Transcriptomic studies in human ageing and longevity

Willemijn M. Passtoors

Transcriptomic studies in human ageing and longevity
Drs Willemijn M. Passtoors

Publication of this thesis was financially supported by the Consortium for Healthy Ageing (NGI 050060810) and Unilever PLC.

PhD thesis with summary in Dutch
ISBN: 978-90-8891-768-4

Cover design: Betty de Ronde
Printed by: Proefschriftmaken.nl || UitgeverijBOXPress
Published by: UitgeverijBOXPress, ‘s-Hertogenbosch

Transcriptomic studies in human ageing and longevity

Proefschrift

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus prof. Mr. C.J.J.M. Stolker,
volgens besluit van het College van Promotores
te verdedigen op donderdag 19 december 2013
klokke 13:45 uur

door

Willemijn Marie-Antoinette Passtoors
Geboren te Haarlem
in 1980

Promotiecommissie

Promotor:	Prof. Dr. P.E. Slagboom
Co-promotores:	Dr. M. Beekman
	Dr. J.M. Boer *Erasmus University Medical Center*
Overige leden:	Prof. Dr. C. Franceschi *University of Bologna, Italy*
	Prof. Dr. A.G. Uitterlinden *Erasmus University Medical Center*
	Prof. Dr. B.J. Zwaan *Wageningen University*

Content

Chapter 1

General introduction

The life expectancy of the general western population has increased dramatically in the last 200 years (1) and is generally expected to continue this rise. Many people reach a high age, but at the same time on average 25% of the lifespan is spent in disability. For virtually all common diseases, age is the primary risk factor. The process of ageing is complex, generally considered as the consequence of accumulation of damage to macromolecules which occurs at a variable rate in different individuals, and affects organs in different ways, ultimately leading to physiological heterogeneity of elderly individuals, populations and patient groups. The scientific human ageing field focuses on identifying biomarkers that monitor the rate of ageing, markers of the biological age rather than just the chronological age of individuals. Either such markers are causally involved in ageing or can be used as tools to discover biological mechanisms that determine the ageing rate and the ability to age healthily and become long-lived.

Longevity determinants

The maximum and mean lifespan of animal models can be manipulated by intervening in a number of pathways (2-4) that indicate a role in lifespan regulation especially for nutrient sensing systems involving insulin/IGF (IIS) and target of rapamycin (TOR) signaling. These determinants have been investigated as candidate genes and pathways in human studies aimed at identifying determinants of human lifespan regulation. In particular, genetic association studies, in which gene frequencies are compared in highly aged cases and younger controls, provided positive associations for genes in the IIS pathway (5). For example, genetic variation in *AKT* and *FOXO3a* (6;7) has been shown to have an effect on human longevity. Until now however, the only locus widely confirmed to affect human mortality risk up to the highest ages is the *APOE* locus on chromosome 19 (8-10).

Biomarkers for human ageing

Studies for biomarkers of human ageing have mainly investigated the relation of physiological and molecular phenotypes with chronological age to identify changes in these phenotypes representative of normative ageing. Since many molecular changes occur as a function of chronological age, a more valuable biomarker of human ageing would also associate with morbidity and mortality in prospective studies (11) and thereby reflecting biological ageing. An example of such a biomarker is leukocyte telomere length (LTL) which associates with metabolic health and mortality in various human cohort studies including prospective analyses (12;13).

Nevertheless, this biomarker may not necessarily be a determinant of ageing since a causal role for telomere length in human normative ageing has not been established yet.

One strategy to find candidate biomarkers for human ageing is to compare molecular traits in normative and healthy ageing groups within the human population. The healthy ageing individuals can be selected from families in which the longevity trait is clustered. In the Longevity Genes Project such families have been collected by selection for centenarian single cases (14) whereas in the Leiden Longevity Study selection has been focused on nonagenarian siblings (15). As compared to normative controls, the middle-aged descendants of these centenarians and nonagenarians have a low prevalence of metabolic diseases such as insulin resistance, type 2 diabetes, hypertension and myocardial infarction (16) and lower mortality risk (15;17) and are therefore regarded as healthy agers.

Biomarker studies comparing the healthy ageing offspring of the long-lived to normative ageing controls have indicated that despite the absence of difference in body mass index and lifestyle between the groups, the offspring displays a healthier intrinsic metabolism reflected by serum lipid profiles (14;18), adiponectin levels (19) and markers of glucose metabolism (20). These measures of intrinsic metabolism are currently considered as candidate biomarkers for human ageing in middle age. For elderly populations however, the predictive power of these biomarkers for morbidity and mortality is less clear.

The transcriptome

Besides the serum parameters as candidate biomarkers for human ageing, an additional source for biomarkers may come from variation in gene expression. Genes and functionally coherent gene sets of which the gene expression changes with chronological age could set the profile for normative ageing, providing a first insight into the mechanisms of ageing processes and age-related disease (21). The last decade gene expression changes with age have regularly been investigated both in animal models and humans, using different designs, phenotypes and techniques. This work has been reviewed for studies until 2008 in chapter 2 of this thesis. In short, in different tissues in humans, the top most differentially expressed genes with chronological age have been enriched for biological terms like immune response, energy metabolism and mitochondrial processes. After 2008, explorative studies on age-related changes in the human transcriptome have fundamentally increased their power by increasing

sample sizes by mainly investigating RNA isolated from blood (21-23). While at the single gene level the mutual overlap of transcriptional changes with chronological age between different studies is limited, the enrichment is similar. The major consistent additional enrichment among the top list of genes of which gene expression changes with age since 2009 is for post-transcriptional regulation. Thus processes involved in immune response, energy metabolism, mitochondrial processes and post-transcriptional regulation may harbor potential biomarkers for the rate of normative ageing.

For practical reasons the largest human studies into gene expression changes with chronological age are naturally performed in blood. Smaller studies indicated that common gene sets are differentially expressed with ageing over different tissues (22;24). The similarities in age-affected pathways across tissues and species suggests that the analysis of the ageing transcriptome by functionally grouped gene sets in blood is a promising approach to discover biomarkers and determinants of biological ageing in humans.

Aim and outline of this thesis

The aim of this thesis was to identify in the human blood transcriptome, potential biomarker profiles that associate with chronological age and discriminate between 'healthy agers' from long-lived families and normative ageing controls. Such profiles may harbour determinants of the biological ageing rate.

The design for our study is based on the Leiden Longevity Study (LLS), a Dutch cohort in which families have been included on the basis of nonagenarian sibling pairs (15). The LLS consists of 420 Caucasian families with at least two long-lived siblings (n = 944), together with their middle-aged offspring (n = 1671) and the partners of this offspring (n = 744) which will act as controls in this study (Figure 2 of chapter 2 of this thesis). We have explored the blood transcriptome to recognize genes of which the expression marks chronological ageing (comparing groups of young and elderly individuals) and to identify which of those genes also mark familial longevity (comparing among middle-aged individuals groups of healthy and normative agers) (Chapter 3). Based on the top hits and most interesting findings from these explorative analyses we extended, within the Leiden Longevity Study, our analysis to a larger set of samples and genes belonging to the *IL7R* pathway (Chapter 4) and to the *MTOR* pathway (Chapter 5). The results of the latter study were followed by studies in primary fibroblasts cultures from donors of the LLS (Chapter 6). Finally to

improve and generate more robust and specific hypotheses for markers of normative ageing we performed an explorative meta-analysis in three large transcriptomic datasets using an integrative co-expressed network approach (Chapter 7).

References

(1) Oeppen J, Vaupel JW. Demography. Broken limits to life expectancy. Science 2002 May 10;296(5570):1029-31.

(2) Harrison DE, Strong R, Sharp ZD, Nelson JF, Astle CM, Flurkey K et al. Rapamycin fed late in life extends lifespan in genetically heterogeneous mice. Nature 2009 July 16;460(7253):392-5.

(3) Kirkwood TL, Kapahi P, Shanley DP. Evolution, stress, and longevity. J Anat 2000 November;197 Pt 4:587-90.

(4) Lin K, Hsin H, Libina N, Kenyon C. Regulation of the Caenorhabditis elegans longevity protein DAF-16 by insulin/IGF-1 and germline signaling. Nat Genet 2001 June;28(2):139-45.

(5) Deelen J, Uh HW, Monajemi R, van HD, Thijssen PE, Bohringer S et al. Gene set analysis of GWAS data for human longevity highlights the relevance of the insulin/IGF-1 signaling and telomere maintenance pathways. Age (Dordr) 2011 November 24.

(6) Flachsbart F, Caliebe A, Kleindorp R, Blanche H, von Eller-Eberstein H, Nikolaus S et al. Association of FOXO3A variation with human longevity confirmed in German centenarians. Proc Natl Acad Sci U S A 2009 February 24;106(8):2700-5.

(7) Willcox BJ, Donlon TA, He Q, Chen R, Grove JS, Yano K et al. FOXO3A genotype is strongly associated with human longevity. Proc Natl Acad Sci U S A 2008 September 16;105(37):13987-92.

(8) Deelen J, Beekman M, Uh HW, Helmer Q, Kuningas M, Christiansen L et al. Genome-wide association study identifies a single major locus contributing to survival into old age; the APOE locus revisited. Aging Cell 2011 August;10(4):686-98.

(9) Christensen K, Johnson TE, Vaupel JW. The quest for genetic determinants of human longevity: challenges and insights. Nat Rev Genet 2006 June;7(6):436-48.

(10) Schachter F, Faure-Delanef L, Guenot F, Rouger H, Froguel P, Lesueur-Ginot L et al. Genetic associations with human longevity at the APOE and ACE loci. Nat Genet 1994 January;6(1):29-32.

(11) Sprott RL. Biomarkers of aging and disease: introduction and definitions. Exp Gerontol 2010 January;45(1):2-4.

(12) Fitzpatrick AL, Kronmal RA, Kimura M, Gardner JP, Psaty BM, Jenny NS et al. Leukocyte telomere length and mortality in the Cardiovascular Health Study. J Gerontol A Biol Sci Med Sci 2011 April;66(4):421-9.

(13) Cawthon RM, Smith KR, O'Brien E, Sivatchenko A, Kerber RA. Association between telomere length in blood and mortality in people aged 60 years or older. Lancet 2003 February 1;361(9355):393-5.

(14) Barzilai N, Gabriely I, Gabriely M, Iankowitz N, Sorkin JD. Offspring of centenarians have a favorable lipid profile. J Am Geriatr Soc 2001 January;49(1):76-9.

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

(16) Westendorp RG, van HD, Rozing MP, Frolich M, Mooijaart SP, Blauw GJ et al. Nonagenarian siblings and their offspring display lower risk of mortality and morbidity than sporadic nonagenarians: The Leiden Longevity Study. J Am Geriatr Soc 2009 September;57(9):1634-7.

(17) Perls TT, Wilmoth J, Levenson R, Drinkwater M, Cohen M, Bogan H et al. Life-long sustained mortality advantage of siblings of centenarians. Proc Natl Acad Sci U S A 2002 June 11;99(12):8442-7.

(18) Heijmans BT, Beekman M, Houwing-Duistermaat JJ, Cobain MR, Powell J, Blauw GJ et al. Lipoprotein particle profiles mark familial and sporadic human longevity. PLoS Med 2006 December;3(12):e495.

(19) Atzmon G, Pollin TI, Crandall J, Tanner K, Schechter CB, Scherer PE et al. Adiponectin levels and genotype: a potential regulator of life span in humans. J Gerontol A Biol Sci Med Sci 2008 May;63(5):447-53.

(20) Rozing MP, Westendorp RG, de Craen AJ, Frolich M, de Goeij MC, Heijmans BT et al. 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 2010 March;58(3):564-9.

(21) Harries LW, Hernandez D, Henley W, Wood AR, Holly AC, Bradley-Smith RM et al. Human aging is characterized by focused changes in gene expression and deregulation of alternative splicing. Aging Cell 2011 October;10(5):868-78.

(22) de Magalhaes JP, Curado J, Church GM. Meta-analysis of age-related gene expression profiles identifies common signatures of aging. Bioinformatics 2009 April 1;25(7):875-81.

(23) Hong MG, Myers AJ, Magnusson PK, Prince JA. Transcriptome-wide assessment of human brain and lymphocyte senescence. PLoS One 2008;3(8):e3024.

(24) Zahn JM, Sonu R, Vogel H, Crane E, Mazan-Mamczarz K, Rabkin R et al. Transcriptional profiling of aging in human muscle reveals a common aging signature. PLoS Genet 2006 July;2(7):e115.

Chapter 2

Genomic studies in ageing research: the need to integrate genetic and gene expression approaches

W.M. Passtoors[1], M. Beekman[1], D. Gunn[2], J.M. Boer[3], B.T. Heijmans[1], R.G.J. Westendorp[4], B.J. Zwaan[5], P.E. Slagboom[1].

1 Section of Molecular Epidemiology, Leiden University Medical Center, Leiden, The Netherlands
2 Unilever Corporate Research, Colworth Laboratory, Sharnbrook, Bedford, United Kingdom
3 Center for Human and Clinical Genetics, Leiden University Medical Center, Leiden, The Netherlands
4 Department of Gerontology and Geriatrics, Leiden University Medical Center, Leiden, The Netherlands
5 Section of Evolutionary Biology, Institute for Biology, Leiden University, Leiden, The Netherlands.

Journal of Internal Medicine, 2008. 263(2): p. 153-66

Abstract

Genome-wide and hypothesis-based approaches to the study of ageing and longevity have been dominated by genetic investigations. To identify essential mechanisms of a complex trait such as ageing in higher species, a holistic understanding of interacting pathways is required. More information on such interactions is expected to be obtained from global gene expression analysis if combined with genetic studies. Genetic sequence variation often provides a functional gene marker for the trait, whereas a gene expression profile may provide a quantitative biomarker representing complex cellular pathway interactions contributing to the trait. Thus far, gene expression studies have associated multiple pathways to ageing including mitochondrial electron transport and the oxidative stress response. However, most of the studies are underpowered to detect small age-changes. A systematic survey of gene expression changes as a function of age in human individuals and animal models is lacking. Well-designed gene expression studies, especially at the level of biological processes, will provide hypotheses on gene-environmental interactions determining biological ageing rate. Cross-sectional studies monitoring the profile as a chronological marker of ageing must be integrated with prospective studies indicating which profiles represent biomarkers preceding and predicting physiological decline and mortality. New study designs such as the Leiden Longevity Study, including two generations of subjects from longevity families, aim to achieve these combined approaches.

Introduction

The most prevalent diseases in Western societies have in common that age is a major risk factor. The process of ageing across different species is marked by the occurrence of molecular and physiological changes in all cells and tissues. It is not clear which of these changes are causal to ageing, but the accumulation of damaged molecules and loss of homeostasis is generally considered to be central to the rate of the ageing process, generating lifespan variation even in genetically identical individuals in a common environment (1;2). Although different age-related diseases may arise by different pathophysiological mechanisms, it is not unlikely that common mechanisms of ageing influence the onset of different age-dependent diseases and the lifespan. This may be explained if these mechanisms affect the accumulation of damaged molecules and eventually the homeostasis of the organism. Indications that such mechanisms exist

come from the observation that the occurrence of multiple age-dependent disease conditions is simultaneously postponed in caloric restricted laboratory rodents (3) and from the offspring of human centenarians. These individuals are expected to have inherited longevity assurance mechanisms that attenuate their ageing rate and protect from disease (4). For more than two decades there is an ongoing search for mechanisms driving the ageing process in different species. Phenotypic features of ageing that are usually investigated include the mean and maximum population lifespan and the occurrence of age-dependent disease. True intermediate phenotypes, or biomarkers which mark the rate of biological ageing and the mortality risk of individuals in a pre-clinical phase of life, have hardly been identified. The availability of such biomarkers, comparable to the way bone mineral density is a biomarker predicting the risk for osteoporosis (5), would help in the search for mechanisms underpinning ageing.

Successfully applied genetic approaches in ageing research mainly involve the identification of genetic variation influencing intra-species differences in lifespan. These genetic studies revealed that some mechanisms of lifespan regulation, such as the insulin/insulin-like growth factor I signalling (IIS) pathway, are common to species ranging from nematodes to mice (see (3;6) for reviews). Genetically modified ageing mutants in these studies have increased life spans by more than two fold. Other mechanisms so far identified by genetic research in model systems include mitochondrial function and histone modification control (7). Genetic association studies of insulin signalling genes and other candidate loci with human longevity have given contradictory results (see (8;9) for reviews). The gene most consistently associating with human longevity in all studies is the apoE locus (10), which is poorly investigated in lifespan studies in animals. In general, the population based cross-sectional analysis of gene frequencies with age has thus far provided few consistent results. This may be explained by the fact that the mean life expectancy in western societies has increased dramatically in the last 100 years, and that the current variation in lifespan in such populations, therefore, is mainly determined by non-genetic factors (75%, (11;12)). Genetic influences on human longevity are much higher in families containing a clustering of long-lived individuals (13-15) and these observations have resulted in new genetic studies based on familial longevity.

Far less effort than on genetics has been put into obtaining an understanding of the gene expression profiles that are associated with lifespan and how this relates to the genetic variation associated with longevity. To achieve this goal, a systematic survey of gene expression

changes that occur in somatic tissues throughout the lifetime of individuals is required. Such profiles may reflect the interaction of the genome with environmental challenges that occur during the ageing process. Increased power in genetic studies must be accomplished, to uncover information on the interactive effects of multiple loci in ageing of higher species, which is expected to be essential in complex traits. It is thus the combination of approaches that promises most success in unravelling ageing as a complex trait.

In transcriptomic research, changes in gene expression profiles in various tissues are being explored mainly as comparisons of diseased and healthy individuals. Gene expression studies have been performed on a large scale in research into cancer (16;17), stroke (18;19), immune diseases (20-23) and other age-related diseases (24;25). In cancer research, expression profiles were indicated to be able to identify specific types of cancer, classify tumour stage or severity, disease progression and the response to certain treatments (26-31). The basic design in most cancer research is relatively simple compared to other disease research, because target and healthy tissue from patients can directly be compared and the changes in gene expression are often large. For a complex trait like human ageing, the primary affected cells or tissues are unknown and the gene expression changes are probably small and thus more difficult to find. Therefore, this paper serves two functions, (i) to review and discuss the gene expression studies performed in the ageing field and the insights into the biological process involved in ageing and longevity they have revealed, and (ii) to draft up future recommendations for this research field.

Transcriptomics in human ageing

Until now, transcriptomics studies into human ageing utilise candidate gene-based and explorative approaches to identify which mRNA expression levels are regulated in an age-dependent fashion. Study designs mainly consist of comparisons between young and elderly individuals, sometimes including centenarians, or a comparison of RNA expression profiles in patients with progeroid syndromes to young and elderly individuals. Transcriptomic studies into human cancers require several tens of samples to identify significant expression patterns changing at least two fold between cancer and healthy tissue from the same patient (32). The expected differences due to ageing are much smaller and more difficult to detect given the large inter-individual genomic variation in humans that is not due to ageing (33). Large experimental groups as well

as a good study designs are required to enable reproducible conclusions to be made from studies of gene expression and ageing.

Candidate gene-based studies

Research based on candidate genes is generally performed in a cross-sectional design using cultured skin fibroblasts (34;35), lymphocytes (36), and peripheral blood mononuclear cells (PBMC) (37). The majority of the studies were performed by quantitative PCR methods (34;35;37), others use Northern blot analysis (36) to investigate specific candidate genes, such as heat shock protein HSP70 (36), osteoclastogenesis inhibitory factor (OCIF) (34), several proteasome subunits (35) and caspases (37). Some statistically significant results were obtained from candidate gene-based studies (34-36) investigating a small number of genes in only 2-5 healthy females per group. Centenarians appeared to have an expression level that resembles that of young (<40 yrs) individuals more than that of elderly individuals (65-80 yrs) (34-36). The authors interpreted this result as an indication that the expression levels of these genes of long-lived individuals as well as in younger ones reflects the healthier condition of these individuals as compared to elderly individuals.

In a relatively large study of 246 individuals aged 10-102 years old, Lacelle et al. (37) investigated gene expression of 4 caspases as a function of chronological age. In general, their results indicated a larger variation within an age category than across categories. Caspase-1 expression is increased in elderly (70-89 yrs) and long-lived (90+ yrs) individuals compared to younger (≤69 yrs) individuals. Caspase-8 expression was increased in the elderly only and caspase-3 in the long-lived individuals only, when compared to the younger age group. Expression of caspase-10 was similar across all age groups. In this cohort, a high level of caspase-8 is the biosignature for elderly individuals, while high levels of caspase-3 and low levels of caspase-8 mRNA expressions were the hallmark biosignature for nonagenarians and centenarians. With regard to both caspase-1 and -3, population profiles revealed a slow shift with age toward higher levels of caspase mRNA expression. The authors hypothesize that these age-related shifts are important, as they may indicate that a subgroup of individuals within the younger population, expressing high levels of caspase-3 mRNA in their PBMC, is favoured to attain longevity. In the cross-sectional design of this study, however, such hypotheses can not be tested, as it is not possible to establish whether the changes observed were the cause or the consequence of an older age. A large prospective study of individuals followed from middle age onwards for medical history and

lifestyle until the date of death would be able to reveal whether expression changes predict a decline of health and mortality risk. Such large human epidemiological studies do exist, and their follow up period is 15 to over 50 years (38;39).

Genome-wide studies

The introduction of the whole genome microarrays for gene expression research provided the opportunity to investigate the expression of the whole human transcriptome. Transcriptomic studies with an explorative approach could provide useful information about co-related genes and biological processes involved in ageing, including those that are not yet considered as candidates. Explorative expression studies have been performed on home-spotted microarrays containing ~4.000-6.000 transcripts (40-42) and on commercial Affymetrix microarrays, with 12.000 (43;44) to ~54.000 (45-49) transcripts on each array. The design commonly used was the cross-sectional comparison of young and elderly healthy individuals, about eight individuals maximum per group or an analysis of individuals across an age-range. Various tissues were investigated in these studies, such as cultured lymphocytes (40), fibroblast cell lines (41;42;46), brain (43;44), kidney (48), and skeletal muscle (45;47;49).

Several pathways were highlighted by differentially expressed genes between young (19-29 yrs) and elderly (65-85 yrs) individuals (40), including energy metabolism, cell cycle, signal transduction, and DNA repair (see Table 1 for a summary)(45;49). Other studies additionally included the comparison of individuals with progeria, which are genetic disorders that cause segmental premature ageing such as Werner syndrome (WS) (41) and Hutchinson-Gilford progeria syndrome (HGPS) (42), to represent ageing mechanisms in the general population. These studies showed that genes from important pathways act similarly in old (87-96 yrs) individuals and progeria patients (7-37 yrs) suggesting that some but not all normal ageing mechanisms are accelerated in progeroid syndromes. There is still debate as to whether progeria is a good model for human ageing (50).

Only four large transcriptomic studies were performed investigating age ranges from 13-106 years (43;44;47;48), using sample sizes from 30-81 individuals. Applying Affymetrix microarrays with ~12.000-54.000 transcripts per array, these groups found 250-985 transcripts that differ significantly in gene expression as a function of age in several tissues of healthy donors, with fold changes below two (43;44). Biological pathways

found to be changed with age are listed in table 1 and includes genes involved in the mitochondrial electron transport chain, cell cycle and extracellular matrix.

In addition to limited sample sizes, a major problem in gene expression studies of healthy humans is the restricted availability of relevant human tissue. Zahn et al. (47) performed a meta-analysis on data of multiple studies investigating RNA from muscle, kidney and brain in a cross-sectional design. There was almost no correlation between the expression levels, or changes with age of individual genes in the different tissues, meaning that in each tissue different individual genes change their expression with age. Importantly, using pathway analysis, gene expression changes influencing six pathways were consistently found in all three investigated tissues (the first six pathways that are listed in Table 1). These results indicate that these pathways, but not particularly individual genes, are common elements in the age-related expression changes of different tissues. This may implicate that in addition to tissue specific effects, a common ageing signature may be found in any tissue that reflects the age of the whole organism. This would give major implications for human epidemiological studies, for which frequently only blood is available.

Table 1. Pathways found to be changed by ageing in explorative human transcriptomic studies.

