

## Identifying the genomic determinants of aging and longevity in human population studies: Progress and challenges

Joris Deelen<sup>1)2)</sup>, Marian Beekman<sup>1)2)</sup>, Miriam Capri<sup>3)</sup>, Claudio Franceschi<sup>3)</sup> and P. Eline Slagboom<sup>1)2)\*</sup>

Human lifespan variation is mainly determined by environmental factors, whereas the genetic contribution is 25–30% and expected to be polygenic. Two complementary fields go hand in hand in order to unravel the mechanisms of biological aging: genomic and biomarker research. Explorative and candidate gene studies of the human genome by genetic, transcriptomic, and epigenomic approaches have resulted in the identification of a limited number of interesting positive linkage regions, genes, and pathways that contribute to lifespan variation. The possibilities to further exploit these findings are rapidly increasing through the use of novel technologies, such as next-generation sequencing. Genomic research is progressively being integrated with biomarker studies on aging, including the application of (noninvasive) deep phenotyping and omics data – generated using novel technologies – in a wealth of studies in human populations. Hence, these studies may assist in obtaining a more holistic perspective on the role of the genome in aging and lifespan regulation.

### Keywords:

■ biomarkers; genomics; healthy aging; human; longevity

### Introduction

Human life expectancy has increased remarkably over the last two centuries worldwide [1], although it is still highly variable between countries [2]. This lifespan extension is mainly due to improvement of health care, hygiene, and nutrition. The healthy life expectancy, however, has not increased at the same rate; in Europe, men spend on average 20.5% and women 25.4% of their life dealing with disability caused through disease or injury (Healthy Life Years; <http://www.healthy-life-years.eu/>) [3]. Although age is the main risk factor for the majority of common diseases contributing to disability, reaching an old age does not necessarily result in a higher degree of age-related disability. This is illustrated by the presence of long-lived individuals from families expressing exceptional longevity that may reach high ages without major disabilities [4, 5]. Moreover, their offspring – considered “decelerated” or “healthy agers” – have a lower prevalence of age-related diseases, such as cancer, cardiovascular disease, hypertension, and type 2 diabetes [6–9], compared to similar-aged controls. Concomitantly, they show beneficial or “youthful” profiles for many metabolic and immune-related parameters [10]. Most of the human aging studies are concentrated around long-lived families, including highly and middle-aged members, sporadic highly aged individuals from

DOI 10.1002/bies.201200148

<sup>1)</sup> Section of Molecular Epidemiology, Leiden University Medical Center, Leiden, The Netherlands

<sup>2)</sup> Netherlands Consortium for Healthy Ageing, Leiden University Medical Center, Leiden, The Netherlands

<sup>3)</sup> Department of Specialized, Diagnostic and Experimental Medicine, University of Bologna, Bologna, Italy

### \*Corresponding author:

P. Eline Slagboom  
E-mail: p.slagboom@lumc.nl

### Abbreviations:

**CHARGE**, Cohorts for Heart and Aging Research in Genomic Epidemiology; **CNV**, copy number variant; **FHS**, Framingham Heart Study; **GEHA**, GEnetics of Healthy Ageing; **GH**, growth hormone; **GWAS**, genome-wide association study; **IDEAL**, Integrated research on Developmental determinants of Aging and Longevity; **IGF-1**, insulin-like growth factor 1; **iPOP**, integrative personal omics profile; **LLFS**, Long Life Family Study; **LLS**, Leiden Longevity Study; **LOD**, logarithm (base 10) of odds; **MRI**, magnetic resonance imaging; **mTOR**, mammalian target of rapamycin; **NECS**, New England Centenarian Study; **NHGRI**, National Human Genome Research Institute; **NMR**, nuclear magnetic resonance; **ROH**, region of homozygosity; **SICS**, Southern Italian Centenarian Study; **SNP**, single nucleotide polymorphism.

the general population or population-based cohorts containing different age groups.

Due to the different study designs (Box 1 and Fig. 1), human aging cohorts provide complementary information and are intensively being studied from a biomarker and genomic perspective. The assumption is that, together, these studies will provide insight into the mechanisms that could (i) drive the biological aging rate, (ii) positively and negatively influence the risk for age-related disease, and (iii) explain the variation in lifespan between individuals. Genomic research, including genetic, epigenetic, and transcriptomic studies, is expected to provide both markers and determinants of aging. The search for biomarkers for human

aging and longevity is aimed at identifying parameters and profiles that reflect the biological age of individuals and predict long-term morbidity and/or mortality [11].

For most diseases, like osteoarthritis, osteoporosis, and type 2 diabetes, standardized phenotypes and diagnostic criteria are used for genomic research. No standardized phenotype or marker, however, is indicating biological aging rate. Hence, genomic studies into aging thus far focus on the determinants of human lifespan variation by using age at death, prospective survival, disease-free survival, or exceptional longevity as outcome. Biomarker research is therefore just as relevant for genomic studies of human aging as the analysis of the genome itself.

## Box 1

### Study designs

The ultimate epidemiological study design to investigate markers and determinants of biological aging and longevity in humans would be to follow a large group of individuals during their entire lifetime. These individuals should be examined at different time points so that changes in markers could be related to the actual lifespan of the individual. However, since this design is not feasible, several other designs are being applied in human studies (Fig. 1).

### Cross-sectional study designs

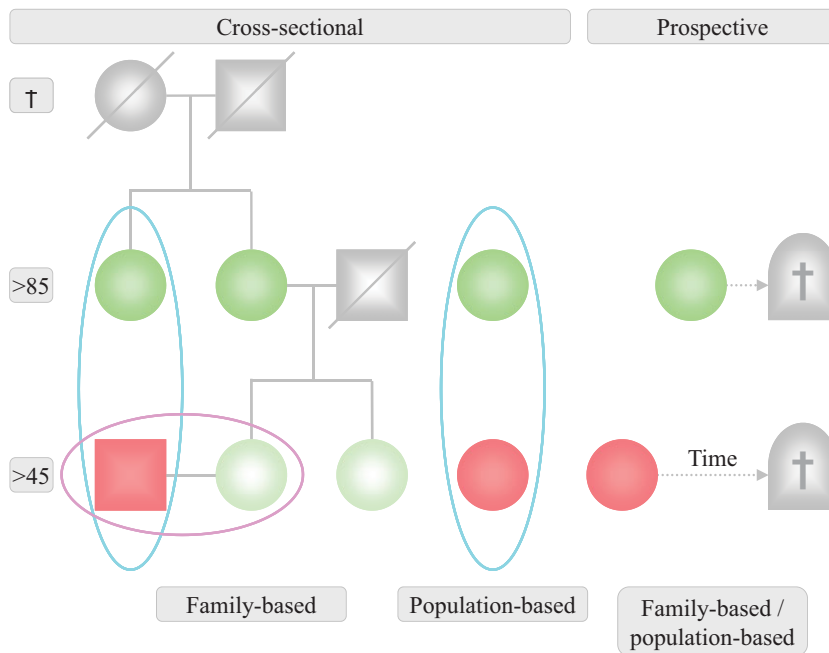
**Population-based cohorts:** Cross-sectional longevity studies typically compare unrelated highly aged individuals (nonagenarians/centenarians) with younger controls or evaluate differences between groups of unrelated individuals in categories of increasing age. Inclusion of individuals for these studies is relatively easy, which is reflected by the large sample sizes of population-based cross-sectional studies. The cross-sectional study usually provides the first level of observation that a parameter is correlated with chronological age or a health condition. However, causality of the genetic and/or genomic parameter on aging and longevity cannot be determined from a cross-sectional design. For cross-sectional studies the long-lived cases should be compared with controls originating from the same birth cohort. However, since these controls usually already died, controls are generally selected from other birth cohorts. Given that these cohorts have a different life expectancy, this could confound the studied association. In addition, structural differences between birth cohorts, caused by, e.g. migration, could also confound the results. Examples of longevity studies used for cross-sectional analysis in unrelated individuals are the New England Centenarian Study (NECS) [90], German long-lived individuals [91], French centenarians [92], and Southern Italian Centenarian Study (SICS) [93]. In addition, various cross-sectional studies are included in the MARK-AGE project, which consists of

2,320 randomly recruited volunteers from the general population (35–74 years).

