The quest for genetic determinants of human longevity: challenges and insights

Kaare Christensen, Thomas E. Johnson‡ and James W. Vaupel§*

Abstract | Twin studies show that genetic differences account for about a quarter of the variance in adult human lifespan. Common polymorphisms that have a modest effect on lifespan have been identified in one gene, *APOE*, providing hope that other genetic determinants can be uncovered. However, although variants with substantial beneficial effects have been proposed to exist and several candidates have been put forward, their effects have yet to be confirmed. Human studies of longevity face numerous theoretical and logistical challenges, as the determinants of lifespan are extraordinarily complex. However, large-scale linkage studies of long-lived families, longitudinal candidate-gene association studies and the development of analytical methods provide the potential for future progress.

Cohort
A designated group of individuals who are studied over a time period.

**Epidemiology, Institute of Public Health, University of Southern Denmark, J.B. Winslows Vej 9B, 5000 Odense C, Denmark.*
‡Institute for Behavioral Genetics, Department of Integrative Physiology, University of Colorado, Boulder, Colorado 80309-0447, USA.
§Max Planck Institute for Demographic Research, Konrad-Zuse-Strasse, 18057 Rostock, Germany.
Correspondence to K.C.
e-mail: KChristensen@health.sdu.dk
doi:10.1038/nrg1871

Human lifespan is of vital importance, both for individuals and society. In the past century, most Western countries have experienced large increases in mean life expectancy, from around 50 years to around 75–80 years. This has been due to a marked reduction in early life mortality during the first half of the twentieth century, followed by a less recognized almost twofold reduction in mortality at ages above 70 in the past 50 years[1]. These changes are rapid on an evolutionary timescale, and suggest that important factors causing variation in lifespan are unlikely to be of genetic origin. However, within a given birth cohort in a given country there is still a large variation in lifespan (FIG. 1). Clarifying to what extent this variation is related to genetic differences among individuals and understanding the roles of specific genetic factors in this variation is central to the understanding of human ageing and lifespan, including exceptionally long lifespan, which is known as longevity. The ultimate aim of this research is to provide targets that can be used in the prevention and treatment of disabilities and diseases that occur with increasing age.

Lifespan is the outcome of complicated processes that might involve thousands of genes and non-genetic factors. Many researchers prefer to study specific diseases that are associated with increased mortality, and some argue that even these outcomes are too broad to study. They suggest instead that the focus should be on intermediate phenotypes that predispose to disease, or on physiological outcomes that vary with age and predict lifespan. However, findings from animal studies have provided evidence that individual genes can have a significant effect on lifespan. Furthermore, human genetic studies have shown that common polymorphisms in one gene — apolipoprotein E (*APOE*) — influence lifespan, probably mainly through their association with disease[2]. This, together with familial recurrence patterns for longevity, has given cause for optimism that it will be possible to identify other genetic variants that affect lifespan. However, although the past few years have seen the identification of many candidate genes for involvement in human lifespan, only the role of *APOE* has so far been consistently confirmed.

Here we provide an overview of the genetics of human lifespan and discuss how the genetic factors that underlie variation in lifespan might be successfully identified in the future. Many challenges face researchers trying to identify genetic variants that are associated with human longevity, and we examine different approaches that can be taken to overcome them, along with a discussion of the most promising candidate genes that have been investigated so far. We begin with a brief overview of the study of longevity in animal models, from which most of our understanding of the biology of longevity has originated, and discuss how this knowledge can be applied to the study of human lifespan.

Theories and insights from model organisms

A short lifespan and the ability to control both environment and genotype have made invertebrates especially useful for studying genetic variants that are associated with longevity[3,4]. A substantial proportion of the variation in invertebrate lifespan is heritable (estimates range from 20% to 50%) under controlled environmental conditions in the laboratory[5], and hundreds of genetic variants that lead to life extension have been identified (see the Science of Aging Knowledge Environment web site and REFS 3,4). Single mutants in *Caenorhabditis elegans* can reduce mortality threefold[6,7] and combinations of variants lead to as much as a sixfold extension in lifespan, increasing to almost eightfold when combined with dietary restriction[8].