Pathways	Age range studies	N total	Young vs. old studies	N case / N control
Electron transport chain	Lu et al. 2004 (43) Erraji-Benckekroun et al. 2005 (44) Rodwell et al. 2004 (48) Zahn et al. 2006 (47)	30 39 74 81	Visala Rao et al. 2003 (40) Kyng et al. 2003 (41)	2 / 2 5 / 6
Cell cycle/cell growth	Lu et al. 2004 (43) Erraji-Benckekroun et al. 2005 (44) Rodwell et al. 2004 (48) Zahn et al. 2006 (47)	30 39 74 81	Visala Rao et al. 2003 (40) Ly et al. 2000 (42) Kyng et al. 2003 (41) Welle et al. 2003 (49)	2 / 2 3 / 2 5 / 6 8 / 8
Extracellular matrix	Lu et al. 2004 (43) Rodwell et al. 2004 (48) Zahn et al. 2006 (47)	30 74 81	Visala Rao et al. 2003 (40) Ly et al. 2000 (42) Csoka et al. 2004 (46)	2 / 2 3 / 2 4 / 3
Chloride transport	Lu et al. 2004 (43) Rodwell et al. 2004 (48) Zahn et al. 2006 (47)	30 74 81		
Complement activation	Lu et al. 2004 (43) Rodwell et al. 2004 (48) Zahn et al. 2006 (47)	30 74 81		

Ribosomes	Lu et al. 2004 (43) Rodwell et al. 2004 (48) Zahn et al. 2006 (47)	30 74 81		
Inflammation	Lu et al. 2004 (43) Erraji-Benckekroun et al. 2005 (44)	30 39	Visala Rao et al. 2003 (40) Kyng et al. 2003 (41) Welle et al. 2003 (49)	2 / 2 5 / 6 8 / 8
Signal transduction	Erraji-Benckekroun et al. 2005 (44) Rodwell et al. 2004 (48)	39 74	Visala Rao et al. 2003 (40) Kyng et al. 2003 (41) Welle et al. 2003 (49)	2 / 2 5 / 6 8 / 8
Protein metabolism	Lu et al. 2004 (43) Erraji-Benckekroun et al. 2005 (44)	30 39	Ly et al. 2000 (42) Welle et al. 2003 (49)	3 / 2 8 / 8
Transcription	Lu et al. 2004 (43) Rodwell et al. 2004 (48)	30 74	Csoka et al. 2004 (46) Welle et al. 2003 (49)	4 / 3 8 / 8
Lipid metabolism	Lu et al. 2004 (43) Erraji-Benckekroun et al. 2005 (44)	30 39	Visala Rao et al. 2003 (40)	2 / 2
Ca homeostasis/signalling	Lu et al. 2004 (43) Erraji-Benckekroun et al. 2005 (44)	30 39		
MAP kinase/PKC family	Lu et al. 2004 (43) Erraji-Benckekroun et al. 2005 (44)	30 39		
Oxidative stress	Lu et al. 2004 (43) Erraji-Benckekroun et al. 2005 (44)	30 39		
Stress response/DNA repair	Lu et al. 2004 (43)	30	Visala Rao et al. 2003 (40) Ly et al. 2000 (42) Kyng et al. 2003 (41) Welle et al. 2004 (45) Welle et al. 2003 (49)	2 / 2 3 / 2 5 / 6 8 / 7 8 / 8
Apoptosis	Erraji-Benckekroun et al. 2005 (44)	39	Kyng et al. 2003 (41)	5 / 6
Energy metabolism	Erraji-Benckekroun et al. 2005 (44)	39	Welle et al. 2003 (49)	8 / 8
Hormonal	Lu et al. 2004 (43)	30	Welle et al. 2003 (49)	8 / 8
Synaptic transmission	Lu et al. 2004 (43)	30		
Vesicular transport	Lu et al. 2004 (43)	30		

A striking observation in the genome-wide studies was that individuals younger than 30 yrs showed very similar expression patterns (43). Also, the age group older than 90 yrs of age showed very similar expression patterns, although this homogenous pattern is different from that of the age group younger than 30 yrs. However, the individuals aged between 30 and

90 showed highly variable expression patterns of which some are more similar to the young age group and others to the old age group. This may coincide with individuals diverging in their rates of ageing as they go through middle age and later their diversity in age at death. The heterogeneity in gene expression that becomes detectable from the age of 30 onwards may be an intermediate phenotype or biomarker, on the condition that it would represent biological differences in the ageing process of individuals that precede and contribute to disease and mortality. Two studies (48)provided proof that this indeed was the case, by showing that the gene expression profiles of human muscle and kidney corresponded to the measured physiological functionality (the 'biological age') of the tissue rather than to the chronological age of the subject (47).

Animal ageing models

A large number of studies have investigated the changes in gene expression in relation to ageing and longevity in laboratory animals using commercially available microarrays. The most frequently used animal model for this research is the mouse (51-54), but nematodes (55), flies (55-57), rats (58) and monkeys (59;60) have also been investigated. A range of tissues were studied such as liver, heart, muscle, aorta and brain. Across all species and in most experimental designs there is a gender effect on ageing features and gene expression (59;61).

The main design for measuring chronological age-changes, as in humans, was a comparison of young and old individuals, using on average three animals per group. Since such studies are usually performed in inbred strains, the variation between individual animals is smaller than among human individuals. A meta-analysis showed that, among many observed effects that were not comparable between species, the mitochondrial electron transport chain genes showed a consistent overall decrease in expression with chronological age in humans, mice, and flies, but not in nematodes (47). Interestingly, this pathway also consistently changed with age in different human tissues as mentioned above (47). Another meta-analysis found that genes involved in the oxidative stress response were similarly changed with age in human, mouse and rat brain as well as in nematodes (43).

Together these studies will eventually provide insight in the gene expression changes just associated with chronological age, those with biological contribution to the ageing process of each species and those pointing at evolutionary conserved mechanisms of longevity. Some short-

lived mutants are expected to represent a model for human progeria, whereas long-lived mutants and caloric restricted animals may represent familial and sporadic (non-familial) human longevity, respectively. Caloric Restriction (CR) refers to a diet with fewer calories than is typical, but which contains all the necessary nutrients and vitamins to support life. The comparison of caloric restricted rodents and littermates that did not receive CR showed that CR resulted in a partial inhibition of age-related changes in gene expression (51-54;58). This indicates that CR, as it prolongs lifespan, also influences the age-related changes in expression profiles in a way that may point to a decreased rate of biological ageing. These results were consistent among mice, flies and nematodes but could not be confirmed in monkeys (60).

The question is whether pathways affected by CR were also affected by mutations inducing increased lifespan in these species. A recent study compared gene expression profiles associated with mutations inducing longevity by reduced IIS signalling in nematodes, flies and mice (62). Conservation of the IIS related longevity regulation was tested at the level of individual orthologous genes, for paralogous gene sets and at the level of biological processes reflected by gene classes. There was hardly any conservation of IIS regulation at the gene level between the species. However, the process analysis revealed that the conserved element in IIS regulation of the long-lived mutants associated with up-regulation of sugar catabolism and energy generation, with cellular detoxification pathways and with down-regulation of protein biosynthesis and translation. The authors suggested that of these IIS regulated processes those that most likely reflect the longevity regulation, because of findings in other ageing studies, are lowered protein biosynthesis and increased cellular detoxification reflected by Glutathione-S-transferase (GST) activity.

Future directions for human research

Clearly, genomic studies in animal models generate many hypotheses that are worth testing in humans, where the possibilities for investigating healthy individuals are very limited. Here we will discuss what improvements may be accomplished.

Sample size

As for genetic association studies, the sample size of a gene expression study in humans greatly determines the credibility of the results, especially concerning the multiple testing performed in genome-wide explorative

analyses. A review from Allison et al. (63) nicely summarised the main statistical issues in expression studies. Minimal sample size calculations depend on design, the total background variation within the data, the magnitude of the expression change that is to be detected, the desired statistical power or probability to detect differentially expressed genes, the specified error rate, and the statistical method being used to detect the changes in gene expression. They advocated that the minimal, and certainly not optimal, sample size to be used in expression studies is five biological cases per group. This amount is only useable when testing for differential expression with fold changes of three and up (64), and then the reproducibility of the results is still questionable. Since fold changes in differential gene expression are expected to be below two in human ageing, it has been suggested that a minimum of 25 individuals per group is required in the analysis of a case control design (65;66). Large sample sizes are thus required. As a consequence, the primary studies aimed at the identification of biological processes relevant for human ageing by gene expression approaches are bound to be performed on easily available tissue from healthy individuals such as blood, buccal cells from the mouth and skin fibroblasts.

The value of studying whole blood in transcriptomics

Because blood is an easily available tissue, several studies have investigated gene expression profiles in whole human blood. Whole blood mRNA was used for studies into Huntington disease, Down syndrome, multiple sclerosis, psoriatic arthritis, Alzheimer's disease and epilepsy (67-70). In a candidate gene study for schizophrenia, 21 out of the 31 candidate genes expressed in the brain were similarly expressed in whole blood (67). Gene expression in blood is also shown to mirror disease state in multiple sclerosis (20). In general, the expression levels of genes as investigated on a 34K Affymetrix microarray show an average correlation of 0.5 when comparing blood and other tissues (67). Blood has also been successfully used for classification of disease subgroups, for progression prediction and to investigate the effects of medical treatment (71-74). It is expected that blood reflects a limited proportion of the biological processes relevant for ageing and longevity. However, ageing does not affect one specific tissue and blood both initiates and reflects processes that affect the whole body. Therefore, studying gene expression in blood cells is a good start for finding the relevant intermediate phenotype that emerges from genetic variation for ageing and longevity. Gene expression profiles in blood were found to be highly similar in monozygotic and moderately similar in dizygotic twin pairs, with a heritability of 0.33 (75-77). Variation in gene

expression in healthy unrelated individuals is found to be limited and detected as a result of blood cell subsets, gender, age, time of day and individual differences (78-80).

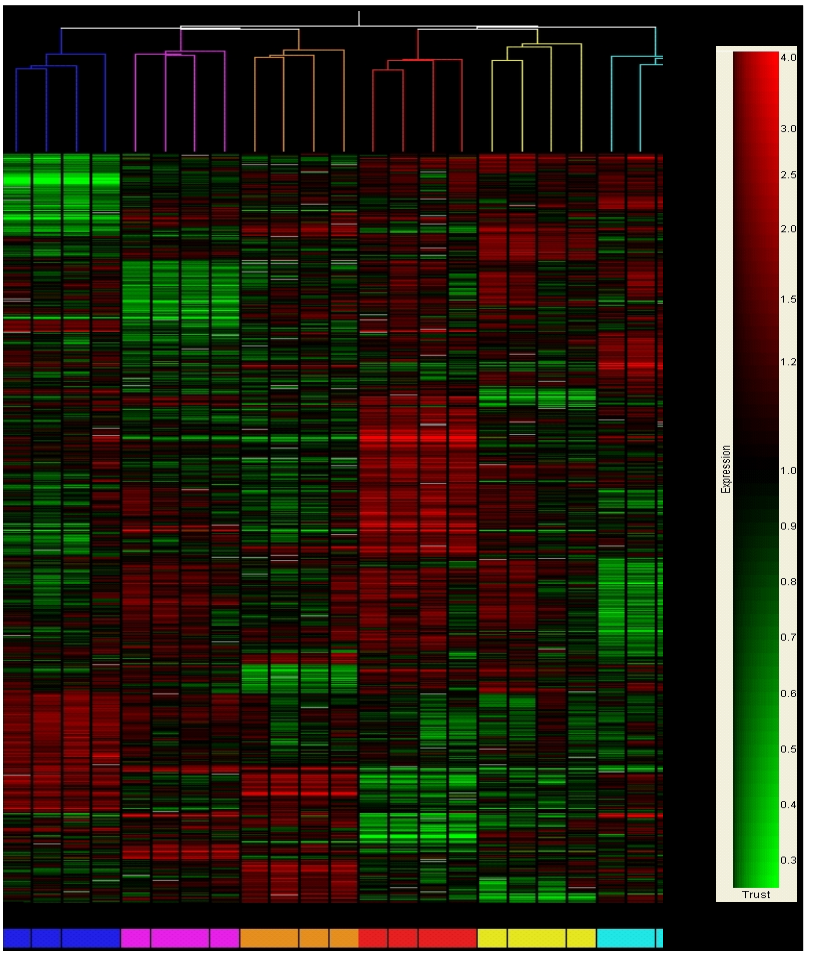

Figure 1. Gene expression in whole blood of volunteers.
Two-dimensional unsupervised hierarchical cluster analysis of 1548 most differentially expressed genes between individuals is shown. Each column represents a sample and each row a gene. The colorgram depicts high (red) and low (green) relative levels of gene expression. Grey means the data is not available. The four different time points of one individual clustered together, and so did the males and females (from left to right: female 1, 2 and 3 and male 1, 2 and 3, each colour on the bottom axis represents one individual).

As an illustration to the latter point, we compared gene expression profiles of mRNA isolated from whole blood samples obtained from three healthy

married volunteer couples (61-76 yrs). Blood was taken on four occasions; on two different days (three weeks apart) and in the morning (10 a.m.) as well as in the afternoon (2 p.m.). Gene expression was measured using CodeLink whole genome microarrays (~54.000 transcripts). Figure 1 provides an overview of the 1548 genes most variable between the individuals, which showed very similar expression patterns at the four time points within an individual, suggesting a stable expression profile over this time interval. In addition, statistical analysis of this pilot comparison revealed that the expression of very few genes was influenced by the time of day, or day at which the blood was collected. Most differentially expressed genes were associated with individual variation, including gender.

All in all, current data justifies the use of whole blood in the study of human ageing. However, as will be the case for including only one tissue, much of the relevant biological context will be missed. Therefore, such investigations should be linked to parallel studies in animal models. These models can show what interactions between biological processes with each other and with environmental stimuli are reflected in blood. As has been shown in animal studies, the human studies will become more informative if study designs allow for a combined analysis of genetic variation, gene expression and environmental variation.

Study design

Cross-sectional designs comparing young and elderly individuals are thus far most frequently used in gene expression studies of human ageing. In these designs there is no possibility to distinguish causal changes that are of major importance in determining biological ageing, mortality and longevity of individuals, and chronological changes that reflect effects thereof. Prospective study designs are available in large epidemiological human cohort studies of 20-30 years to follow up, that may indicate which RNA profiles (if cells are available in the study) precede and predict changes in physiological health and mortality risk. Twin studies will add vital information as to how genetic variation and environmental factors relevant for human ageing influence gene expression profiles throughout life. Comparable to long-lived strains in animal models, human lifespan variation is studied also in families of exceptional longevity. Especially the offspring of long-lived individuals may reflect which RNA profiles associate to the familial longevity trait at middle age (13). The middle aged offspring of centenarians do express part of the beneficial phenotype already in the sense that they have a life-long decreased risk of mortality from coronary

heart disease, diabetes and cancer. This can be viewed as secondary to a decreased rate of biological ageing (4) occurring in longevity families.

As an illustration of how new designs can be used in genomic investigations of human ageing and longevity we will now discuss the concerted genetic and gene expression studies in the Leiden Longevity Study. The aim of this research is to identify genomic and biochemical parameters associated with human longevity by combining research in familial long-lived individuals and their family members and in sporadic long-lived individuals from the general population. The Leiden Longevity Study (Figure 2) consists of 420 Caucasian families with at least two long-lived siblings (men aged 89 years or above; women aged 91 years or above, n=960), the middle aged offspring (n=1764) and the partners of this offspring (n=757). In 2001, less than 0.5% of the Dutch population fulfilled these sex-specific criteria (14). The survival benefit of the long-lived families is marked by a 30% decreased mortality risk observed in the survival analysis of three generations, the parents of the sibling pairs, their deceased siblings and their offspring (14). The familial survival advantage thereby exceeds the increased lifespan expectancy in the last generations of western societies due to non-genetic factors (improved nutrition, hygiene and health care). Therefore, this selected population is enriched for heritable familial longevity. An ongoing genome-wide association study using Affymetrix 500k SNP arrays should reveal which genetic variation is associated with familial longevity and which individuals carry this variation.

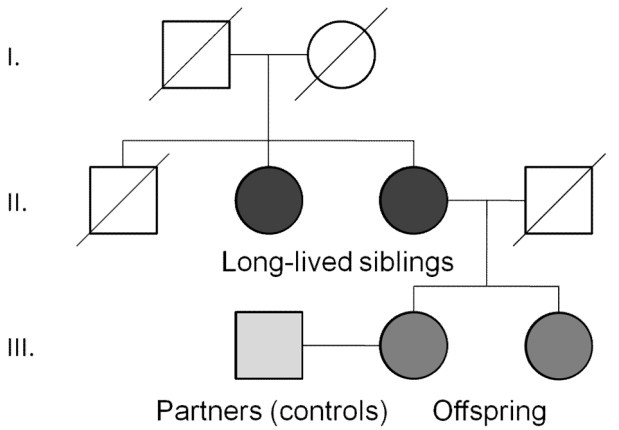

Figure 2. Design of the Leiden Longevity study
Long-lived families were recruited if they included a long-lived sibling pair with males aged 89 y or older and women aged 91 y or older.

For the purpose of biomarker identification, including gene expression profiling, a cross-sectional comparison of young controls (partners of offspring) and elderly cases (long-lived siblings) may be performed. Chronological age changes will also be observed in an analysis of the age range of each subgroup: long-lived siblings (89-103 yrs), offspring (34-80 yrs) and partners (30-79 yrs). Biologically relevant age-changes in biomarkers will emerge from a comparison of offspring and partners. The offspring can be viewed as cases with a higher susceptibility to become long-lived and their partners, representatives of the general population, as controls (14). The key advantage of this design is that the offspring and their partners are of similar age and have a similar environment. Therefore, differentially expressed genes in the comparison of these two groups could represent processes associating to familial longevity detectable at middle age and thereby preceding the long-lived phenotype. Some of these genes will become even more marked in their differential expression when studied in a comparison of the elderly long-lived cases and controls. This methodology has revealed metabolic markers (81;82) as potential biomarkers and is now being used for gene expression studies. Whole blood RNA samples from 150 individuals of the Leiden Longevity study are being analysed (50 long-lived siblings, 50 offspring and 50 partners). At the level of the individual, the expression profile or expression level of specific genes can be linked to the genetic variation that was found to associate to familial longevity in the genetic studies. This analysis is expected to reveal causal changes in biological processes associated with longevity. As stated before, to obtain an understanding of the biology of ageing, it is highly relevant to investigate interactions of genomic (genetic and gene expression) variation with the environmental component. For this purpose data on nutrition, lifestyle and socioeconomic status are being collected. Although in human studies the biological depth of such information is limited, those are the factors that may be beneficially manipulated to improve health.

The proof of the pudding for a parameter to be a biomarker of biological ageing will be established only if it predicts decline of physiological health and mortality risk in a prospective survey of the middle aged offspring and partners. A study of such a marker in a prospective population based cohort study, such as the Leiden 85+ study (83), will reveal whether it marks ageing mechanisms relevant only for familial longevity or also for population wide longevity. To illustrate the value of this methodology, we have recently observed that low concentrations of small LDL (Low Density Lipoprotein) particles is a feature of familial longevity, detectable at middle age and is expressed especially in healthy familial nonagenarians (health

based on scores on the Activity of Daily Living questionnaire), but also in sporadic nonagenarians in the population (82).

Concluding remarks

Genetic and transcriptomic approaches have both provided insight into the biological processes involved in ageing. Transcriptomic studies identified a number of promising pathways that may be affected by ageing in different species, including humans, and across different tissues within a species, among which the electron transport chain and oxidative stress response promising. These pathways are connected because when synthesizing ATP, the mitochondrial electron transport chain releases electrons which can reduce oxygen, forming reactive oxygen species (ROS) such as superoxide. These ROS can induce oxidative stress and may contribute to the decline in mitochondrial function associated with the ageing process (84). Caloric restriction has been found to not only increase the life-span of rodents, but also to reduce ROS production (85). This can be explained when high amounts of energy are available, mitochondria do not operate very efficiently and generate more ROS, while when calorie restricted, energy is conserved and there is less free radical generation. Glutathione-S-transferases are involved in cellular detoxification and have recently been linked to longevity across species (62). Glutathione-S-transferase (GST) can act as an antioxidative enzyme by protecting against damaging ROS and thereby decreasing oxidative stress in cells (86).

Gene expression studies in animal models have identified various biological processes that can serve as hypotheses in human studies. However, for a true testing of such hypotheses and the identification of human pathways, improvements in methodology are vital. These include larger sample sizes, the concerted application of population/patients based designs and familial designs that can distinguish chronological age effects from biological ageing. Highly relevant is also the combined analysis of genetic variation, gene expression variation and environmental factors. Figure 3 represents a scheme depicting how genomic studies may be integrated in research aimed at disentangling the complex biology of ageing. Any input study of lifespan variation whether associated with naturally occurring genetic variation (human cohorts, or different animal strains), induced by mutations or by modulation of the environment (such as in caloric restriction) may be studied for gene expression changes.

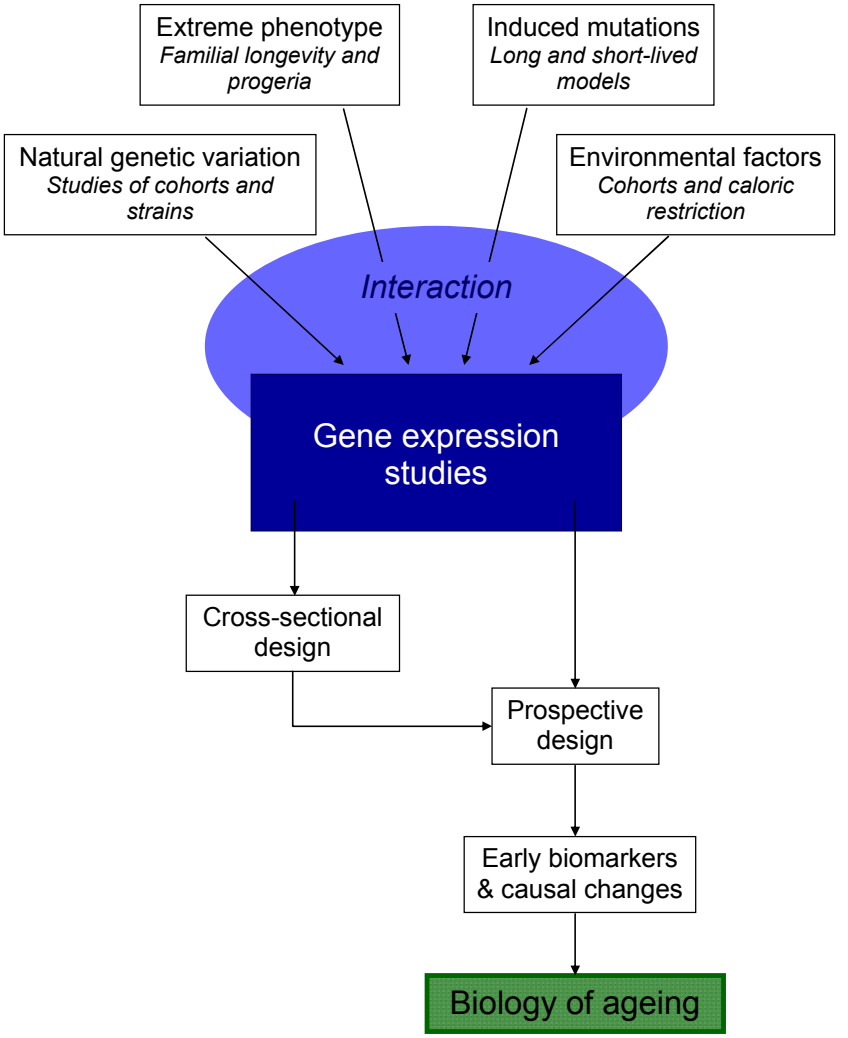

Figure 3. Scheme depicting the integration of genomic studies into the biology of ageing

Gene expression studies may directly reveal leads to relevant mechanisms in a species, but especially the integration of gene expression studies at the level of biological processes (such as in (62)) will provide hypotheses as to whether and which gene-environmental interactions associate to common processes determining biological ageing rate. It is also necessary to integrate cross-sectional studies monitoring the profile as a chronological marker of ageing, and prospective studies indicating which of these is a biomarker preceding and predicting physiological decline and mortality. The

integration of such information with knowledge emerging from gene-environmental interactions will reveal causal mechanisms for ageing that can be tested simultaneously in animal models and in humans.

Human studies should include larger sample sizes to increase the power of the analyses and reliability of the results, better systematic surveys of gene expression changes over a lifetime and designs that allow a good integration with genetic studies. Understanding the aetiology of a complex trait such as longevity will require more than just transcriptomics. There will be a need to combine genetics, transcriptomics, proteomics and metabolomics in order to gain further knowledge about biological processes involved in ageing and longevity.

