy aging in Europe.* Annals of the New York Academy of Sciences, 2007. **1100**(1): p. 21-45.

121. Franceschi, C., M. Capri, D. Monti, S. Giunta, F. Olivieri, F. Sevini, M.P. Panourgia, L. Invidia, L. Celani, and M. Scurti, *Inflammaging and anti-inflammaging: a systemic perspective on aging and longevity emerged from studies in humans.* Mechanisms of ageing and development, 2007. **128**(1): p. 92-105.

122. Franceschi, C., *Inflammaging as a major characteristic of old people: can it be prevented or cured?* Nutrition reviews, 2007. **65**(s3): p. S173-S176.

123. Frisard, M.I., A. Broussard, S.S. Davies, L.J. Roberts, 2nd, J. Rood, L. de Jonge, X. Fang, S.M. Jazwinski, W.A. Deutsch, E. Ravussin, and S. Louisiana Healthy Aging, *Aging, resting metabolic rate, and oxidative damage: results from the*

Louisiana Healthy Aging Study. J Gerontol A Biol Sci Med Sci, 2007. **62**(7): p. 752-9.

124. Gong, J.-S., J. ZHANG, M. YONEDA, K. SAHASHI, H. MIYAJIMA, K. YAMAUCHI, K. YAGI, and M. TANAKA, *Mitochondrial genotype frequent in centenarians predisposes resistance to adult-onset diseases.* Journal of clinical biochemistry and nutrition, 1998. **24**(2): p. 105-111.

125. Guerin, J.C., *Emerging area of aging research: long-lived animals with "negligible senescence".* Ann N Y Acad Sci, 2004. **1019**: p. 518-20.

126. Guo, X., K.Y. Popadin, N. Markuzon, Y.L. Orlov, Y. Kraytsberg, K.J. Krishnan, G. Zsurka, D.M. Turnbull, W.S. Kunz, and K. Khrapko, *Repeats, longevity and the sources of mtDNA deletions: evidence from 'deletional spectra'.* Trends Genet, 2010. **26**(8): p. 340-3.

127. Guo, X., I. Dombrovsky, T. Leibman, Y. Kraytsberg, K. Popadin, and K. Khrapko, *Mitochondrial DNA, direct repeats, deletions, and centenarians: An unfinished puzzle.* Biochimica et Biophysica Acta (BBA)-Bioenergetics, 2010. **1797**: p. 74.

128. Iannitti, T. and B. Palmieri, *Inflammation and genetics: an insight in the centenarian model.* Hum Biol, 2011. **83**(4): p. 531-59.

129. Ivanova, R., V. Lepage, D. Charron, and F. Schachter, *Mitochondrial genotype associated with French Caucasian centenarians.* Gerontology, 1998. **44**(6): p. 349.

130. Iwata, N., J. Zhang, G. Atzmon, S. Leanza, J. Cho, A. Chomyn, R.D. Burk, N. Barzilai, and G. Attardi, *Aging-related occurrence in Ashkenazi Jews of leukocyte heteroplasmic mtDNA mutation adjacent to replication origin frequently remodeled in Italian centenarians.* Mitochondrion, 2007. **7**(4): p. 267-272.

131. Kasabri, V. and N. Bulatova, *Current View on Biology of Aging.* Jordan Medical Journal, 2010. **44**(2).

132. Khrapko, K., N. Bodyak, W.G. Thilly, N.J. van Orsouw, X. Zhang, H.A. Coller, T.T. Perls, M. Upton, J. Vijg, and J.Y. Wei, *Cell-by-cell scanning of whole*

mitochondrial genomes in aged human heart reveals a significant fraction of myocytes with clonally expanded deletions. Nucleic Acids Res, 1999. **27**(11): p. 2434-41.

133. Klemba, A., M. Kowalewska, W. Kukwa, K. Tonska, A. Szybinska, M. Mossakowska, A. Scinska, P. Golik, K. Koper, J. Radziszewski, A. Kukwa, A.M. Czarnecka, and E. Bartnik, *Mitochondrial genotype in vulvar carcinoma - cuckoo in the nest.* J Biomed Sci, 2010. **17**: p. 73.

134. Kokaze, A., M. Ishikawa, N. Matsunaga, M. Yoshida, Y. Sekine, K. Sekiguchi, M. Satoh, M. Harada, K. Teruya, N. Takeda, Y. Uchida, T. Tsunoda, and Y. Takashima, *Longevity-associated mitochondrial DNA 5178 A/C polymorphism modulates effects of daily drinking and cigarette consumption on serum triglyceride levels in middle-aged Japanese men.* Experimental Gerontology, 2003. **38**(10): p. 1071-1076.

