

Published in final edited form as:

Ageing Res Rev. 2011 April ; 10(2): 191–200. doi:10.1016/j.arr.2010.07.005.

Contemporary Human Genetic Strategies in Aging Research

Cinnamon S. Bloss, Ph.D.^a, Ludmila Pawlikowska, Ph.D.^b, and Nicholas J. Schork, Ph.D.^c

^aThe Scripps Translational Science Institute & Scripps Health 3344 North Torrey Pines Court, Suite 300 La Jolla, CA 92037, U.S. 858-554-5737 858-546-9280 cbloss@scripps.edu

^bInstitute for Human Genetics & Department of Anesthesia University of California, San Francisco 513 Parnassus Ave, HSW 901A, Box 0793 San Francisco, CA 94143-0793, U.S. 415-514-3876 415-476-2956 (fax) pawlikowskal@anesthesia.ucsf.edu

^cThe Scripps Translational Science Institute & Department of Molecular and Experimental Medicine, The Scripps Research Institute 3344 North Torrey Pines Court, Suite 300 La Jolla, CA 92037, U.S. 858-554-5705 (admin) 858-546-9284 (fax) nschork@scripps.edu

Abstract

Human aging is a complex, multifactorial process influenced by a number of genetic and non-genetic factors. This article first reviews genetic strategies for human aging research and considers the advantages and disadvantages of each. We then discuss the issue of phenotypic definition for genetic studies of aging, including longevity/life span, as well as disease-free survival and other endophenotypes. Finally, we argue that extensions of this area of research, including incorporation of gene \times environment interactions, multivariate phenotypes, integration of functional genomic annotations, and exploitation of orthology – many of which are already initiated and ongoing – are critical to advancing this field.

1. Introduction

Human aging is a complex, multifactorial process influenced by a number of genetic and non-genetic factors. The identification of genetic variations that influence age-related phenomena such as disease-free survival, post-disease survival, and lifespan in general has received considerable attention and will likely receive even more attention in the future as genetic technologies become more refined and cost-efficient. In this review, we describe the basic approaches to the identification of genetic variations that influence human longevity and longevity-related phenotypes. Although we focus on truly genetic strategies (i.e., strategies involving the transmission of genetic information from generation to generation) for human aging research, as opposed to evolutionary, model organism, and purely genomic and/or somatic cell-based approaches, we draw on these other approaches when appropriate. We also consider the advantages and disadvantages of various approaches as well as directions for future research. We note that some recent reviews have also evaluated the merits of different genetic approaches to the study of human aging and we refer the reader to these reviews for additional insights and references [1–3]. Table 1 also provides a list of genetic studies investigating human aging and age-related phenotypes that the reader might

© 2010 Elsevier Ireland Ltd. All rights reserved

Correspondence to: Nicholas J. Schork.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

consider for further insight into the application and interpretation of genetic research paradigms in studies focusing on human aging.

1.1. Genetics vs. genomics

Although at times subtle, the distinction between 'genetics' as a scientific discipline replete with specific research paradigms, and 'genomics' as a scientific discipline is an important one since it puts into context the orientation, rationale and required resources for particular approaches. Genetics is essentially the study of inherited variations that populate DNA sequence. Genetics considers phenomena such as recombination, gene conversion, aneuploidy, and mutation. These are phenomena that influence inheritance patterns as well as the frequency of variations in the population. Many mathematical constructs and theorems have been developed to describe phenomena of relevance to the study of inheritance, such as Mendel's laws and Hardy-Weinberg equilibrium, which describe the probability that an individual transmits one of the two variations within his or her homologous chromosomes to an offspring, and the behavior of variations in the population at large under the assumption of random mating, respectively.

Genomics, on the other hand, involves the study of the composition, structure and function of DNA sequence. Genomics researchers often seek to identify patterns in DNA sequence, such as the typical nucleotide composition of coding regions in the genome as opposed to non-coding and/or regulatory regions of the genome, without regard to the transmission of that DNA sequence from one generation to the next. Of course any structural or functional characterization of the genome could be considered in studies of the transmission of variations in those regions from parents to offspring. Thus, genetics and genomics are complementary disciplines and, as discussed in sections 4.3 and 4.4, are often integrated to make sense of inherited factors and the genetic basis of a particular phenotype.

A great deal of contemporary genetics research is devoted to the identification of inherited DNA sequence variations that influence the expression of a particular phenotype such as a disease (Figure 1 provides a schematic depicting the transmission of genetic variations both at and neighboring a particular phenotypically-relevant genomic position or 'locus' across different numbers of generations). Using genetic analyses to identify specific variations that causally influence a particular phenotype is often referred to as 'genetic mapping' or, more historically, 'positional cloning' (i.e., as the ultimate isolation of a causal variation often involved cloning the DNA harboring the variation). Although it is possible that non-inherited *de novo* DNA sequence mutations could influence a particular phenotype, since one would have to rule out such mutations as inherited by using genetics principles, the study of *de novo* mutations is considered a part of genetic research. There are many genetic strategies and study designs that a researcher can utilize in order to identify DNA sequence variations that influence a particular phenotype, and virtually all of these strategies and designs have been considered in the study of human aging, lifespan, and related phenotypes. The development and refinement of these strategies has followed a very rapid progression that reflects both an eagerness to pursue genetic studies as organizing principles for virtually all cell biology, physiology, evolutionary biology, and biomedical science, as well as technological advances, as discussed in the section 2.1.

1.2. The evolution of genetics research paradigms

For any given individual, in order to understand which DNA sequence variations influence which phenotypic features (such as long life span), sequence variations that can be studied in this context must first be identified. This involves determining the nucleotide content of individuals' DNA via DNA sequencing technologies. It is no surprise then that the development and refinement of genetic mapping approaches to the identification of

phenotypically-relevant inherited variations parallels technological developments in DNA sequencing. Two other developments, the creation of genetic 'maps' of markers that detail the positions in the genome of variable sites, as well as the creation of efficient technologies for genotyping (i.e., rapidly determining the variations an individual possesses at a site in DNA sequence that is known to be polymorphic or vary among individuals in the population at large), also closely parallel developments in genetic mapping. This history is elegantly described in a review by Lander and Weinberg [4] and we capture a bit of this history in the following, with greater focus on genetic strategies applied to human aging in section 3.

Initial genetic mapping approaches to the identification of phenotypically-relevant DNA sequence variations used large pedigrees and related individuals in order to trace the 'co-segregation' of chromosomal regions marked by specific genetic variations – i.e., those that could be genotyped across the family members – with a particular trait or disease from generation to generation. These approaches required knowledge of the genomic locations of relatively few genetic variations since the size of the chromosomal regions that remained intact after recombination events over a small number of generations was likely to be large. Once variations within a chromosomal region have been identified as likely to be 'linked' to a variant segregating with the phenotype in a family, DNA sequencing technologies can be used to identify the offending causal variations. Lander and Schork describe some of the issues in family-based linkage analysis methods for gene mapping studies [5]. Figure 1 depicts a nuclear family, denoted by the number 3, where the two individuals with a phenotype both received variations ('A' and 'C') flanking the causal variant ('T') that are intact from the chromosomal segment harboring the causal variant from their father. Note that this particular haplotype, 'A-T-C,' is not the same in other families and individuals harboring the causative 'T' allele, due to the greater number of recombination events that occurred during the transmission of the chromosome harboring the causative 'T' allele to those individuals from their common ancestor, denoted by '1' in Figure 1.