Acknowledgements

We acknowledge Unilever Colworth for their help in arraying the blood samples.
The efforts in the Leiden Longevity Study are supported by a grant from Innovation Orientated research Program IOP on Genomics (SenterNovem IGE5007), by the stimulation grant from the Netherlands Genomics Initiative/Netherlands Organisation for Scientific Research (NGI/NWO) (grant number 05040202; Netherlands Initiative into Healthy Ageing) and by the Centre for Medical Systems Biology (CMSB).

References

(1) Herndon LA, Schmeissner PJ, Dudaronek JM, Brown PA, Listner KM, Sakano Y, et al. Stochastic and genetic factors influence tissue-specific decline in ageing C. elegans. Nature 2002 Oct 24;419(6909):808-14.

(2) Vijg J, Suh Y. Functional genomics of ageing. Mech Ageing Dev 2003 Jan;124(1):3-8.

(3) Kenyon C. The plasticity of aging: insights from long-lived mutants. Cell 2005 Feb 25;120(4):449-60.

(4) Terry DF, Wilcox MA, McCormick MA, Pennington JY, Schoenhofen EA, Andersen SL, et al. Lower all-cause, cardiovascular, and cancer mortality in centenarians' offspring. J Am Geriatr Soc 2004 Dec;52(12):2074-6.

(5) Orwoll ES, Bauer DC, Vogt TM, Fox KM. Axial bone mass in older women. Study of Osteoporotic Fractures Research Group. Ann Intern Med 1996 Jan 15;124(2):187-96.

(6) Tatar M. Diet restriction in Drosophila melanogaster. Design and analysis. Interdiscip Top Gerontol 2007;35:115-36.

(7) Guarente L, Kenyon C. Genetic pathways that regulate ageing in model organisms. Nature 2000 Nov 9;408(6809):255-62.

(8) Christensen K, Johnson TE, Vaupel JW. The quest for genetic determinants of human longevity: challenges and insights. Nat Rev Genet 2006 Jun;7(6):436-48.

(9) Heijmans BT, Westendorp RG, Slagboom PE. Common gene variants, mortality and extreme longevity in humans. Exp Gerontol 2000 Sep;35(6-7):865-77.

(10) Corder EH, Lannfelt L, Viitanen M, Corder LS, Manton KG, Winblad B, et al. Apolipoprotein E genotype determines survival in the oldest old (85 years or older) who have good cognition. Arch Neurol 1996 May;53(5):418-22.

(11) Skytthe A, Pedersen NL, Kaprio J, Stazi MA, Hjelmborg JV, Iachine I, et al. Longevity studies in GenomEUtwin. Twin Res 2003 Oct;6(5):448-54.

(12) 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 Mar;97(3):319-23.

(13) Perls TT, Bubrick E, Wager CG, Vijg J, Kruglyak L. Siblings of centenarians live longer. Lancet 1998 May 23;351(9115):1560.

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

(15) Perls TT, Wilmoth J, Levenson R, Drinkwater M, Cohen M, Bogan H, et al. Life-long sustained mortality advantage of siblings of centenarians. Proc Natl Acad Sci U S A 2002 Jun 11;99(12):8442-7.

(16) Alizadeh AA, Eisen MB, Davis RE, Ma C, Lossos IS, Rosenwald A, et al. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. Nature 2000 Feb 3;403(6769):503-11.

(17) Perou CM, Sorlie T, Eisen MB, van de RM, Jeffrey SS, Rees CA, et al. Molecular portraits of human breast tumours. Nature 2000 Aug 17;406(6797):747-52.

(18) Baird AE. Blood genomics in human stroke. Stroke 2007 Feb;38(2 Suppl):694-8.

(19) Sharp FR, Xu H, Lit L, Walker W, Pinter J, Apperson M, et al. Genomic profiles of stroke in blood. Stroke 2007 Feb;38(2 Suppl):691-3.

(20) Achiron A, Gurevich M. Peripheral blood gene expression signature mirrors central nervous system disease: the model of multiple sclerosis. Autoimmun Rev 2006 Oct;5(8):517-22.

(21) Izuhara K, Saito H. Microarray-based identification of novel biomarkers in asthma. Allergol Int 2006 Dec;55(4):361-7.

(22) Heller RA, Schena M, Chai A, Shalon D, Bedilion T, Gilmore J, et al. Discovery and analysis of inflammatory disease-related genes using cDNA microarrays. Proc Natl Acad Sci U S A 1997 Mar 18;94(6):2150-5.

(23) Laprise C, Sladek R, Ponton A, Bernier MC, Hudson TJ, Laviolette M. Functional classes of bronchial mucosa genes that are differentially expressed in asthma. BMC Genomics 2004 Mar 23;5(1):21.

(24) Hauser MA, Li YJ, Xu H, Noureddine MA, Shao YS, Gullans SR, et al. Expression profiling of substantia nigra in Parkinson disease, progressive supranuclear palsy,

and frontotemporal dementia with parkinsonism. Arch Neurol 2005 Jun;62(6):917-21.

(25) Nagasaka Y, Dillner K, Ebise H, Teramoto R, Nakagawa H, Lilius L, et al. A unique gene expression signature discriminates familial Alzheimer's disease mutation carriers from their wild-type siblings. Proc Natl Acad Sci U S A 2005 Oct 11;102(41):14854-9.

(26) Chang JC, Wooten EC, Tsimelzon A, Hilsenbeck SG, Gutierrez MC, Elledge R, et al. Gene expression profiling for the prediction of therapeutic response to docetaxel in patients with breast cancer. Lancet 2003 Aug 2;362(9381):362-9.

(27) Golub TR, Slonim DK, Tamayo P, Huard C, Gaasenbeek M, Mesirov JP, et al. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science 1999 Oct 15;286(5439):531-7.

(28) 't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AA, Mao M, et al. Gene expression profiling predicts clinical outcome of breast cancer. Nature 2002 Jan 31;415(6871):530-6.

(29) Roepman P, Wessels LF, Kettelarij N, Kemmeren P, Miles AJ, Lijnzaad P, et al. An expression profile for diagnosis of lymph node metastases from primary head and neck squamous cell carcinomas. Nat Genet 2005 Feb;37(2):182-6.

(30) Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc Natl Acad Sci U S A 2001 Sep 11;98(19):10869-74.

(31) Goeman JJ, van de Geer SA, de Kort F, van Houwelingen HC. A global test for groups of genes: testing association with a clinical outcome. Bioinformatics 2004 Jan 1;20(1):93-9.

(32) Michiels S, Koscielny S, Hill C. Interpretation of microarray data in cancer. Br J Cancer 2007 Apr 23;96(8):1155-8.

(33) Han E, Hilsenbeck SG. Array-based gene expression profiling to study aging. Mech Ageing Dev 2001 Jul 31;122(10):999-1018.

(34) Chondrogianni N, de CMS, Franceschi C, Gonos ES. Cloning of differentially expressed genes in skin fibroblasts from centenarians. Biogerontology 2004;5(6):401-9.

(35) Chondrogianni N, Petropoulos I, Franceschi C, Friguet B, Gonos ES. Fibroblast cultures from healthy centenarians have an active proteasome. Exp Gerontol 2000 Sep;35(6-7):721-8.

(36) Ambra R, Mocchegiani E, Giacconi R, Canali R, Rinna A, Malavolta M, et al. Characterization of the hsp70 response in lymphoblasts from aged and centenarian subjects and differential effects of in vitro zinc supplementation. Exp Gerontol 2004 Oct;39(10):1475-84.

(37) Lacelle C, Xu S, Wang E. Identification of high caspase-3 mRNA expression as a unique signature profile for extremely old individuals. Mech Ageing Dev 2002 Apr 30;123(8):1133-44.

(38) Hofman A, Grobbee DE, de Jong PT, van den Ouweland FA. Determinants of disease and disability in the elderly: the Rotterdam Elderly Study. Eur J Epidemiol 1991 Jul;7(4):403-22.

(39) Dawber TR, Meadors GF, Moore FE, Jr. Epidemiological approaches to heart disease: the Framingham Study. Am J Public Health Nations Health 1951 Mar;41(3):279-81.

(40) Visala RD, Boyle GM, Parsons PG, Watson K, Jones GL. Influence of ageing, heat shock treatment and in vivo total antioxidant status on gene-expression profile and protein synthesis in human peripheral lymphocytes. Mech Ageing Dev 2003 Jan;124(1):55-69.

(41) Kyng KJ, May A, Kolvraa S, Bohr VA. Gene expression profiling in Werner syndrome closely resembles that of normal aging. Proc Natl Acad Sci U S A 2003 Oct 14;100(21):12259-64.

(42) Ly DH, Lockhart DJ, Lerner RA, Schultz PG. Mitotic misregulation and human aging. Science 2000 Mar 31;287(5462):2486-92.

(43) Lu T, Pan Y, Kao SY, Li C, Kohane I, Chan J, et al. Gene regulation and DNA damage in the ageing human brain. Nature 2004 Jun 24;429(6994):883-91.

(44) Erraji-Benchekroun L, Underwood MD, Arango V, Galfalvy H, Pavlidis P, Smyrniotopoulos P, et al. Molecular aging in human prefrontal cortex is selective and continuous throughout adult life. Biol Psychiatry 2005 Mar 1;57(5):549-58.

(45) Welle S, Brooks AI, Delehanty JM, Needler N, Bhatt K, Shah B, et al. Skeletal muscle gene expression profiles in 20-29 year old and 65-71 year old women. Exp Gerontol 2004 Mar;39(3):369-77.

(46) Csoka AB, English SB, Simkevich CP, Ginzinger DG, Butte AJ, Schatten GP, et al. Genome-scale expression profiling of Hutchinson-Gilford progeria syndrome reveals widespread transcriptional misregulation leading to mesodermal/mesenchymal defects and accelerated atherosclerosis. Aging Cell 2004 Aug;3(4):235-43.

(47) Zahn JM, Sonu R, Vogel H, Crane E, Mazan-Mamczarz K, Rabkin R, et al. Transcriptional profiling of aging in human muscle reveals a common aging signature. PLoS Genet 2006 Jul;2(7):e115.

(48) Rodwell GE, Sonu R, Zahn JM, Lund J, Wilhelmy J, Wang L, et al. A transcriptional profile of aging in the human kidney. PLoS Biol 2004 Dec;2(12):e427.

(49) Welle S, Brooks AI, Delehanty JM, Needler N, Thornton CA. Gene expression profile of aging in human muscle. Physiol Genomics 2003 Jul 7;14(2):149-59.

(50) Kyng KJ, Bohr VA. Gene expression and DNA repair in progeroid syndromes and human aging. Ageing Res Rev 2005 Nov;4(4):579-602.

(51) Cao SX, Dhahbi JM, Mote PL, Spindler SR. Genomic profiling of short- and long-term caloric restriction effects in the liver of aging mice. Proc Natl Acad Sci U S A 2001 Sep 11;98(19):10630-5.

(52) Lee CK, Allison DB, Brand J, Weindruch R, Prolla TA. Transcriptional profiles associated with aging and middle age-onset caloric restriction in mouse hearts. Proc Natl Acad Sci U S A 2002 Nov 12;99(23):14988-93.

(53) Lee CK, Weindruch R, Prolla TA. Gene-expression profile of the ageing brain in mice. Nat Genet 2000 Jul;25(3):294-7.

(54) Lee CK, Klopp RG, Weindruch R, Prolla TA. Gene expression profile of aging and its retardation by caloric restriction. Science 1999 Aug 27;285(5432):1390-3.

(55) McCarroll SA, Murphy CT, Zou S, Pletcher SD, Chin CS, Jan YN, et al. Comparing genomic expression patterns across species identifies shared transcriptional profile in aging. Nat Genet 2004 Feb;36(2):197-204.

(56) Lai CQ, Parnell LD, Lyman RF, Ordovas JM, Mackay TF. Candidate genes affecting Drosophila life span identified by integrating microarray gene expression analysis and QTL mapping. Mech Ageing Dev 2007 Mar;128(3):237-49.

(57) Girardot F, Lasbleiz C, Monnier V, Tricoire H. Specific age-related signatures in Drosophila body parts transcriptome. BMC Genomics 2006;7:69.

(58) Sreekumar R, Unnikrishnan J, Fu A, Nygren J, Short KR, Schimke J, et al. Effects of caloric restriction on mitochondrial function and gene transcripts in rat muscle. Am J Physiol Endocrinol Metab 2002 Jul;283(1):E38-E43.

(59) Qiu H, Tian B, Resuello RG, Natividad FF, Peppas A, Shen YT, et al. Sex-specific regulation of gene expression in the aging monkey aorta. Physiol Genomics 2007 Apr 24;29(2):169-80.

(60) Kayo T, Allison DB, Weindruch R, Prolla TA. Influences of aging and caloric restriction on the transcriptional profile of skeletal muscle from rhesus monkeys. Proc Natl Acad Sci U S A 2001 Apr 24;98(9):5093-8.

(61) Tower J. Sex-specific regulation of aging and apoptosis. Mech Ageing Dev 2006 Sep;127(9):705-18.

(62) McElwee JJ, Schuster E, Blanc E, Piper MD, Thomas JH, Patel DS, et al. Evolutionary conservation of regulated longevity assurance mechanisms. Genome Biol 2007 Jul 5;8(7):R132.

(63) Allison DB, Cui X, Page GP, Sabripour M. Microarray data analysis: from disarray to consolidation and consensus. Nat Rev Genet 2006 Jan;7(1):55-65.

(64) Pawitan Y, Michiels S, Koscielny S, Gusnanto A, Ploner A. False discovery rate, sensitivity and sample size for microarray studies. Bioinformatics 2005 Jul 1;21(13):3017-24.

(65) Dobbin K, Simon R. Sample size determination in microarray experiments for class comparison and prognostic classification. Biostatistics 2005 Jan;6(1):27-38.

(66) Garge NR, Page GP, Sprague AP, Gorman BS, Allison DB. Reproducible clusters from microarray research: whither? BMC Bioinformatics 2005 Jul 15;6 Suppl 2:S10.

(67) Sullivan PF, Fan C, Perou CM. Evaluating the comparability of gene expression in blood and brain. Am J Med Genet B Neuropsychiatr Genet 2006 Apr 5;141(3):261-8.

(68) van Baarsen LG, van der Pouw Kraan TC, Kragt JJ, Baggen JM, Rustenburg F, Hooper T, et al. A subtype of multiple sclerosis defined by an activated immune defense program. Genes Immun 2006 Sep;7(6):522-31.

(69) Fjaerli HO, Bukholm G, Krog A, Skjaeret C, Holden M, Nakstad B. Whole blood gene expression in infants with respiratory syncytial virus bronchiolitis. BMC Infect Dis 2006;6:175.

(70) Sharp FR, Xu H, Lit L, Walker W, Apperson M, Gilbert DL, et al. The future of genomic profiling of neurological diseases using blood. Arch Neurol 2006 Nov;63(11):1529-36.

(71) Borovecki F, Lovrecic L, Zhou J, Jeong H, Then F, Rosas HD, et al. Genome-wide expression profiling of human blood reveals biomarkers for Huntington's disease. Proc Natl Acad Sci U S A 2005 Aug 2;102(31):11023-8.

(72) Batliwalla FM, Li W, Ritchlin CT, Xiao X, Brenner M, Laragione T, et al. Microarray analyses of peripheral blood cells identifies unique gene expression signature in psoriatic arthritis. Mol Med 2005 Jan;11(1-12):21-9.

(73) Tang Y, Xu H, Du X, Lit L, Walker W, Lu A, et al. Gene expression in blood changes rapidly in neutrophils and monocytes after ischemic stroke in humans: a microarray study. J Cereb Blood Flow Metab 2006 Aug;26(8):1089-102.

(74) Tang Y, Gilbert DL, Glauser TA, Hershey AD, Sharp FR. Blood gene expression profiling of neurologic diseases: a pilot microarray study. Arch Neurol 2005 Feb;62(2):210-5.

(75) Fraga MF, Ballestar E, Paz MF, Ropero S, Setien F, Ballestar ML, et al. Epigenetic differences arise during the lifetime of monozygotic twins. Proc Natl Acad Sci U S A 2005 Jul 26;102(30):10604-9.

(76) Sharma A, Sharma VK, Horn-Saban S, Lancet D, Ramachandran S, Brahmachari SK. Assessing natural variations in gene expression in humans by comparing with monozygotic twins using microarrays. Physiol Genomics 2005 Mar 21;21(1):117-23.

(77) Tan Q, Christensen K, Christiansen L, Frederiksen H, Bathum L, Dahlgaard J, et al. Genetic dissection of gene expression observed in whole blood samples of elderly Danish twins. Hum Genet 2005 Jul;117(2-3):267-74.

(78) Whitney AR, Diehn M, Popper SJ, Alizadeh AA, Boldrick JC, Relman DA, et al. Individuality and variation in gene expression patterns in human blood. Proc Natl Acad Sci U S A 2003 Feb 18;100(4):1896-901.

(79) Palmer C, Diehn M, Alizadeh AA, Brown PO. Cell-type specific gene expression profiles of leukocytes in human peripheral blood. BMC Genomics 2006;7:115.

(80) Tang Y, Lu A, Ran R, Aronow BJ, Schorry EK, Hopkin RJ, et al. Human blood genomics: distinct profiles for gender, age and neurofibromatosis type 1. Brain Res Mol Brain Res 2004 Dec 20;132(2):155-67.

(81) Barzilai N, Gabriely I, Gabriely M, Iankowitz N, Sorkin JD. Offspring of centenarians have a favorable lipid profile. J Am Geriatr Soc 2001 Jan;49(1):76-9.

(82) Heijmans BT, Beekman M, Houwing-Duistermaat JJ, Cobain MR, Powell J, Blauw GJ, et al. Lipoprotein particle profiles mark familial and sporadic human longevity. PLoS Med 2006 Dec;3(12):e495.

(83) Lagaay AM, van Asperen IA, Hijmans W. The prevalence of morbidity in the oldest old, aged 85 and over: a population-based survey in Leiden, The Netherlands. Arch Gerontol Geriatr 1992 Sep;15(2):115-31.

(84) Huang H, Manton KG. The role of oxidative damage in mitochondria during aging: a review. Front Biosci 2004 May 1;9:1100-17.

(85) Lopez-Lluch G, Hunt N, Jones B, Zhu M, Jamieson H, Hilmer S, et al. Calorie restriction induces mitochondrial biogenesis and bioenergetic efficiency. Proc Natl Acad Sci U S A 2006 Feb 7;103(6):1768-73.

(86) Hayes JD, Flanagan JU, Jowsey IR. Glutathione transferases. Annu Rev Pharmacol Toxicol 2005;45:51-88.

Chapter 3

Transcriptional profiling of human familial longevity indicates a role for *ASF1A* and *IL7R*

W.M. Passtoors[1], J.M. Boer[2], J.J. Goeman[3], E.B. van den Akker[1,4], J. Deelen[1], B.J. Zwaan[5], A. Scarborough[6], R. van der Breggen[1], R.H.A.M. Vossen[2,7], J.J. Houwing-Duistermaat[3], G.J.B. van Ommen[2], R.G.J. Westendorp[8], D. van Heemst[8], A.J.M. de Craen[8], A.J. White[9], D.A. Gunn[10], M. Beekman[1], P.E. Slagboom[1,11]

1 Section of Molecular Epidemiology, Leiden University Medical Center, Leiden, The Netherlands
2 Center for Human and Clinical Genetics, Leiden University Medical Center, Leiden, The Netherlands
3 Department of Medical Statistics, Leiden University Medical Center, Leiden, The Netherlands
4 Information and Communication Theory Group – Bioinformatics, Faculty of Electrical Engineering, Mathematics and Computer Science, Delft University of Technology, Delft, The Netherlands
5 Section of Evolutionary Biology, Institute for Biology, Leiden University, Leiden, The Netherlands
6 Measurement Science, Unilever Colworth Laboratory, Sharnbrook, Bedfordshire, United Kingdom
7 Leiden Genome Technology Center, Leiden University Medical Center, Leiden, The Netherlands
8 Department of Gerontology and Geriatrics, Leiden University Medical Center, Leiden, The Netherlands
9 Safety Environmental Assurance Centre, Unilever Colworth Laboratory, Sharnbrook, Bedfordshire, United Kingdom
10 Unilever Discover, Colworth, Sharnbrook, Bedfordshire, United Kingdom, 11 Netherlands Consortium for Healthy Ageing, The Netherlands

PLoS One, 2012. 7(1): p. e27759

Abstract

The Leiden Longevity Study consists of families that express extended survival across generations, decreased morbidity in middle-age, and beneficial metabolic profiles. To identify which pathways drive this complex phenotype of familial longevity and healthy aging, we performed a genome-wide gene expression study within this cohort to screen for mRNAs whose expression changes with age and associates with longevity. We first compared gene expression profiles from whole blood samples between 50 nonagenarians and 50 middle-aged controls, resulting in identification of 2,953 probes that associated with age. Next, we determined which of these probes associated with longevity by comparing the offspring of the nonagenarians (50 subjects) and the middle-aged controls. The expression of 360 probes was found to change differentially with age in members of the long-lived families. In a RT-qPCR replication experiment utilizing 312 controls, 332 offspring and 79 nonagenarians, we confirmed a nonagenarian specific expression profile for 21 genes out of 25 tested. Since only some of the offspring will have inherited the beneficial longevity profile from their long-lived parents, the contrast between offspring and controls is expected to be weak. Despite this dilution of the longevity effects, reduced expression levels of two genes, *ASF1A* and *IL7R,* involved in maintenance of chromatin structure and the immune system, associated with familial longevity already in middle-age. The size of this association increased when controls were compared to a subfraction of the offspring that had the highest probability to age healthily and become long-lived according to beneficial metabolic parameters. In conclusion, an “aging-signature” formed of 21 genes was identified, of which reduced expression of *ASF1A* and *IL7R* marked familial longevity already in middle-age. This indicates that expression changes of genes involved in metabolism, epigenetic control and immune function occur as a function of age, and some of these, like *ASF1A* and *IL7R,* represent early features of familial longevity and healthy ageing.

Introduction

Nonagenarians and centenarians delay or escape age-related diseases (1), their first degree family members have a life-long survival advantage (2;3) and their middle-aged offspring have a decreased prevalence of and mortality from coronary heart disease, type 2 diabetes, and cancer (4;5). In addition, the offspring of long-lived individuals have beneficial physiological characteristics for lipid and lipoprotein particle profiles (6;7), glucose

metabolism and insulin sensitivity (8;9). However, they do not differ from controls with respect to body mass index, serum IGF-1 levels, height and lifestyle factors such as physical activity levels and smoking behavior (8;10). Although familial longevity is a complex phenotype, identifying transcriptional targets that may contribute to the physiological benefits observed in long-lived families will increase our understanding of which pathways can influence susceptibility to and protection from age-related disease.

Previous studies have investigated whether there are gene expression changes that occur with age in brain, lymphocyte, kidney and skeletal muscle tissues (11-16). The expression of some genes were found not only to change with age but also to reflect the biological function of the source organs (13;14). These studies have demonstrated that gene expression levels are not only markers of chronological age but also of tissue function. However, these studies cannot discriminate between genes showing expression changes in mid-life that may contribute to the aging process, from those showing expression changes later in life as a consequence of the aging process.

Here we report transcriptional profiling of whole blood samples from participants of the Leiden Longevity Study which is based in the Netherlands and comprises nonagenarian sibling pairs, their middle-aged offspring and the partners of the offspring as population controls. A comparison of gene expression profiles between the nonagenarian siblings and controls identified profiles that associate with age. The subsequent comparison of the middle-aged offspring and the controls enabled the identification of genes that mark the propensity to become long-lived in middle-age.

Results

Whole genome microarrays and analysis design

Gene expression profiles were generated from 150 whole blood total RNA samples collected from 50 families belonging to the Leiden Longevity Study (LLS). From each family, one nonagenarian sibling, one of their offspring and the offspring's partner (Table 1) were profiled. We identified 47,209 probes (88% of the total number of probes) which were expressed in at least 10% of the samples. Of these, 45,164 probes (containing at least 17,896 unique genes) could be mapped to a chromosomal position and were, therefore, used for further analyses.