Several general facts about the biology of longevity have become clear from animal models (BOX 1): most of the genes involved are pleiotropic, specify increased stress resistance and result in increased robustness in older animals. Animal studies have also revealed a down side to increased lifespan: many long-lived mutants are slow-growing, with reduced fecundity and fertility, and fail to compete in a changing environment[9]. There seem to be direct trade-offs between higher fertility and rapid development on one hand, and increased stress resistance and longer lifespan on the other[10]. Finally, animal studies have revealed that there is a major stochastic component to lifespan, such that genetically identical individuals that are grown in a common environment do not have the same lifespan, which is one reason why the heritability of lifespan is moderate[11,12].

Animal studies have also provided insights into the types of gene that can be involved in the regulation of lifespan. The first longevity mutant to be identified was the *C. elegans* gene *age-1* (REF. 13) that encodes phosphatidylinositol 3-kinase (PI3K) (REF. 14), which has a key role in a signalling pathway that is homologous to the mammalian insulin–IGF1 (insulin-like growth factor 1) pathway (FIG. 2). This pathway ultimately targets the transcription factor DAF-16 (FOXO), which regulates the expression of numerous downstream genes that mediate stress resistance, innate immunity, metabolic processes and toxin degradation[15,16]. Mutations that affect this pathway show notable effects on longevity in both invertebrates and mammals; several mouse longevity mutants alter key components of the insulin–IGF1 pathway, with one of the strongest lines of evidence being the increased lifespan of mice that are heterozygous for the IGF1 receptor knockout[17].

A second large class of life-extension mutants in the nematode affect mitochondrial function, the so-called Mit mutants. Starting with the identification of *clk-1*, and now involving about a hundred distinct loci, numerous Mit mutations result in life extension, typically of 20–40% and sometimes more[3,18]. Many of these mutants interact with the insulin–IGF1 pathway mutants to cause life extension beyond that observed in single-gene mutants alone[3,4].

Longevity genes have also been identified in other animal models. Two key examples are *sir-2* and *Tor* (*Target of rapamycin*), which were identified in yeast and *Drosophila melanogaster*, respectively. *sir-2* encodes an NAD-dependent protein deacetylase, which might mediate the lifespan-extending effects of dietary restriction, whereas *Tor* encodes a protein that is involved in sensing amino-acid availability (FIG. 2). The effects of these genes on lifespan indicate a link between nutrient intake and longevity, and both genes might be involved in the life-extension effects that are mediated by dietary restriction. Other longevity genes, including *methuselah* (*mth*) and *I'm not dead yet* (*Indy*) (originally identified in flies), and klotho (*Kl*) (the only such gene identified first in mice), are the subjects of intensive investigation, although their specific roles in modulating lifespan have yet to be determined.

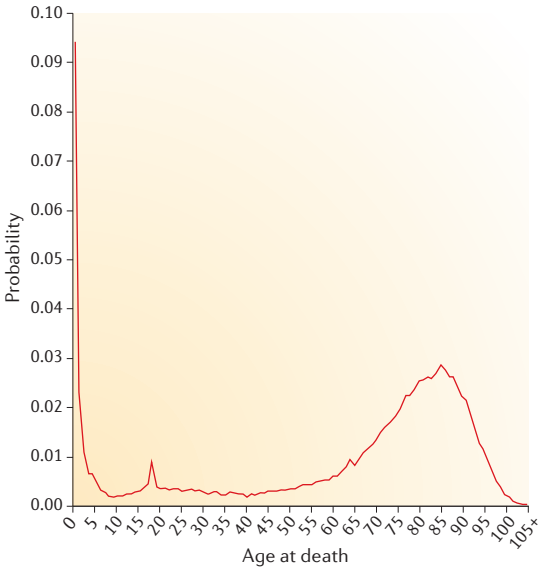

Figure 1 | **Large variation in lifespan within a birth cohort.** The distribution of age at death is shown for a twentieth century Western population that did not experience any world wars — the Swedish female 1900 birth cohort. Data are from The Human Mortality Database.