The use of unrelated individuals for genetic mapping has many advantages, not the least of which is that they are easier to collect than families (see section 2.2). However, in order to identify variations that segregate with a phenotype among individuals separated by an arbitrary number of generations from a common ancestor or set of ancestors who introduced variations into the population at large, the interrogation of a greater set of genetic variations that might capture chromosomal segments not broken up by recombination since their introduction into the population would have to be pursued. Recognition of this fact led to the International HapMap Project, which was essentially a consortium effort to identify as many genetic variations as possible and detail their linkage disequilibrium (LD) relationships in order to identify strings of contiguous nucleotides that harbor variations that have been essentially inherited together among arbitrarily related individuals [6]. The motivation and yield of the International HapMap Project are described in numerous publications as well as on the HapMap website (www.hapmap.org). As an example of the intuitions motivating the HapMap consortium effort, the small black shaded rectangle in Figure 1 within individuals' chromosomes reflects chromosomal material – i.e., a string of nucleotides – that is intact from the individual who introduced the 'T' nucleotide causing the phenotype into the population (individual 1) that all individuals with the phenotype possess. If this shaded material contains other variations that could be genotyped, then all the individuals with the phenotype, barring any mutational events during the transmission of the alleles from generation to generation, would have those variations. As a result, those variations would be in LD and form a 'haplotype block' containing the offending 'T' causal variation. Collins described a plan for the development of appropriate technologies and resources to facilitate genetic studies that leverage the International HapMap resources and the biological insights such studies would likely yield, and this kind of vision, among other things, motivated the

development of genotyping technologies that would allow a researcher to interrogate hundreds of thousands of variations in a single assay for use in a genetic study [7].

The use of the Hapmap Project results and resources in genetic mapping studies has led to some noticeable successes (see section 2.3 for more detail), but has been called into question due to an overarching emphasis on common variations and a consequent downplaying of rare variations that might contribute to phenotypic expression [8]. Thus, it has been argued that phenotypes might be influenced by many infrequent variations that collectively contribute to phenotypes that are relatively frequent in the population at large. Thus, although common variations might be easier to identify, catalog, and ultimately equate with the frequency of more common phenotypes, they may not actually be responsible for those phenotypes [9–11]. Figure 2 provides a graphical representation of this notion and it is discussed in sections 3.4, 4.3 and 4.4). In order to test associations involving rare variations with a particular phenotype, however, one must identify those variations in individuals with the phenotype of interest and this requires efficient, high-throughput DNA sequencing technologies, which are indeed emerging to accommodate studies of rare variations [9–11]. In fact, DNA sequencing technologies are now so advanced that the entire genomes of individuals have been sequenced, and it is thought that at some point individuals will be completely sequenced in genetic studies and not just genotyped at representative or at a minimal set of variations [12].

1.3. Candidate genes vs. genome wide strategies

In addition to differences and motivations behind different genetic research strategies (e.g., those that leverage families, DNA sequencing methods, etc.) there are also differences in the overall focus of genetic studies. Many studies consider the role of variations in specific genes with regard to influencing phenotypic expression. These studies are often referred to as 'candidate' gene studies since they assume that the gene under study is an appropriate one to study for the phenotype in question. Obviously, a candidate gene is only as good as the logic behind its consideration as a candidate. There are many good candidate genes and candidate gene sets to be considered in human genetic studies of aging as described in other articles in this volume. One of the first candidate gene studies of human aging considered specific variations in the ACE and APOE genes and their association with centenarian status [13]. Another prominent candidate gene study in the literature found an association with a favorable lipoprotein phenotype and a CETP variant in healthy nonagenarians [14]. Candidate gene studies are often plagued by the need to limit the studies to previously identified variations in those genes; e.g., those variations that had been catalogued as part of large-scale human genetic variation studies such as the International HapMap Project (www.hapmap.org). This will likely change as sequencing technologies become more efficient, as researchers are likely to simply sequence the candidate genes in individuals with and without a phenotype to identify as many variations as possible that might influence phenotypic expression. There is increasing precedent for such studies [9–11].

In comparison to candidate gene studies, 'genome-wide' studies do not necessarily emphasize the role of any one gene or set of genes but rather more or less 'anonymize' the genetic variations populating the genome and test each one for linkage or association with a phenotype of interest. Although these studies have intrinsic appeal, they are generally more costly than candidate gene studies (though this cost is rapidly decreasing) and require very strict assumptions about what might be considered a statistically significant linkage or association given that hundreds-of-thousands to millions of variations might be tested for linkage or association with a phenotype of interest [5;15]. It is possible to combine biological insights about genes and genomic regions with genome-wide approaches to the discovery of variations that influence a particular phenotype and hence avoid purely

statistical criteria for identifying linkages and associations. Aspects of these studies are considered in section 4.3.

2. Basic Genetic Research Strategies

2.1 Establishing heritability

In order to establish the potential that genetic strategies might have to identify DNA sequence variations that influence a particular phenotype, it is important to assess the heritability of the phenotype. Measures of heritability estimate the fraction of the variability in a phenotype's expression that can be attributed to inherited genetic variations and are often based on the similarity of genetically related individuals (e.g., twins, siblings, parents and offspring, etc.) relative to the similarity of unrelated or very distantly related individuals. If a gradient in phenotypic similarity can be found that appears to be a function of the relatedness of the individuals studied, then one can infer that genetic factors influence the phenotype. A number of studies investigating the heritability of human age-related phenotypes have been pursued. For example, classical twin study designs applied to human longevity suggest that a substantial fraction of human longevity is genetically determined [16–17], although there are issues surrounding the definition of longevity and aging-related phenotypes in general [18].

A study of individuals in large pedigrees from Iceland also suggests that human lifespan does indeed have a sizable familial component [19], some fraction of which is likely to be genetic. Essentially, they determined the probability that individuals related to target long-lived individuals would themselves be long-lived relative to someone who is unrelated to the long-lived target individuals. They found, for example, that the probability that a second sibling is long-lived (based on being in the 95th age percentile) given that the first sibling is long-lived, divided by the probability that the second sibling is long-lived without any information about the first sibling (the population frequency of longevity) is 1.8 [19]. Thus, siblings of long-lived individuals are 1.8 times more likely to be long-lived than randomly chosen individuals. As noted, there are shortcomings of heritability estimates of human aging-related phenotypes, such as the role of gene - environment interactions, sex-specific factors, epigenetic mechanisms, and, importantly, the consistency and replicability of findings [18;20]. Despite this, the consensus evidence of published articles on the heritability of human longevity and age-related phenotypes clearly indicates a role for genetic factors in human aging.

2.2. Family-based approaches

Family-based approaches to the identification of genomic regions harboring variations that influence lifespan and lifespan-related phenotypes can be problematic since it is probable that long-lived individuals identified for a study will not have living parents or grandparents. Sibling studies, however, offer one family-based method that has been used for estimating a genetic component for longevity related phenotypes [21–22]. For example, Perls and colleagues compared survival rates of over 2,000 siblings of centenarians with 1900 birth cohort survival data. Compared with the U.S. 1900 birth cohort, male siblings of centenarians were at least 17 times as likely to attain age 100 themselves, while female siblings were at least 8 times as likely. In addition, it is also possible to use other family-based designs to study age-related diseases or longevity 'endophenotypes'. Briefly, endophenotypes, also known as intermediate phenotypes, are heritable, quantitative traits hypothesized to represent genetic risk for polygenic traits at more biologically tractable levels than distal behavioral or clinical phenotypes [23]. Endophenotypes are discussed in greater depth in sections 3.2 and 3.3.

One of the best examples of family-based linkage studies applied to an aging-related phenotype are a series of studies investigating the genetic basis of Werner's syndrome, a remarkable syndrome whose hallmark is extremely premature aging, such that individuals with Werner's syndrome who are in their teen years essentially have bodies and age-related features consistent with most 50–80 year olds. Initial family-based studies suggested that a region of chromosome 8p12 likely harbors an autosomal recessive gene variant [24] responsible for Werner's syndrome. This led to a series of studies looking at patterns of homozygosity among genetic variations in the 8p12 region that were likely to be in linkage disequilibrium with the variation causing the syndrome in individuals with the syndrome across different families. The variations showed a very clear pattern indicative of the presence of a causal variation [25–26] and ultimately led to pursuit of sequencing studies revealing the variation responsible for Werner's syndrome [27].

2.3. Genome-wide association (GWA) studies and linkage disequilibrium

As introduced above, genome-wide association (GWA) studies interrogate the entire genome by leveraging data assembled by studies such as the HapMap Project, which characterized the LD relationships among common variants in several human reference populations. The data provided by the HapMap project, coupled with technological advances, have led to the commercialization of array technologies that enable genotyping of hundreds of thousands to around a million single nucleotide polymorphisms (SNP) across the genome on a single individual. Refinements of these technologies led to the use of selected SNP probes to accurately represent the vast majority of known common variants that populate the human genome (typically cut off at 5% or sometimes 1% minor allele frequency in reference population sets). Alleles at loci not directly interrogated on the arrays, but known in the reference populations, can be inferred in genotyped cohorts by statistical means (i.e., via imputation methods) based on knowledge of LD relationships. GWA study approaches compare the frequencies of genotypes obtained using these platforms between subjects with and without the phenotype of interest (i.e., cases and controls) or across cohorts exhibiting different levels of a quantifiable phenotype (i.e., quantitative trait analysis). These types of comparisons are done in order to identify variants or genomic regions associated with a trait assuming suitable statistical correction for the large multiple testing undertaken (this includes the number of variants tested as well as the number of hypotheses or models) and consideration and correction for underlying population structure (variation in ethnic ancestry) in the cohort.