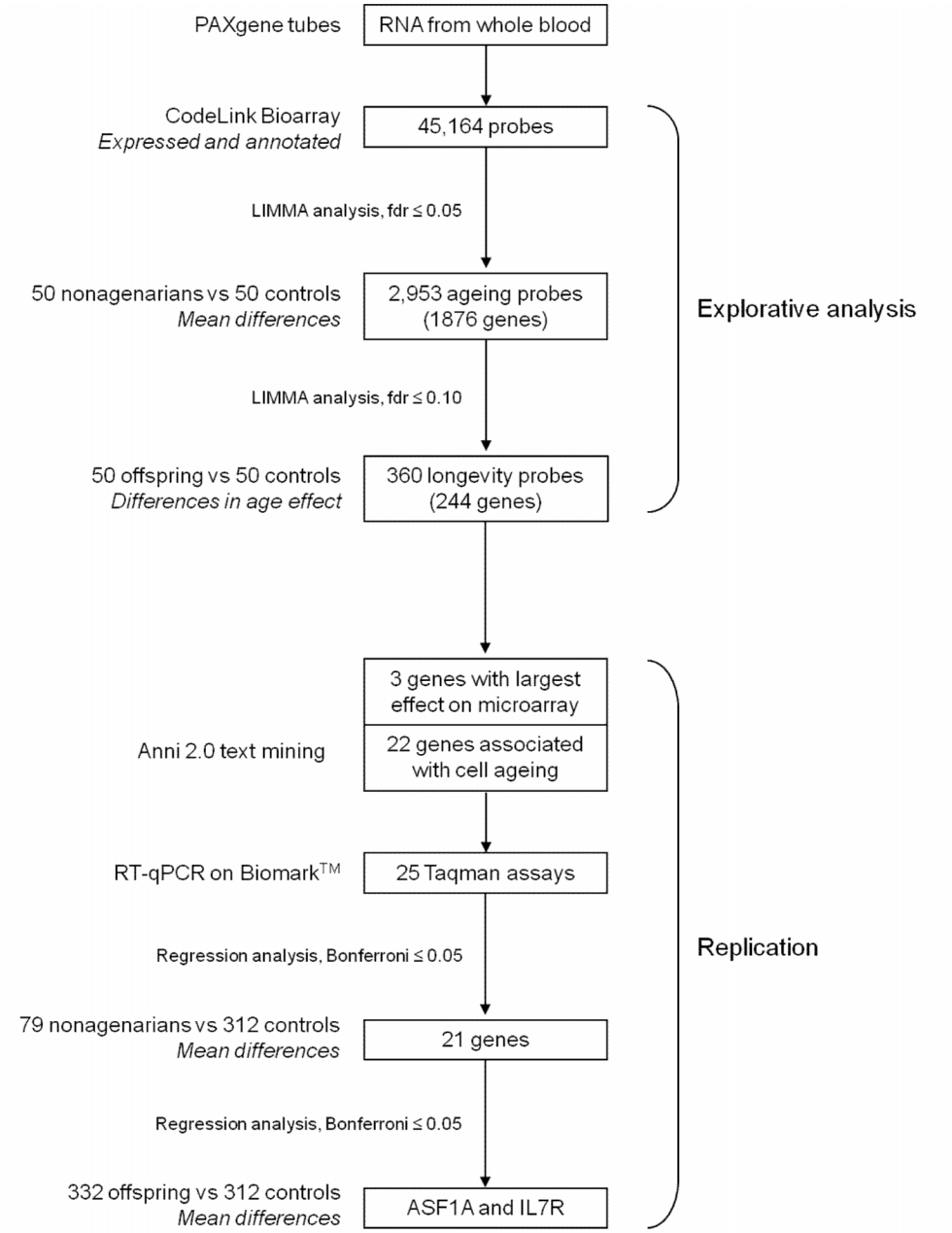

Figure 1. Flowchart of gene expression analyses.
The order of analyses is shown for the explorative analysis (top half of the figure) and the replication analysis (bottom half of the figure). The probes/genes are depicted in the boxes and to the left thereof are the techniques or analyses used.

The explorative analysis is divided in two parts (Figure 1). The first analysis focused on the comparison between the long-lived nonagenarians and the population controls (the partners of the offspring). Using this design we aimed to find genes whose expression changed with increasing age and among these, those that were differentially expressed in long-lived families. In the second analysis we investigated which of the differentially expressed genes emerging from the first analysis were already differentially expressed between long-lived family members and controls in middle-age. Therefore we compared the offspring to the controls for mean gene expression differences and also for the interaction between the two groups with age to identify genes whose expression changed differentially with age between offspring and controls. Following the explorative analysis, we performed replication analyses in an extended group of the LLS using RT-qPCR on a selected subset of the differentially expressed genes.

Table 1. Description of the participants of the LLS in the microarray and validation population.

		Controls	Offspring	Nonagenarians
A	*Microarray experiment*			
	N	50	50	50
	Males / females (% males)	24 / 26 (48%)	25 / 25 (50%)	26 / 24 (52%)
	Mean age in years (range)	61.9 (43.7 – 78.8)	60.8 (42.8 – 74.8)	93.4 (89.3 – 102.2)
B	*RT-qPCR experiment*			
	N	312	332	79
	Males / females (% males)	143 / 169 (45.8%)	190 / 142 (57.2%)	34 /45 (43.0%)
	Mean age in years (range)	61.3 (40.9 – 81.4)	61.3 (33.6 – 78.3)	94.1 (89.0 – 101.2)

Differential gene expression associating with age

To investigate the differences in gene expression levels between the nonagenarian subjects and the middle-aged controls a linear regression model was applied to the probe data. With adjustments for gender and batch effects (Model 1, see Materials and Methods) 2,953 probes (of which 1,853 represented known genes) were found to be differentially expressed at a false discovery rate (FDR) of 0.05 (Figure 2 and Supplementary Table S1). The expression levels of 1,046 probes were increased and 1,907 decreased in the nonagenarians compared to the younger controls. The probe that associated with age with the highest significance (FDR adjusted p-value = 5×10^{-10}) is located in the leucine rich repeat neuronal 3 gene (*LRRN3*, 7q31.1) and showed a 3.1-fold decreased expression in nonagenarians, which was also the largest difference in expression level between the two groups. The largest increase in expression level was 2-

fold for a probe targeting the interferon, alpha-inducible protein 27 isoform (*IFI27*, 14q32.13) gene locus. This probe could also target *SYTL1*, but since the levels of a specific probe for *SYTL1* did not correlate with the *IFI27* probe (Pearson correlation = -0.14), while a unique probe for *IFI27* did correlate (Pearson correlation = 0.96), the result suggests that *IFI27* mRNA is responsible for the association. Average expression levels of the 2,953 probes show a distribution over the whole range of measured intensities (Supplementary Figure S1) and 89.5% of these probes were detected in ≥ 95% of the samples.

If a group of differentially expressed genes act in the same pathway or biological function this could lead to an additive or even synergistic effect on cellular function. Furthermore, the interpretation of transcriptional data in a pathway context is a more robust signal to noise measurement. The Gene Ontology (GO) consortium (17) provides structured vocabularies and classification of genes, covering several domains of molecular and cellular biology. Hence, GO terms were tested for differences between nonagenarians and controls using the Globaltest methodology (18;19). Globaltest determines whether the expression pattern of genes within a set as a whole is associated with an outcome, in this case being nonagenarian or not, without testing single probes. We assayed 1,808 GO gene sets, representing groups of genes closely related in their biological function or process, containing at least 10 probes per GO term, again using Model 1. The Globaltest showed that 109 GO term gene sets were significantly differentially expressed between nonagenarians and controls at a family-wise error rate (FWER) of ≤ 0.05 (Supplementary Table S2). FWER is used because the gene sets for GO terms are partly overlapping and therefore not independent. These 109 GO terms include 73 biological processes, 31 molecular functions and 5 cellular components. The biological processes identified under the higher level GO classifications were ‘lymphocyte activation’, ‘anatomical structure development’, ‘response to stimulus’ (including ‘immune response’), ‘regulation of gene expression’, and ‘regulation of signal transduction’. For the molecular functions list, pathways involved in protein binding were the most abundant.

Differential gene expression in middle-age associating with longevity

The nonagenarian participants in the LLS each exhibit the longevity phenotype and their offspring, as a group, carry the potential to become long-lived as demonstrated by beneficial physiological characteristics and decreased morbidity of that generation (5;6;20). The physiological differences between offspring and controls are small since not all individual

offspring may have inherited the longevity trait. To distinguish which of the 2,953 probes that associated with age also reflect familial longevity already in middle-age, gene expression levels were compared between the 50 offspring of the nonagenarians and the 50 controls (Table 1). In a linear regression model adjusting for age, gender and batch effects (Model 2, see Materials and Methods), we observed no significant differences in the average expression levels between the offspring and controls. Next, we investigated differential expression changes as a function of age between the offspring and the controls by testing the interaction between group and age (Model 3, see Materials and Methods). Differential expression (FDR ≤ 0.1) between offspring and controls across the age-range (43 – 79 years) was observed for 360 probes, representing 244 unique genes (Supplementary Table S3). Of these probes, 359 had a fold change below 1, indicating that expression levels had either a weaker increase with age or a stronger decrease with age in the offspring compared to controls. These age-related expression differences may represent early characteristics of human longevity.

The most significant differentially expressed probe (FDR adjusted p-value = 0.050, 1.6-fold decrease every ten years) corresponded to the zinc finger protein 331 gene (*ZNF331*, 19q13.33) whereas the largest decrease in expression level was for a probe targeting a mRNA at chromosome 1q43 (no known corresponding gene); this had a 2.1-fold decreased expression in offspring every ten years relative to controls. The only probe that demonstrated a significant increase in expression with age in the offspring (a 1.3-fold change every ten years) targeted an EST at 20q13.2 (no known corresponding gene, NCBI Build 36). We observed no differential expression between offspring and controls for *LRRN3* and *IFI27*, which had the most extreme changes with age in the first analysis in which nonagenarians and controls were compared.

To investigate differential expression changes with age in a set of genes acting in the same pathway or having the same function, we implemented Model 3 in GlobalAncova (21); a method similar to the Globaltest for GO terms but additionally suited for models including interaction terms. The GO biological process that was most prominent among the longevity associated gene expression profile was the Rho protein signal transduction pathway (GO:0007266) (FWER = 0.079).

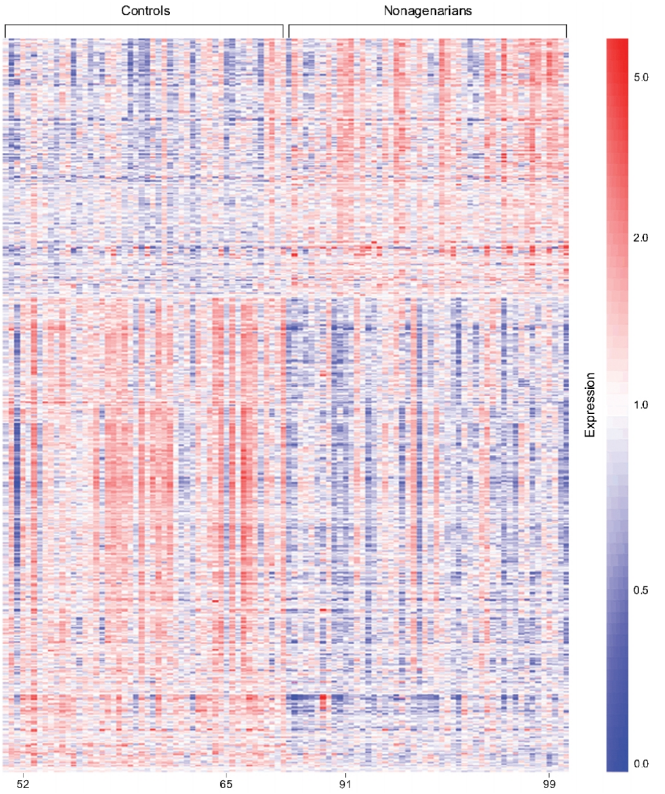

Figure 2. Expression profiles of 2,953 probes that differed between nonagenarians and middle-aged controls.
Expression intensities of the 2,953 probes were analyzed by one-dimensional hierarchical clustering. Each probe is represented by a row; each subject by a column. Samples are organized left to right by increasing age which is indicated for a few individuals for reference. The largest cluster of probes exhibits reduced expression (transition from red to blue), and another cluster exhibits increased expression (transition from blue to red) in nonagenarians compared to controls. Mean centered expression values of probes are displayed according to the color scale in which red represents above average expression levels and blue below average expression levels. Fold changes of individual probes are given in Supplementary Table S1.

Extensive replication study

To replicate our results we measured the expression levels of a subset of genes via RT-qPCR in 79 nonagenarians, 332 offspring and 312 controls (Table 1 and Materials and Methods). We first selected two genes on the basis of their effect size and significance in the analysis between nonagenarians and controls (i.e. *LRRN3* and *IFI27*) and the gene different

between offspring and controls (i.e. *ZNF331*). We added 22 genes that were differentially expressed in the offspring compared to controls which additionally associated with the concept “cell aging” in literature using the text-mining tool Anni 2.0 (22, http://biosemantics.org/anni/). In total 25 genes (Figure 1), including the *WRN* progeria gene, the *MYC* cancer gene and the longevity *MTOR* (also known as *FRAP1*) gene, were selected for replication analyses.

Table 2. RT-qPCR results.

			Nonagenarians vs. controls		Offspring vs. controls	
		N =	79	312	332	312
Gene name		Assay	FC	p	FC	p
	Top genes					
1	IFI27	Hs01086373_g1	1.41	**$2.6x10^{-4}$**	NA	NA
2	LRRN3	Hs00539582_s1	0.56	**$<10^{-6}$**	NA	NA
3	ZNF331	Hs00218578_m1	0.93	**$1.0x10^{-4}$**	0.99	0.574
	Cell aging associated genes					
4	ADAMTS5	Hs00199841_m1	1.01	0.008	1.00	0.872
5	ASF1A	Hs01011627_m1	0.85	**$<10^{-6}$**	0.88	**0.002**
6	CCR6	Hs01890706_s1	0.68	**$<10^{-6}$**	1.06	0.265
7	CD248	Hs00535586_s1	>10	0.322	1.04	0.736
8	CDK6	Hs00608037_m1	0.95	0.053	0.95	0.015
9	ENO2	Hs00157360_m1	0.99	**$<10^{-6}$**	0.99	0.044
10	FLT3LG	Hs00181740_m1	0.25	0.029	0.63	0.202
11	HK3	Hs01092843_g1	1.17	**$<10^{-6}$**	1.00	0.929
12	IL7R	Hs00902338_g1	0.76	**$<10^{-6}$**	0.89	**0.001**
13	LEF1	AI6Q1P7	0.64	**$<10^{-6}$**	0.97	0.565
14	MLLT3	Hs00971090_m1	0.80	**$<10^{-6}$**	0.95	0.087
15	MTOR (FRAP1)	Hs00234508_m1	0.97	**$6.0x10^{-6}$**	0.99	0.337
16	MYC	Hs00153408_m1	0.78	**$<10^{-6}$**	0.98	0.547
17	NOLC1	Hs01102319_g1	0.89	**$1.0x10^{-6}$**	0.95	0.082
18	NR3C2	Hs00230906_m1	0.78	**$<10^{-6}$**	0.95	0.059
19	RUVBL2	AI7ZZWF	0.79	**$6.0x10^{-5}$**	1.07	0.469
20	SIDT1	Hs00214475_m1	0.70	**$5.0x10^{-6}$**	0.91	0.301
21	SMAD3	Hs00706299_s1	0.87	**$<10^{-6}$**	0.99	0.825
22	SMYD5	Hs00300181_m1	0.93	**$<10^{-6}$**	0.98	0.229
23	TCF12	Hs00918972_m1	0.90	**$<10^{-6}$**	0.97	0.127
24	TCF4	Hs00972428_g1	0.88	**$1.5x10^{-5}$**	1.01	0.872
25	WRN	Hs02561119_s1	0.76	**$<10^{-6}$**	1.02	0.791

FC: fold change between groups; a FC above 1 indicates an increase in expression and a FC below 1 indicates a decrease in expression compared to the controls. p: unadjusted p

values. Bold indicate p values are below the significant level of 0.05 after Bonferroni correction for multiple testing (threshold p = 0.002).

As replication of the first analysis, the comparison between nonagenarians and controls, the expression of 21 out of 25 genes was in concordance with the observations in the microarray dataset and hence the RT-qPCR results confirmed the microarray findings (Table 2). *LRRN3* again showed the largest significant decreased expression in the nonagenarian siblings and *IFI27* the largest increased expression. The analysis to test for replication of the novel samples only (58 nonagenarians and 281controls) resulted in the same observations (Supplementary Table S4).

To identify longevity associated genes in middle-age, we replicated the second analysis by comparing the mean expression levels of 23 genes between the offspring of the nonagenarians and the controls (Model 2, see Materials & Methods). We excluded *LRRN3* and *IFI27* from this analysis since expression of these two genes was not different between offspring and controls in the microarray dataset. Two genes showed significant differential expression after Bonferroni correction for multiple testing: *ASF1A* and *IL7R* (Table 2, $p < 0.0022$). These two genes showed a significant decrease in expression in samples from the long-lived families compared to controls, in concordance with the microarray results.

Analysis of healthy offspring

We noticed that the effect sizes of the differential expressed genes in the comparison between offspring and controls were small. Since the offspring are composed of individuals who carry the longevity trait and those that do not, the detectable effect of the longevity trait is diluted in the comparison between offspring and controls. We hypothesize that the effect of the longevity trait will increase if we compare the same controls to a subfraction of offspring with the highest probability to age healthily and become long-lived. Therefore we selected the offspring with the most beneficial metabolic profile.

Offspring of long-lived parents exhibit as a group at least six beneficial serum characteristics including: low levels of glucose (8;10), triglycerides (23), total cholesterol over HDL cholesterol ratio (23;24), free triiodothyronine (25), and large low-density lipoprotein (LDL) particles (23), complemented by high adiponectin (26). The beneficial metabolic profile is also reflected by the lower Framingham risk scores (24) which indicates a lower risk of cardiovascular disease over the course of 10 years.

We selected a subfraction of the offspring, consisting of the 5% beneficial tail of the distribution of each metabolic parameter, separately for men and women, resulting in 78 offspring which we consider the best proxy for the true long-lived cases. For these offspring *ASF1A* and *IL7R* gene expression was compared to that of 312 controls: this comparison resulted respectively in a 5.0% and 1.2% additional decrease of mean gene expression relative to the effect observed in the comparison between all offspring and controls (Table 3).

Table 3. Gene expression levels of ASF1A and IL7R in all as well as subfraction of offspring compared to controls.

			ASF1A		*IL7R*	
		n	Mean	p	Mean	p
Controls	All	312	1.85		6.18	
Offspring	All	332	1.65	**0.002**[2]	5.98	**0.001**[2]
	Subfraction[1]	78	1.57	**0.009**[3]	5.91	**0.010**[3]

Mean: relative expression in fold change to reference value. [1]: subfraction of offspring most probable to age healthily (for more details, see Materials and Methods), [2]: p value of comparison between controls and all offspring, [3]: p value of comparison between controls and subfraction of offspring (best 5% men and women per parameter taken together). Bold indicated p values are below the significant level of 0.05 after Bonferroni correction for multiple testing.

Discussion

We have identified a transcriptional profile of 244 genes that represent a potential “longevity-signature”. In an extensive biological replication study of nonagenarians, middle-aged offspring and controls, we focused on a subset of 25 genes in RT-PCR experiments. An “aging-signature” was formed by the expression pattern of 21 genes. Two genes, *ASF1A* and *IL7R*, represented a “longevity-signature” since members of long-lived families expressed at middle-age a 1.14-fold and 1.12-fold lower level of these genes compared to controls respectively. The effect size of this association with longevity was stronger for offspring of nonagenarians with a beneficial metabolic profile characterized by 6 serum parameters and the Framingham Risk Score, which we considered to be the best proxy for the true long-lived case group. The decreased expression of these two genes is therefore likely to mark metabolic health and it might precede or even contribute to human longevity. ASF1A is a histone chaperone that is important in the remodeling of chromatin structure during replication, DNA repair, and cellular senescence (27-29). Interestingly, histone acetylation is among the processes regulated by signaling through the IL7 receptor, which is required for development and maintenance of the immune system

(30). Our results may point at the importance of interactions between immune response, metabolic state, and epigenetic control for human aging and longevity.

The genome-wide expression analysis between nonagenarians and controls resulted in the identification of 2,953 probes associated with age. RT-qPCR replication experiments with a large sample size resulted in replication of 21 out of 25 genes identified by the microarray analysis as differentially expressed with age. Five of these genes (*LRRN3*, *ZNF331*, *ASF1A*, *MLLT3* and *SIDT1*) have previously been reported to associate with age in a microarray study of T cell mRNA (31), although this study had a relatively small number of samples (25 male subjects). The largest age-related effects observed in our study were for the *IFI27* and *LRRN3* genes. *IFI27* encodes interferon alpha-inducible protein 27 whose biological function is currently unknown. Leucine rich repeat neuronal 3 (*LRRN3*) is mainly involved in activation of MAPK activity and endocytosis (17) and, further supporting our findings, also showed the largest age-related decrease in expression in a large study on lymphocytes (15). This “aging-signature” included decreased expression of several well-known genes involved in aging and lifespan, like *MYC*, *WRN* and *MTOR*. The MYC protein is a transcription factor that regulates transcription of specific target genes and overexpression of the *MYC* gene has been associated with a variety of cancers (32-34). Defects in the *WRN* gene cause the Werner progeria syndrome, an autosomal recessive disorder characterized by premature aging and genetic variation in or near this gene have been associated with several age-related diseases and survival (35-38). The *MTOR* gene encodes a serine/threonine kinase and its downregulation is associated with extended lifespan in model organisms (39-41) and elevated mTOR activity has been implicated in different forms of human cancer (reviewed in 42;43). We conclude from our data that, from all significantly differentially expressed genes at least 21 genes distinguish nonagenarians from middle-aged controls. In addition, our study is the first to include an extensive biological replication sample set validating these results.

The main cellular pathways that changed with age in the microarray dataset were ‘response to stimulus’ (including ‘immune response’ and ‘response to stress’), ‘signal transduction’, ‘gene expression’ and ‘protein binding’. These GO terms have previously been found to associate with aging, suggesting that these are systemic age-related processes. The only pathway that associated with longevity at middle-age in the offspring-control comparison was the ‘Rho protein signal transduction’ pathway, which is part of the GO term ‘signal transduction’. The Rho family of GTPases are small GTPases

that regulate a wide variety of processes in the cell including growth, cytoskeletal organization, transcription and lipid metabolism (44;45). Rho signaling is regulated by the mTOR complex 2, a part of the mTOR pathway which is shown to influence lifespan and health. All associated pathways are general processes, indicating that regulation of the system seem to be an important process involved in aging and longevity.

Research into human familial longevity and healthy ageing is complex in the sense that there are no controls from the same birth cohort to compare to long-lived persons since such controls would have died twenty years ago. We investigated offspring as a proxy for the nonagenarian case group since these can be compared to controls from the same birth cohort. However, in the offspring the longevity phenotype will undoubtedly be diluted as compared to the nonagenarians, since only a part of the offspring will become long-lived and a part will age comparable to controls. The consequence of the dilution of the longevity cases is that effect sizes are underestimated. Indeed the effect size of the association with longevity increased when the healthiest offspring was compared to the controls. Future follow-up data on age of death will reveal which offspring carries the longevity phenotype.

In this study we investigated expression profiles in whole blood samples of participants because blood is an easily accessible tissue. This allows us to investigate the large sample sizes required to detect small effect sizes. An advantage of using blood compared to other tissues is that cell subsets can easily be measured and used to select samples or used as covariates in analyses whereas different cell subsets present in other tissues are difficult to quantify and can not be taken into account in any analyses. In our microarray study we selected samples from offspring and controls with similar cell counts. Furthermore, since aging affects the whole organism and since blood is in contact with all tissues, blood may reflect in part the physical health of the whole body. Disease state is mirrored by gene expression profiles in blood (summarized in 16), which at least partial overlap with expression profiles in other tissues (46). Thus, although tissue-specific effects will undoubtedly be missed, investigating blood is valuable and practical for researching human aging.

In conclusion, we identified a transcriptional signature in whole blood consisting of 21 genes that repeatedly differentiated between nonagenarians and middle-aged controls. The expression level of two of these genes, *ASF1A* and *IL7R* marked familial longevity already in middle-age and the effect size was enhanced in the subset of longevity family

members with a beneficial metabolic marker profile. Functional and longitudinal studies are necessary to establish which of these genes are true biomarkers for healthy ageing and which contribute causally to this trait. Our findings illustrate that gene expression changes occurring as a function of age may partly represent early detectable features of human longevity and healthy ageing.