**Family-based cohorts:** Family-based longevity studies consist of nonagenarians/centenarians (siblings) and their middle-aged offspring. The controls used in these studies are either (age-matched) random individuals from the general population or spouses of the offspring of the long-lived individuals. Due to the common genetic background among family members, family-based longevity studies are enriched for familial and genetic effects on longevity and are more robust against population substructure. However, these studies generally have a small sample size since it is quite difficult to collect long-lived families. To determine which age-related phenotypes associate with human familial longevity, the offspring of long-lived individuals, which are predisposed to longevity, can be compared to geographically- and age-matched population controls. This design allows analysis of molecular and clinical parameters specific for long-lived family members in multiple generations. Examples of family-based longevity studies are the Ashkenazi Jews cohort [71], GEnetics of Healthy Ageing (GEHA) study (of which the offspring is collected in the MARK-AGE project) [94], Long Life Family Study (LLFS) [7], and Leiden Longevity Study (LLS) [15].

### Prospective studies

Most prospective longevity studies consist of highly (>85 years of age) or middle-aged (>55 years of age) individuals (related or unrelated) that are followed over time and sampled at multiple time points. This design is most often applied to provide more evidence for causality of determinants or markers detected in cross-sectional studies. In this design an (unbiased) baseline parameter may show to precede a functional aspect of aging. Several large population-based prospective studies have been initiated. However, the main disadvantage of these studies is that the number of individuals that will become long-lived is usually very small. Examples of prospective longevity studies are the Leiden 85-plus study [95, 96], Newcastle 85+ study [97, 98], Danish 1905 cohort [99], the population-based Rotterdam Study [100], and Framingham Heart Study (FHS) (consisting of three generations) [101].



**Figure 1.** Study designs applied in studies of healthy aging and longevity. Family- or population-based cross-sectional designs usually compare highly aged individuals with younger controls (blue ovals). Alternatively, the offspring of long-lived individuals is compared to age-matched controls (their spouses or random population controls) (purple oval). Thirdly, prospective studies are performed in highly or middle-aged individuals (unrelated or from (long-lived) families) which are followed over time (ranging from 10 to 30 years, depending on the study). Highly aged individuals are depicted in green, their offspring in light green and middle-aged individuals in red.

The possibility to study causal determinants and quantitative biomarkers of biological aging and longevity in humans strongly depends on the study designs that are available (Box 1 and Fig. 1). Using these designs, we determined four relevant phases in aging studies in order to establish whether a quantitative parameter (or profile) is a biomarker of biological age; (i) Determine the change in a quantitative parameter with chronological age in cross-sectional studies and, preferably, by repeated measures in longitudinal studies. Parameters reflecting biological age are expected to show an increased variance with age. (ii) Determine whether a marker of chronological age also discriminates individuals with a youthful or old level relative to their age category in the general population, which would indicate that the quantitative parameter potentially marks biological age (Fig. 2). The comparison between offspring of long-lived individuals and age-matched population controls is also part of this phase. (iii) Determine whether the potential marker for biological age associates to known parameters of health, such as blood pressure, serum levels of glucose, insulin, and cholesterol. (iv) Determine whether the potential marker for biological age associates with morbidity (based on clinical endpoints) and/or mortality in prospective studies.

In this review we will give an overview of the main genomic approaches and discuss the concept of biomarker approaches used in the research field of human aging and longevity. In addition, we will discuss the progress and challenges of integration of data that has been generated using these approaches.

## Genomic research

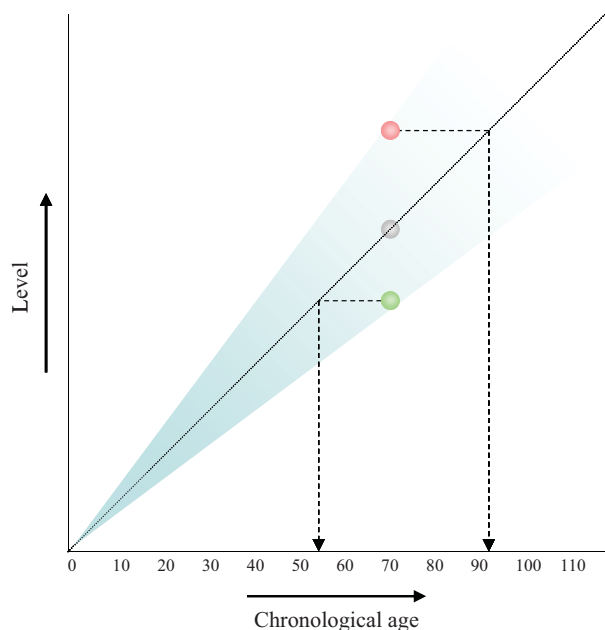
### Human longevity is not just explained by the absence of disease-susceptibility alleles

Genomic research into human lifespan regulation could be subdivided into genetic, epigenetic, and transcriptomic research.

Studies of mono- and dizygous twins have revealed that the genetic contribution to the variation in human lifespan is about 25–30% [12, 13], and is most prominent in families clustered for longevity [14, 15]. This genetic contribution is mainly apparent after the age of 60 years and seems to increase with age [13, 16]. Furthermore, human lifespan is a complex trait which is assumed to be determined by many genes with small individual effects [17], although the polygenic architecture still needs to be characterized [18, 19]. The diverse health features of long-lived families illustrate that different age-related diseases have common determinants and implicate that pathways can be identified that attenuate aging and delay age-related disease. From a genomic perspective, individuals from long-lived families are assumed to be characterized by a decreased prevalence of disease-promoting variants (referred to as disease-susceptibility alleles) and an increased prevalence of variants conferring maintenance of health and protection from disease, when compared to population controls. In the last 5 years, many disease-susceptibility alleles have been identified (National Human Genome Research Institute (NHGRI) genome-wide association study (GWAS) Catalog; <http://www.genome.gov/gwastudies/>) [20]. A first comparison between long-lived individuals, selected from both long-lived families (LLS) and the general population (Leiden 85-plus study), and young controls showed no difference in the distribution or frequency of disease-susceptibility alleles identified in cancer, coronary artery disease and type 2 diabetes [21]. The search for lifespan regulating loci – contributing to longevity and population mortality – must therefore extend beyond a focus on disease-susceptibility alleles. We will first discuss the efforts to identify longevity loci by genetics approaches.