To date, a small number of GWA studies have been conducted with longevity-related phenotypes. For example, an early GWA study focused on age at death and morbidity-free survival at age 65 years (albeit an age that is by current standards too young to be considered “long-lived” or reflective of healthy aging) with 1,345 individuals from 330 families [28]. This study utilized a SNP array that is by today's standards quite sparse (i.e., the Affymetrix 100k SNP GeneChip), and findings did not indicate any associations that reached genome-wide significance. Further, a more recent meta-analysis of four genome-wide association studies of survival to age 90 years or older was performed [29]. In total, this study included 1,836 individuals in the 90+ age group and 1,955 individuals who died between the ages of 55 and 80 years as the comparison group. Although the study identified some regions of interest, they also did not find any associations that reached genome-wide significance [29]. These studies and others like them have led some to evaluate the power of GWA study case-control designs for longevity-related phenotypes. For example, Tan and colleagues used the Danish life tables and simulations to assess the empirical power for different sample sizes when cases are defined as centenarians or as nonagenarians [1]. Their results indicated that power is drastically reduced when nonagenarians are considered cases, with a more than 5-fold difference in the size of the case sample required to achieve comparable power as that

found with centenarians. This suggests that phenotypic definition of aging-related phenotypes (as discussed further in section 3) is of paramount importance.

GWA studies clearly allow for examination of human genetic variation at an unprecedented level of resolution and afford researchers the opportunity to conduct an unbiased “scan” of the genome to find variants associated with common complex traits and diseases such as longevity. This is a major advantage over traditional candidate gene studies, which in essence require an a priori “educated guess” about which genes, as well as which variations within those genes, are associated with the phenotype of interest. These advantages, however, come with a price, which is to some degree exemplified by the studies cited above. Namely, because of the multiple testing burden, large sample sizes (i.e., a thousand or more depending on how the phenotype is defined) are required for well-powered GWA studies, and there is still an enormous potential for false positive results. Further, studies with individuals from ancestral groups not well-represented among HapMap reference populations can be problematic, as can issues related to population stratification in general. Finally, current GWA study technologies are also not optimal for examining forms of variation beyond SNPs, and in some cases allele copy number variation. These include rare variants that are by definition not measured with current SNP array technologies. Indeed, as discussed previously in section 1.2, recent studies have called into question the ‘common-disease common-variant’ hypothesis in favor of the notion that phenotypes might be influenced by many rare variants that collectively contribute to phenotypes that are relatively frequent in the population at large [30]. The inability of current SNP arrays to measure these rare variants, coupled with the likelihood that they may underlie many salient human diseases and traits, suggests the need for genetic strategies with even higher resolution such as DNA sequencing.

2.4. DNA sequencing: the ultimate resolution for genetics research

DNA sequencing is another approach for investigating the genetic basis of disease and has, with recent advances in relevant technology, become more refined, cost-effective, and available. While this approach is also not without limitations (e.g., it is currently limited to searching candidate genes), it has the enormous benefit of providing information about all of the different forms of common, as well as rare, genetic variation within the genome, including SNPs, insertion/deletion variations, copy number variations, inversion polymorphisms, and unique *de novo* single base mutations. In the aging literature, although there is little precedent for DNA sequencing on a large scale, one recent example of this approach is a study by Halaschek-Wiener and colleagues in which they sequenced 24 candidate aging genes in healthy adults aged 85 years or older. Of note, 41% of the variants they identified were not previously recorded in existing reference databases [31]. This suggests that previous genetic strategies such as GWAS and candidate gene studies are likely unable to detect much of the genetic variation that underlies complex diseases and phenotypes, and that DNA sequencing will be critical in this regard. As mentioned in section 1.2, DNA sequencing technologies are now so advanced that the entire genomes of a small number of individuals have been sequenced; it is thought that at some point individuals will be completely sequenced as part of large-scale genetic studies and not just genotyped at representative or at a minimal set of variations [12]. Finally, although with DNA sequencing researchers will likely be better equipped to pinpoint the genetic variation responsible for specific phenotypes, this will still leave the need for functional characterization of the variants identified. Indeed, advances in this area will be necessary for drug and other treatment development efforts.

3. Defining the aging phenotype for genetic studies

3.1. Longevity, life span, and disease

As noted, aging is complex process, influenced by the interaction of genetic, environmental, epigenetic, and purely stochastic factors [32]. This complexity makes it difficult to identify each and every genetic factor contributing to the aging process since the effects of any one genetic factor may be obscured by the effects of other factors. This complexity raises questions about how to study genetic factors that influence human aging since it is difficult to identify phenotypes that capture all the important elements of the aging process. Thus, for example, a genetic researcher focusing on human aging may consider linkage and/or association studies that focus on contrasting long-lived individuals with individuals who died at an early age, which raises questions about what age one should use to define 'long-lived.' In this context, Martin and colleagues review three basic approaches to the genetic analysis of the biology of human aging [33]. The first approach they describe involves the search for genes that contribute to the susceptibilities to certain aging-related disorders, including a range of progeroid syndromes such as Werner's syndrome. The second approach seeks to identify genetic factors that contribute to exceptionally long life spans. The third involves longitudinal studies of the genetic contribution to the decline in various physiological functions among related individuals. Each of these study designs, they argue, gets at different sets of genetic factors contributing to the aging process and human longevity.

The complexity of the aging process also raises questions about why certain individuals reach advanced ages and whether there should be an intrinsic interest not in genes that influence human longevity, but rather genes that influence human 'health span,' or individuals who have lived an exceptionally long *and* healthy life. Nadeau and Topol [34] and Glatt [35] discuss the genetic basis of health, health maintenance, and 'successful' aging as something that should be studied not just for its intrinsic interest, but also in order to arrive at insights into genetic mechanisms that enhance overall longevity. It is important, however, to recognize that long-lived individuals are long-lived for a number of reasons, some genetic and some not. For example, long-lived individuals may be long-lived because: 1. they do not possess disease predisposing genetic variations and hence do not develop life-threatening diseases to the same degree as other individuals; 2. they live in health-enhancing environments and/or engage in health-enhancing lifestyles; 3. they do possess disease-predisposing variations or live in unhealthy environments but also possess 'protective' genetic variations that mitigate the deleterious effects of other genes or environmental conditions; or 4. they have some combination of these [36–37].

Longevity association studies of unrelated individuals also present the need to use an appropriate control group. When selecting a control group, a number of issues are important to consider, perhaps most importantly, matching for environmental exposures, as well as for genetic background or ancestry [38]. One aspect of matching for environmental exposure often involves selecting controls matched with cases based on birth year, but who died due to non-accidental causes prior to reaching the stated "long-lived" cutoff age. Using younger cohorts or random population controls (versus age-matched) can lead to the confounding effects of cohort-specific characteristics. When the phenotype being studied is disease-free survival (discussed below) rather than simply longevity, some studies have utilized age-matched spouses of cases (assuming these individuals are not disease free) as an even more precise way of matching for environmental exposure. Regarding population stratification, when cases and controls are not matched based on genetic ancestry, this well known problem can occur and can produce misleading results. In addition, genetic background differences between different centenarian studies is an important reason why individual

studies can produce valid results that are not reproducible between studies (i.e., in different cohorts).

3.2. Disease-free survival

Since individuals who have lived a long and healthy life have clearly been able to avoid disease, it is arguable that by identifying disease-predisposing genetic variations one has, in effect, identified variations contributing to 'longevity' since the wild-type or alternative alleles or variations to those associated with disease are by default associated with health. As noted, however, individuals that have avoided disease could possess genetic enhancements or protective variants that allow them to exhibit 'disease free survival.' Such variations may simply mediate disease susceptibility and/or mediate basic or intrinsic mechanisms of aging (i.e., senescence [34]). There may also be variants that predispose to healthy aging that display more complex modalities of action, such as antagonistic pleiotropy [39].