Materials and methods

Study population

The individuals investigated in this study are participants of the Leiden Longevity Study. The families participating in this study have at least two siblings with a minimum age for men of 89 years and for women of 91 years (2). The offspring of these long-lived individuals, who have an increased potential to become long-lived individuals, were also included. In addition, the partners of the offspring were included as population controls of similar age and environmental exposures as the offspring, and as a young control group for the nonagenarian siblings. Blood samples were taken from all the participants. The Leiden Longevity Study was approved by the medical ethical committee of the Leiden University Medical Centre and all participants gave written informed consent.

Sample collection and RNA preparation

One long-lived sibling, one of their offspring and the partner of the offspring were selected from 50 families for the current study (Table 1). These trios were randomly selected, but in such a way that age and gender were balanced between the groups and the age range for the offspring and partners was as large as possible. Additionally, individuals with outlying cell counts (beyond 3 SDs below or above the standard error of the mean) were excluded. From the 150 selected non-fasted individuals, peripheral blood was harvested using PAXgene^{TM} tubes (Qiagen, Venlo, The Netherlands). The tubes were frozen and kept at -20°C for ~3-5 years. After thawing at room temperature for at least 2 hours, total RNA was extracted from the approximately 2.5 ml of peripheral blood in each tube following the manufacturer's recommended protocol (PAXgene Blood RNA Kit Handbook, Qiagen, Venlo, The Netherlands). The quality and integrity of the total RNA was evaluated on the 2100 Bioanalyzer (Agilent Technologies, Amstelveen, The Netherlands) and the concentration was measured using a NanoDrop spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA). Quality criteria included a 28S/18S ratio as

measured by the Bioanalyzer of at least 1.2, and a total RNA yield of at least 3 µg.

Oligonucleotide microarrays

The 150 samples that met the RNA quality criteria were hybridized onto 54k CodeLink Human Whole Genome Bioarrays (GE Healthcare, Bucks, UK, cat. No. 300026, currently of Applied Microarrays). cDNA synthesis, amplification, biotin labeling and hybridization onto the microarrays were performed according to the manufacturer's instructions using the Codelink iExpress reagent kit (cat. No. 67601000). The slides were scanned with a MicroArray Scanner G2505B scanner (Agilent Technologies, South Queensferry, UK) and the image was quantified with the CodeLink Expression software (version 4.2).

Microarray data pre-processing

Raw intensities were background subtracted, set to 0.5 when results were negative and normalized using the Cyclic Loess method in the Codelink software package (47) of the Bioconductor R software (48) (http://www.bioconductor.org). After normalization, we used $\log_2$-transformed expression intensities for all subsequent analyses. Raw and normalized microarray data are stored in the GEO online database record GSE16717 in compliance to MIAME guidelines. A principal component analysis (PCA) was performed on all samples (GeneSpring software, Agilent Technologies, South Queensferry, UK) and hybridization date was identified as a confounding factor causing a deviation in the data, which was attributed to a scanner maintenance check during the measurements of the samples (data not shown). Therefore, all subsequent analyses were adjusted for hybridization date coded as the 14 days of hybridization as categorical variable, which was sufficient to adjust for this technically induced variation. Since samples were randomly hybridized, no confounding with group was present. None of the other tested parameters, like RNA quality, isolation date and time of blood draw appeared to be a significant confounder of the expression data.

Probe annotation and filtering

The 54,243 probes on the CodeLink Bioarray were newly annotated to Entrez Gene ID's and GO identifiers in two steps. First, all probe sequences were mapped to Unigene and dbEst sequences with BLAT while allowing for, at most, one mismatching nucleotide. Subsequently, all

Probe to Unigene annotations were transformed to probe to Entrez Gene ID and probe to GO ID annotations using Entrez Gene-to-Unigene and Entrez Gene-to-GO ID mappings available on the ftp server of NCBI. Probe-to-EST annotations were treated in a similar way, except that ESTs were first mapped to RefSeq Gene IDs by aligning ESTs to RefSeq exons using galaxy and the genomic alignments of ESTs and RefSeq genes from UCSC (hg18). All EST to RefSeq mappings with a sequence similarity >95% were maintained for further mappings of ESTs-to-Entrez Gene IDs using the Entrez Gene-to-RefSeq Gene ID mappings available at NCBI. All information used was downloaded in October 2008, using versions NCBI Build 36.1 or UCSC hg18.

Probes without a "Good" flag indicating that the mRNA is detected, as determined by the CodeLink Expression software, in at least 10% of the samples (7,034 probes) and/or probes without at least a known chromosome band location according to the new annotation (an additional 2,045) were excluded from the analysis, resulting in 45,164 remaining probes.

Single gene analysis

All single gene analyses were performed using the Limma (Linear models for microarray data) package in R (49;50). To determine changes in expression levels of each probe with age, we used the following linear regression model:

$$Y_{ij} = \beta_{0j} + \beta_{1j} Group_i + \beta_{2j} Gender_i + \beta_{3j} Hyb_i + \varepsilon_{ij} \qquad \textit{(model 1)}$$

where Y_{ij} is the base 2 logarithm of the expression level of probe *j* in sample i, $Group_i$ is the group (0 for control or 1 for long-lived nonagenarian) of subject contributing sample *i*, $Gender_i$ corresponds to the gender of the *i*th sample (0 for male, or 1 for female), Hyb_i is the categorical term of hybridization day on which the sample *i* was measured and ε_{ij} represents an error term. The coefficients β_{1j} represents by how much gene expression increases between the groups, β_{2j} represents the change in expression for a female in comparison with a male sample, β_{3j} represents the change of expression across hybridization dates, and β_{0j} represents the baseline regression level of the probe, for male control samples on the first hybridization date. Resulting p-values were adjusted for multiple testing using Benjamini and Hochberg's False Discovery Rate (FDR) method (51). One-dimensional hierarchical clustering of probes was performed using

GeneSpring (version GX 7.3.1) gene tree clustering, using Pearson correlation and average linkage.
To find longevity-related differences between offspring and controls, the following model was used:

$$Y_{ij} = \beta_{0j} + \beta_{1j} Group_i + \beta_{2j} Age_i + \beta_{3j} Gender_i + \beta_{4j} Hyb_i + \varepsilon_{ij} \qquad \textit{(model 2)}$$

where $Group_i$ corresponds to the offspring/control status of the *i*th sample (0 for control, 1 for offspring) and β_{1j} is the change of expression with group status. For each probe *j*, we determined the coefficient with respect to group status (β_{1j}).
To test differences in biological aging rate in offspring and controls, we used the following model:

$$\begin{aligned} Y_{ij} &= \beta_{0j} + \beta_{1j} Group_i + \beta_{2j} Age_i + \beta_{3j} Gender_i + \beta_{4j} Group_i \times Age_i \\ &+ \beta_{5j} Age_i \times Gender_i + \beta_{6j} Group_i \times Gender_i + \beta_{7j} Hyb_i + \varepsilon_{ij} \end{aligned} \qquad \textit{(model 3)}$$

where $Group_i{*}Age_i$ corresponds to the slope of expression with age of the *j*th probe set for offspring or controls, $Age_i{*}Gender_i$ corresponds to the slope of expression with age of the *j*th probe set for males or females, $Group_i{*}Gender_i$ indicates the interaction of group and gender of the *j*th probe set for the *i*th sample, β_{4j} represents the change of expression with age for offspring in comparison to control, β_{5j} represents the change of expression with age for a female in comparison with a male sample, β_{6j} represents the change of expression with gender between offspring and controls. The interaction between group and offspring in comparison to control (β_{2j} for the intercept, β_{4j} for the other groups) was determined and resulting p-values were corrected for multiple testing using the FDR method.

Pathway analysis

The Globaltest methodology was designed to determine whether the common expression pattern of genes within a pre-defined set is significantly related to clinical outcome (18;19;52). A generalized linear model is used to estimate a "Q-statistic" for each gene set, which describes the correlation between gene expression profiles, X, and clinical outcomes, Y. The Q-statistic for a gene set is the average of the Q-statistics for each gene in the set. The Globaltest method was used to perform pathway analysis on Model 1.

When performing pathway analysis on Model 3 we used GlobalAncova, which is a method similar to Globaltest suited for models including interaction terms (21;53). Resulting p-values from both methods were corrected for multiple testing using Holm's procedure for controlling the Family-Wise Error Rate (FWER) method (54).

RT-qPCR

To confirm the accuracy of the measured expression profiles, we compared the expression level of 25 probes from the CodeLink Bioarrays with corresponding Taqman® assay (Applied Biosystems, Table 2). Samples included 18 nonagenarians, 16 offspring and 21 controls that have been measured on the microarray and additional novel replication samples of randomly chosen 61 nonagenarians, 316 offspring and 291 controls. Reverse transcription was performed by using total RNA from blood of in total 723 samples, excluding individuals with outlying cell counts (outside 3SD of the mean), which passed QC and processed with the First Strand cDNA Synthesis Kit according to the manufacturer's protocol (Roche Applied Science). cDNA was amplified using the DNA Engine Tetrad® 2 Peltier Thermal Cycler (Bio-Rad). qPCR was then performed with the Taqman® method using the Biomark™ 48.48 and 96.96 Dynamic Arrays (Fluidigm). Relative gene expressions of the BioMark™ Array data were calculated by using the $2^{-\Delta\Delta Ct}$ method, in which Ct indicates cycle threshold, the fractional cycle number where the fluorescent signal reaches detection threshold (55). YKT6 was used as internal control and commercially available human total reference RNA (Clontech Laboratories, Mountain View, CA, USA) as reference RNA. Differences in expression level between long-lived siblings, their offspring and the partners of their offspring were assessed using linear regression. In these analyses, expression level was the dependent variable and the two groups of individuals (either nonagenarians vs. controls or offspring vs. controls) were included in the model as a categorical variable together with age (in offspring vs. controls only) and gender and their interaction as covariates. To take into account dependencies within sibships, robust standard errors were used, i.e. the variance was computed from the between family variation. P-values were also based on these robust standard errors. Analyses were performed using the software package STATA/SE 11.0 (DPC Software, StataCorp 2009).

Analysis of healthy offspring

To further investigate the candidate genes, their expression level was again tested for association with longevity, but only including the offspring with most beneficial profile of seven published longevity markers: a low level of non-fasted serum glucose (mmol/L) (8;10), triglycerides (mmol/L) (23) and free triiodothyronine levels (pmol/L) (25), a small ratio of total cholesterol (mmol/L) over HDL cholesterol (mmol/L) (23;24), small low-density lipoprotein (LDL) particle sizes (nm) (23), a high level of adiponectin (mg/L) [26], and a low Framingham risk score (FRS) which is based on the factors age, sex, total cholesterol level, HDL cholesterol level, systolic blood pressure (mm/Hg), and whether the person smokes (24). The FRS is a well known test reflecting the risk of cardiovascular disease over the course of 10 years.

All serum measurements were performed with fully automated equipment. For glucose, triglycerides, total cholesterol, HDL-cholesterol, adiponectin and free triiodothyronine, the Hitachi Modular or the Cobas Integra 800, both from Roche, Almere, the Netherlands were applied. CVs of these measurements were all below 5 %. Lipoprotein particle sizes have been analyzed in 165 families from the Leiden Longevity Study using a 400-MHz proton NMR analyzer at LipoScience.

To select that subfraction of offspring with the highest probability to age healthily and become long-lived because of their metabolic risk profile in middle-age, we identified the subjects, separately for men and women, within the lower 5% tail of the distribution for glucose, triglycerides, free triiodothyronine, ratio of total cholesterol over HDL cholesterol and the FRS. Additionally we identified for LDL particle size and adiponectin those subjects, separately for men and women, within the upper 5% tail of the distribution. This resulted in a total of 78 offspring out of 332 from which gene expression levels of *ASF1A* and *IL7R* was compared to the levels in 312 controls using linear regression. In this analysis, expression level was the dependent variable and the two groups of individuals (offspring vs. controls) were included in the model as a categorical variable together with age and gender and their interaction as covariates. To take into account dependencies within sibships, robust standard errors were used, i.e. the variance was computed from the between family variation. P-values were also based on these robust standard errors. Analyses were performed using the software package STATA/SE 11.0 (DPC Software, StataCorp 2009).

Acknowledgements

The Leiden Longevity study was supported by a grant from Innovation-Oriented Research Program IOP on Genomics (SenterNovem IGE05007), a stimulation grant (05040202, Healthy Ageing), the Centre for Medical Systems Biology and the National Institute for Healthy Aging (Grant 05060810), all in the framework of the Netherlands Genomics Initiative/Netherlands Organization for Scientific Research and by Unilever PLC (http://www.unilever.com/). The research leading to these results has also received funding from the European Union's Seventh Framework Programme (FP7/2007-2011) under grant agreement n° 259679 (IDEAL).

References

(1) Hitt R, Young-Xu Y, Silver M, Perls T. Centenarians: the older you get, the healthier you have been. Lancet 1999 Aug 21;354(9179):652.

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

(3) Perls TT, Wilmoth J, Levenson R, Drinkwater M, Cohen M, Bogan H, et al. Life-long sustained mortality advantage of siblings of centenarians. Proc Natl Acad Sci U S A 2002 Jun 11;99(12):8442-7.

(4) Terry DF, Wilcox M, McCormick MA, Lawler E, Perls TT. Cardiovascular advantages among the offspring of centenarians. J Gerontol A Biol Sci Med Sci 2003 May;58(5):M425-M431.

(5) Westendorp RG, van HD, Rozing MP, Frolich M, Mooijaart SP, Blauw GJ, et al. Nonagenarian siblings and their offspring display lower risk of mortality and morbidity than sporadic nonagenarians: The Leiden Longevity Study. J Am Geriatr Soc 2009 Sep;57(9):1634-7.

(6) Heijmans BT, Beekman M, Houwing-Duistermaat JJ, Cobain MR, Powell J, Blauw GJ, et al. Lipoprotein particle profiles mark familial and sporadic human longevity. PLoS Med 2006 Dec;3(12):e495.

(7) Barzilai N, Gabriely I, Gabriely M, Iankowitz N, Sorkin JD. Offspring of centenarians have a favorable lipid profile. J Am Geriatr Soc 2001 Jan;49(1):76-9.

(8) Rozing MP, Westendorp RG, Frolich M, de Craen AJ, Beekman M, Heijmans BT, et al. Human insulin/IGF-1 and familial longevity at middle age. Aging (Albany NY) 2009 Aug;1(8):714-22.

(9) Wijsman CA, Rozing MP, Streefland TC, le CS, Mooijaart SP, Slagboom PE, et al. Familial longevity is marked by enhanced insulin sensitivity. Aging Cell 2011 Feb;10(1):114-21.

(10) Rozing MP, Westendorp RG, de Craen AJ, Frolich M, de Goeij MC, Heijmans BT, et al. 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 2010 Mar;58(3):564-9.

(11) Erraji-Benchekroun L, Underwood MD, Arango V, Galfalvy H, Pavlidis P, Smyrniotopoulos P, et al. Molecular aging in human prefrontal cortex is selective and continuous throughout adult life. Biol Psychiatry 2005 Mar 1;57(5):549-58.

(12) Ly DH, Lockhart DJ, Lerner RA, Schultz PG. Mitotic misregulation and human aging. Science 2000 Mar 31;287(5462):2486-92.

(13) Zahn JM, Sonu R, Vogel H, Crane E, Mazan-Mamczarz K, Rabkin R, et al. Transcriptional profiling of aging in human muscle reveals a common aging signature. PLoS Genet 2006 Jul;2(7):e115.

(14) Rodwell GE, Sonu R, Zahn JM, Lund J, Wilhelmy J, Wang L, et al. A transcriptional profile of aging in the human kidney. PLoS Biol 2004 Dec;2(12):e427.

(15) Hong MG, Myers AJ, Magnusson PK, Prince JA. Transcriptome-wide assessment of human brain and lymphocyte senescence. PLoS One 2008;3(8):e3024.

(16) Passtoors WM, Beekman M, Gunn D, Boer JM, Heijmans BT, Westendorp RG, et al. Genomic studies in ageing research: the need to integrate genetic and gene expression approaches. J Intern Med 2008 Feb;263(2):153-66.

(17) Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet 2000 May;25(1):25-9.

(18) Globaltest, R package, version 4.10.0 [computer program]. 2008.

(19) Goeman JJ, van de Geer SA, de Kort F, van Houwelingen HC. A global test for groups of genes: testing association with a clinical outcome. Bioinformatics 2004 Jan 1;20(1):93-9.

(20) Slagboom PE, Beekman M, Passtoors WM, Deelen J, Vaarhorst AA, Boer JM, et al. Genomics of human longevity. Philos Trans R Soc Lond B Biol Sci 2011 Jan 12;366(1561):35-42.

(21) Hummel M, Meister R, Mansmann U. GlobalANCOVA: exploration and assessment of gene group effects. Bioinformatics 2008 Jan 1;24(1):78-85.

(22) Jelier R, Schuemie MJ, Veldhoven A, Dorssers LC, Jenster G, Kors JA. Anni 2.0: a multipurpose text-mining tool for the life sciences. Genome Biol 2008;9(6):R96.

(23) Vaarhorst AA, Beekman M, Suchiman EH, van HD, Houwing-Duistermaat JJ, Westendorp RG, et al. Lipid metabolism in long-lived families: the Leiden Longevity Study. Age (Dordr) 2010 Sep 3.

(24) Terry DF, Evans JC, Pencina MJ, Murabito JM, Vasan RS, Wolf PA, et al. Characteristics of Framingham offspring participants with long-lived parents. Arch Intern Med 2007 Mar 12;167(5):438-44.

(25) Rozing MP, Westendorp RG, de Craen AJ, Frolich M, Heijmans BT, Beekman M, et al. Low serum free triiodothyronine levels mark familial longevity: the Leiden Longevity Study. J Gerontol A Biol Sci Med Sci 2010 Apr;65(4):365-8.

(26) Atzmon G, Pollin TI, Crandall J, Tanner K, Schechter CB, Scherer PE, et al. Adiponectin levels and genotype: a potential regulator of life span in humans. J Gerontol A Biol Sci Med Sci 2008 May;63(5):447-53.

(27) Chen CC, Tyler J. Chromatin reassembly signals the end of DNA repair. Cell Cycle 2008 Dec 15;7(24):3792-7.

(28) Adams PD. Remodeling of chromatin structure in senescent cells and its potential impact on tumor suppression and aging. Gene 2007 Aug 1;397(1-2):84-93.

(29) Das C, Lucia MS, Hansen KC, Tyler JK. CBP/p300-mediated acetylation of histone H3 on lysine 56. Nature 2009 May 7;459(7243):113-7.

(30) DeKoter RP, Schweitzer BL, Kamath MB, Jones D, Tagoh H, Bonifer C, et al. Regulation of the interleukin-7 receptor alpha promoter by the Ets transcription factors PU.1 and GA-binding protein in developing B cells. J Biol Chem 2007 May 11;282(19):14194-204.

(31) Remondini D, Salvioli S, Francesconi M, Pierini M, Mazzatti DJ, Powell JR, et al. Complex patterns of gene expression in human T cells during in vivo aging. Mol Biosyst 2010 Oct 1;6(10):1983-92.

(32) Lenz G, Nagel I, Siebert R, Roschke AV, Sanger W, Wright GW, et al. Aberrant immunoglobulin class switch recombination and switch translocations in activated B cell-like diffuse large B cell lymphoma. J Exp Med 2007 Mar 19;204(3):633-43.

(33) Bernard O, Cory S, Gerondakis S, Webb E, Adams JM. Sequence of the murine and human cellular myc oncogenes and two modes of myc transcription resulting from chromosome translocation in B lymphoid tumours. EMBO J 1983;2(12):2375-83.

(34) Ia-Favera R, Bregni M, Erikson J, Patterson D, Gallo RC, Croce CM. Human c-myc onc gene is located on the region of chromosome 8 that is translocated in Burkitt lymphoma cells. Proc Natl Acad Sci U S A 1982 Dec;79(24):7824-7.

(35) Castro E, Edland SD, Lee L, Ogburn CE, Deeb SS, Brown G, et al. Polymorphisms at the Werner locus: II. 1074Leu/Phe, 1367Cys/Arg, longevity, and atherosclerosis. Am J Med Genet 2000 Dec 11;95(4):374-80.

(36) Ye L, Miki T, Nakura J, Oshima J, Kamino K, Rakugi H, et al. Association of a polymorphic variant of the Werner helicase gene with myocardial infarction in a Japanese population. Am J Med Genet 1997 Feb 11;68(4):494-8.

(37) Luo Y, Ji L, Han X. [The role of Werner's syndrome gene in the genetic susceptibility to the type 2 diabetes in Chinese population]. Zhonghua Yi Xue Za Zhi 2000 May;80(5):360-1.

(38) Lunetta KL, D'Agostino RB, Sr., Karasik D, Benjamin EJ, Guo CY, Govindaraju R, et al. 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.

(39) Vellai T, Takacs-Vellai K, Zhang Y, Kovacs AL, Orosz L, Muller F. Genetics: influence of TOR kinase on lifespan in C. elegans. Nature 2003 Dec 11;426(6967):620.

(40) Kapahi P, Zid BM, Harper T, Koslover D, Sapin V, Benzer S. Regulation of lifespan in Drosophila by modulation of genes in the TOR signaling pathway. Curr Biol 2004 May 25;14(10):885-90.

(41) Stanfel MN, Shamieh LS, Kaeberlein M, Kennedy BK. The TOR pathway comes of age. Biochim Biophys Acta 2009 Jun 16.

(42) Guertin DA, Sabatini DM. An expanding role for mTOR in cancer. Trends Mol Med 2005 Aug;11(8):353-61.

(43) Stanfel MN, Shamieh LS, Kaeberlein M, Kennedy BK. The TOR pathway comes of age. Biochim Biophys Acta 2009 Jun 16.

(44) Jaffe AB, Hall A. Rho GTPases: biochemistry and biology. Annu Rev Cell Dev Biol 2005;21:247-69.

(45) van Aelst L, Souza-Schorey C. Rho GTPases and signaling networks. Genes Dev 1997 Sep 15;11(18):2295-322.

(46) Sullivan PF, Fan C, Perou CM. Evaluating the comparability of gene expression in blood and brain. Am J Med Genet B Neuropsychiatr Genet 2006 Apr 5;141(3):261-8.

(47) Diez D, Alvarez R, Dopazo A. Codelink: an R package for analysis of GE healthcare gene expression bioarrays. Bioinformatics 2007 May 1;23(9):1168-9.

(48) Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, et al. Bioconductor: open software development for computational biology and bioinformatics. Genome Biol 2004;5(10):R80.

(49) Smyth GK. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. Stat Appl Genet Mol Biol 2004;3:Article3.

(50) Smyth GK. Limma: linear models for microarray data. In: Gentleman R, Carey V, Dudoit R, Irizarry R, Huber W, editors. Bioinformatics and Computational Biology Solutions using R and Bioconductor.New York: Springer; 2005. p. 397-420.

(51) Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society Series B (Methodological) 1995;57(1):289-300.

(52) Goeman JJ, Oosting J, Cleton-Jansen AM, Anninga JK, van Houwelingen HC. Testing association of a pathway with survival using gene expression data. Bioinformatics 2005 May 1;21(9):1950-7.

(53) GlobalAncova R package version 3.6.0 [computer program]. 2007.

(54) Holm S. A simple sequentially rejective multiple test procedure. Scandinavian Journal of Statistics 1979;6:65-70.

(55) Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 2001 Dec;25(4):402-8.

Supplementary information

Supplementary figure and tables can be found online: http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0027759#s5

Figure S1. *Distribution of expression intensities of 2,953 age-related probes.*

Table S1. *Significant gene expression changes with chronological age and/or familial longevity.*

Table S2. *GO terms found significantly differentially expressed between nonagenarians and controls.*

Table S3. *List of probes of which the gene expression changes with age significantly differed between offspring and controls.*

Table S4. *RT-qPCR results of replication samples only.*

Chapter 4

IL7R gene expression network associates with human familial longevity

W.M. Passtoors[1], A.B. Maier[2], R. van der Breggen[1], J. Deelen[1,3], S. Trompet[2,4], J. W. Jukema[4,5,6], E. Derhovanessian[7], G. Pawelec[7], G.J.B. van Ommen[8,9], P.E. Slagboom[1,3], M. Beekman[1,3*]*

1 Section of Molecular Epidemiology, Leiden University Medical Center, Leiden, the Netherlands
2 Section of Gerontology and Geriatrics, Department of Internal Medicine, VU Medical Center, Amsterdam, the Netherlands
3 Netherlands Consortium for Healthy Ageing, the Netherlands
4 Department of Cardiology, Leiden University Medical Center, Leiden, the Netherlands
5 Durrer Center for Cardiogenetic Research, Amsterdam, the Netherlands
6 The Interuniversity Cardiology Institute (ICIN), Utrecht, the Netherlands
7 Center for Medical Research, University of Tübingen, Tübingen, Germany
8 Center for Human and Clinical Genetics, Leiden University Medical Center, Leiden, the Netherlands
9 Center for Medical Systems Biology, the Netherlands

[*]Both authors contributed equally to the studies

Manuscript in preparation.