### Candidate gene studies identified *APOE* and *FOXO3A* as human longevity genes

The first genetic longevity studies mainly focused on lifespan regulating loci that emerged from animal models [22]. Lifespan



**Figure 2.** Interpretation of the potential relationship between a marker of chronological age and biological age using categories of increasing age. The blue zone indicates the increasing variance of the marker with age. Individuals can be assigned to having a marker level which matches (i) the expected level for their age in the population (gray dot, 75 years in this example), (ii) the level of a younger age group (green dot, biological age may be lower than chronological age), or (iii) the level of an older age group (red dot, biological age may be higher than chronological age).

extension in animal models was obtained by applying caloric restriction or by modifying gene functions (mutagenesis) using RNA interference, knock-out or overexpression of single genes (GenAge; <http://genomics.senescence.info/genes/>) [23]. The most interesting pathways identified using these models are the growth hormone (GH)/insulin/insulin-like growth factor 1 (IGF-1) signaling and mammalian target of rapamycin (mTOR) signaling pathways [24]. Thus far, lifespan has been the main phenotype investigated in animal models. In order to make these models more translatable to human studies research should focus on defining the parameters that reflect the physiology and pathology of aging in both animals and humans [25, 26].

Most of the human candidate gene studies were performed in cross-sectional designs (Box 1 and Fig. 1), comparing allele frequencies of potential longevity loci between highly aged individuals and young controls. The candidate gene studies based on single genes have pointed a role for genes involved in, e.g., GH/insulin/IGF-1 signaling, immune regulation, and lipoprotein metabolism (Supporting Information Table S1), although most of these results have not (yet) been confirmed in sufficient independent studies. The most convincing human longevity loci today are *APOE* and *FOXO3A* which have frequently been associated with longevity in cross-sectional studies (see for a review [26]) and survival in prospective studies [27–29] (Fig. 3). *APOE* encodes the protein apolipoprotein E

which seems to play a role in e.g., lipoprotein metabolism, cognitive function, and immune regulation [30]. *FOXO3A* encodes the protein forkhead box O3 which acts as a transcription factor for many different genes involved in processes like apoptosis and oxidative stress [31].

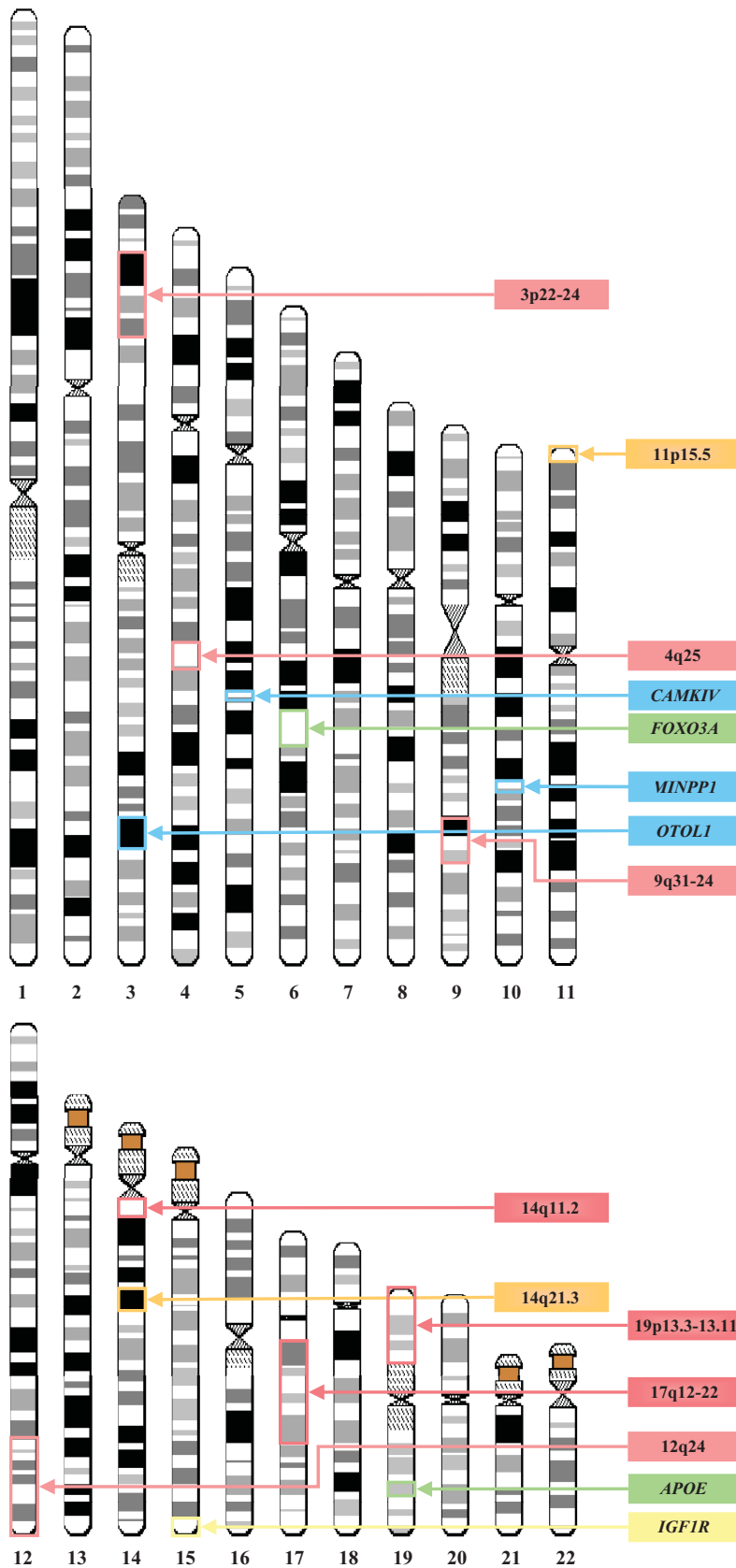
In addition to single gene studies, several candidate gene studies based on whole pathways have been performed. These pathway-based candidate gene studies showed a role for genes within the DNA damage signaling and repair, GH/insulin/IGF-1 signaling, immune regulation, pro/antioxidant, and telomere maintenance pathways [32–36] (Supporting Information Table S1). Most of these pathway-based studies tested for effects of individual single nucleotide polymorphisms (SNPs) on prospective mortality or longevity [32, 34, 35], and, so far, only a limited number of studies determined the joint effect of SNPs within a pathway [33, 36].

### Large meta-GWAS are required for identification of novel human longevity loci

As an alternative to hypothesis-based candidate gene studies, hypothesis-free or explorative approaches could also be applied to studies of the genome. These methods should initially be aimed at prioritizing the location of regions linked to longevity and, subsequently, identifying the genetic variation causal to the trait. One example of an explorative approach is the GWAS. In this cross-sectional approach, in which long-lived individuals are compared with young or shorter-lived controls, the usually small-effect of common variants can be identified. Typically, genotype distributions of 300,000–2,500,000 SNPs are assessed for association with the trait in GWAS. Since longevity is assumed to be determined by many genes with small effects it could be a successful method to identify novel longevity loci. However, so far, GWAS for longevity in the LLS [37], Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) [38, 39], NECS [40], German long-lived individuals [41], and SICS [42] have only identified one genome-wide significant ( $p < 5 \times 10^{-8}$ ) locus: *APOE*, which has long been established as a longevity gene. Several other loci showed suggestive association with longevity ( $p < 5 \times 10^{-6}$ ), namely *MINPP1* [38], *OTOL1* [39], and *CAMKIV* [42] (Fig. 3).

However, the effect of these loci on prospective mortality is not yet known. All GWAS-identified suggestive longevity loci are deleterious, i.e., the minor allele is associated with a decreased probability to become long-lived, and, as expected, their effects are small (odds ratio's  $> 0.5$ ).