135. Lio, D., L. Scola, A. Crivello, G. Colonna-Romano, G. Candore, M. Bonafe, L. Cavallone, C. Franceschi, and C. Caruso, *Gender-specific association between−1082 IL-10 promoter polymorphism and longevity.* Genes and immunity, 2002. **3**(1): p. 30-33.

136. Lorenc, A., K. Tonska, D. Kabzinska, and E. Bartnik, *Mitochondrial DNA in polish centenarians.* Toxicol Mech Methods, 2004. **14**(1-2): p. 91-5.

137. Maruszak, A., B. Pepłońska, K. Safranow, M. Chodakowska-Żebrowska, M. Barcikowska, and C. Żekanowski, *TOMM40 rs10524523 polymorphism's role in late-onset Alzheimer's disease and in longevity.* Journal of Alzheimer's Disease, 2012. **28**(2): p. 309-322.

138. Mcnerlan, S., O. Ross, M. Breen, and M. Rea, *Defining the cytokine and immunoglobulin phenotype associated with mitochondrial J haplotype carriers in the BELFAST elderly longitudinal free-living ageing study (BELFAST).* European Geriatric Medicine, 2012. **3**: p. S65-S66.

139. Munakata, K., M. Tanaka, K. Mori, S. Washizuka, M. Yoneda, O. Tajima, T. Akiyama, S. Nanko, H. Kunugi, K. Tadokoro, N. Ozaki, T. Inada, K. Sakamoto, T. Fukunaga, Y. Iijima, N. Iwata, M. Tatsumi, K. Yamada, T. Yoshikawa, and T. Kato, *Mitochondrial DNA 3644T→C mutation associated with bipolar disorder.* Genomics, 2004. **84**(6): p. 1041-1050.

140. Pan, H., Q.-p. Kong, Y.-t. Cheng, S.-g. Lian, J. Yang, S.-j. Gao, L.-y. Xu, and Y.-p. Zhang, *Absence of association between mitochondrial DNA C150T polymorphism and longevity in a Han Chinese population.* Experimental Gerontology, 2011. **46**(6): p. 511-515.

141. Passarino, G., P.A. Underhill, L.L. Cavalli-Sforza, O. Semino, G.M. Pes, C. Carru, L. Ferrucci, M. Bonafe, C. Franceschi, and L. Deiana, *Y chromosome binary markers to study the high prevalence of males in Sardinian centenarians and the genetic structure of the Sardinian population.* Human heredity, 2001. **52**(3): p. 136-139.

142. Pinos, T., G. Nogales-Gadea, J.R. Ruiz, G. Rodriguez-Romo, C. Santiago-Dorrego, C. Fiuza-Luces, F. Gomez-Gallego, A. Cano-Nieto, N. Garatachea, M. Moran, M. Angel Martin, J. Arenas, A.L. Andreu, and A. Lucia, *Are mitochondrial haplogroups associated with extreme longevity? A study on a Spanish cohort.* Age (Dordr), 2011. **34**(1): p. 227-33.

143. Raule, N., F. Sevini, A. Santoro, S. Altilia, and C. Franceschi, *Association studies on human mitochondrial DNA: methodological aspects and results in the most common age-related diseases.* Mitochondrion, 2007. **7**(1-2): p. 29-38.

144. Rea, I.M., S.E. McNerlan, G.P. Archbold, D. Middleton, M.D. Curran, I.S. Young, and O.A. Ross, *Mitochondrial J haplogroup is associated with lower blood pressure and anti-oxidant status: findings in octo/nonagenarians from the BELFAST Study.* Age (Dordr), 2012.

145. Ren, W.H., X.H. Li, H.G. Zhang, F.M. Deng, W.Q. Liao, Y. Pang, Y.H. Liu, M.J. Qiu, G.Y. Zhang, and Y.G. Zhang, *Mitochondrial DNA haplogroups in a Chinese Uygur population and their potential association with longevity.* Clin Exp Pharmacol Physiol, 2008. **35**(12): p. 1477-81.

146. Rose, G., G. Romeo, S. Dato, P. Crocco, A.C. Bruni, A. Hervonen, K. Majamaa, F. Sevini, C. Franceschi, G. Passarino, and G.E.o.H.A.P. Consortium, *Somatic point mutations in mtDNA control region are influenced by genetic background and associated with healthy aging: a GEHA study.* PLoS One, 2010. **5**(10): p. e13395.