Abstract

The level of expression of the interleukin 7 receptor (*IL7R*) gene in blood was recently found to be associated with familial longevity and healthy ageing. The IL7R is crucial for T cell development and important for immune competence. To further investigate the IL7R pathway in ageing we identified the closest interacting genes to construct an *IL7R* gene network that consisted of *IL7R* and six interacting genes: *IL2RG*, *IL7*, *TSLP*, *CRLF2*, *JAK1* and *JAK3*. This network was explored for association with chronological age, familial longevity and immune-related diseases (type 2 diabetes, COPD and rheumatoid arthritis) in 87 nonagenarians, 337 of their middle-aged offspring and 321 middle-aged controls from the Leiden Longevity Study (LLS). We observed that expression level of the *IL7R* gene network as a whole was significantly different between the nonagenarians and middle aged controls ($P=4.6 \times 10^{-4}$) being driven by a significantly lower level of expression of *IL7R*, *IL2RG* and *IL7* in the elderly. After correction for multiple testing and blood cell subsets, only *IL7R* was found to be associated with familial longevity exhibiting lower levels in middle aged offspring of nonagenarian siblings than in age matched controls (P=0.006). Higher *IL7R* gene expression in these groups associated with a higher prevalence of immune-related disease (P=0.001). Together our results indicate a positive correlation of *IL7R* gene expression with biological age. Paradoxically however a higher *IL7R* gene expression level associated with better prospective survival both in nonagenarians (HR=0.63, P=0.037) and the middle aged individuals ($HR=0.33$, $P=1.9 \times 10^{-4}$). We conclude that the *IL7R* network reflected by gene expression levels in blood may influence the biological age and health status of elderly individuals.

Introduction

Ageing is considered to be the consequence of an accumulation of physiological changes over time eventually increasing the mortality risk. Although ageing is the major risk factor for reduced health status and the most common human diseases of Western societies, treatment of elderly patients may be improved by understanding their biological age, the rate at which a person ages, and health status better than just by indicating their chronological age.

Many potential biomarkers of biological age have been suggested and are being tested for their association with chronological age, diseases and prospective mortality (1). In a search for new transcriptomic biomarkers of

ageing we found that the expression in blood of 1853 genes was associated with chronological age, 244 of which were associated with familial longevity, being differentially expressed between middle aged offspring of nonagenarian siblings and controls in the Leiden Longevity Study (LLS). Of the latter group of genes, a low expression level of the interleukin 7 (IL7) receptor (*IL7R*) associates with low disease prevalence among these offspring (2). The IL7 receptor is important for the body's innate and adaptive immune responses, and plays a role in regulating development, differentiation and survival of T cells (3;4). The IL7R is required for IL7 signaling, which in mice was shown to be crucial for early T cell development, as well as for homeostasis of naïve and memory CD8+ T cells (5;6). Previously, we observed that the offspring of nonagenarian siblings avoid the usual age-related reduction of percentages and numbers of naïve T cells in the periphery (7). Reduction of proinflammatory IL7 signaling may contribute to this better retention of naïve T cells and may thereby influence the biological age and health status of elderly individuals, given that possessing a fuller naive T cell repertoire would be expected to protect better against new pathogens.

To explore whether the expression profiles of other genes in the close vicinity of IL7R may exhibit even better ageing biomarker properties, we first identified the six interaction partners of *IL7R* using the STRING protein-protein-interaction database. Next, gene expression levels of all seven genes in the *IL7R* network were determined in 87 nonagenarians also having a nonagenarian sibling, 337 of their offspring and 321 controls of the LLS. We tested whether the level of expression of the *IL7R* network genes associated with chronological age by comparing the nonagenarians with the middle-aged controls. Differential expression in this comparison may be explained by the age difference between these groups, early environmental factors, or by the longevity trait in these nonagenarians which is not present in controls. To further investigate whether the expression of the age-associated genes also associates with biological age and disease prevalence, we compared the controls with the similarly aged offspring of the nonagenarians, representing individuals with a reduced biological age marked by lower prevalence of common age-related disease and beneficial metabolic profiles (8;9). Additionally, we explored whether the expression of the IL7R interaction partners is associated with common immune-related diseases in middle-aged individuals. Finally, we performed survival analysis in the nonagenarians to determine the relationship between expression of these genes and mortality. Because whole blood consists of multiple cell types, all associations have been adjusted for the distribution of the cell types.

Results

IL7R network

To identify proteins interacting with the IL7 receptor, we searched for known and predicted protein-protein interactions using the freely-accessible STRING 9.0 data base (http://string-db.org/) in May 2012. Interactions based on text mining were excluded, while predicted interactions based on experimental data with the highest confidence (score > 0.900) were taken into account. This approach resulted in an IL7R network that consists of IL7R and the following six interacting proteins: IL2RG, IL7, TSLP, CRLF2, JAK1 and JAK3 (Figure 1).

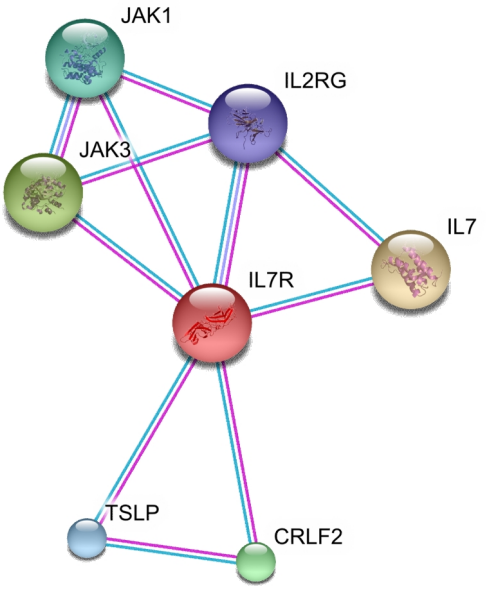

Figure 1. IL7R STRING network

Gene expression in whole blood

We measured gene expression levels for each of the seven genes using RT-qPCR on whole blood samples from 87 nonagenarians, 337 of their middle-aged offspring and 321 middle-aged controls (Table 1). To test for associations between expression of genes in the IL7R network and chronological age, we compared long-lived individuals to the middle-aged controls for differential expression of the total gene set of the *IL7R* network using a global test (10;11). We observed that the expression level of the

IL7R gene set as a whole was significantly different between the groups (P-value=4.6 x 10^{-4}).

Table 1. Subject characteristics

	Nonagenarians	Offspring	Controls
Number	87	337	321
Mean Age	94.3	61.3	61.2
Age Range	89.0 - 101.7	33.6 - 78.3	32.4 - 81.4
Women (%)	47 (54.0%)	143 (42.4%)	175 (54.5%)
N T2D (%)	NA	17 (5.0%)	26 (8.1%)
N COPD (%)	NA	19 (5.6%)	10 (3.1%)
N RA (%)	NA	3 (0.9%)	2 (0.6%)

N T2D: number of known patients with Type 2 Diabetes, N COPD: number of known patients with chronic obstructive pulmonary disease, N RA: number of known patients with rheumatoid arthritis, NA: data not available.

To investigate which genes were primarily responsible for the association of the IL7R network with age, we tested single gene expression levels using linear regression of the seven genes in the IL7R network (Table 2). After Bonferroni correction for multiple testing, three genes showed significant differential expression with at least a 5% difference between nonagenarians and middle aged controls, namely *IL7R*, *IL2RG* and *IL7*. We observed that expression of the IL receptor complex/ligand genes *IL7R*, *IL2RG*, *IL7*, *TSLP* and *CRLF2* were all lower in nonagenarians, while expression of *JAK1* and *JAK3* was increased.

Table 2. Gene expression of nonagenarians compared to controls by linear regression analysis.

Gene	Coef	FC	P	Bonferroni
IL7R	**-0.40**	**0.76**	**<10-6**	**0.000**
IL2RG	**-2.09**	**0.23**	**0.007**	**0.049**
IL7	**-0.81**	**0.57**	**<10-6**	**0.000**
TSLP	-0.01	0.99	0.013	0.091
CRLF2	-0.15	0.90	0.040	0.280
JAK1	0.21	1.16	0.046	0.322
JAK3	1.68	3.20	0.009	0.063

Coef: coefficient from linear regression model. FC: fold change, above one indicated higher expression in long-lived individuals. p: raw p value from linear regression model. Bonferroni: p value after adjustment for multiple testing (n=7) by the Bonferroni method. Genes significantly differentially expressed with at least 5% are depicted in bold.

Because differential expression in these comparisons may be explained by the age difference between the two groups, cohort effects, or by the

longevity trait in these families which is not present in controls, we investigated whether the differences in expression of the seven IL7R network genes was characteristic for these long-lived families, exhibiting a lower biological age, and not just a marker for chronological age. Therefore, we compared the expression in the middle-aged offspring of nonagenarians to their partners serving as controls matched for age and adult lifestyle and environmental conditions. Expression of all the IL7R complex/ligands *IL7R*, *IL2RG*, *IL7*, and *TSLP* and of *JAK3* was found to be lower in offspring of nonagenarians than in controls. Of these, *IL7R* remained significant after Bonferroni correction for multiple testing (Table 3), indicating its association with familial longevity and biological age. Similar results were obtained after adjustment for blood cell counts (Supplementary Table S1). Expression of *JAK1* was higher in offspring as compared to controls. The direction of the differential gene expression was the same as for nonagenarians, except for *JAK3* for which the expression was increased in nonagenarians and decreased in their offspring relative to controls.

Table 3. Gene expression of offspring from nonagenarians compared to controls by linear regression analysis.

Gene	Coef	FC	P	Bonferroni
IL7R	**-0.18**	**0.89**	**0.001**	**0.006**
IL2RG	-0.72	0.61	0.563	1.000
IL7	-0.39	0.77	0.050	0.350
TSLP	-0.01	0.99	0.140	0.980
CRLF2	-0.04	0.97	0.742	1.000
JAK1	0.08	1.05	0.557	1.000
JAK3	-0.30	0.81	0.650	1.000

Coef: coefficient from linear regression model. FC: fold change, above one indicated higher expression in offspring from long-lived individuals. p: raw p value from linear regression model. Bonferroni: p value after adjustment for multiple testing (n=7) by the Bonferroni method. Genes significantly differentially expressed with at least 5% are depicted in bold.

Relation of IL7R gene expression in whole blood with membrane-bound IL7R protein in PBMCs

Since decreased *IL7R* gene expression associates with familial longevity already in middle-age, the question arises whether this reflects soluble or membrane-bound IL7R protein, also known as cd127. We were able to investigate membrane-bound cd127 levels in PBMCs of 71 offspring and 73 controls. We observed no correlation between *IL7R* gene expression in whole blood and cd127 levels in PBMCs of the same individuals (Table 4) and no difference in cd127 between offspring and controls (mean level offspring=38.1, mean level controls=37.5, p-value=0.91). It is therefore less

likely that the difference in gene expression levels reflect differences in levels of membrane-bound cd127.

Table 4. Correlation of IL7R gene expression and cd127 expression in offspring from nonagenarians and controls.

	Offspring (n = 53)		Controls (n = 53)	
Subset	Cor	p	Cor	p
PBMC	0.161	0.250	-0.171	0.221
Lymphocytes	0.169	0.226	-0.135	0.336
T cells	0.174	0.212	-0.159	0.256
Non-T cells	0.152	0.278	-0.130	0.353
CD4+ cells	0.169	0.227	-0.151	0.281
CD8+ cells	0.167	0.231	-0.135	0.336

Cor: Pearson Correlation. p: raw p value from Pearson Correlation.

Association of IL7R gene expression with immune-related disease

Because *IL7R* and IL7 signaling have been implicated in the etiology of immune-related disease (12-15), the observed differences in gene expression with biological age might also be associated with different immune-related disease prevalence between the groups (Table 1). We investigated expression levels of *IL7R* relative to disease status for type 2 diabetes (T2D), chronic obstructive pulmonary disease (COPD) and rheumatoid arthritis (RA) in the offspring of long-lived individuals and controls. Table 5 shows that higher *IL7R* expression is associated with a higher prevalence of immune-related diseases. However, the difference in *IL7R* expression between offspring and controls did not change when adjusted for prevalence of T2D, COPD and RA (Supplementary Table S2). Thus, the *IL7R* expression levels in blood associate with immune-related disease on the one hand and familial longevity on the other, independent of these immune-related diseases.

Table 5. Association of IL7R gene expression with immune-related diseases in middle aged subjects

	T2D (n=43)		COPD (n=29)		RA (n=5)		Sumscore (n=70)	
Gene	Coef	p	Coef	p	Coef	P	Coef	p
IL7R	**0.22**	**0.029**	**0.35**	**9.9E-05**	-0.09	0.391	**0.21**	**0.001**

Total: total prevalence of T2D, COPD or RA. Coef: coefficient from linear regression model. FC: fold change, above one indicated positive association between gene expression and disease prevalence (patient = 1, control = 0). p: raw p value from linear regression model. Number of known patients is mentioned in table, the total number of offspring and controls is 658. Significant associations are depicted in bold.

Association of IL7R gene expression with mortality

Many markers indicating health status in middle age (such as blood pressure) associate also with mortality at higher ages, albeit not always in the expected direction (16). We examined whether the *IL7R* gene expression level is associated with prospective mortality in the subset of LLS participants for which we have measured expression levels. We performed a survival analysis using a Cox-proportional hazard model for low versus high gene expression levels in 81 nonagenarians (Model 1) and 619 offspring and partners (Model 2). The hazard ratio among the nonagenarians was 0.63 (95%CI 0.40-0.97) with a P-value of 0.037, and the hazard ratio among the middle aged individuals was 0.33 (95%CI 0.18-0.59) with a P-value of 1.91×10^{-4}. Unexpectedly, higher IL7R gene expression levels in blood associated with better survival in both age groups.

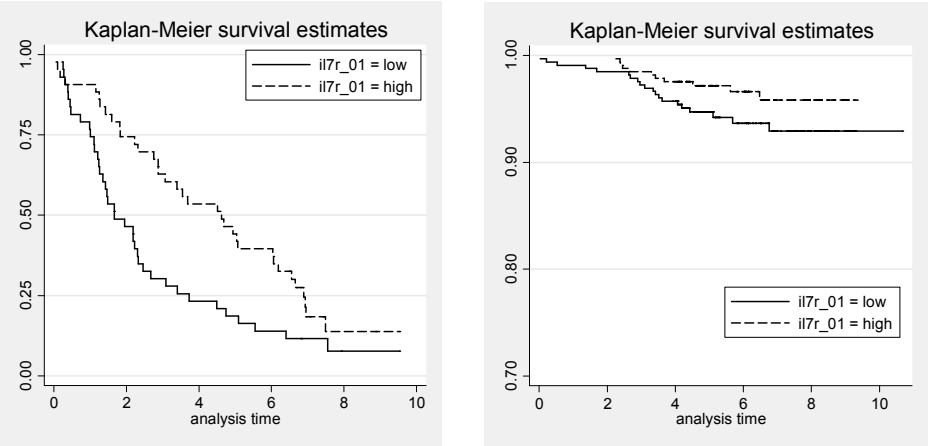

Figure 2. *Kaplan-Meier curve for high (dotted line) and low (solid line) IL7R gene expression in nonagenarians (left panel) and middle aged subjects (right panel).*

Discussion

Gene expression levels in blood of the IL7 geneset as a whole correlates negatively with chronological age. The expression of the *IL7R* gene in particular was found to be indicative for biological age and immune-related diseases. Relative to controls, reduced *IL7R* expression characterized members of longevity families, both middle aged offspring and nonagenarian siblings, whereas increased expression associated with prevalence of immune-related diseases in these groups (T2D and COPD). Paradoxically, however, higher *IL7R* gene expression levels associated

with better survival in a prospective analysis of both the nonagenarians and middle aged subjects. The mechanism underlying these observations requires further investigation, and especially whether regulation of *IL7R* gene expression has a causal role as determinant of biological ageing.

The notion that low *IL7R* expression levels are beneficial for reaching old age healthily corresponds with previous observations that autoimmune disease patients express increased levels of the IL receptor/ligand complex genes (12;13;15) and that antagonizing IL7 or IL7R may offer possible treatment (14;15). An increase in systemic inflammation has generally been reported with increasing age, so-called “inflamm-ageing” (17). Long-lived individuals such as LLS nonagenarians can be seen as slow or healthy agers who do not show the commonly observed age-related characteristics of “immunosenescence” and display relatively low levels of pro-inflammatory markers (7). Our results may suggest that nonagenarian members of long-lived families have more efficient IL7 signaling since they seem to require less compensation because of their remaining naïve T cell population, resulting in greater reserve capacity to cope with infections in old age. On the other hand, very low IL7R signaling has been observed in severe combined immunodeficiency (SCID) (18) and in HIV infection (19;20). We conclude that, besides the troublesome effect of the absence of IL7 signaling, a somewhat lower baseline level of *IL7R* gene expression, and potentially IL7 signaling, may contribute to healthy aging.

Despite these observations, the expression of the *IL7R* gene network was negatively correlated with age suggesting that low expression levels correlate with decreasing health, consistent with our observations that low *IL7R* gene expression associates with higher prospective mortality. Clearly this finding is inconsistent with the observation that in the same group of subjects low *IL7R* gene expression level is associated with familial longevity and decreased disease risk. A similar paradox has been found in PBMCs, where a lower frequency of naïve T cells (and a higher frequency of differentiated T cells) was associated with better survival in elderly individuals while a higher frequency was found in offspring of nonagenarians compared to the controls suggesting an association with lower biological age (7;21). This might be interpreted as follows: if the individual is able to counter the reduction of *IL7R* gene expression with age by keeping the level of differentiated T cells high as memory cells to control disease, this may be associated with healthy old age.

Contributing to the paradox may be the fact that there are two forms of IL7R exerting different functions, membrane-bound and soluble IL7R.

Membrane-bound IL7R may transduce IL7 signaling, while soluble IL7R may represent a negative compensatory mechanism regulating IL7 signaling (20;22;23). Because we used the recommended Taqman assay for the measurement of *IL7R* gene expression that cannot distinguish between the *IL7R* splice forms, we are unable to interpret how IL7 signaling is affected by the gene expression changes. Since *IL7R* gene expression levels and membrane-bound IL7R protein did not correlate in offspring or controls, and also the protein levels did not show a difference between the two groups, our results might suggest that the decrease in *IL7R* gene expression reflects mainly a decrease in the mRNA coding for soluble IL7R, resulting in more efficient IL7 signaling.

We showed that *IL7R* gene expression in blood is associated with immune-related disease. Previous meta-analysis and GWAS studies showed that genetic variation in the *IL7R* gene is associated with ulcerative colitis (24), multiple sclerosis (25-27), primary biliary cirrhosis (28), and type 1 diabetes (29). These results may suggest that genetic variation in and the expression of the *IL7R* gene is involved in auto-immune and chronic inflammation disease. Hence optimal response of the immune system may contribute to human longevity.

Further evidence that IL7 signaling may contribute to biological ageing and longevity is that it is closely connected to mTOR signaling, a pathway known for its effects on lifespan in animal models and which is now also implicated in human aging and longevity (30;31). Several studies in mice provided evidence for the connection between IL7 and mTOR signaling. IL7 induces phosphorylation of the mTOR complex 1 (mTORC1) downstream targets S6 and 4EBP1, an effect antagonized by the mTOR inhibitor rapamycin. The reverse may occur as well: rapamycin inhibits proliferation and induces apoptosis of pre-B acute lymphoblastic leukemia (ALL), effects which are abrogated by IL7 (32). Functional studies validated *IL7R* as a FoxO1 target gene (33). Acute deletion of FoxO1 induced a rapid and profound downregulation of IL7R expression associated with a significant reduction of *IL7R* mRNA (34). Also, cytokine (including IL7) stimulation induces FoxO1 phosphorylation and decreased transcription of target genes (34). In B cells, it has been shown that the mTOR complex 2 (mTORC2) suppresses *IL7R* gene expression by regulating FoxO1 phosphorylation (35). Taken together, decreased IL7 levels and decreased mTORC1 activity seem to go hand-in-hand, further implying that decreased IL7 levels are beneficial because decreased mTORC1 increases lifespan. The same is true for decreased IL7R and increased mTORC2, the latter we have recently shown to be associated with human longevity (31). Gene

expression levels of several of the mTOR-related genes are indeed positively and significantly correlated with *IL7R* expression in the LLS, which included *RPTOR*, *FOXO1* and *MTOR*. Considering the clear connections with IL7 signaling and similar findings on the level of gene expression variations, mTOR signaling might also be involved in “inflamm-ageing” as part of its lifespan-regulating effect.

In a previous study within the LLS families Dekker et al (36) found in the cultured skin fibroblasts of the offspring of the nonagenarians more apoptotic activity and less senescence than in those of controls. Flow cytometric analysis in HIV-infected individuals demonstrated that CD8 cells expressing high levels of *IL7R* also expressed slightly higher levels of anti-apoptotic markers, whereas nearly all apoptotic cells had low levels of *IL7R* (37). Our finding of decreased *IL7R* expression in members of long-lived families fits well with these previous results.

In conclusion, an overall lower expression level in blood of genes belonging to the IL7R network was found to be associated with higher chronological age. Yet, low *IL7R* gene expression was significantly associated with familial longevity in middle-age independent of blood cell counts and high expression level with prevalence of T2D, COPD and RA. Intriguingly nonetheless, higher *IL7R* gene expression associates with better prospective survival. The level of expression of the *IL7R* gene in the blood is a very promising marker for healthy ageing in long lived families, although further research is required to understand how *IL7R* gene regulation contributes to biological ageing.

Materials & Methods

Study population

Leiden Longevity Study
The individuals investigated in this study are participants of the Leiden Longevity Study. The families participating in this study have at least two siblings with a minimum age for men of 89 years and for women of 91 years (38). The offspring of these long-lived individuals, who have an increased chance of becoming long-lived (30% reduced standardized mortality rate), were also included. In addition, the partners of the offspring were included as population controls of similar age and environmental exposures as the offspring, and as a young control group for the nonagenarian siblings. Blood samples were taken from all the participants. The Leiden Longevity Study was approved by the Medical Ethical

Committee of Leiden University Medical Centre and all participants gave written informed consent.

Sample collection and RNA preparation

87 non-related long-lived siblings, 337 offspring and 321 partners belonging to 281 nuclear families were selected for the current study (Table 1 and Supplementary Table S5). These samples were randomly selected, but in such a way that age and gender were balanced between the groups and age range was as large as possible. Only individuals without outlying cell counts (beyond 3 SDs below or above the standard error of the mean) were included. This subpopulation is representative for the whole LLS regarding disease prevalence and parameters involved in metabolic syndrome (Supplementary Table S5) (9;39). From these non-fasted individuals, peripheral blood was harvested using PAXgeneTM tubes (Qiagen, Venlo, The Netherlands). The tubes were frozen and kept at -20°C for ~3-5 years. After thawing at room temperature for at least 2 hours, total RNA was extracted from the approximately 2.5 ml of peripheral blood in each tube following the manufacturer's recommended protocol (PAXgene Blood RNA Kit Handbook, Qiagen, Venlo, The Netherlands). The quality and integrity of the total RNA was evaluated on the 2100 Bioanalyzer (Agilent Technologies, Amstelveen, The Netherlands) and the concentration was measured using a NanoDrop spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA). Quality criteria included a 28S/18S ratio as measured by the Bioanalyzer of at least 1.2, and a total RNA yield of at least 3 µg.