Caution should be used when investigating human candidate genes that are identified by their orthology to those that have been highlighted by animal studies, as it is unlikely that every longevity gene found in model organisms will correspond to a human longevity locus. For example, in *C. elegans* even disruption of homologues of genes that cause severe human disease, such as frataxin (*FXN*), can result in life extension[19].

Importantly, animal studies have shown that mortality is affected at some ages but not all[7]: the *age-1* mutant reduces late-life mortality as much as 14-fold, but *spe-10* mutants only lower reproductive phase mortality and *clk-1* only affects late-life mortality[7], so human longevity genes might also be age-specific.

In providing these insights, invertebrate studies have motivated the search for human genes that are involved in longevity and have provided candidate genes, while also revealing challenges that must be kept in mind when carrying out these studies.

Pleiotropy
The action of a single gene on two or more distinct phenotypic characters.

Box 1 | Generalities from studies of lifespan-increasing invertebrate genes

- All the longevity genes that have been identified have primary roles in other physiological processes and especially in signal transduction (FIG. 2). It therefore seems that natural selection does not select for genes that cause ageing, but rather ageing occurs as a result of pleiotropic effects of genes that specify other processes[24].
- Most life-extension effects have been found to result from hypomorphic or nullomorphic mutations, which can be interpreted to mean that the wild-type gene shortens lifespan under laboratory conditions. Such genes specify a process that has a negative effect on longevity, and therefore blocking their expression increases longevity. These genes might be called 'gerontogenes' and should be distinguished from 'longevity-assurance genes' for which nullomorphs result in life shortening.
- Where tested, gerontogene mutants show decreased 'fitness' and fail to compete with wild-type animals. These mutants show trade-offs between fitness components such as speed of adaptation to a new environment or fertility schedule[9,10].
- Most longevity mutations also increase the ability to handle stress, such as oxidative stress and starvation[102]. Stress resistance therefore seems to be a public mechanism of ageing[24,102], that is, one that is shared by different species.
- The longest-lived individuals and strains are also the most robust and disease resistant, showing extensions not only of life but also of health (ability to move and react) well into ages at which wild-type controls are dead[3,11].
- Where mortality has been ascertained using populations of several hundred if not thousands of individuals, longevity mutants can alter either or both initial mortality and the slope of the age-dependent increase in mortality[3,6,7], and mutations can affect mortality at some ages, but not at others. All aspects of longevity and mortality seem to be modulated by genes and contribute to variation in longevity.
- Manipulations of more than 100 genes have been found to increase longevity in *Caenorhabditis elegans*. This is in contrast to initial expectations that a few rate-limiting targets modulate ageing[103].

Human lifespan as a heritable trait

Lifespan phenotypes. Studies of life duration can focus on several phenotypes, which are important to understand before considering genetic studies of human lifespan. The most direct measure is individual lifespan, but this can only be studied directly in extinct (or nearly extinct) cohorts, meaning that cohorts should be born at least 100 years ago. Individual lifespan after adolescence is often studied because infant and childhood deaths are likely to have distinct causes (for example, prematurity or congenital malformations). Lifespan researchers are also usually interested in excluding the effects of sex and cohort — that is, the similarity in lifespan for two same-sex twins that arises from having the same sex and year of birth is not the focus. Therefore, the deviation from the sex-specific and cohort-specific mean is often used.

Early deaths (adult deaths before 50–60 years old, depending on the study) are of particular interest because they represent a loss of many years of life and often have significant social consequences. Because of genetic diseases, such as those that are associated with early-onset cardiovascular disease, one might expect a stronger genetic component to early death compared with death at older ages. On the other hand, violent deaths comprise a higher proportion of early deaths than later deaths, which could reduce the genetic component.

Late deaths (after 90–100 years old, depending on the study) are of interest because they could be a marker of successful ageing. The clustering of late deaths in families with many extremely long-living individuals has provided support for a familial component to longevity[20,21]. But is this genetic? On one hand, the accumulation of unique environmental exposures during a long life might be the main determinant of lifespan and health at older ages, predicting decreased heritability at older ages[22]. Alternatively, evolutionary biologists have argued that the reduced selective pressure against deleterious genetic mutations that are expressed only late in life predicts an increase in genetic variance among the oldest[23,24].