147. Rose, G., G. Passarino, V. Scornaienchi, G. Romeo, S. Dato, D. Bellizzi, V. Mari, E. Feraco, R. Maletta, A. Bruni, C. Franceschi, and G. De Benedictis, *The*

mitochondrial DNA control region shows genetically correlated levels of heteroplasmy in leukocytes of centenarians and their offspring. BMC Genomics, 2007. **8**: p. 293.

148. Rose, G., G. Passarino, G. Carrieri, K. Altomare, V. Greco, S. Bertolini, M. Bonafe, C. Franceschi, and G. De Benedictis, *Paradoxes in longevity: sequence analysis of mtDNA haplogroup J in centenarians.* Eur J Hum Genet, 2001. **9**(9): p. 701-7.

149. Salvioli, S., F. Olivieri, F. Marchegiani, M. Cardelli, A. Santoro, E. Bellavista, M. Mishto, L. Invidia, M. Capri, S. Valensin, F. Sevini, E. Cevenini, L. Celani, F. Lescai, E. Gonos, C. Caruso, G. Paolisso, G. De Benedictis, D. Monti, and C. Franceschi, *Genes, ageing and longevity in humans: problems, advantages and perspectives.* Free Radic Res, 2006. **40**(12): p. 1303-23.

150. Salvioli, S., M. Capri, A. Santoro, N. Raule, F. Sevini, S. Lukas, C. Lanzarini, D. Monti, G. Passarino, G. Rose, G. De Benedictis, and C. Franceschi, *The impact of mitochondrial DNA on human lifespan: a view from studies on centenarians.* Biotechnol J, 2008. **3**(6): p. 740-9.

151. Santoro, A., S. Salvioli, N. Raule, M. Capri, F. Sevini, S. Valensin, D. Monti, D. Bellizzi, G. Passarino, G. Rose, G. De Benedictis, and C. Franceschi, *Mitochondrial DNA involvement in human longevity.* Biochim Biophys Acta, 2006. **1757**(9-10): p. 1388-99.

152. Schächter, F., D. Cohen, and T. Kirkwood, *Prospects for the genetics of human longevity.* Human Genetics, 1993. **91**(6): p. 519-526.

153. Schmuczerova, J., R. Brdicka, M. Dostal, R.J. Sram, and J. Topinka, *Genetic variability of HVRII mtDNA in cord blood and respiratory morbidity in children.* Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis, 2009. **666**(1–2): p. 1-7.

154. Sultan, S. and N. Hynes, *Cardiovascular disease: primary prevention, disease modulation and regenerative therapy.* Vascular, 2012. **20**(5): p. 243-50.

155. Takagi, K., Y. Yamada, J.-S. Gong, T. Sone, M. Yokota, and M. Tanaka, *Association of a 5178C→A (Leu237Met) polymorphism in the mitochondrial DNA*

with a low prevalence of myocardial infarction in Japanese individuals. Atherosclerosis, 2004. **175**(2): p. 281-286.

156. Takasaki, S., *Japanese Alzheimer's disease and other complex disorders diagnosis based on mitochondrial SNP haplogroups.* Int J Alzheimers Dis, 2012. **2012**: p. 245038.

157. Takasaki, S., *Mitochondrial SNPs associated with Japanese centenarians, Alzheimer's patients, and Parkinson's patients.* Comput Biol Chem, 2008. **32**(5): p. 332-7.

158. Takasaki, S., *Mitochondrial haplogroups associated with Japanese centenarians, Alzheimer's patients, Parkinson's patients, type 2 diabetic patients and healthy non-obese young males.* J Genet Genomics, 2009. **36**(7): p. 425-34.

159. Tanaka, M., J. Gong, J. Zhang, Y. Yamada, H.J. Borgeld, and K. Yagi, *Mitochondrial genotype associated with longevity and its inhibitory effect on mutagenesis.* Mech Ageing Dev, 2000. **116**(2-3): p. 65-76.

160. Tanaka, M., N. Fuku, T. Takeyasu, L.-J. Guo, R. Hirose, M. Kurata, H.-J.W. Borgeld, Y. Yamada, W. Maruyama, Y. Arai, N. Hirose, Y. Oshida, Y. Sato, N. Hattori, Y. Mizuno, S. Iwata, and K. Yagi, *Golden mean to longevity: Rareness of mitochondrial cytochrome b variants in centenarians but not in patients with Parkinson's disease.* Journal of Neuroscience Research, 2002. **70**(3): p. 347-355.

161. Tanaka, M., *Mitochondrial genotypes and cytochrome b variants associated with longevity or Parkinson's disease.* J Neurol, 2002. **249 Suppl 2**: p. II11-8.