RT-qPCR

For all 7 IL7R network genes the suggested Taqman® assay (Applied Biosystems) was selected. Reverse transcription was performed with total RNA from blood of in total 745 samples, which passed QC using the First Strand cDNA Synthesis Kit according to the manufacturer's protocol (Roche Applied Science). cDNA was amplified using the DNA Engine Tetrad® 2 Peltier Thermal Cycler (Bio-Rad). qPCR was then performed with the Taqman® method using the BiomarkTM 48.48 and 96.96 Dynamic Arrays (Fluidigm). Relative gene expression of the BioMark™ Array data were calculated by using the 2-ΔΔCt method, in which Ct indicates cycle threshold, the fractional cycle number where the fluorescent signal reaches the detection threshold (40). YKT6 was used as internal control and commercially available human total reference RNA (Clontech Laboratories, Mountain View, CA, USA) as reference RNA.

Geneset analysis of gene expression data

The Globaltest methodology was designed to determine whether the common expression pattern of genes within a pre-defined set is significantly related to clinical outcome (10;41). A generalized linear model is used to estimate a "Q-statistic" for each gene set, which describes the correlation between gene expression profiles, X, and clinical outcomes, Y. The Q-statistic for a gene set is the average of the Q-statistics for each gene in the set. The Globaltest method was used to perform geneset analysis comparing two groups of individuals (either nonagenarians vs. controls or offspring vs. controls) including age (in offspring vs. controls only) and gender and their interaction as covariates. Global test package in R (10) has been used to perform analyses.

Single gene analysis of gene expression data

Differences in expression level between nonagenarians, their offspring and the partners of their offspring were assessed using linear regression. In these analyses, expression level was the dependent variable and the two groups of individuals (either nonagenarians vs. controls or offspring vs. controls) were included in the model as a categorical variable together with age (in offspring vs. controls only), gender and their interaction as covariates. To take into account dependencies within sibships, robust standard errors were used, i.e. the variance was computed from the between family variation. P-values were also based on these robust standard errors. Analyses were performed using the software package STATA/SE 11.0 (DPC Software, StataCorp 2009).

Blood cell subtypes

To further investigate the candidate genes, their expression level was associated with several blood cell counts. In the whole blood samples of the participants the following cell subtypes were counted using the automated Siemens ADVIA 1200 system (SMSD, Tarrytown, NY) in the Leiden Medical Diagnostics Center: leukocytes, thrombocytes, neutrophils, lymphocytes, monocytes, basophils and eosinophils. Next, both the comparison between nonagenarians and controls and between offspring and controls were adjusted for each of these cell counts separately.

CD127 staining

Cryopreserved peripheral blood mononuclear cells (PBMC) were thawed and treated with human immunoglobulin (GAMUNEX; TalecrisBiotherapeuthics) and ethidiummonoazide (EMA, Invitrogen) to block Fc receptors and stain dead cells followed by an indirect staining for CD3 using an OKT3 supernatant and a pacific-orange-cojugated anti-mouse IgG (Invitrogen). After blocking unbound secondary antibodies with mouse serum (Chemicon, Millipore), cells were surface-stained with CD4-Pacifid Blue, CD127-Alexa Fluor 647 (BioLegend, San Diego, USA) and CD8-PerCP (BD Biosciences, Heidelberg, Germany). Cells were measured immediately using an LSR-II (BD).

For data analysis, EMA+ dead cells were excluded and lymphocytes were gated using an FSC-vs.-SSC dot plot based on their size and granularity. T-cells and non-T-cells were characterized as CD3+ and CD3- cells within the lymphocyte gate. In the CD3+ gate, CD4 and CD8 cells were characterized in an CD4-vs.-CD8 dot plot as CD4+CD8- and CD4-CD8+ cells, respectively. The mean fluorescence intensity (MFI) of Alexa-Fluor647 (CD127) was determined on total lymphocytes, CD3-, CD3+, CD4+ and CD8+ cells. To standardize for fluctuation of the instrument over the measurement period of a few weeks, MFI of each studied population was standardized against the MFI of an unstained control PBMC for each experimental day, by dividing the MFI of each population by that of the cells in the lymphocyte gate of the unstained control. Flow cytometry data analysis was performed using the FlowJo software (Tristar, San Diego, USA).

Association of gene expression with immune-related diseases

Information on medical history was requested from the participants' general practitioners. Gene expression of *IL7R* and *IL7* was associated with prevalence of T2D, COPD and RA and the sum score of these diseases. The sum score indicates the number of patients with T2D, COPD or RA or a combination thereof. Next, the comparison between offspring and controls for *IL7R* and *IL7* expression association with longevity was adjusted for the prevalence of these diseases, using the linear regression model described above.

Prospective survival analysis

Prospective analysis of IL7R-related genes was performed with 81 nonagenarians (Model 1), 313 offspring and 306 controls (Model 2). After a mean follow-up time of 7.40 (nonagenarians) and 6.15 years (offspring and controls), 87.7% (n = 71, nonagenarians) and 4.5% (n = 28, offspring and controls) of the individuals had died. Mortality analyses were performed with STATA/SE 11.2 (StataCorp LP, TX, USA) using an age at inclusion, sex and white blood cell count-adjusted, left-truncated Cox proportional hazards model to adjust for late entry into the dataset according to age.

Model 1: $\lambda(t) \sim \lambda 0(t) \exp(\beta 1 * \text{age at inclusion} + \beta 2 * \text{sex} + \beta 3 * \text{lymphocyte count} + \beta 4 * \text{neutrophil count} + \beta 5 * \text{monocyte count} + \beta 6 * \text{eosinophil count} + \beta 7 * \text{basophil count} + \beta 8 * \text{Gene})$

Model 2: $\lambda(t) \sim \lambda 0(t) \exp(\beta 1 * \text{age at inclusion} + \beta 2 * \text{sex} + \beta 3(\text{age} * \text{sex}) + \beta 4 * \text{group} + \beta 5 * \text{lymphocyte count} + \beta 6 * \text{neutrophil count} + \beta 7 * \text{monocyte count} + \beta 8 * \text{eosinophil count} + \beta 9 * \text{basophil count} + \beta 10 * \text{Gene})$

Age at inclusion was coded in years, sex was coded as 1 (male) or 2 (female) and group was coded as 1 (control) or 2 (offspring). Robust standard errors were used to account for sibship relations.

Acknowledgements

We thank all participants of the Leiden Longevity Study. The research leading to these results has received funding from the European Union's Seventh Framework Programme (FP7/2007-2011) under grant agreement n° 259679 (IDEAL) and n° HEALTH-F2-2009-223004. This study was supported by a grant from the Innovation-Oriented Research Program on Genomics (SenterNovem IGE05007), the Centre for Medical Systems Biology (CMSB) [http://www.cmsb.nl], the Netherlands Consortium for Healthy Ageing (NCHA, Grant 050-060-810) [http://www.healthy-ageing.nl], all in the framework of the Netherlands Genomics Initiative, Netherlands Organization for Scientific Research (NWO) and by Unilever Colworth. This work was furthermore funded by grants from the Interuniversity Cardiology Institute of the Netherlands (ICIN) http://www.icin.nl/ and JWJ is an established clinical investigator of the Netherlands Heart Foundation (2001D032) [http://www.hartstichting.nl/]. The funders had no role in study design, data collection and analysis, decision to publish or the preparation of the manuscript.

References

(1) Deelen J, Beekman M, Capri M, Franceschi C, Slagboom PE. Identifying the genomic determinants of aging and longevity in human population studies: progress and challenges. Bioessays 2013 April;35(4):386-96.

(2) Passtoors WM, Boer JM, Goeman JJ, van den Akker EB, Deelen J, Zwaan BJ et al. Transcriptional Profiling of Human Familial Longevity Indicates a Role for ASF1A and IL7R. PLoS One 2012;7(1):e27759.

(3) DeKoter RP, Schweitzer BL, Kamath MB, Jones D, Tagoh H, Bonifer C et al. Regulation of the interleukin-7 receptor alpha promoter by the Ets transcription factors PU.1 and GA-binding protein in developing B cells. J Biol Chem 2007 May 11;282(19):14194-204.

(4) Gregory SG, Schmidt S, Seth P, Oksenberg JR, Hart J, Prokop A et al. Interleukin 7 receptor alpha chain (IL7R) shows allelic and functional association with multiple sclerosis. Nat Genet 2007 September;39(9):1083-91.

(5) von Freeden-Jeffry U, Vieira P, Lucian LA, McNeil T, Burdach SE, Murray R. Lymphopenia in interleukin (IL)-7 gene-deleted mice identifies IL-7 as a nonredundant cytokine. J Exp Med 1995 April 1;181(4):1519-26.

(6) Peschon JJ, Morrissey PJ, Grabstein KH, Ramsdell FJ, Maraskovsky E, Gliniak BC et al. Early lymphocyte expansion is severely impaired in interleukin 7 receptor-deficient mice. J Exp Med 1994 November 1;180(5):1955-60.

(7) Derhovanessian E, Maier AB, Beck R, Jahn G, Hahnel K, Slagboom PE et al. Hallmark features of immunosenescence are absent in familial longevity. J Immunol 2010 October 15;185(8):4618-24.

(8) Heijmans BT, Beekman M, Houwing-Duistermaat JJ, Cobain MR, Powell J, Blauw GJ et al. Lipoprotein particle profiles mark familial and sporadic human longevity. PLoS Med 2006 December;3(12):e495.

(9) Westendorp RG, van HD, Rozing MP, Frolich M, Mooijaart SP, Blauw GJ et al. Nonagenarian siblings and their offspring display lower risk of mortality and morbidity than sporadic nonagenarians: The Leiden Longevity Study. J Am Geriatr Soc 2009 September;57(9):1634-7.

(10) Globaltest, R package, version 4.10.0 [computer program]. 2008.

(11) Goeman JJ, van de Geer SA, de Kort F, van Houwelingen HC. A global test for groups of genes: testing association with a clinical outcome. Bioinformatics 2004 January 1;20(1):93-9.

(12) Lee SK, Jeon EK, Kim YJ, Seo SH, Kim CD, Lim JS et al. A global gene expression analysis of the peripheral blood mononuclear cells reveals the gene expression signature in psoriasis. Ann Dermatol 2009 August;21(3):237-42.

(13) McKinney EF, Lyons PA, Carr EJ, Hollis JL, Jayne DR, Willcocks LC et al. A CD8+ T cell transcription signature predicts prognosis in autoimmune disease. Nat Med 2010 May;16(5):586-91, 1p.

(14) Liu X, Leung S, Wang C, Tan Z, Wang J, Guo TB et al. Crucial role of interleukin-7 in T helper type 17 survival and expansion in autoimmune disease. Nat Med 2010 February;16(2):191-7.

(15) Bikker A, Hack CE, Lafeber FP, van Roon JA. Interleukin-7: a key mediator in T cell-driven autoimmunity, inflammation, and tissue destruction. Curr Pharm Des 2012;18(16):2347-56.

(16) van BT, Gussekloo J, Westendorp RG, Blauw GJ. In a population-based prospective study, no association between high blood pressure and mortality after age 85 years. J Hypertens 2006 February;24(2):287-92.

(17) Franceschi C, Bonafe M, Valensin S, Olivieri F, De LM, Ottaviani E et al. Inflamm-aging. An evolutionary perspective on immunosenescence. Ann N Y Acad Sci 2000 June;908:244-54.

(18) Puel A, Ziegler SF, Buckley RH, Leonard WJ. Defective IL7R expression in T(-)B(+)NK(+) severe combined immunodeficiency. Nat Genet 1998 December;20(4):394-7.

(19) Pellegrini M, Calzascia T, Toe JG, Preston SP, Lin AE, Elford AR et al. IL-7 engages multiple mechanisms to overcome chronic viral infection and limit organ pathology. Cell 2011 February 18;144(4):601-13.

(20) Crawley AM, Angel JB. The influence of HIV on CD127 expression and its potential implications for IL-7 therapy. Semin Immunol 2012 June;24(3):231-40.

(21) Derhovanessian E, Maier AB, Hahnel K, Zelba H, de Craen AJ, Roelofs H et al. Lower proportion of naive peripheral CD8+ T cells and an unopposed pro-inflammatory response to human Cytomegalovirus proteins in vitro are associated with longer survival in very elderly people. Age (Dordr) 2013 August;35(4):1387-99.

(22) Fernandez-Botran R. Soluble cytokine receptors: basic immunology and clinical applications. Crit Rev Clin Lab Sci 1999 June;36(3):165-224.

(23) Fernandez-Botran R, Crespo FA, Sun X. Soluble cytokine receptors in biological therapy. Expert Opin Biol Ther 2002 August;2(6):585-605.

(24) Anderson CA, Boucher G, Lees CW, Franke A, D'Amato M, Taylor KD et al. Meta-analysis identifies 29 additional ulcerative colitis risk loci, increasing the number of confirmed associations to 47. Nat Genet 2011 March;43(3):246-52.

(25) Sawcer S, Hellenthal G, Pirinen M, Spencer CC, Patsopoulos NA, Moutsianas L et al. Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. Nature 2011 August 11;476(7359):214-9.

(26) De Jager PL, Jia X, Wang J, de Bakker PI, Ottoboni L, Aggarwal NT et al. Meta-analysis of genome scans and replication identify CD6, IRF8 and TNFRSF1A as new multiple sclerosis susceptibility loci. Nat Genet 2009 July;41(7):776-82.

(27) Hafler DA, Compston A, Sawcer S, Lander ES, Daly MJ, De Jager PL et al. Risk alleles for multiple sclerosis identified by a genomewide study. N Engl J Med 2007 August 30;357(9):851-62.

(28) Nakamura M, Nishida N, Kawashima M, Aiba Y, Tanaka A, Yasunami M et al. Genome-wide association study identifies TNFSF15 and POU2AF1 as susceptibility loci for primary biliary cirrhosis in the Japanese population. Am J Hum Genet 2012 October 5;91(4):721-8.

(29) Todd JA, Walker NM, Cooper JD, Smyth DJ, Downes K, Plagnol V et al. Robust associations of four new chromosome regions from genome-wide analyses of type 1 diabetes. Nat Genet 2007 July;39(7):857-64.

(30) Harries LW, Fellows AD, Pilling LC, Hernandez D, Singleton A, Bandinelli S et al. Advancing age is associated with gene expression changes resembling mTOR inhibition: evidence from two human populations. Mech Ageing Dev 2012 August;133(8):556-62.

(31) Passtoors WM, Beekman M, Deelen J, van der BR, Maier AB, Guigas B et al. Gene expression analysis of mTOR pathway: association with human longevity. Aging Cell 2012 October 13.

(32) Brown VI, Hulitt J, Fish J, Sheen C, Bruno M, Xu Q et al. Thymic stromal-derived lymphopoietin induces proliferation of pre-B leukemia and antagonizes mTOR inhibitors, suggesting a role for interleukin-7Ralpha signaling. Cancer Res 2007 October 15;67(20):9963-70.

(33) Ouyang W, Beckett O, Flavell RA, Li MO. An essential role of the Forkhead-box transcription factor Foxo1 in control of T cell homeostasis and tolerance. Immunity 2009 March 20;30(3):358-71.

(34) Kerdiles YM, Beisner DR, Tinoco R, Dejean AS, Castrillon DH, DePinho RA et al. Foxo1 links homing and survival of naive T cells by regulating L-selectin, CCR7 and interleukin 7 receptor. Nat Immunol 2009 February;10(2):176-84.

(35) Lazorchak AS, Liu D, Facchinetti V, Di LA, Sessa WC, Schatz DG et al. Sin1-mTORC2 suppresses rag and il7r gene expression through Akt2 in B cells. Mol Cell 2010 August 13;39(3):433-43.

(36) Dekker P, Maier AB, van HD, de Koning-Treurniet C, Blom J, Dirks RW et al. Stress-induced responses of human skin fibroblasts in vitro reflect human longevity. Aging Cell 2009 September;8(5):595-603.

(37) Paiardini M, Cervasi B, Albrecht H, Muthukumar A, Dunham R, Gordon S et al. Loss of CD127 expression defines an expansion of effector CD8+ T cells in HIV-infected individuals. J Immunol 2005 March 1;174(5):2900-9.

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

(39) Rozing MP, Westendorp RG, de Craen AJ, Frolich M, de Goeij MC, Heijmans BT et al. 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 2010 March;58(3):564-9.

(40) Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 2001 December;25(4):402-8.

(41) Goeman JJ, Oosting J, Cleton-Jansen AM, Anninga JK, van Houwelingen HC. Testing association of a pathway with survival using gene expression data. Bioinformatics 2005 May 1;21(9):1950-7.

Supplementary information

Table S1. Linear regression results of gene expression of offspring from long-lived individuals compared to controls, adjusted for blood cell numbers.

	No adjustment		Neutrophils		Lymphocytes		Monocytes		Eosinophils		Basophils	
Gene	Coef	p	Coef	p	Coef	p	Coef	p	Coef	p	Coef	p
IL7R	**-0.18**	**0.001**	**-0.16**	**0.003**	**-0.15**	**0.004**	**-0.17**	**0.001**	**-0.18**	**0.001**	**-0.17**	**0.001**
IL2RG	-0.72	0.563	-0.34	0.787	-0.19	0.877	-0.19	0.882	-0.51	0.687	-0.52	0.681
IL7	-0.39	0.050	-0.38	0.049	-0.35	0.070	-0.44	0.032	-0.45	0.024	-0.44	0.030
TSLP	-0.01	0.140	-0.01	0.121	-0.01	0.134	-0.01	0.103	-0.01	0.112	-0.01	0.092
CRLF2	-0.04	0.742	-0.05	0.715	-0.04	0.765	-0.05	0.713	-0.06	0.651	-0.09	0.515
JAK1	0.08	0.557	0.08	0.530	0.08	0.541	0.15	0.245	0.10	0.471	0.08	0.530
JAK3	-0.30	0.650	-0.36	0.593	-0.37	0.593	-0.08	0.905	-0.30	0.668	-0.29	0.680

Coef: coefficient from linear regression model. FC: fold change, above one indicated higher expression in offspring from long-lived individuals. p: raw p value from linear regression model. Genes significantly differentially expressed with at least 5% are depicted in bold.

Table S2. Linear regression results of IL7R gene expression of offspring from long-lived individuals compared to controls, adjusted for prevalence of immune-related diseases.

	No adjustment		Adjusted for T2D		Adjusted for COPD		Adjusted for RA		Adjusted for total	
Gene	Coef	p	Coef	p	Coef	p	Coef	p	Coef	p
IL7R	**-0.18**	**0.001**	**-0.17**	**0.001**	**-0.19**	**2.4E-4**	**-0.18**	**0.001**	**-0.18**	**0.001**

Total: total prevalence of T2D, COPD or RA. Coef: coefficient from linear regression model. FC: fold change, above one indicated higher expression in offspring from long-lived individuals. p: raw p value from linear regression model. Genes significantly differentially expressed with at least 5% are depicted in bold.

Chapter 5

Gene expression analysis of mTOR pathway: association with human longevity

W.M. Passtoors[1], M. Beekman[1,6], J. Deelen[1,6], R. van der Breggen[1], A.B. Maier[2], B. Guigas[3], E. Derhovanessian[4], D. van Heemst[2], A.J.M. de Craen[2], D.A. Gunn[5], G. Pawelec[4], P.E. Slagboom[1,6]

1 Section of Molecular Epidemiology, Leiden University Medical Center, Leiden, The Netherlands
2 Department of Gerontology and Geriatrics, Leiden University Medical Center, Leiden, The Netherlands
3 Department of Molecular Cell Biology, Leiden University Medical Center, Leiden, The Netherlands
4 Center for Medical Research, University of Tübingen, Tübingen, Germany
5 Unilever Discover, Colworth, Sharnbrook, Bedfordshire, United Kingdom
6 Netherlands Consortium for Healthy Ageing, The Netherlands.

Aging Cell, 2013, Feb;12(1): p. 24-31

Abstract

mTOR signalling is implicated in the development of disease and in lifespan extension in model organisms. This pathway has been associated with human diseases such as diabetes and cancer, but has not been investigated for its impact on longevity *per se*. Here, we investigated whether transcriptional variation within the mTOR pathway is associated with human longevity using whole blood samples from the Leiden Longevity Study (LLS). This is an unique cohort of Dutch families with extended survival across generations, decreased morbidity and beneficial metabolic profiles in middle-age. By comparing mRNA levels of nonagenarians and middle-aged controls, the mTOR signalling gene set was found to associate with old age ($p=4.6 \times 10^{-7}$). Single gene analysis showed that seven out of 40 mTOR pathway genes had a significant differential expression of at least 5%. Of these, the *RPTOR (Raptor)* gene was found to be differentially expressed also when the offspring of nonagenarians was compared to their spouses, indicating association with familial longevity in middle-age. This association was not explained by variation between the groups in the prevalence of type 2 diabetes and cancer or glucose levels. Thus, the mTOR pathway not only plays a role in the regulation of disease and aging in animal models, but also in human health and longevity.

Introduction

Mammalian target of rapamycin (mTOR) is an evolutionarily conserved nutrient-sensing protein kinase that belongs to the phosphoinositide 3-kinase (PI3K)-related protein kinase family and acts as a central regulator of growth and cell division (1). The core kinase mTOR is present in two distinct multiprotein complexes called mTOR complex 1 (mTORC1) and mTOR complex 2 (mTORC2), each with different functions (2), although their exact respective roles remain to be elucidated. The main known functions of mTORC1 include translation initiation, protein synthesis and autophagy, while those of mTORC2 include cytoskeletal organization (3). In addition, the two complexes are connected in that mTORC2 regulates Akt phosphorylation, part of an upstream pathway that controls activation of mTORC1 (1). Inhibition of the mTOR signalling pathway in yeast (4), worms (5), flies (6) and mice (7) results in lifespan extension. In mouse models, inhibition of the mTOR signalling pathway provides cardiovascular benefits and improved metabolic function (1). Additionally, a primary hallmark of calorie restriction (CR) in rodents, which leads to inhibition of the mTORC1 pathway, is a dramatic reduction in age-associated cancer incidence and

growth rate (8), implying that enhanced mTOR signalling plays a central role in cancer progression. In rats, the mTOR pathway has also been implicated in both Type 1 and Type 2 diabetes, with chronic activity of mTORC1 contributing to obesity and insulin insensitivity (9). Because reduced mTOR signalling in model organisms extends lifespan and is associated with better metabolic health, we postulated that a similar reduction in this signalling pathway may also contribute to longevity and healthy ageing in humans.

Thus far, mTOR signalling has been investigated in humans mainly by studying disease. Its activity has been found to be elevated in different cancers including lymphomas, melanomas, breast and prostate (1;10). Elevated mTOR signalling in human tumours is correlated with poor tumor prognosis, and inhibitors of this pathway are showing promising results in clinical trials (11-13). Additionally, intermediate phenotype insulin-resistance in human diabetes can be prevented by mTOR inhibitors (14). Metformin, a commonly-prescribed anti-diabetic drug which inhibits mTORC1 downstream pathways through activation of AMP kinase, has also been shown to reduce tumour growth (15). Human CR studies used as model for down-regulation of mTOR are much more short-term in comparison to animal studies, but nevertheless CR may reduce the risk of age-related diseases even in non-obese humans (16;17). Thus, accumulating evidence is showing that mTOR signalling can play an important role in human disease. In contrast, it has not been investigated whether natural variation in mTOR signalling could also contribute to disease risk.

Recently, we have identified 360 genes, the expression of which was associated with human longevity (18). *MTOR* (also known as *FRAP1*) was among these, suggesting a role for the mTOR signalling pathway in human longevity. In the current study, we investigated whether common natural genetic and transcriptional variation among the 40 genes of the mTOR pathway (Figure 1) was associated with human longevity. To do this, we took advantage of the Leiden Longevity Study (LLS), a Dutch cohort in which families are selected on the basis of nonagenarian sibling pairs (19). The middle-aged offspring of these nonagenarians have a decreased prevalence of myocardial infarction and type 2 diabetes (20), and display healthier intrinsic metabolism reflected by lipid profiles (21;22), glucose metabolism and preservation of insulin sensitivity (23;24) than aged matched controls. They also have different immune profiles in that they fail to show a lower number of CD8+ naive T cells and a higher number of CD8+ late-stage differentiated memory cells dependent on cytomegalovirus

infection often seen as a hallmark of immune aging (25). To investigate whether expression in mTOR signalling genes is associated with familial longevity, we compared the levels of mRNA for these genes in peripheral blood of nonagenarians from families enriched for longevity and younger controls. In addition, genetic variation of mTOR pathway genes of LLS nonagenarians is compared to that of younger controls. Next, mRNA levels of the top longevity-associated genes were compared between the offspring of the nonagenarians and similarly aged controls to exclude the possibility that differential expression of these genes did not simply mark old age, but are a defining characteristic of the long-lived families in middle-age. Finally, we examined whether the association of the expression of mTOR signalling genes with longevity is independent of metabolic health parameters, namely plasma glucose levels, and the prevalence of type 2 diabetes and cancer, because these factors are known characteristics of the long-lived families and are suggested to be influenced by mTOR signalling.