In general, to have sufficient power to detect significant effects, GWAS require much larger sample sizes than thus far accomplished for human longevity. One of the challenges of GWAS for longevity is that the lifespan variation induced by the genetic component is expected to be small relative to that induced by the environmental component (i.e., health care and nutrition). A large sample size, acquired through meta-analysis of GWAS (meta-GWAS), may cope with the so-called “phenocopies” and could potentially detect genome-wide significant loci besides *APOE*. Currently, two initiatives for meta-GWAS for longevity are on-going. One consists of ~8,000 long-lived individuals ( $\geq 85$  years of age) from all over Europe (Integrated research on DEvelopmental determinants



**Figure 3.** Karyogram (adapted from <http://hapmap.ncbi.nlm.nih.gov/karyogram/gwas.html>) containing candidate genes whose association with longevity has been replicated in multiple association studies (green), candidate genes with interesting results from sequencing studies (yellow), interesting loci from linkage ( $LOD \geq 2.95$ ) (red), and CNV (orange) studies and loci that showed suggestive association with longevity ( $p \leq 5 \times 10^{-6}$ ) in GWAS (blue).

of Aging and Longevity (IDEAL) GWAS Longevity Study), while the other consists of ~6,000 long-lived individuals ( $\geq 90$  years of age), collected in Northern America and Europe, from the CHARGE consortium. If these meta-GWAS lead to the identification of new loci that significantly associate with longevity, they should consequently be tested for an effect on prospective survival in middle and old age.

### Copy number variant (CNV) studies identified potential longevity regions

Besides SNP analysis, several other methods have been applied to study the genetics of longevity, mainly using a prospective design (Box 1 and Fig. 1). One study determined the effect of CNVs, which are deletions or duplications of stretches of DNA, on longevity in the Rotterdam Study and FHS. The meta-analysis of these cohorts showed an association between the burden of large ( $\geq 500$  kb) CNVs and mortality at old age. In addition, they showed an association of common CNV regions on 11p15.5 and 14q21.3 [43] (Fig. 3). However, to qualify them as longevity-regions, these associations still need to be replicated in several larger independent cross-sectional and prospective studies.

The same group also studied the effect of regions of homozygosity (ROHs), which are uninterrupted stretches of homozygous SNPs, on longevity in the Rotterdam Study and found no association between ROHs and survival into old age [44]. However, to rule out effects of ROHs on longevity larger cross-sectional and prospective studies should be performed.

### Linkage studies have discovered chromosomal regions linked to human longevity

The explorative studies of the genome for longevity effects actually started with linkage analysis in family-based designs (Box 1 and Fig. 1). For this approach, the excess sharing of alleles between siblings identical by descent at 6,000–12,000 loci not in linkage disequilibrium over sharing by chance provides a likelihood for the presence of a longevity locus in any region on the genome. There have been several small-scale genome-wide linkage studies of long-lived sibling pairs ( $n_{\text{cases}} < 300$ ) that showed inconsistent results [45–48] (Fig. 3). Recently, a large linkage analysis for longevity has been performed in 2,118 nonagenarian Caucasian sibling pairs from the GEHA study. In this study, linkage with longevity was observed at chromosome 14q11.2 (logarithm (base 10) of odds (LOD) = 3.47), chromosome 17q12-22 (LOD = 2.95), chromosome 19p13.3-13.11 (LOD = 3.76), and chromosome 19q13.11-13.32 (LOD = 3.57) (Fig. 3), of which the latter was explained by the ApoE  $\epsilon 4$  and ApoE  $\epsilon 2$  alleles [49]. Since the linkage at the remaining loci could not be explained by association of common variants, human familial longevity at these loci may be explained by rare variants.

### Next-generation sequencing studies may reveal rare longevity-associated variants

Rare variants can be identified by applying next-generation (whole-genome or exome) sequencing. In the case of Mendelian disorders and strong familial traits, next-generation sequenc-

ing of a limited number of well-selected individuals may reveal relevant alleles with functional consequences. Analysis of sequencing data is a bioinformatic challenge and good sample selection is therefore extremely important. The most informative individuals for next-generation sequencing in longevity research would be individuals from long-lived families with a long family history of longevity. One candidate gene study analyzed the complete coding region of *IGF1* and *IGF1R* using 2D gene scanning and DNA sequencing in centenarians and their offspring. Two rare nonsynonymous SNPs in *IGF1R* associated with both longevity and decreased IGF-1 signaling. This further indicates a role for GH/insulin/IGF-1 signaling genes in human longevity [50] (Fig. 3).

For exploratory analyses, the whole genome can be analyzed. Up to now, this has been published for one female and one male supercentenarian [51]. To identify variants relevant for longevity, analysis on the genomes of many more of such individuals must be performed. Various initiatives are ongoing in which larger amounts of genomes of population and family-based centenarians are being sequenced, e.g., the Welllderly Study (consisting of ~1,000 individuals  $\geq 80$  years of age) and the LLS (consisting of 220 individuals  $\geq 90$  years of age).

### Explorative studies identify transcriptomic profiles marking longevity

Since the genetic approaches have thus far provided little robust evidence for loci contributing to human aging and longevity, attempts have been made to identify such loci by exploration of the human transcriptome. The transcriptome of an individual reflects the influence of genetic variation, as well as the response to the environment. As an approach to find determinants of aging and longevity, transcriptomic studies require specific designs to disentangle primary and causal changes in gene expression from the consequences of aging.

Most studies of the transcriptome try to identify genes that show a differential change with chronological age and mainly use cross-sectional designs (Box 1 and Fig. 1). In these designs highly aged individuals are compared to young controls or categories of increasing age are examined. The larger studies are performed in whole blood, since this is the most accessible tissue. However, whole blood contains different cell populations, which may confound observed differences in gene expression. If possible, observations of differential gene expression should thus be adjusted for proportions of blood cell subsets, which is not always done. One study partly circumvented this problem by investigating the transcriptome of T cells from healthy individuals with ages ranging from 25 to over 95 years and highlighted similarities in gene expression profiles between young and “successfully aged” individuals [52]. This illustrates that cross-sectional transcriptome studies may be used to identify genes potentially indicative of the biological age of an individual by comparing the expression level of the gene for an individual to the average expression of individuals of his/her chronological age.

The transcriptomic studies focused on chronological age revealed that genes and microRNAs involved in many different processes, e.g., oxidative phosphorylation, complement activation, and synaptic transmission, change with age [53–58]. The pathways that have been associated with chronological

age include peroxisome proliferator-activated receptor, glucose and glutathione metabolism, and mTOR signaling [52]. The relevance of mTOR pathway genes for human aging has been further illustrated by associations of gene expression changes with chronological age in a candidate gene study of two independent human cohorts [56]. Most of the gene expression associations with chronological age in human populations have not yet been validated and replicated with comparable technology platforms in independent studies. In addition, transcriptomic studies on chronological age cannot rate which changes are causal and which are consequential to aging.

One way to overcome (part of) this problem is by using a family-based study design (Box 1 and Fig. 1), in which the offspring of long-lived individuals – representing “healthy agers” – are compared to similar-aged controls from the general population. The differential gene expression profiles identified using this design may represent markers of healthy aging and familial longevity. This approach has been applied in the LLS to explore the transcriptome in whole blood for association with human familial longevity. Genes belonging to the mTOR pathway, as well as *ASF1A* and *IL7R*, were differentially expressed between offspring and controls [59, 60]. In addition, the expression of mTOR genes in blood associated to prevalent diabetes and serum glucose. However, the association with familial longevity was not dependent on this. Thus, gene expression profiles in blood mark human longevity in middle age and potentially provide information on the pathways that contribute to healthy aging and longevity.