162. Tanaka, M., T. Takeyasu, N. Fuku, G.U.O. Li-Jun, and M. Kurata, *Mitochondrial Genome Single Nucleotide Polymorphisms and Their Phenotypes in the Japanese.* Annals of the New York Academy of Sciences, 2004. **1011**(1): p. 7-20.

163. Tanaka, M., *Common, rare, and individual variations of mitochondrial DNA associated with diseases or longevity.* Journal of clinical biochemistry and nutrition, 2000. **28**(3): p. 191-199.

164. Tomita-Mitchell, A., B.P. Muniappan, P. Herrero-Jimenez, H. Zarbl, and W.G. Thilly, *Single nucleotide polymorphism spectra in newborns and centenarians: identification of genes coding for rise of mortal disease.* Gene, 1998. **223**(1–2): p. 381-391.

165. von Wurmb-Schwark, N., T. Schwark, A. Caliebe, C. Drenske, S. Nikolaus, S. Schreiber, and A. Nebel, *Low level of the mtDNA4977 deletion in blood of exceptionally old individuals.* Mechanisms of Ageing and Development, 2010. **131**(3): p. 179-184.

166. Willcox, D.C., B. Willcox, W.-C. Hsueh, and M. Suzuki, *Genetic determinants of exceptional human longevity: insights from the Okinawa Centenarian Study.* AGE, 2006. **28**(4): p. 313-332.

167. Yang, X., X. Wang, H. Yao, J. Deng, Q. Jiang, Y. Guo, G. Lan, D.J. Liao, and H. Jiang, *Mitochondrial DNA polymorphisms are associated with the longevity in the Guangxi Bama population of China.* Mol Biol Rep, 2012. **39**(9): p. 9123-31.

168. Yashin, A.I., G. De Benedictis, J.W. Vaupel, Q. Tan, K.F. Andreev, I.A. Iachine, M. Bonafe, S. Valensin, M. De Luca, L. Carotenuto, and C. Franceschi, *Genes and longevity: lessons from studies of centenarians.* J Gerontol A Biol Sci Med Sci, 2000. **55**(7): p. B319-28.

169. Zhang, J., J. Asin-Cayuela, J. Fish, Y. Michikawa, M. Bonafe, F. Olivieri, G. Passarino, G. De Benedictis, C. Franceschi, and G. Attardi, *Strikingly higher frequency in centenarians and twins of mtDNA mutation causing remodeling of replication origin in leukocytes.* Proc Natl Acad Sci U S A, 2003. **100**(3): p. 1116-21.

170. Yao, Y.G., Q.P. Kong, C.Y. Wang, C.L. Zhu, and Y.P. Zhang, *Different matrilineal contributions to genetic structure of ethnic groups in the silk road region in china.* Mol Biol Evol, 2004. **21**(12): p. 2265-80.

171. Cochran, w.g., *The combination of estimates from different experiments.* Biometrics, 1954. **10**: p. 101-129.

172. Villanueva, E.V. and S. Zavarsek, *Evaluating heterogeneity in cumulative meta-analyses.* BMC Med Res Methodol, 2004. **4**: p. 18.

173. Jylhava, J., *Genetic Determinants of Human Life-span* 2012, Chichester: John Wiley & Sons, Ltd.

174. Bonner JD, Fisher RA , Klein J , Lu Q , Wilch E, Friderici KH, Elfenbein JL , Schutte DL, and S. BC, *PEDIGREE STRUCTURE AND KINSHIP MEASUREMENTS IN A MID-MICHIGAN COMMUITY: A NEW NORTH AMERICAN POPULATION ISOLATE IDENTIFIED.* Human Biology (In Press), 2013.

175. Byrne, E.M., A.F. McRae, Z.Z. Zhao, N.G. Martin, G.W. Montgomery, and P.M. Visscher, *The use of common mitochondrial variants to detect and characterise population structure in the Australian population: implications for genome-wide association studies.* Eur J Hum Genet, 2008. **16**(11): p. 1396-403.

176. Wallace, D.C., *Diseases of the mitochondrial DNA.* Annual review of biochemistry, 1992. **61**(1): p. 1175-1212.

177. Wallace, D.C., G. Singh, M.T. Lott, J.A. Hodge, T.G. Schurr, A. Lezza, L.J. Elsas, and E.K. Nikoskelainen, *Mitochondrial DNA mutation associated with Leber's hereditary optic neuropathy.* Science, 1988. **242**(4884): p. 1427-1430.

178. Dowling, D.K., U. Friberg, and J. Lindell, *Evolutionary implications of non-neutral mitochondrial genetic variation.* Trends in Ecology & Evolution, 2008. **23**(10): p. 546-554.