The study of disease-free survival as opposed to lifespan is not trivial, since technically one would have to screen the eligible older individuals for a study with respect to subclinical signs of disease. Important insights, however, into the connections between disease-free survival, longevity, and basic mechanisms of aging could be obtained by focusing on individuals related to long-lived individuals, as suggested in Martin and colleagues third approach to the study of human aging [33]. For example, a study was recently conducted to examine the disease burden among offspring of centenarians [40]. They found that relative to non-centenarian offspring, centenarian offspring had a 78% lower risk of myocardial infarction ($P < .04$), 83% lower risk of stroke ($P < .004$), and 86% lower risk of developing diabetes mellitus ($P < .005$). They also found that centenarian offspring were 81% less likely to die than the referent cohort during the follow-up. The authors concluded that centenarian offspring possess a physiologic profile that offers them 'cardiovascular advantages' over non-centenarian offspring.

3.3. Endophenotypes

An alternative approach to the identification of genetic factors that overtly influence human lifespan or aging involves the identification of genetic factors that influence phenotypes and processes known to contribute to, or be associated with, aging. There are many 'endophenotypes' of aging whose genetic bases have been considered in the literature. Four promising human aging endophenotypes are telomere length [41–42], bone-related measures [43], age-associated diseases that might not be life-threatening [44], and aspects of cognitive function [45]. The recent study by Kulminski [44] combined aspects of the disease-free survival phenotype and endophenotype approach to the assessment of genetic factors that contribute to human longevity. The authors focused on 32 geriatric diseases among aging Danish twins. Interestingly, the authors found that geriatric diseases measured in some family members can predict life span in the other family members.

The use of endophenotypes has been pursued in the study of the genetic basis of many diseases, such as the study of lipid levels and heart disease [46], as well as glucose, insulin levels and diabetes [47]. This suggests that there is potential for their use in the identification of genetic factors that contribute to human lifespan and aging.

4. Extending genetic studies

4.1. Gene × Environment interactions

Environmental factors, such as smoking, diet, use of medication, and others, all contribute to human longevity. Taking them into consideration in genetic studies of human lifespan and aging is therefore of extreme importance. It may also be the case that certain genetic

variations exacerbate, or have effects that are exacerbated by, the effects of environmental factors on aging. Thus, for example, individuals with a genetically compromised system for handling oxidative stress could age at higher rates than individuals without a compromised system when chronically exposed to environmental or dietary oxidants [48]. Phenomena such as this suggest the existence of an interaction between the environmental exposure and a genetic variation or set of genetic variations. Evidence for interactions like this have been found for lung cancer, whereby smokers with a particular genetic variation are at markedly elevated non-additive risk for lung cancer than either individuals without the variation who smoke or individuals with the variation who do not smoke [49]. Because of the large sample sizes needed to detect, with sufficient statistical power, an interaction effect, consortia large-scale studies have been initiated to facilitate the identification of gene \times environment interactions in aging, such as the GENEVA consortium [50] and the AGES-Reykjavik study, which focuses on gene - environment interactions that specifically influence longevity and aging [51].

4.2. Multivariate phenotypes

Aging and many age-related diseases involve multiple physiological systems and molecular pathways. The ability to probe and characterize multiple phenotypes that capture these systems and pathways could be of extreme value in genetic studies of aging. Recent methodological developments in the genetic analysis of multiple phenotypes suggest that the use of multiple phenotypes to define a condition can add power to genetic studies (e.g., [52]). One obvious way in which this increase in power is realized is that multivariate analyses reduce the number of statistical tests performed (i.e., because many phenotypes are being tested simultaneously rather than individually). The development of multivariate profiles for aging research is in its infancy, but one promising area involves the examination of multiple phenotypic trajectories that reflect healthy aging [53].

4.3. Integrating functional genomic annotations

Statistical evidence for linkages or association of a particular genetic variation or set of variations to a phenotype does not imply a causal relationship between that variation and the phenotype. Solidifying the causal biological relationship between a sequence variation and a phenotype is complicated and can involve a great deal of laboratory experimentation. Such experimentation would be even more costly if a number of genetic variations showed evidence for linkage or association in a study. Therefore, leveraging information that has been previously obtained on the biological functions of particular genomic regions and/or specific genetic variations could expedite the search for biologically meaningful genetic linkages and associations. Resources such as the University of California Santa Cruz (UCSC) human genome browser (<http://genome.ucsc.edu>) provide extensive information about the functional elements that populate the human genome and can be used to make sense of linkages and associations. Figure 3 provides a snapshot of the UCSC genome browser representation of the AKT1 gene region implicated in longevity [54] and includes information about the location of exons within the gene, locations of previously identified genetic variations (SNPs), empirically-identified protein-DNA interaction sites (e.g., GMI26 H3K4me3; K562 CTCF), and regions of high evolutionary conservation, which might be indicative of function. There are many additional resources and programs for assessing the likely functional significance of individual DNA sequence variations that should be leveraged in linkage and association studies of aging (e.g., [55]).

4.4. Exploiting orthology

One way of identifying candidate genes or genomic regions for a human genetic linkage or association study is to consider orthologous genes and genomic regions in model organisms, which are implicated in the expression of phenotypes analogous to the one of interest in

humans. Determining if a genomic region shown to be linked or associated in a human genetic study has an ortholog that has been implicated in a model organism study can help solidify the biological basis for the linkage or association. There are a growing number of resources and strategies for leveraging non-human species in studies of human aging [56], such as the ANAGE database (<http://genomics.senescence.info/species/>), as well as many examples of such studies [57]. One potential problem in the use of orthologous genomic regions is that, depending on how remote the evolutionary connection with humans is, a certain species may have many genes or genomic regions that show sufficient homology with a human gene or genomic region to be considered orthologs. Thus, it may not be clear which genes or genomic regions provide a reasonable counterpart for the human gene or genomic region of interest.

5. Concluding remarks and the future of genetic research into aging

Research strategies for exploring the genetic basis of human aging will increase in sophistication as technological advances in relevant assays are made. However, sophisticated assays for identifying and querying DNA sequence variations in a large number of individuals will not be sufficient to unlock the secrets of human aging, as it is becoming increasingly clear that greater attention to phenotypes that reflect different aspects of the aging process will also be needed. It is also generally acknowledged that the number of genetic pathways that contribute to the aging process is quite large [58] suggesting that specialists, including systems biologists, with sufficient understanding of each of these pathways will be necessary to understand the biological basis of a linkage or association. These facts call for more integrated approaches to the study of human aging, such as the approach(es) considered by large consortia like the Longevity Consortium (<http://www.longevityconsortium.org/>).

Research Highlights for: Contemporary Human Genetic Strategies in Aging Research

- Human aging is a complex, multifactorial process influenced by a number of genetic and non-genetic factors.
- There are many strategies for elucidating the genetic underpinnings of aging and aging-related phenotypes.
- Extensions of this area of research, including incorporation of gene \times environment interactions, multivariate phenotypes, integration of functional genomic annotations, and exploitation of orthology are critical to advancing this field.

Acknowledgments

CSB and NJS are supported in part by the following research grants: The National Institute on Aging Longevity Consortium [grant number U19 AG023122-01]; The NIMH-funded Genetic Association Information Network Study of Bipolar Disorder National [grant number 1 R01 MH078151-01A1]; National Institutes of Health grants: grant numbers N01 MH22005, U01 DA024417-01, P50 MH081755-01; the Scripps Translational Sciences Institute Clinical Translational Science Award [grant number UL1 RR025447], the Price Foundation and Scripps Genomic Medicine.