### Epigenomic studies are at hand

Another molecular level that could provide additional insight in the processes of aging is the epigenome, the intermediate layer of genomic information between the genome and transcriptome. Epigenetic regulation of transcription is mediated by histon modification, DNA methylation, and microRNAs. Changes in the epigenome with chronological age have been explored and show that methylation patterns of genes involved in e.g., development and morphogenesis, DNA binding and regulation of transcription [61–63] tend to change with age. A recent remarkable finding in a small study sample, confirmed in a cohort of 501 individuals ranging from birth to 99 years, was the progressive linear increase in methylation with age at the *ELOVL2* gene [64]. Because the epigenomic field recently became more accessible for the screening of large study populations, the identification of a new range of epigenetic biomarkers is at hand. To consider such epigenetic measures as markers for biological age, confounding of cell type distributions should be accounted for – like in transcriptomic studies – and effects should be established using various study designs.

In conclusion, up to now, genomic research to identify drivers of healthy aging and longevity in humans has not yet delivered many robust longevity loci and pathways. However, larger studies, new methodologies and the consistent use of different study designs to follow up results might help to unravel the genomic component of healthy aging and longevity.

## Phenotypes that reflect biological aging

In addition to focusing on lifespan as primary phenotype, genomic studies into aging may profit from insights into phenotypes that reflect biological age. One can think of parameters or profiles reflecting immunosenescence or metabolic health established as pre-clinical measures in middle-aged individuals. In addition, phenotyping by novel noninvasive technologies, such as imaging (e.g., functional magnetic resonance imaging (MRI)) and longitudinal and ambulatory measurements using electronic devices (e.g., gait speed, 24-hour glucose, blood pressure), will improve the monitoring of the physiology of aging in epidemiological studies. Such research is often referred to as biomarker research and is aimed at finding parameters and profiles predicting long-term morbidity and/or mortality. Classical examples are blood pressure and hypertension as markers for clinical events in cardiovascular disease, joint-space width as marker for osteoarthritis and bone mineral density and risk of fracture as markers for osteoporosis. Comparable to the genomic research of the transcriptome and epigenome, the main problem with biomarker research is that it is hard to disentangle the changes causal to aging and longevity from those that are a consequence of normative aging. For classical (e.g., leukocyte telomere length) and novel potential biomarker of aging the four relevant phases to establish whether a quantitative parameter (or profile) is a biomarker of biological aging should be taken into account (see Introduction section).

### Clinical biomarkers for biological age hint at metabolic processes

Several prospective studies investigated the effect of clinical, physical, and cognitive parameters on mortality. Many different parameters have been shown to influence mortality after 55 years of age in the general population [65–70]. To determine whether these parameters potentially contribute to longevity from middle age onwards, family-based studies have been performed (Box 1 and Fig. 1), whereby the offspring of long-lived individuals is compared with similar-aged controls from the general population. Of the parameters that associate with mortality after 55 years of age, cortisol levels, digit symbol substitution test score, fasting glucose levels, free triiodothyronine levels, and gait speed also mark familial longevity in middle age [7, 71–78] (Table 1). Together, these biomarkers for biological age suggest the involvement of metabolic processes in healthy aging and longevity.

### Metabolic profiles seem promising predictive biomarkers

Instead of testing single quantitative parameters from a clinical perspective, the development of novel technologies and methodologies has made it possible to study age-related changes in the whole glycome and metabolome [79, 80]. These novel explorative omics studies could potentially be much more informative on physiological aspects of aging than the single parameters studied so far, since a single-point measurement contains a wealth of information. A cross-sectional comparison of “healthy agers” and similar-aged controls has

**Table 1. Potential biomarkers which have been shown to be associated with mortality in prospective studies of different age categories and their association with familial longevity in middle age**

Parameter	Effect on mortality	Reference	Effect on familial longevity	Reference
Ankle-arm index	>65 years	[66]	No	[7]
Blood pressure (diastolic)	>55 years/>90 years	[65, 69, 70]	No	[7]
Blood pressure (systolic)	>55 years/>90 years	[65, 66, 69, 70]	No	[7]
Body mass index	>55 years/>90 years	[69, 70]	No	[7, 77]
Cholesterol (high-density lipoprotein)	>85 years	[65, 69, 70]	No	[7, 71, 77]
Cholesterol (total)	>55 years/>90 years	[65, 69, 70]	No	[7, 71, 77]
Cortisol	>85 years	[65]	Yes	[73]
C-reactive protein	>55 years	[65, 66, 70]	No	[76]
Creatinine	>55 years/>65 years	[65, 66, 70]	No	[7, 78]
Digit symbol substitution test score	>65 years	[66]	Yes	[7]
Fasting glucose	>65 years	[65, 66]	Yes	[7, 74]
Forced vital capacity	>65 years	[66]	No	[7]
Free triiodothyronine levels	>85 years	[65]	Yes	[75]
Gait speed	>65 years	[68]	Yes	[7]
Grip strength	>92 years	[65, 67]	No	[7, 72]
Instrumental activities of daily living impairment	>65 years	[66, 70]	No	[7]

shown that decreased levels of bisecting GlcNAc glycoforms of IgG and higher levels of specific *N*-glycan features mark healthy aging and familial longevity [81, 82]. Datasets generated by metabolomic platforms provide information on biogenic amines, central metabolism, and lipids and can give insight into their relevance for morbidity and/or mortality, as was previously shown for cardiovascular disease [83]. In a recent study, using a prospective design, it was shown that a single-point nuclear magnetic resonance (NMR) measurement could also predict incident risk of coronary heart disease, comparable to the gold standard (the Framingham risk score) (unpublished results). However, additional prospective studies into morbidity and/or mortality, preferably on the basis of repeated measures, need to be performed to provide more information about the usefulness of metabolomic and glycomic profiles as biomarkers of biological age and longevity.

### Integrating genomics and biomarker research

Once the use of established biomarkers of biological age is standardized, the biomarker information can be integrated into studies aimed at finding causal determinants of aging and longevity. An example of an integrated approach to identify lifespan regulating loci is represented by testing whether genetic variants associated with potential biomarkers also associate with longevity. To date, GWAS have identified many genetic variants that associate with age-associated traits, such as leukocyte telomere length and features from glycome and metabolome profiles [84–86]. The joint effect of the majority of these variants on aging and longevity still needs to be determined. One study identified a haplotype in the *TERT* gene that was associated with increased telomere length and longevity, which indicates that genetic variants associated with telomere length regulation might also play a role in longevity [87].

### Conclusions and prospects

Over the past two decades the human aging field has built up the necessary resources to study the biology of aging and

longevity by establishing human populations with a diversity of designs. Meta-analyses integrating genetic and phenotypic datasets have successfully identified variants associated with a range of age-related traits and diseases. Despite these accomplishments, the number of novel leads contributing to human lifespan regulation is limited. Although positive regions of linkage and suggestive GWAS hits have been reported, the field has not yet identified the loci that explain the clustering of longevity in families and the variation in biological aging rate in the population. As for animal models, down-signaling of the IIS and mTOR pathway appeared to be relevant in humans. These findings are being followed up by molecular and physiological profiling using skin, fat and muscle tissue of long-lived family members and controls. Human studies now also include the response of nutrient sensing systems to the application of dietary and physical challenges.

The ongoing whole genome sequencing of centenarians and their families may provide novel genes contributing to longevity. Relevant variations may include gain-of-function mutations or heterozygous loss-of-function mutations in genes with deleterious effect late in life. Novel biomarkers represented by omics profiles and ambulatory measures to establish the biological aging rate (such as 24-hour glucose [88]) will be used in integrated analyses. It has already become feasible to study the integrative personal omics profiles (iPOP), the combination of the genetic, transcriptomic, proteomic, metabolomic, and autoantibody profile of individuals [89].