Finally, age-specific susceptibility to death, known as frailty, can be studied. Frailty is likely to have a higher heritability than lifespan *per se* as it is more plausible that one inherits a level of susceptibility to death than a fixed lifespan[25,26].

Genetic epidemiology of human lifespan. Twin studies have consistently found that for cohorts born around 100 years ago, approximately 25% of the variation in lifespan is caused by genetic differences[27,28] (FIG. 3a). Recent combined analyses of ~20,000 twins born in Nordic countries between 1870 and 1910 confirm this, but they also show that the genetic influences on lifespan are minimal before the age of 60 and only increase after that age. This finding provides support for the search for genes that affect longevity in humans, especially at advanced ages[29]. The results are comparable in the various Nordic countries, but other settings lack similar data to provide heritability estimates. Countries with larger socio-economic differences might be expected to have lower heritability estimates owing to larger environmental variance.

The genetic contribution to phenotypes other than lifespan *per se* has also been studied. Adoption studies have suggested a genetic component to some causes of premature death. The only large adoption study that has been published shows a correlation between Danish adoptees and their biological parents, especially for death that is due to vascular causes[30]. However, a later extension of this study found smaller effects[31], and twin data[29] indicate that the overall genetic effect on premature death is minimal.

There is also evidence that longevity clusters in some families. Perls and co-workers found that the chances of survival until 80–94 years old for siblings of centenarians were about four times as high as those for siblings of individuals who died at 73 years of age[20] (FIG. 3b), and even higher values were reported later by the same group[32]. In addition, a study that was based on Mormon genealogies found an increased recurrence risk for siblings for surviving to extreme ages, although the estimate was lower than those from the studies by Perls and colleagues[33]. Similarly, an investigation using the population-based genealogy in Iceland found that first-degree relatives (parents, siblings and offspring) of probands who live to extreme old age (≥95 percentile) are twice as likely as controls to survive to the same age[21]. Finally, Schoenmaker *et al.*[34] found mortality rates to be about 30% lower than in the general population for first-degree relatives of exceptionally long-lived siblings in Holland. However, such studies can only provide an upper limit for the genetic influence, because clustering can be due to both genetic factors and a shared family environment.

Recurrence risk
The likelihood that a given condition that is diagnosed in one or more family members will recur in other family members or in subsequent generations.

Proband
A subject who is ascertained on the basis of their phenotype; probands are often used to identify affected families for genetic studies.

In terms of frailty, a study of Nordic twin pairs estimated the heritability of this trait as approximately 50% (REF. 25). Further analyses[1] suggested that about half the variation in lifespan after 30 years old might be due to survival attributes that are fixed for individuals by the time they reach this age; a third to a half of this effect is predicted to be due to genetic factors, and a half to two-thirds to non-genetic survival attributes (related to, for example, socio-economic status or nutritional and disease history). The model indicates that the importance of survival attributes might increase with life expectancy.

The genetic architecture of human longevity

There are definitely many rare mutations that have large negative effects on lifespan, which are best illustrated by segmental progeriod syndromes that mimic premature ageing, such as the monogenic disorders Werner syndrome[35] and Hutchinson–Gilford disease[36,37]. However, the epidemiological studies described above indicate that common genetic variants with large effects on human longevity are unlikely to exist, as revealed by the low recurrence risk for exceptional longevity within families. One important question is whether families that show clustering of exceptional longevity have rare mutations that are unique to the family and that increase their chances of living to very old ages, and whether these mutations tend to be in the same genes in different families that show this type of clustering. Another key question is how many common genetic variants with moderate effects on lifespan might be identified. The strong evidence for effects of common *APOE* variants on lifespan has generated considerable optimism for finding other common variants. However, the genetic architecture of lifespan probably involves many rare variants with small effects, and the complexity of this phenotypic trait is likely to have contributed significantly to the slow progress in this area.