References

- [1]. Tan Q, Zhao JH, Zhang D, Kruse TA, Christensen K. Power for genetic association study of human longevity using the case-control design. *Am J Epidemiol* 2008;168:890–896. [PubMed: 18756013]

- [2]. Manolio TA. Study designs to enhance identification of genetic factors in healthy aging. *Nutr Rev* 2007;65:S228–233. [PubMed: 18240554]
- [3]. Fallin MD, Mattheini A. Genetic epidemiology in aging research. *J Gerontol A Biol Sci Med Sci* 2009;64:47–60. [PubMed: 19168782]
- [4]. Lander ES, Weinberg RA. Genomics: journey to the center of biology. *Science* 2000;287:1777–1782. [PubMed: 10755930]
- [5]. Lander ES, Schork NJ. Genetic dissection of complex traits. *Science* 1994;265:2037–2048. [PubMed: 8091226]
- [6]. Frazer KA, Ballinger DG, Cox DR, Hinds DA, Stuve LL, Gibbs RA, Belmont JW, Boudreau A, Hardenl P, Leal SM, Pasternak S, Wheeler DA, Willis TD, Yu F, Yang H, Zeng C, Gao Y, Hu H, Hu W, Li C, Lin W, Liu S, Pan H, Tang X, Wang J, Wang W, Yu J, Zhang B, Zhang Q, Zhao H, Zhou J, Gabriel SB, Barry R, Blumenstiel B, Camargo A, Defelice M, Faggart M, Goyette M, Gupta S, Moore J, Nguyen H, Onofrio RC, Parkin M, Roy J, Stahl E, Winchester E, Ziaugra L, Altshuler D, Shen Y, Yao Z, Huang W, Chu X, He Y, Jin L, Liu Y, Sun W, Wang H, Wang Y, Xiong X, Xu L, Wayne MM, Tsui SK, Xue H, Wong JT, Galver LM, Fan JB, Gunderson K, Murray SS, Oliphant AR, Chee MS, Montpetit A, Chagnon F, Ferretti V, Leboeuf M, Olivier JF, Phillips MS, Roumy S, Sallee C, Verner A, Hudson TJ, Kwok PY, Cai D, Koboldt DC, Miller RD, Pawlikowska L, Taillon-Miller P, Xiao M, Tsui LC, Mak W, Song YQ, Tam PK, Nakamura Y, Kawaguchi T, Kitamoto T, Morizono T, Nagashima A, Ohnishi Y, Sekine A, Tanaka T, Tsunoda T, Deloukas P, Bird CP, Delgado M, Dermitzakis ET, Gwilliam R, Hunt S, Morrison J, Powell D, Stranger BE, Whittaker P, Bentley DR, Daly MJ, de Bakker PI, Barrett J, Chretien YR, Maller J, McCarroll S, Patterson N, Pe'er I, Price A, Purcell S, Richter DJ, Sabeti P, Saxena R, Schaffner SF, Sham PC, Varilly P, Stein LD, Krishnan L, Smith AV, Tello-Ruiz MK, Thorisson GA, Chakravarti A, Chen PE, Cutler DJ, Kashuk CS, Lin S, Abecasis GR, Guan W, Li Y, Munro HM, Qin ZS, Thomas DJ, McVean G, Auton A, Bottolo L, Cardin N, Eyheramendy S, Freeman C, Marchini J, Myers S, Spencer C, Stephens M, Donnelly P, Cardon LR, Clarke G, Evans DM, Morris AP, Weir BS, Mullikin JC, Sherry ST, Feolo M, Skol A, Zhang H, Matsuda I, Fukushima Y, Macer DR, Suda E, Rotimi CN, Adebamowo CA, Ajayi I, Aniagwu T, Marshall PA, Nkwodimmah C, Royal CD, Leppert MF, Dixon M, Peiffer A, Qiu R, Kent A, Kato K, Niikawa N, Adewole IF, Knoppers BM, Foster MW, Clayton EW, Watkin J, Muzny D, Nazareth L, Sodergren E, Weinstock GM, Yakub I, Birren BW, Wilson RK, Fulton LL, Rogers J, Burton J, Carter NP, Clee CM, Griffiths M, Jones MC, McLay K, Plumb RW, Ross MT, Sims SK, Willey DL, Chen Z, Han H, Kang L, Godbout M, Wallenburg JC, L'Archeveque P, Bellemare G, Saeki K, An D, Fu H, Li Q, Wang Z, Wang R, Holden AL, Brooks LD, McEwen JE, Guyer MS, Wang VO, Peterson JL, Shi M, Spiegel J, Sung LM, Zacharia LF, Collins FS, Kennedy K, Jamieson R, Stewart J. A second generation human haplotype map of over 3.1 million SNPs. *Nature* 2007;449:851–861. [PubMed: 17943122]
- [7]. Collins FS, Green ED, Guttmacher AE, Guyer MS. A vision for the future of genomics research. *Nature* 2003;422:835–847. [PubMed: 12695777]
- [8]. Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorf LA, Hunter DJ, McCarthy MI, Ramos EM, Cardon LR, Chakravarti A, Cho JH, Guttmacher AE, Kong A, Kruglyak L, Mardis E, Rotimi CN, Slatkin M, Valle D, Whittemore AS, Boehnke M, Clark AG, Eichler EE, Gibson G, Haines JL, Mackay TF, McCarroll SA, Visscher PM. Finding the missing heritability of complex diseases. *Nature* 2009;461:747–753. [PubMed: 19812666]
- [9]. Bodmer W, Bonilla C. Common and rare variants in multifactorial susceptibility to common diseases. *Nat Genet* 2008;40:695–701. [PubMed: 18509313]
- [10]. Schork NJ, Murray SS, Frazer KA, Topol EJ. Common vs. rare allele hypotheses for complex diseases. *Curr Opin Genet Dev* 2009;19:212–219. [PubMed: 19481926]
- [11]. Frazer KA, Murray SS, Schork NJ, Topol EJ. Human genetic variation and its contribution to complex traits. *Nat Rev Genet* 2009;10:241–251. [PubMed: 19293820]
- [12]. Mir KU. Sequencing genomes: from individuals to populations. *Brief Funct Genomic Proteomic* 2009;8:367–378. [PubMed: 19808932]
- [13]. Schachter F, Faure-Delanef L, Guenot F, Rouger H, Froguel P, Lesueur-Ginot L, Cohen D. Genetic associations with human longevity at the APOE and ACE loci. *Nat Genet* 1994;6:29–32. [PubMed: 8136829]