In conclusion, novel methodologies, comprehensively applied to multiple studies of well-phenotyped (middle and highly aged) individuals from long-lived families and large prospective cohort studies, will help to connect human molecular epidemiology and biology in aging research. Ultimately, this will provide leads that can be followed up in animal studies.

### Acknowledgments

The research leading to these results has received funding from the European Union's Seventh Framework Programme (FP7/2007-2011) under grant agreement no. 259679. This study



was supported by a grant from the Innovation-Oriented Research Program on Genomics (SenterNovem IGE05007), the Centre for Medical Systems Biology, and the Netherlands Consortium for Healthy Ageing (grant 050-060-810), all in the framework of the Netherlands Genomics Initiative, Netherlands Organization for Scientific Research (NWO), and by Unilever Colworth.

## References

- Oeppen J, Vaupel JW. 2002. Demography. Broken limits to life expectancy. *Science* **296**: 1029–31.
- World Health Organization. 2012. *World Health Statistics 2012*. Geneva: WHO.
- Jagger C, Gillies C, Moscone F, Cambois E, et al. 2008. Inequalities in healthy life years in the 25 countries of the European Union in 2005: a cross-national meta-regression analysis. *Lancet* **372**: 2124–31.
- Christensen K, McGue M, Petersen I, Jeune B, et al. 2008. Exceptional longevity does not result in excessive levels of disability. *Proc Natl Acad Sci USA* **105**: 13274–9.
- Terry DF, Sebastiani P, Andersen SL, Perls TT. 2008. Disentangling the roles of disability and morbidity in survival to exceptional old age. *Arch Intern Med* **168**: 277–83.
- Atzmon G, Schechter C, Greiner W, Davidson D, et al. 2004. Clinical phenotype of families with longevity. *J Am Geriatr Soc* **52**: 274–7.
- Newman AB, Glynn NW, Taylor CA, Sebastiani P, et al. 2011. Health and function of participants in the Long Life Family Study: A comparison with other cohorts. *Aging (Albany NY)* **3**: 63–76.
- Terry DF, Wilcox MA, McCormick MA, Pennington JY, et al. 2004. Lower all cause, cardiovascular, and cancer mortality in centenarians' offspring. *J Am Geriatr Soc* **52**: 2074–6.
- Westendorp RG, van Heemst D, Rosing MP, Frolich M, et al. 2009. Nonagenarian siblings and their offspring display lower risk of mortality and morbidity than sporadic nonagenarians: the Leiden Longevity Study. *J Am Geriatr Soc* **57**: 1634–7.
- Slagboom PE, Beekman M, Passtoors WM, Deelen J, et al. 2011. Genomics of human longevity. *Philos Trans R Soc Lond B Biol Sci* **366**: 35–42.
- Sprott RL. 2010. Biomarkers of aging and disease: introduction and definitions. *Exp Gerontol* **45**: 2–4.
- Herskind AM, McGue M, Holm NV, Sorensen TI, et al. 1996. The heritability of human longevity: a population-based study of 2872 Danish twin pairs born 1870–1900. *Hum Genet* **97**: 319–23.
- Hjelmborg JV, Iachine I, Skytthe A, Vaupel JW, et al. 2006. Genetic influence on human lifespan and longevity. *Hum Genet* **119**: 312–21.
- Perls T, Shea-Drinkwater M, Bowen-Flynn J, Ridge SB, et al. 2000. Exceptional familial clustering for extreme longevity in humans. *J Am Geriatr Soc* **48**: 1483–5.
- Schoenmaker M, de Craen AJ, de Meijer PH, Beekman M, et al. 2006. Evidence of genetic enrichment for exceptional survival using a family approach: the Leiden Longevity Study. *Eur J Hum Genet* **14**: 79–84.
- Gavrilova NS, Gavrilov LA, Evdokushkina GN, Semyonova VG, et al. 1998. Evolution, mutations, and human longevity: European royal and noble families. *Hum Biol* **70**: 799–804.
- Finch CE, Tanzi RE. 1997. Genetics of aging. *Science* **278**: 407–11.
- Stahl EA, Wegmann D, Trynka G, Gutierrez-Achury J, et al. 2012. Bayesian inference analyses of the polygenic architecture of rheumatoid arthritis. *Nat Genet* **44**: 483–9.
- Yang J, Manolio TA, Pasquale LR, Boerwinkle E, et al. 2011. Genome partitioning of genetic variation for complex traits using common SNPs. *Nat Genet* **43**: 519–25.
- Hindorf LA, Sethupathy P, Junkins HA, Ramos EM, et al. 2009. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc Natl Acad Sci USA* **106**: 9362–7.
- Beekman M, Nederstigt C, Suchiman HE, Kremer D, et al. 2010. Genome-wide association study (GWAS)-identified disease risk alleles do not compromise human longevity. *Proc Natl Acad Sci USA* **107**: 18046–9.
- Christensen K, Johnson TE, Vaupel JW. 2006. The quest for genetic determinants of human longevity: challenges and insights. *Nat Rev Genet* **7**: 436–48.
- de Magalhães JP, Budovsky A, Lehmann G, Costa J, et al. 2009. The Human Ageing Genomic Resources: online databases and tools for biogerontologists. *Aging Cell* **8**: 65–72.
- Fontana L, Partridge L, Longo VD. 2010. Extending healthy life span—from yeast to humans. *Science* **328**: 321–6.
- Pettan-Brewer C, Treuting PM. 2011. Practical pathology of aging mice. *Pathobiol Aging Age Relat Dis* **1**: 1.
- Crimmins EM, Finch CE. 2012. Special issue on genetics of aging. *J Gerontol A Biol Sci Med Sci* **67**: 467–519.
- Soerensen M, Dato S, Christensen K, McGue M, et al. 2010. Replication of an association of variation in the FOXO3A gene with human longevity using both case-control and longitudinal data. *Aging Cell* **9**: 1010–7.
- Soerensen M, Dato S, Tan Q, Thinggaard M, et al. 2012. Evidence from case-control and longitudinal studies supports associations of genetic variation in APOE, CETP, and IL6 with human longevity. *Age (Dordr)* in press DOI: 10.1007/s11357-011-9373-7.
- Zeng Y, Cheng L, Chen H, Cao H, et al. 2010. Effects of FOXO genotypes on longevity: a biodemographic analysis. *J Gerontol A Biol Sci Med Sci* **65**: 1285–99.
- Mahley RW, Rall SC Jr. 2000. Apolipoprotein E: far more than a lipid transport protein. *Annu Rev Genomics Hum Genet* **1**: 507–37.
- van der Horst A, Burgering BM. 2007. Stressing the role of FoxO proteins in lifespan and disease. *Nat Rev Mol Cell Biol* **8**: 440–50.
- Broer L, Demerath EW, Garcia ME, Homuth G, et al. 2012. Association of heat shock proteins with all-cause mortality. *Age (Dordr)* in press DOI: 10.1007/s11357-012-9417-7.
- Deelen J, Uh HW, Monajemi R, van Heemst D, et al. 2013. Gene set analysis of GWAS data for human longevity highlights the relevance of the insulin/IGF-1 signaling and telomere maintenance pathways. *Age (Dordr)* **35**: 235–49.
- Pawlikowska L, Hu D, Huntsman S, Sung A, et al. 2009. Association of common genetic variation in the insulin/IGF1 signaling pathway with human longevity. *Aging Cell* **8**: 460–72.
- Soerensen M, Dato S, Tan Q, Thinggaard M, et al. 2012. Human longevity and variation in GH/IGF-1/insulin signaling, DNA damage signaling and repair and pro/antioxidant pathway genes: cross sectional and longitudinal studies. *Exp Gerontol* **47**: 379–87.
- van Heemst D, Beekman M, Mooijaart SP, Heijmans BT, et al. 2005. Reduced insulin/IGF-1 signalling and human longevity. *Aging Cell* **4**: 79–85.
- Deelen J, Beekman M, Uh HW, Helmer Q, et al. 2011. Genome-wide association study identifies a single major locus contributing to survival into old age; the APOE locus revisited. *Aging Cell* **10**: 686–98.
- Newman AB, Walter S, Lunetta KL, Garcia ME, et al. 2010. A meta-analysis of four genome-wide association studies of survival to age 90 years or older: the Cohorts for Heart and Aging Research in Genomic Epidemiology Consortium. *J Gerontol A Biol Sci Med Sci* **65**: 478–87.
- Walter S, Atzmon G, Demerath EW, Garcia ME, et al. 2011. A genome-wide association study of aging. *Neurobiol Aging* **32**: 2109–28.
- Sebastiani P, Solovieff N, Dewan AT, Walsh KM, et al. 2012. Genetic signatures of exceptional longevity in humans. *PLoS One* **7**: e29848.
- Nebel A, Kleindorp R, Caliebe A, Nothnagel M, et al. 2011. A genome-wide association study confirms APOE as the major gene influencing survival in long-lived individuals. *Mech Ageing Dev* **132**: 324–0.
- Malovini A, Illario M, Iaccarino G, Villa F, et al. 2011. Association study on long-living individuals from Southern Italy identifies rs10491334 in the CAMKIV gene that regulates survival proteins. *Rejuvenation Res* **14**: 283–91.
- Kuningas M, Estrada K, Hsu YH, Nandakumar K, et al. 2011. Large common deletions associate with mortality at old age. *Hum Mol Genet* **20**: 4290–6.
- Kuningas M, McQuillan R, Wilson JF, Hofman A, et al. 2011. Runs of homozygosity do not influence survival to old age. *PLoS One* **6**: e22580.
- Beekman M, Blauw GJ, Houwing-Duistermaat JJ, Brandt BW, et al. 2006. Chromosome 4q25, microsomal transfer protein gene, and human longevity: novel data and a meta-analysis of association studies. *J Gerontol A Biol Sci Med Sci* **61**: 355–62.
- Boyden SE, Kunkel LM. 2010. High-density genomewide linkage analysis of exceptional human longevity identifies multiple novel Loci. *PLoS One* **5**: e12432.