Genetic study designs in longevity research

Challenges to genetic studies of human lifespan. Epidemiological and demographical analyses have identified numerous factors that are associated with survival at all ages. But why some humans live to extreme ages — some even in relatively good health — is largely unknown. One of the most astonishing results from studies of centenarians is how diverse they are[38]. The few environmental factors that have been shown to be associated with extreme survival are avoidance of heavy smoking and severe obesity, and relatively high educational attainment. In addition, American and Japanese studies have indicated that psychological factors are important, helping centenarians to cope with morbidity and disability. Some studies have indicated that centenarians have escaped major diseases[39], but others have shown that centenarians often have multimorbidity[38,40], indicating that there are multiple ways to achieve exceptional longevity. These factors, together with the probable genetic architecture of human longevity, have influenced the approaches that are used to identify genetic variants that affect this trait. Here we discuss the implications of these

Segmental progeroid syndromes
Syndromes that mimic normal ageing and affect multiple organs and tissues.

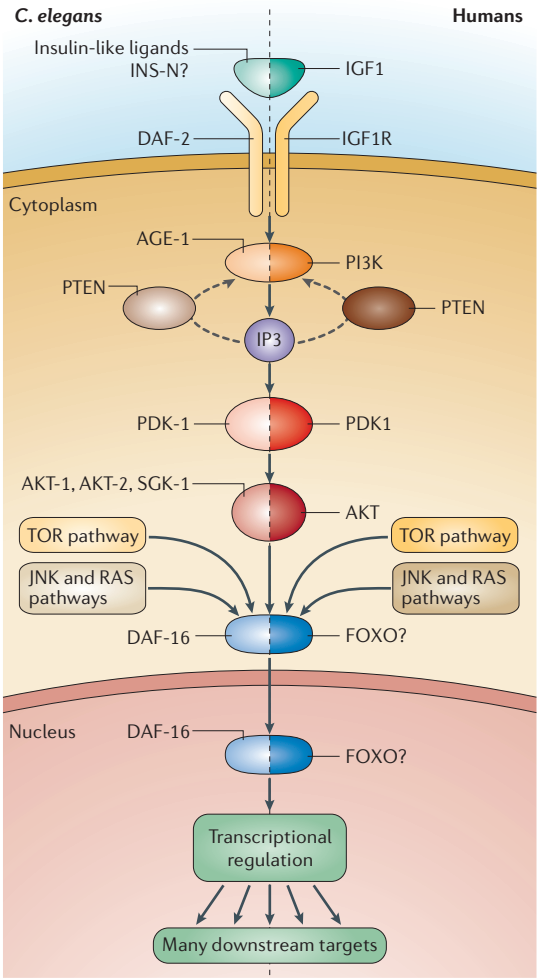

Figure 2 | **Some of the molecular pathways that lengthen lifespan in *Caenorhabditis elegans* and the corresponding components in humans.** The *C. elegans* insulin–IGF-1 (insulin-like growth factor 1) like signal-transduction pathway is shown on the left, and human homologues of the proteins that are involved are shown on the right. This pathway involves a cascade of phosphorylation events that ultimately regulate the nuclear translocation of DAF-16 (REF. 10). INS-N is an unknown insulin-like peptide, and DAF-2 is its cell-surface receptor, which has tyrosine kinase activity. AGE-1 encodes a phosphatidylinositol 3-kinase (PI3K). IP3 is phosphatidylinositol-3,4,5-trisphosphate (for simplicity, it is shown here as part of the pathway, although it is actually a membrane component), which is produced as a result of AGE-1 activity and activates PDK-1. PTEN is a phosphatase with IP3 substrate activity and suppresses AGE-1. PDK-1 is an IP3-dependent kinase that activates AKT-1, AKT-2 and SGK-1, which are serine/threonine kinases. DAF-16 is a forkhead class transcription factor that is homologous to the FOXO class of human transcription factors, and is probably orthologous to human FOXO3A. The target of rapamycin (TOR), JNK and RAS pathways also feed into the insulin-like signalling pathway at the level of DAF-16 regulation. TOR is a kinase that responds to intracellular amino acids, especially leucine, among other activities; RAS and JNK are involved in numerous signal-transduction cascades in mammals. Numerous other genes in which mutations lead to life extension in *C. elegans* (for example, *sir-2* and mitochondrial genes) are not shown.