- [14]. Barzilai N, Atzmon G, Schechter C, Schaefer EJ, Cupples AL, Lipton R, Cheng S, Shuldiner AR. Unique lipoprotein phenotype and genotype associated with exceptional longevity. *Jama* 2003;290:2030–2040. [PubMed: 14559957]
- [15]. Manolio TA, Brooks LD, Collins FS. A HapMap harvest of insights into the genetics of common disease. *J Clin Invest* 2008;118:1590–1605. [PubMed: 18451988]
- [16]. Ljungquist B, Berg S, Lanke J, McClearn GE, Pedersen NL. The effect of genetic factors for longevity: a comparison of identical and fraternal twins in the Swedish Twin Registry. *J Gerontol A Biol Sci Med Sci* 1998;53:M441–446. [PubMed: 9823748]
- [17]. Herskind AM, McGue M, Holm NV, Sorensen TI, Harvald B, Vaupel JW. The heritability of human longevity: a population-based study of 2872 Danish twin pairs born 1870–1900. *Hum Genet* 1996;97:319–323. [PubMed: 8786073]
- [18]. Guo SW. Does higher concordance in monozygotic twins than in dizygotic twins suggest a genetic component? *Hum Hered* 2001;51:121–132. [PubMed: 11173963]
- [19]. Gudmundsson H, Gudbjartsson DF, Frigge M, Gulcher JR, Stefansson K. Inheritance of human longevity in Iceland. *Eur J Hum Genet* 2000;8:743–749. [PubMed: 11039573]
- [20]. Cournil A, Kirkwood TB. If you would live long, choose your parents well. *Trends Genet* 2001;17:233–235. [PubMed: 11335017]
- [21]. Perls TT, Wilmoth J, Levenson R, Drinkwater M, Cohen M, Bogan H, Joyce E, Brewster S, Kunkel L, Puca A. Life-long sustained mortality advantage of siblings of centenarians. *Proc Natl Acad Sci U S A* 2002;99:8442–8447. [PubMed: 12060785]
- [22]. Kerber RA, O'Brien E, Smith KR, Cawthon RM. Familial excess longevity in Utah genealogies. *J Gerontol A Biol Sci Med Sci* 2001;56:B130–139. [PubMed: 11253150]
- [23]. Bearden CE, Freimer NB. Endophenotypes for psychiatric disorders: ready for primetime? *Trends Genet* 2006;22:306–313. [PubMed: 16697071]
- [24]. Goto M, Rubenstein M, Weber J, Woods K, Drayna D. Genetic linkage of Werner's syndrome to five markers on chromosome 8. *Nature* 1992;355:735–738. [PubMed: 1741060]
- [25]. Schellenberg GD, Martin GM, Wijsman EM, Nakura J, Miki T, Ogihara T. Homozygosity mapping and Werner's syndrome. *Lancet* 1992;339:1002. [PubMed: 1348795]
- [26]. Nakura J, Wijsman EM, Miki T, Kamino K, Yu CE, Oshima J, Fukuchi K, Weber JL, Piussan C, Melaragno MI, et al. Homozygosity mapping of the Werner syndrome locus (WRN). *Genomics* 1994;23:600–608. [PubMed: 7851888]
- [27]. Yu CE, Oshima J, Fu YH, Wijsman EM, Hisama F, Alisch R, Matthews S, Nakura J, Miki T, Ouais S, Martin GM, Mulligan J, Schellenberg GD. Positional cloning of the Werner's syndrome gene. *Science* 1996;272:258–262. [PubMed: 8602509]
- [28]. Lunetta KL, D'Agostino RB, Karasik D Sr, Benjamin EJ, Guo CY, Govindaraju R, Kiel DP, Kelly-Hayes M, Massaro JM, Pencina MJ, Seshadri S, Murabito JM. Genetic correlates of longevity and selected age-related phenotypes: a genome-wide association study in the Framingham Study. *BMC Med Genet* 2007;8(Suppl 1.):S13. [PubMed: 17903295]
- [29]. Newman AB, Walter S, Lunetta KL, Garcia ME, Slagboom PE, Christensen K, Arnold AM, Aspelund T, Aulchenko YS, Benjamin EJ, Christiansen L, D'Agostino RB, Fitzpatrick AL Sr, Franceschini N, Glazer NL, Gudnason V, Hofman A, Kaplan R, Karasik D, Kelly-Hayes M, Kiel DP, Launer LJ, Marcianti KD, Massaro JM, Miljkovic I, Nalls MA, Hernandez D, Psaty BM, Rivadeneira F, Rotter J, Seshadri S, Smith AV, Taylor KD, Tiemeier H, Uh HW, Uitterlinden AG, Vaupel JW, Walston J, Westendorp RG, Harris TB, Lumley T, van Duijn CM, Murabito JM. A Meta-analysis of Four Genome-Wide Association Studies of Survival to Age 90 Years or Older: The Cohorts for Heart and Aging Research in Genomic Epidemiology Consortium. *J Gerontol A Biol Sci Med Sci*. 2010
- [30]. Dickson SP, Wang K, Krantz I, Hakonarson H, Goldstein DB. Rare variants create synthetic genome-wide associations. *PLoS Biol* 2010;8:e1000294. [PubMed: 20126254]
- [31]. Halaschek-Wiener J, Amirabbasi-Beik M, Monfared N, Pieczyk M, Sailer C, Kollar A, Thomas R, Agalaridis G, Yamada S, Oliveira L, Collins JA, Meneilly G, Marra MA, Madden KM, Le ND, Connors JM, Brooks-Wilson AR. Genetic variation in healthy oldest-old. *PLoS One* 2009;4:e6641. [PubMed: 19680556]

- [32]. Martin GM. Keynote lecture: an update on the what, why and how questions of ageing. *Exp Gerontol* 2006;41:460–463. [PubMed: 16621387]
- [33]. Martin GM, Bergman A, Barzilai N. Genetic determinants of human health span and life span: Progress and new opportunities. *Plos Genetics* 2007;3:1121–1130.
- [34]. Nadeau JH, Topol EJ. The genetics of health. *Nat Genet* 2006;38:1095–1098. [PubMed: 17006459]
- [35]. Glatt SJ, Chayavichitsilp P, Depp C, Schork NJ, Jeste DV. Successful aging: from phenotype to genotype. *Biol Psychiatry* 2007;62:282–293. [PubMed: 17210144]
- [36]. Perls T, Kunkel LM, Puca AA. The genetics of exceptional human longevity. *J Am Geriatr Soc* 2002;50:359–368. [PubMed: 12028221]
- [37]. Perls T, Terry D. Understanding the determinants of exceptional longevity. *Ann Intern Med* 2003;139:445–449. [PubMed: 12965974]
- [38]. Christensen K, Johnson TE, Vaupel JW. The quest for genetic determinants of human longevity: challenges and insights. *Nat Rev Genet* 2006;7:436–448. [PubMed: 16708071]
- [39]. Martin GM. Modalities of gene action predicted by the classical evolutionary biological theory of ageing. *Ann N Y Acad Sci* 2007;1100:14–20. [PubMed: 17460162]
- [40]. Adams ER, Nolan VG, Andersen SL, Perls TT, Terry DF. Centenarian offspring: start healthier and stay healthier. *J Am Geriatr Soc* 2008;56:2089–2092. [PubMed: 18811609]
- [41]. Kappei D, Londono-Vallejo JA. Telomere length inheritance and aging. *Mech Ageing Dev* 2008;129:17–26. [PubMed: 18054991]
- [42]. Atzmon G, Barzilai N, Surks MI, Gabriely I. Genetic predisposition to elevated serum thyrotropin is associated with exceptional longevity. *J Clin Endocrinol Metab* 2009;94:4768–4775. [PubMed: 19837933]
- [43]. Cauley JA, Lui LY, Barnes D, Ensrud KE, Zmuda JM, Hillier TA, Hochberg MC, Schwartz AV, Yaffe K, Cummings SR, Newman AB. Successful skeletal aging: a marker of low fracture risk and longevity. The Study of Osteoporotic Fractures (SOF). *J Bone Miner Res* 2009;24:134–143. [PubMed: 18715137]
- [44]. Kulminski AM, Arbeev KG, Culminkaya IV, Ukraintseva SV, Christensen K, Yashin AI. Health-related phenotypes and longevity in danish twins. *J Gerontol A Biol Sci Med Sci* 2009;64:1–8. [PubMed: 19211549]
- [45]. Yashin AI, Arbeev KG, Kulminski A, Borecki I, Christensen K, Barmada M, Hadley E, Rossi W, Lee JH, Cheng R, Elo IT. “Predicting” parental longevity from offspring endophenotypes: data from the Long Life Family Study (LLFS). *Mech Ageing Dev* 131:215–222. [PubMed: 20184914]
- [46]. Clarke R, Peden JF, Hopewell JC, Kyriakou T, Goel A, Heath SC, Parish S, Barlera S, Franzosi MG, Rust S, Bennett D, Silveira A, Malarstig A, Green FR, Lathrop M, Gigante B, Leander K, de Faire U, Seedorf U, Hamsten A, Collins R, Watkins H, Farrall M. Genetic variants associated with Lp(a) lipoprotein level and coronary disease. *N Engl J Med* 2009;361:2518–2528. [PubMed: 20032323]
- [47]. Dupuis J, Langenberg C, Prokopenko I, Saxena R, Soranzo N, Jackson AU, Wheeler E, Glazer NL, Bouatia-Naji N, Gloyn AL, Lindgren CM, Magi R, Morris AP, Randall J, Johnson T, Elliott P, Rybin D, Thorleifsson G, Steinthorsdottir V, Henneman P, Grallert H, Dehghan A, Hottenga JJ, Franklin CS, Navarro P, Song K, Goel A, Perry JR, Egan JM, Lajunen T, Grarup N, Sparso T, Doney A, Voight BF, Stringham HM, Li M, Kanoni S, Shrader P, Cavalcanti-Proenca C, Kumari M, Qi L, Timpson NJ, Gieger C, Zabena C, Rocheleau G, Ingelsson E, An P, O’Connell J, Luan J, Elliott A, McCarroll SA, Payne F, Roccascocca RM, Pattou F, Sethupathy P, Ardlie K, Ariyurek Y, Balkau B, Barter P, Beilby JP, Ben-Shlomo Y, Benediktsson R, Bennett AJ, Bergmann S, Bochud M, Boerwinkle E, Bonnefond A, Bonnycastle LL, Borch-Johnsen K, Bottcher Y, Brunner E, Bumpstead SJ, Charpentier G, Chen YD, Chines P, Clarke R, Coin LJ, Cooper MN, Cornelis M, Crawford G, Crisponi L, Day IN, de Geus EJ, Delplanque J, Dina C, Erdos MR, Fedson AC, Fischer-Rosinsky A, Forouhi NG, Fox CS, Frants R, Franzosi MG, Galan P, Goodarzi MO, Graessler J, Groves CJ, Grundy S, Gwilliam R, Gyllenstein U, Hadjadj S, Hallmans G, Hammond N, Han X, Hartikainen AL, Hassanali N, Hayward C, Heath SC, Hercberg S, Herder C, Hicks AA, Hillman DR, Hingorani AD, Hofman A, Hui J, Hung J, Isomaa B, Johnson PR, Jorgensen T, Jula A, Kaakinen M, Kaprio J, Kesaniemi YA, Kivimaki M, Knight B, Koskinen S, Kovacs P, Kyvik KO, Lathrop GM, Lawlor DA, Le Bacquer O, Lecoeur C, Li Y,