47. **Puca AA, Daly MJ, Brewster SJ, Matise TC**, et al. 2001. A genome-wide scan for linkage to human exceptional longevity identifies a locus on chromosome 4. *Proc Natl Acad Sci USA* **98**: 10505–8.
48. **Reed T, Dick DM, Uniacke SK, Foroud T**, et al. 2004. Genome-wide scan for a healthy aging phenotype provides support for a locus near D4S1564 promoting healthy aging. *J Gerontol A Biol Sci Med Sci* **59**: 227–32.
49. **Beekman M, Blanche H, Perola M, Hervonen A**, et al. 2013. Genome-wide linkage analysis for human longevity: Genetics of Healthy Ageing Study. *Aging Cell* in press, DOI: 10.1111/accel.12039.
50. **Suh Y, Atzmon G, Cho MO, Hwang D**, et al. 2008. Functionally significant insulin-like growth factor I receptor mutations in centenarians. *Proc Natl Acad Sci USA* **105**: 3438–42.
51. **Sebastiani P, Riva A, Montano M, Pham P**, et al. 2011. Whole genome sequences of a male and female supercentenarian, ages greater than 114 years. *Front Genet* **2**: 90.
52. **Remondini D, Salvioi S, Francesconi M, Pierini M**, et al. 2010. Complex patterns of gene expression in human T cells during in vivo aging. *Mol Biosyst* **6**: 1983–92.
53. **de Magalhaes JP, Curado J, Church GM**. 2009. Meta-analysis of age-related gene expression profiles identifies common signatures of aging. *Bioinformatics* **25**: 875–81.
54. **Elsharawy A, Keller A, Flachsbarth F, Wendschlag A**, et al. 2012. Genome-wide miRNA signatures of human longevity. *Aging Cell* **11**: 607–16.
55. **Harries LW, Hernandez D, Henley W, Wood AR**, et al. 2011. Human aging is characterized by focused changes in gene expression and deregulation of alternative splicing. *Aging Cell* **10**: 868–78.
56. **Harries LW, Fellows AD, Pilling LC, Hernandez D**, et al. 2012. Advancing age is associated with gene expression changes resembling mTOR inhibition: evidence from two human populations. *Mech Ageing Dev* **133**: 556–62.
57. **Hong MG, Myers AJ, Magnusson PK, Prince JA**. 2008. Transcriptome-wide assessment of human brain and lymphocyte senescence. *PLoS One* **3**: e3024.
58. **Passtoors WM, Beekman M, Gunn D, Boer JM**, et al. 2008. Genomic studies in ageing research: the need to integrate genetic and gene expression approaches. *J Intern Med* **263**: 153–66.
59. **Passtoors WM, Boer JM, Goeman JJ, van den Akker EB**, et al. 2012. Transcriptional profiling of human familial longevity indicates a role for ASF1A and IL7R. *PLoS One* **7**: e27759.
60. **Passtoors WM, Beekman M, Deelen J, van der Breggen R**, et al. 2012. Gene expression analysis of mTOR pathway: association with human longevity. *Aging Cell* in press DOI: 10.1111/accel.12015.
61. **Bell JT, Tsai PC, Yang TP, Pidsley R**, et al. 2012. Epigenome-wide scans identify differentially methylated regions for age and age-related phenotypes in a healthy ageing population. *PLoS Genet* **8**: e1002629.
62. **Hernandez DG, Nalls MA, Gibbs JR, Arepalli S**, et al. 2011. Distinct DNA methylation changes highly correlated with chronological age in the human brain. *Hum Mol Genet* **20**: 1164–72.
63. **Rakyan VK, Down TA, Maslau S, Andrew T**, et al. 2010. Human aging-associated DNA hypermethylation occurs preferentially at bivalent chromatin domains. *Genome Res* **20**: 434–9.
64. **Garagnani P, Bacalini MG, Pirazzini C, Gori D**, et al. 2012. Methylation of ELOVL2 gene as a new epigenetic marker of age. *Aging Cell* **11**: 1132–4.
65. **Martin-Ruiz C, Jagger C, Kingston A, Collerton J**, et al. 2011. Assessment of a large panel of candidate biomarkers of ageing in the Newcastle 85+ study. *Mech Ageing Dev* **132**: 496–502.
66. **Newman AB, Sachs MC, Arnold AM, Fried LP**, et al. 2009. Total and cause specific mortality in the cardiovascular health study. *J Gerontol A Biol Sci Med Sci* **64**: 1251–61.
67. **Oksuzyan A, Maier H, McGue M, Vaupel JW**, et al. 2010. Sex differences in the level and rate of change of physical function and grip strength in the Danish 1905-cohort study. *J Aging Health* **22**: 589–610.
68. **Studenski S, Perera S, Patel K, Rosano C**, et al. 2011. Gait speed and survival in older adults. *JAMA* **305**: 50–8.
69. **van Vliet P, Oleksik AM, van Heemst D, de Craen AJ**, et al. 2010. Dynamics of traditional metabolic risk factors associate with specific causes of death in old age. *J Gerontol A Biol Sci Med Sci* **65**: 488–94.
70. **Walter S, Mackenbach J, Voko Z, Lhachimi S**, et al. 2012. Genetic, physiological, and lifestyle predictors of mortality in the general population. *Am J Public Health* **102**: e3–10.
71. **Barzilai N, Atzmon G, Schechter C, Schaefer EJ**, et al. 2003. Unique lipoprotein phenotype and genotype associated with exceptional longevity. *JAMA* **290**: 2030–40.
72. **Ling CH, de Craen AJ, Slagboom PE, Westendorp RG**, et al. 2012. Handgrip strength at midlife and familial longevity: the Leiden Longevity Study. *Age (Dordr)* **34**: 1261–8.
73. **Noordam R, Gunn DA, Tomlin CC, Rozing MP**, et al. 2012. Cortisol serum levels in familial longevity and perceived age: the Leiden Longevity Study. *Psychoneuroendocrinology* **37**: 1669–75.
74. **Rozing MP, Westendorp RG, de Craen AJ, Frolich M**, et al. 2010. Favorable glucose tolerance and lower prevalence of metabolic syndrome in offspring without diabetes mellitus of nonagenarian siblings: the Leiden longevity study. *J Am Geriatr Soc* **58**: 564–9.
75. **Rozing MP, Westendorp RG, de Craen AJ, Frolich M**, et al. 2010. Low serum free triiodothyronine levels mark familial longevity: the Leiden Longevity Study. *J Gerontol A Biol Sci Med Sci* **65**: 365–8.
76. **Rozing MP, Mooijaart SP, Beekman M, Wijsman CA**, et al. 2011. C-reactive protein and glucose regulation in familial longevity. *Age (Dordr)* **33**: 623–30.
77. **Vaarhorst AA, Beekman M, Suchiman EH, van Heemst D**, et al. 2011. Homocysteine in long-lived families: the Leiden Longevity Study. *Age (Dordr)* **33**: 219–27.
78. **Wijsman CA, van Heemst D, Rozing MP, Slagboom PE**, et al. 2011. Homocysteine and familial longevity: the Leiden Longevity Study. *PLoS One* **6**: e17543.
79. **Dall'olio F, Vanhooren V, Chen CC, Slagboom PE**, et al. 2012. N-glycomic biomarkers of biological aging and longevity: a link with inflammaging. *Ageing Res Rev* in press DOI: 10.1016/j.arr.2012.02.002.
80. **Yu Z, Zhai G, Singmann P, He Y**, et al. 2012. Human serum metabolic profiles are age dependent. *Aging Cell* **11**: 960–7.
81. **Ruhaak LR, Uh HW, Beekman M, Koeleman CA**, et al. 2010. Decreased levels of bisecting GlcNAc glycoforms of IgG are associated with human longevity. *PLoS One* **5**: e12566.
82. **Ruhaak LR, Uh HW, Beekman M, Hokke CH**, et al. 2011. Plasma protein N-glycan profiles are associated with calendar age, familial longevity and health. *J Proteome Res* **10**: 1667–74.
83. **Shah SH, Bain JR, Muehlbauer MJ, Stevens RD**, et al. 2010. Association of a peripheral blood metabolic profile with coronary artery disease and risk of subsequent cardiovascular events. *Circ Cardiovasc Genet* **3**: 207–14.
84. **Codd V, Nelson CP, Albrecht E, Mangino M**, et al. 2013. Identification of seven loci affecting mean telomere length and their association with disease. *Nat Genet* in press.
85. **Huffman JE, Knezevic A, Vitart V, Kattila J**, et al. 2011. Polymorphisms in B3GAT1, SLC9A9 and MGAT5 are associated with variation within the human plasma N-glycome of 3533 European adults. *Hum Mol Genet* **20**: 5000–11.
86. **Kettunen J, Tukiainen T, Sarin AP, Ortega-Alonso A**, et al. 2012. Genome-wide association study identifies multiple loci influencing human serum metabolite levels. *Nat Genet* **44**: 269–76.
87. **Atzmon G, Cho M, Cawthon RM, Budagov T**, et al. 2010. Evolution in health and medicine Sackler colloquium: genetic variation in human telomerase is associated with telomere length in Ashkenazi centenarians. *Proc Natl Acad Sci USA* **107**: 1710–7.
88. **Wijsman CA, van Heemst D, Hoogeveen ES, Slagboom PE**, et al. 2013. Ambulant 24-hour glucose rhythms mark calendar and biological age in apparently healthy individuals. *Aging Cell* in press, DOI: 10.1111/accel.12042.
89. **Chen R, Mias GI, Li-Pook-Than J, Jiang L**, et al. 2012. Personal omics profiling reveals dynamic molecular and medical phenotypes. *Cell* **148**: 1293–307.
90. **Perls TT, Bochen K, Freeman M, Alpert L**, et al. 1999. Validity of reported age and centenarian prevalence in New England. *Age Ageing* **28**: 193–7.
91. **Nebel A, Croucher PJ, Stiegeler R, Nikolaus S**, et al. 2005. No association between microsomal triglyceride transfer protein (MTP) haplotype and longevity in humans. *Proc Natl Acad Sci USA* **102**: 7906–9.
92. **Blanche H, Cabanne L, Sahbatou M, Thomas G**. 2001. A study of French centenarians: are ACE and APOE associated with longevity? *C R Acad Sci III* **324**: 129–35.
93. **Anselmi CV, Malovini A, Roncarati R, Novelli V**, et al. 2009. Association of the FOXO3A locus with extreme longevity in a southern Italian centenarian study. *Rejuvenation Res* **12**: 95–104.
94. **Skytthe A, Valensin S, Jeune B, Cevenini E**, et al. 2011. Design, recruitment, logistics, and data management of the GEHA (Genetics of Healthy Ageing) project. *Exp Gerontol* **46**: 934–45.
95. **Bootsma-van der Wiel A, van Exel E, de Craen AJ, Gussekloo J**, et al. 2002. A high response is not essential to prevent selection bias: results from the Leiden 85-plus study. *J Clin Epidemiol* **55**: 1119–25.

96. **Weverling-Rijnsburger AW, Blauw GJ, Lagaay AM, Knook DL**, et al. 1997. Total cholesterol and risk of mortality in the oldest old. *Lancet* **350**: 1119–23.
97. **Collerton J, Barrass K, Bond J, Eccles M**, et al. 2007. The Newcastle 85+ study: biological, clinical and psychosocial factors associated with healthy ageing: study protocol. *BMC Geriatr* **7**: 14.
98. **Collerton J, Davies K, Jagger C, Kingston A**, et al. 2009. Health and disease in 85 year olds: baseline findings from the Newcastle 85+ cohort study. *BMJ* **339**: b4904.
99. **Nybo H, Gaist D, Jeune B, Bathum L**, et al. 2001. The Danish 1905 cohort: a genetic-epidemiological nationwide survey. *J Aging Health* **13**: 32–46.
100. **Hofman A, Breteler MM, van Duijn CM, Janssen HL**, et al. 2009. The Rotterdam Study: 2010 objectives and design update. *Eur J Epidemiol* **24**: 553–72.
101. **Kannel WB, Feinleib M, McNamara PM, Garrison RJ**, et al. 1979. An investigation of coronary heart disease in families. The Framingham offspring study. *Am J Epidemiol* **110**: 281–90.