Lyssenko V, Mahley R, Mangino M, Manning AK, Martinez-Larrad MT, McAteer JB, McCulloch LJ, McPherson R, Meisinger C, Melzer D, Meyre D, Mitchell BD, Morken MA, Mukherjee S, Naitza S, Narisu N, Neville MJ, Oostra BA, Orru M, Pakyz R, Palmer CN, Paolisso G, Pattaro C, Pearson D, Peden JF, Pedersen NL, Perola M, Pfeiffer AF, Pichler I, Polasek O, Posthuma D, Potter SC, Pouta A, Province MA, Psaty BM, Rathmann W, Rayner NW, Rice K, Ripatti S, Rivadeneira F, Roden M, Rolandsson O, Sandbaek A, Sandhu M, Sanna S, Sayer AA, Scheet P, Scott LJ, Seedorf U, Sharp SJ, Shields B, Sigurethsson G, Sijbrands EJ, Silveira A, Simpson L, Singleton A, Smith NL, Sovio U, Swift A, Syddall H, Syvanen AC, Tanaka T, Thorand B, Tichet J, Tonjes A, Tuomi T, Uitterlinden AG, van Dijk KW, van Hoek M, Varma D, Visvikis-Siest S, Vitart V, Vogelzangs N, Waeber G, Wagner PJ, Walley A, Walters GB, Ward KL, Watkins H, Weedon MN, Wild SH, Willemsen G, Witteman JC, Yarnell JW, Zeggini E, Zelenika D, Zethelius B, Zhai G, Zhao JH, Zillikens MC, Borecki IB, Loos RJ, Meneton P, Magnusson PK, Nathan DM, Williams GH, Hattersley AT, Silander K, Salomaa V, Smith GD, Bornstein SR, Schwarz P, Spranger J, Karpe F, Shuldiner AR, Cooper C, Dedoussis GV, Serrano-Rios M, Morris AD, Lind L, Palmer LJ, Hu FB, Franks PW, Ebrahim S, Marmot M, Kao WH, Pankow JS, Sampson MJ, Kuusisto J, Laakso M, Hansen T, Pedersen O, Pramstaller PP, Wichmann HE, Illig T, Rudan I, Wright AF, Stumvoll M, Campbell H, Wilson JF, Bergman RN, Buchanan TA, Collins FS, Mohlke KL, Tuomilehto J, Valle TT, Altshuler D, Rotter JI, Siscovick DS, Penninx BW, Boomsma DI, Deloukas P, Spector TD, Frayling TM, Ferrucci L, Kong A, Thorsteinsdottir U, Stefansson K, van Duijn CM, Aulchenko YS, Cao A, Scuteri A, Schlessinger D, Uda M, Ruokonen A, Jarvelin MR, Waterworth DM, Vollenweider P, Peltonen L, Mooser V, Abecasis GR, Wareham NJ, Sladek R, Froguel P, Watanabe RM, Meigs JB, Groop L, Boehnke M, McCarthy MI, Florez JC, Barroso I. New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. *Nat Genet* 2010;42:105–116. [PubMed: 20081858]

- [48]. Migliore L, Coppede F. Environmental-induced oxidative stress in neurodegenerative disorders and aging. *Mutat Res* 2009;674:73–84. [PubMed: 18952194]
- [49]. Amos CI, Wu X, Broderick P, Gorlov IP, Gu J, Eisen T, Dong Q, Zhang Q, Gu X, Vijayakrishnan J, Sullivan K, Matakidou A, Wang Y, Mills G, Doheny K, Tsai YY, Chen WV, Shete S, Spitz MR, Houlston RS. Genome-wide association scan of tag SNPs identifies a susceptibility locus for lung cancer at 15q25.1. *Nat Genet* 2008;40:616–622. [PubMed: 18385676]
- [50]. Cornelis MC, Agrawal A, Cole JW, Hansel NN, Barnes KC, Beaty TH, Bennett SN, Bierut LJ, Boerwinkle E, Doheny KF, Feenstra B, Feingold E, Fornage M, Haiman CA, Harris EL, Hayes MG, Heit JA, Hu FB, Kang JH, Laurie CC, Ling H, Manolio TA, Marazita ML, Mathias RA, Mirel DB, Paschall J, Pasquale LR, Pugh EW, Rice JP, Udren J, van Dam RM, Wang X, Wiggs JL, Williams K, Yu K. The gene, environment association studies consortium (GENEVA): maximizing the knowledge obtained from GWAS by collaboration across studies of multiple conditions. *Genet Epidemiol.* 2010
- [51]. Harris TB, Launer LJ, Eiriksdottir G, Kjartansson O, Jonsson PV, Sigurdsson G, Thorgeirsson G, Aspelund T, Garcia ME, Cotch MF, Hoffman HJ, Gudnason V. Age, Gene/Environment Susceptibility-Reykjavik Study: multidisciplinary applied phenomics. *Am J Epidemiol* 2007;165:1076–1087. [PubMed: 17351290]
- [52]. Ferreira MA, Purcell SM. A multivariate test of association. *Bioinformatics* 2009;25:132–133. [PubMed: 19019849]
- [53]. Franco OH, Karnik K, Osborne G, Ordovas JM, Catt M, van der Ouderaa F. Changing course in ageing research: The healthy ageing phenotype. *Maturitas* 2009;63:13–19. [PubMed: 19282116]
- [54]. Pawlikowska L, Hu D, Huntsman S, Sung A, Chu C, Chen J, Joyner AH, Schork NJ, Hsueh WC, Reiner AP, Psaty BM, Atzmon G, Barzilay N, Cummings SR, Browner WS, Kwok PY, Ziv E. Association of common genetic variation in the insulin/IGF1 signaling pathway with human longevity. *Aging Cell* 2009;8:460–472. [PubMed: 19489743]
- [55]. Plumptre M.; Barnes MR. Predictive functional analysis of polymorphisms: An overview. In: Barnes MR., editor. *Bioinformatics for Geneticists*. Second ed.. John Wiley and Sons; New York: 2007.
- [56]. Austad SN. Methusaleh's Zoo: how nature provides us with clues for extending human health span. *J Comp Pathol* 2010;142(Suppl 1):S10–21. [PubMed: 19962715]

- [57]. Curran SP, Ruvkun G. Lifespan regulation by evolutionarily conserved genes essential for viability. *PLoS Genet* 2007;3:e56. [PubMed: 17411345]
- [58]. Bostock CV, Soiza RL, Whalley LJ. Genetic determinants of ageing processes and diseases in later life. *Maturitas* 2009;62:225–229. [PubMed: 19203848]
- [59]. Mangino M, Richards JB, Soranzo N, Zhai G, Aviv A, Valdes AM, Samani NJ, Deloukas P, Spector TD. A genome-wide association study identifies a novel locus on chromosome 18q12.2 influencing white cell telomere length. *J Med Genet* 2009;46:451–454. [PubMed: 19359265]
- [60]. Liu YZ, Guo YF, Wang L, Tan LJ, Liu XG, Pei YF, Yan H, Xiong DH, Deng FY, Yu N, Zhang YP, Zhang L, Lei SF, Chen XD, Liu HB, Zhu XZ, Levy S, Papasian CJ, Drees BM, Hamilton JJ, Recker RR, Deng HW. Genome-wide association analyses identify SPOCK as a key novel gene underlying age at menarche. *PLoS Genet* 2009;5:e1000420. [PubMed: 19282985]
- [61]. Gravina S, Lescai F, Hurteau G, Brock GJ, Saramaki A, Salvioli S, Franceschi C, Roninson IB. Identification of single nucleotide polymorphisms in the p21 (CDKN1A) gene and correlations with longevity in the Italian population. *Ageing (Albany NY)* 2009;1:470–480. [PubMed: 20126416]
- [62]. Costa MD, Cherni L, Fernandes V, Freitas F, Ammar El Gaaied AB, Pereira L. Data from complete mtDNA sequencing of Tunisian centenarians: testing haplogroup association and the “golden mean” to longevity. *Mech Ageing Dev* 2009;130:222–226. [PubMed: 19133286]
- [63]. Atzmon G, Rincon M, Schechter CB, Shuldiner AR, Lipton RB, Bergman A, Barzilai N. Lipoprotein genotype and conserved pathway for exceptional longevity in humans. *PLoS Biol* 2006;4:e113. [PubMed: 16602826]
- [64]. Flachsbar F, Caliebe A, Kleindorp R, Blanche H, von Eller-Eberstein H, Nikolaus S, Schreiber S, Nebel A. Association of FOXO3A variation with human longevity confirmed in German centenarians. *Proc Natl Acad Sci U S A* 2009;106:2700–2705. [PubMed: 19196970]

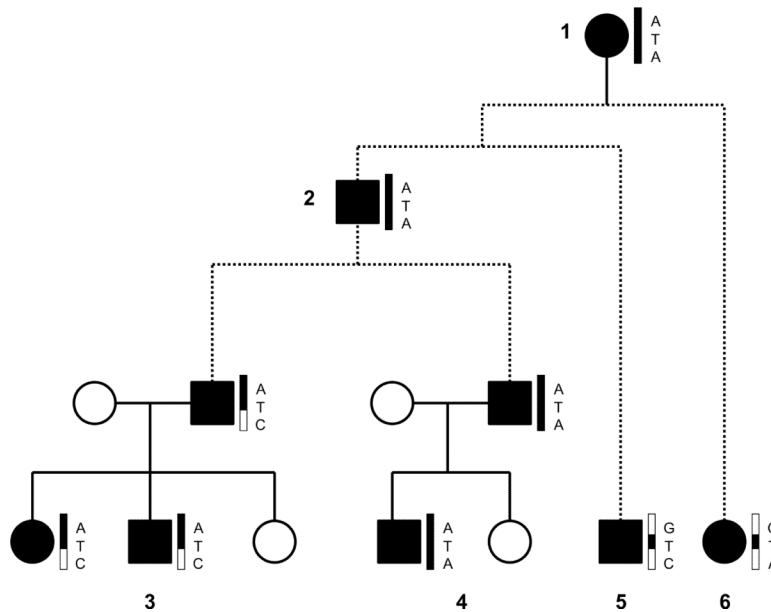


Figure 1.

Diagrammatic representation of the hypothetical transmission of a genetic variant influencing lifespan from a distant relative, as well as chromosomal markers flanking the variant. Solid lines denote exact generational links, whereas dotted lines denote lines of descent involving a large number of generations. Shaded pedigree members have a specific lifespan phenotype. Vertical rectangles next to pedigree member symbols reflect a single chromosome and letters next to the rectangle reflect nucleotides at SNP loci. Shading within the rectangles indicates chromosomal material from the founder (denoted by the number 1) who introduced the life span-influencing nucleotide variation 'T' into the population. This founder had 'A' nucleotides at SNP positions flanking the site of the life span-influencing locus. The entire founder haplotype 'A-T-A' was preserved during transmission to individual 2. However, recombination events during meioses in individuals with the line of descent from individual 1 to individuals 5 and 6 caused individuals 5 and 6 to only inherit the 'T' allele associated with the life span-influencing phenotype. Due to a recombination event in the line of descent from individual 2 to the father in nuclear family 3, the father in nuclear family 3 only inherited the 'A-T' haplotype encompassing the 1st two markers from ancestor 2 but not the ancestral founder allele 'A' at the third locus. However, this 'A-T' haplotype was preserved in the transmission of the chromosome harboring the 'T' life span-influencing variant to two offspring in the nuclear family. The father in nuclear family 5 inherited the entire founder 'A-T-A' haplotype and transmitted it in its entirety to one of his offspring due to a lack of recombination. Thus, although two siblings share part ('A-T') of the ancestral founder chromosomal haplotype harboring the 'T' life span-influencing nucleotide ('A-T-A'), and an offspring in nuclear family 4 inherited the entire 'A-T-A' haplotype, all the individuals in the latest generation with the life span phenotype only share the 'T' ancestral functional nucleotide inherited from individual 1.

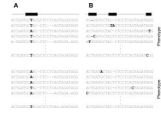
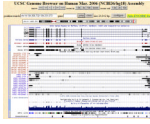


Figure 2.

Two hypothetical situations arising from sequencing studies pursued in order to identify variations associated with a particular phenotype. Letters refer to adjacent nucleotides in haploid sequence obtained from individuals with and without a particular longevity-related phenotype. Dashes indicate deletions or absent nucleotides. Shaded rectangles above the sequence denote functional elements within the sequence. In Figure 2a, a common variant, in bold, is associated with the presence/absence of the phenotype. Note that due to phenocopies, incomplete penetrance, and other factors influencing the phenotype, some, but not all individuals with/without the phenotype have exactly the same nucleotide at the position of the common functional variant. In Figure 2b a series of rare variations (in bold) in functional regions of the sequence are greater in frequency among the individuals with the phenotype. Note that some individuals without the phenotype also possess rare variations but they are not in functional regions and hence are not likely to influence the phenotype.

**Figure 3.**

Functional annotation of the AKT1 gene obtained from the UCSC genome browser (<http://genome.ucsc.edu/>; chromosome 14, base positions: 104,306,732-104,331,573) recently shown to harbor variations associated with human longevity [54]. The figure reveals the amount of information amassed about the AKT1 gene. The base locations and positions of coding regions and known mRNAs and spliced transcripts are depicted in the upper portions of the figure (e.g., 'Refseq genes' and 'Human ESTs'). The middle portions of the figure provide the locations of empirically identified regulatory elements in the region including enhancers and silencers obtained from protein-DNA binding studies involving different transcription factors (e.g., 'CTCF' and 'H3K4me1'). Cross species nucleotide conservation levels, the locations of known single nucleotide polymorphisms, and the locations of repeated sequences are provided in the lower portions of the figure.

Table 1

Example Human Genetic Studies of Aging, Longevity, Aging Syndromes and Relevant Endophenotypes

Phenotype	Reference	Study Type	Results
Telomere length	[42]	GWAS	Telomerase variations associated with telomere length
Healthy aging	[31]	Gene sequencing	Many variations identified in candidate genes
Telomere length	[59]	GWAS	Variations in the 18q12.2 region marginally associated
Healthy Nonagenarians	[14]	Candidate gene	Association with favorable lipoprotein phenotype and CETP variant
Nonagenarians	[54]	Candidate gene	AKT1 & FOXO3A variants associated/replicated in 3 cohorts
Age at Menarche	[60]	GWAS	SPOCK gene associated and replicated in 2 cohorts
Nonagenarians	[61]	Gene sequencing	Rare alleles in CDKN1A underrepresented in oldest
Centenarians	[62]	mtDNA sequencing	No lack of mtDNA mutations among Centenarians
Centenarians	[13]	Candidate gene	APOE and ACE variants associated with longevity
Centenarians	[63]	Candidate gene	APOC3 variant associated with longevity
Centenarians	[64]	Candidate gene	FOXO3A associated/replicated in 2 cohorts
Centenarian families	[42]	Candidate gene	TSHR variants/TSH levels associated with longevity
Lifespan in families	[28]	GWAS	No variations associated with lifespan/age-related traits
Werner's syndrome	[27]	Positional cloning	Identification of the Werner's syndrome gene
Werner's syndrome	[24]	Linkage analysis	Strong homozygosity-based linkage to chromosome 8