Results

Selection of mTOR pathway genes

For the mTOR pathway, we selected 40 genes encoding proteins that belong to the well-described core of the mammalian mTOR pathway. First we selected the two mTOR complexes mTORC1 and mTORC2 including their downstream targets often described in literature (2;3). Next, we included the direct activators and inhibitors of these two complexes. Figure 1 shows the mTOR pathway as we investigated.

Gene expression in whole blood: nonagenarians vs controls

We measured gene expression levels for each of the 40 mTOR signalling genes using RT-qPCR on whole blood samples from 87 LLS nonagenarians, 337 of their middle-aged offspring and 321 middle-aged LLS controls (Table 1). When comparing long-lived individuals to younger controls differentially expressed genes may associate either with calendar age, familial longevity, or both. We tested the total gene set for such association using the globaltest method (26). We observed that the gene expression levels of the mTOR signalling gene set did associate significantly with age and/or familial longevity in these families ($p=4.6 \times 10^{-7}$).

Second, to investigate which genes are primarily responsible for this association, single gene analysis using linear regression was performed on all 40 mTOR signalling genes (Table 2). After Bonferroni correction for multiple testing, seven genes showed significant differential expression with at least a 5% expression difference: *FOXO1* and *RPTOR* were expressed at a lower level, and *EIF4EBP1*, *LAMTOR2*, *AKT1S1*, *PRR5L*, and *RHOA* at a higher level in nonagenarians.

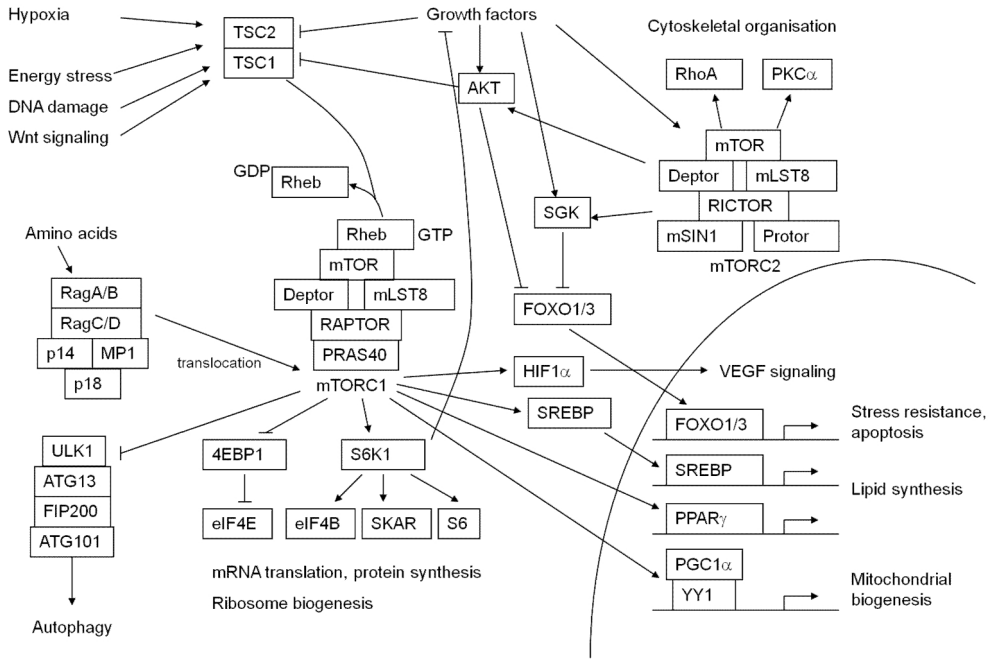

Figure 1. mTOR signalling pathway of which genetic variation and expression of 40 genes are investigated for their association with human ageing and longevity.
mTORC1 receives signals from growth factors, nutrients, energy status, and a range of stressors and positively regulates cell growth and proliferation by promoting many anabolic processes, including biosynthesis of proteins, lipids and organelles, and by limiting catabolic processes such as autophagy (44). mTORC2 phosphorylates and activates Akt, serum- and glucocorticoid-regulated kinase (sGK), RhoA and protein kinase C (PKC), which regulate cell survival, cell cycle progression and cytoskeleton organization (45-47).

Table 1 Analyzed samples: gene expression

Samples	Number	Mean Age	Age Range	Men/Women
Long-lived individuals	87	94.3	89-102	40/47
Offspring	337	61.3	34-78	194/143
Controls	321	61.2	32-81	146/175

Genetic analysis: nonagenarians vs controls

Since we had GWAS data available in the LLS study, we analyzed the set of GWAS-based SNPs within a 10-kb window around the 40 mTOR signalling genes for differences in variation between 417 unrelated nonagenarian participants and 476 younger controls from the Leiden Longevity Study (LLS) (Supplementary Table S1) (27). By using the PLINK set-based test we investigated 1,018 SNPs in the 40 mTOR signalling genes simultaneously in one test (p-value <0.05 for significance) (Supplementary Table S2). We observed a significant association of genetic variation in the mTOR pathway as a whole with familial longevity (p=0.009). However, the single gene analysis (of 40 SNP sets), for which the significance level for p values after Bonferroni correction for multiple testing is 0.00125, did not result in significant associations. (Supplementary Table S3), suggesting that effects of single common variation in mTOR signalling genes on familial longevity is very small.

Table 2. Linear regression results of gene expression of long-lived individuals compared to controls.

Protein	Genename	N cases	N controls	Mean	SD	FC	p
4EBP1	***EIF4EBP1***	**76**	**303**	**2.37**	**0.048**	**1.19**	**2.1×10^{-8}**
AKT	*AKT1*	76	303	2.27	0.103	1.18	0.030
ATG101	*C12orf44*	71	296	1.02	0.020	1.00	0.644
ATG13	*ATG13*	71	296	1.10	0.019	0.99	0.095
Deptor	*DEPTOR*	70	296	0.05	0.001	1.00	0.001
eIF4B	*EIF4B*	70	295	0.47	0.018	1.00	0.572
eIF4E	*EIF4E*	71	296	1.10	0.045	0.94	0.012
FIP200	*RB1CC1*	71	295	2.41	0.153	1.29	0.005
FOXO1	***FOXO1***	**71**	**296**	**1.50**	**0.034**	**0.90**	**1.0×10^{-11}**
FOXO3	*FOXO3*	71	296	1.73	0.057	1.05	0.166
HIF1a	*HIF1A*	71	296	4.05	0.322	1.90	0.005
mLST8	*MLST8*	71	296	0.42	0.010	1.00	0.446
MP1	*MAPKSP1*	71	296	1.64	0.039	0.97	0.112
mSIN1	*MAPKAP1*	70	294	1.42	0.018	0.98	0.002
mTOR	*MTOR*	68	294	0.71	0.013	0.97	6.0E-06
p14	***LAMTOR2***	**71**	**296**	**1.50**	**0.049**	**1.08**	**9.1×10^{-5}**
p18	*CDKN2C*	70	295	0.71	0.020	1.00	0.630
PGC1a	*PPARGC1A*	66	287	0.02	0.003	1.00	0.762
PKCa	*PRKCA*	71	296	0.57	0.016	0.95	5.4×10^{-9}
PPARg	*PPARG*	70	294	0.03	0.002	1.01	8.4×10^{-5}
PRAS40	***AKT1S1***	**70**	**296**	**0.87**	**0.049**	**1.14**	**1.9×10^{-4}**
Protor1	*PRR5*	71	296	4.47	0.197	0.82	0.035

Protor2	***PRR5L***	**71**	**296**	**7.03**	**0.211**	**1.81**	**4.9×10^{-5}**
RagA	*RRAGA*	71	295	1.09	0.038	0.95	0.009
RagB	*RRAGB*	71	296	0.67	0.019	0.99	0.028
RagC	*RRAGC*	71	296	1.69	0.041	1.02	0.231
RagD	*RRAGD*	70	296	0.66	0.015	1.02	0.035
Raptor	***RPTOR***	**79**	**312**	**0.70**	**0.037**	**0.92**	**0.001**
Rheb	*RHEB*	71	296	0.73	0.028	0.97	0.060
RhoA	***RHOA***	**71**	**296**	**2.15**	**0.037**	**1.08**	**5.6×10^{-5}**
Rictor	*RICTOR*	71	296	6.24	0.182	0.99	0.874
S6	*RPS6*	71	295	1.92	0.073	0.93	0.055
S6K1	*RPS6KB1*	76	303	3.05	0.110	0.94	0.648
SGK	*SGK1*	71	296	0.79	0.019	1.01	0.247
SKAR	*POLDIP3*	67	260	1.00	0.020	1.05	1.3×10^{-4}
SREBP	*SREBF1*	71	296	0.72	0.031	1.00	0.758
TSC1	*TSC1*	71	296	1.02	0.020	0.98	0.019
TSC2	*TSC2*	71	296	0.78	0.021	1.04	3.8×10^{-4}
ULK1	*ULK1*	76	303	2.47	0.081	1.19	0.002
YY1	*YY1*	71	296	1.73	0.037	1.01	0.740

Mean: mean relative expression level. SD: standard deviation of relative expression level. FC: fold change, above one indicated higher expression in long-lived individuals. p: raw p value from linear regression model. Significance level for p value after Bonferroni correction for multiple testing is 0.00125. Genes significantly differentially expressed with at least 5% are depicted in bold.

Gene expression in whole blood: offspring of nonagenarians vs controls

To investigate whether the differences in expression of the seven genes differentially expressed between nonagenarians and controls are a characteristic of the long-lived families and not just a marker of old age, we compared their expression in the middle-aged offspring of nonagenarians to the controls of similar age (Supplementary Table S5). While the association for six genes was not significant after correction for multiple testing, as compared to the controls, the offspring expressed significantly less *RPTOR* mRNA (Table 3). Thus in two generations the *RPTOR* gene expression was decreased as compared to controls. Since association of *PRR5L* was borderline significant and displayed the largest effect size (FC = 0.74), we also included *PRR5L* for further follow up in this paper. Gene expression levels of *PRR5L* were increased in the nonagenarians, but decreased in their offspring. The gene expression differences between offspring and control groups may be due to variation in gene expression

level indeed or could (partially) be explained by differences in between the groups in blood cell subsets or in disease prevalence.

Table 3. Linear regression results of gene expression of offspring compared to controls.

Protein	Genename	N cases	N controls	Mean	SD	FC	p
4EBP1	*EIF4EBP1*	320	303	2.03	0.295	1.03	0.311
FOXO1	*FOXO1*	311	296	2.18	0.238	1.00	0.878
p14	*LAMTOR2*	311	296	1.93	0.397	0.99	0.852
PRAS40	*AKT1S1*	311	296	1.67	0.278	0.96	0.282
Protor2	*PRR5L*	311	296	5.59	1.143	0.74	0.041
Raptor	***RPTOR***	**332**	**312**	**0.99**	**0.184**	**0.92**	**0.001**
RhoA	*RHOA*	311	296	2.15	0.265	1.00	0.927

Mean: mean relative expression level. SD: standard deviation of relative expression level. FC: fold change, above one indicated higher expression in offspring of nonagenarians. p: raw p value from linear regression model. Significance level for p value after Bonferroni correction for multiple testing is 0.00714. Genes significantly differentially expressed with at least 5% are depicted in bold.

Relation of gene expression in whole blood with cell subtypes

Whole blood consists of several cell subtypes and subsets of differentiated immune cells that may have different mRNA expression profiles. To investigate whether the proportions of different blood cell types could explain differential expression levels of *RPTOR* and *PRRL5*, the comparisons between offspring and controls were corrected for the relative counts of leukocytes, thrombocytes, neutrophils, lymphocytes, monocytes, basophils or eosinophils present in the whole blood samples (see Experimental procedures). We observed that variation between cell counts in offspring and controls did not influence the associations of the gene expression with familial longevity.

Recently an increasingly important role for mTOR in directing T cell activation and differentiation has become apparent (25). We therefore examined whether the difference in expression levels of *RPTOR* and *PRR5L* between offspring and partners is influenced by the level of T cell differentiation reflected in the distribution of different T cell phenotypes. To this end, the frequency of naïve (CD45RA+CCR7+CD27+CD28+), central memory (CD45RA-CCR7+CD27+CD28+), effector memory (CD45RA-CCR7-CD27-CD28-) and late-stage differentiated (CD45RA+CCR7-CD27-CD28-) T cells was assessed in 71 offspring and 73 controls. The largest effect was seen when correcting for the relative amount of effector memory T cells present in this sample; it reduced the expression difference between

offspring and controls for *PRR5L* by about 60% (the beta changed from -0.44 to -0.17, Supplementary Table S4). This indicated that differences in the proportion of effector memory T cells in these groups explained about half of the differential gene expression effect for *PRR5L*. No remarkable changes in effect size were found for *RPTOR* when correcting for the above-mentioned T cell differentiation markers (Supplementary Table S4).

Association of gene expression in whole blood with disease and glucose phenotypes

Since in literature expression levels of mTOR pathway proteins have been implicated in the pathogenesis of type 2 diabetes (9) and progression of cancer (1;10) and since the prevalence of diabetes differed between LLS offspring and controls (20)) we tested whether *RPTOR* and *PRR5L* expression were simply marking differences in disease prevalence. We therefore performed the same analysis between offspring and controls without known diabetics and cancer patients, which yielded similar associations to those described above (Supplementary Table S6). Because glucose levels differed between offspring and controls (Supplementary Table S5), we investigated whether the mRNA differences of *RPTOR* and *PRR5L* depended on glucose levels in non-diabetic participants. The association of familial longevity with *RPTOR* was not affected by adjustment for glucose levels and the association with *PRR5L* expression gained significance (Supplementary Table S7). Thus, the expression level of the *RPTOR* gene in blood mark familial longevity independent of the prevalence of type 2 diabetes and cancer and glucose levels. Due to the influence of T cell differentiation, glucose levels and diabetes prevalence on its gene expression, the association of *PRRL5* with familial longevity is more complex.

Discussion

By comparing mRNA levels in blood of nonagenarians and middle-aged controls from the Leiden Longevity Study, we found that the mTOR signalling gene is differentially expressed in old age. Single gene analysis showed that seven out of 40 mTOR pathway genes had a significant differential expression of at least 5%, up to 1.81-fold. The expression of the genes *EIF4EBP1*, *LAMTOR2*, *AKT1S1*, *PRR5L* and *RHOA* were higher in nonagenarians, whereas *FOXO1* and *RPTOR* were lower. The directions of these differential gene expressions indicate separate effects on mTORC1 and mTORC2 complexes. Two of the seven differentially expressed genes, *RPTOR* and *PRR5L,* were expressed to a lower level in middle-aged

members of the longevity families as compared to similarly aged controls. The significant differential expression of *RPTOR* and suggestive differential *PRR5L* expression indicate that expression of these genes may not just mark old age but familial longevity *per se*.

We also compared nonagenarians and controls in a genetic approach. Variation in the mTOR gene set associated with longevity in the LLS due to the joint contribution of small effects in several genes. This association could not pinpoint the main effector genes or SNPs, so we conclude that this finding must be replicated in much larger studies to test whether genetic variation in the pathway and specific genes truly associate with familial longevity.

Decreased mTORC1 protein activity is associated with lower glucose levels and a lower prevalence of diabetes in humans (14;28), while impaired mTORC2 is associated with glucose intolerance and insulin resistance (29). The offspring of the nonagenarian siblings in the Leiden Longevity Study have a lower prevalence of diabetes and lower glucose levels than similarly aged controls. Since adjustment for glucose levels and excluding diabetic patients provides similar associations, differences in the prevalence of type 2 diabetes or plasma glucose levels between offspring and controls did not explain the association of *RPTOR* and *PRR5L* mRNA level with familial longevity.

Our results could be explained if *RPTOR* gene expression in blood marks familial longevity and metabolic health in middle-age and would thus be an early marker of familial longevity. The altered gene expression levels in the longevity family members may be the consequence of ageing, or merely a trait shared by the long-lived families or may truly contribute to human longevity. In view of the findings in animal models (4-7), it is not unlikely that a familial and genetic background contributing to low mTORC1 signalling activity (and preservation of low mTOR throughout life) contributes to metabolic health and an extended life expectancy across species.

In our study we found *RPTOR*, *AKT1S1* and *EIF4EBP1* to be associated with old age and/or familial longevity, and the directions of association indicate a transcriptional down-regulation of mTORC1. Even in middle-age a down-regulation of *RPTOR* mRNA was associated with familial longevity, providing another example of the highly conserved role mTOR signalling has in lifespan regulation across different species.

Not only mTORC1, but also mTORC2 activity has been linked to health and lifespan. Rictor, an important component of mTORC2, contributes to glucose homeostasis in murine muscle tissue (30), TORC2 *C. elegans* mutants show modulation of lifespan by affecting feeding and metabolism of different diets (31), TORC2 is a mediator of proliferative and survival signals in cancer cells (32) and a study in mouse embryonic fibroblasts showed that mTORC2 regulates the TLR-mediated inflammatory response via FoxO1 (33). In addition, *PRR5L*, which is part of mTORC2, has been suggested to play a role in apoptosis in HeLa cells, but whether it is pro-apoptotic or anti-apoptotic remains unclear (34). Recently, Lamming et al. showed that disruption of mTORC2 can result in glucose intolerance and insulin resistance in rodents and may be relevant in the pathogenesis of type 2 diabetes and metabolic syndrome (29). Our results regarding the association of *PRR5L*, *RHOA* and *FOXO1* with old age and/or familial longevity suggest a transcriptional up-regulation of mTORC2 which would be in concordance with these earlier findings in animal models. As we provide the first evidence that mTORC2 is beneficial for healthy ageing in humans, further research is required to replicate these findings.

We observed that *PRR5L* expression was higher in nonagenarians, but lower in their offspring as compared to the controls. If expression of a gene is associated with familial longevity in this population, we would expect to find a similar direction in both long-lived nonagenarians as well as their offspring. Late-differentiated memory T cells accumulate with age and this hallmark feature of immunosenescence is believed to contribute to the weakened immune status in the elderly. It has also been shown that offspring of the long-lived nonagenarians have less accumulation of these effector memory T cells in comparison to controls (25). Therefore, varying proportions of T cell subsets might influence observed *PRR5L* expression levels in the different groups of the LLS. When testing for association with cell subtype measurements in a subset of middle-aged LLS participants, we found that *PRR5L* expression is positively associated with the amount of CCR7-CD45RA-CD27-CD28- late-differentiated effector memory cells. Data on T cell subset counts are not available for the LLS nonagenarians, but Koch and colleagues showed the same subset of effector memory cells to be increased in the elderly (35). This may explain the higher *PRR5L* expression levels we found in the nonagenarians while lower amounts of effector memory cells (25) and also lower *PRR5L* expression was found in the middle-aged offspring compared to controls. From this we conclude that *PRR5L* expression may be a marker for the differentiation state of T cells, indicating a more late-differentiated T cell compartment in nonagenarians as compared to younger controls, but a less-differentiated phenotype in the

offspring in middle-age. The link between the mTOR pathway and T cell activation and differentiation is perhaps not surprising when it is considered that rapamycin was first identified and exploited as a T cell-directed immunosuppressive agent in transplantation (36). The differential distribution of T cell phenotypes contributed approximately to 60% of the *PRR5L* effect, suggesting that *PRR5L* might have another role in longevity other than marking/influencing T cell population subtypes; further research is required to identify such a role.

In conclusion, the expression level in blood of genes belonging to the two mTOR complexes 1 and 2 associated with old age showing down- and up-regulation in the long-lived respectively. *RPTOR* gene expression was further significantly associated with familial longevity in middle-age independent of glucose levels and the prevalence of type 2 diabetes and cancer. At the level of suggestive evidence this was also the case for *PRRL5*. These results suggest that the mTOR pathway may not only be involved in lifespan regulation in animal models, but also with the metabolic health and extended lifespan of human longevity families.

Materials & Methods

Study population

The individuals investigated in this study are participants of the Leiden Longevity Study. The families participating in this study have at least two siblings with a minimum age for men of 89 years and for women of 91 years (19). The offspring of these long-lived individuals, who have an increased potential to become long-lived individuals (30% reduced standardized mortality rate), were also included. In addition, the partners of the offspring were included as population controls of similar age and environmental exposures as the offspring, and as a young control group for the nonagenarian siblings. Blood samples were taken from all the participants. The Leiden Longevity Study was approved by the Medical Ethical Committee of Leiden University Medical Centre and all participants gave written informed consent.

Selection of mTOR pathway

For the mTOR pathway, we selected 40 genes encoding proteins that belong to the well-described core of the mammalian mTOR pathway. First we selected the two mTOR complexes mTORC1 and mTORC2 including their downstream targets often described in literature (2;3). Next, we

included the direct activators and inhibitors of these two complexes. Figure 1 shows the mTOR pathway as we investigated.

Sample collection and RNA preparation

87 non-related long-lived siblings, 337 offspring and 321 partners belonging to 281 nuclear families were selected for the current study (Table 1 and Supplementary Table S5). These samples were randomly selected, but in such a way that age and gender were balanced between the groups and age range was as large as possible. Additionally, individuals with outlying cell counts (beyond 3 SDs below or above the standard error of the mean) were excluded. This subpopulation is representative for the whole LLS regarding disease prevalence and parameters involved in metabolic syndrome (Supplementary Table S5) (20;37). From these non-fasted individuals, peripheral blood was harvested using PAXgeneTM tubes (Qiagen, Venlo, The Netherlands). The tubes were frozen and kept at -20°C for ~3-5 years. After thawing at room temperature for at least 2 hours, total RNA was extracted from the approximately 2.5 ml of peripheral blood in each tube following the manufacturer's recommended protocol (PAXgene Blood RNA Kit Handbook, Qiagen, Venlo, The Netherlands). The quality and integrity of the total RNA was evaluated on the 2100 Bioanalyzer (Agilent Technologies, Amstelveen, The Netherlands) and the concentration was measured using a NanoDrop spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA). Quality criteria included a 28S/18S ratio as measured by the Bioanalyzer of at least 1.2, and a total RNA yield of at least 3 µg.

RT-qPCR

For all 40 mTOR signalling genes the suggested Taqman® assay (Applied Biosystems, Bleiswijk, the Netherlands) was selected. Reverse transcription was performed with total RNA from blood of in total 790 samples, which passed QC using the First Strand cDNA Synthesis Kit according to the manufacturer's protocol (Roche Applied Science, Almere, the Netherlands). cDNA was amplified using the DNA Engine Tetrad® 2 Peltier Thermal Cycler (Bio-Rad, Veenendaal, the Netherlands). qPCR was then performed with the Taqman® method using the BiomarkTM 48.48 and 96.96 Dynamic Arrays (Fluidigm, Amsterdam, the Netherlands). Relative gene expression of the BioMark™ Array data were calculated by using the 2-ΔΔCt method, in which Ct indicates cycle threshold, the fractional cycle number where the fluorescent signal reaches the detection threshold (38). YKT6 was used as internal control and commercially available human total

reference RNA (Clontech Laboratories, Mountain View, CA, USA) as reference RNA.

Geneset analysis of gene expression data

The Globaltest methodology was designed to determine whether the common expression pattern of genes within a pre-defined set is significantly related to clinical outcome (39;40). A generalized linear model is used to estimate a "Q-statistic" for each gene set, which describes the correlation between gene expression profiles, X, and clinical outcomes, Y. The Q-statistic for a gene set is the average of the Q-statistics for each gene in the set. The Globaltest method was used to perform geneset analysis comparing two groups of individuals (either nonagenarians vs. controls or offspring vs. controls) including age (in offspring vs. controls only) and gender and their interaction as covariates. Global test package in R (41) has been used to perform analyses.

Single gene analysis of gene expression data

Differences in expression level between long-lived siblings, their offspring and the partners of their offspring were assessed using linear regression. In these analyses, expression level was the dependent variable and the two groups of individuals (either nonagenarians vs. controls or offspring vs. controls) were included in the model as a categorical variable together with age (in offspring vs. controls only), gender and their interaction as covariates. To take into account dependencies within sibships, robust standard errors were used, i.e. the variance was computed from the between family variation. P-values were also based on these robust standard errors. Analyses were performed using the software package STATA/SE 11.0 (DPC Software, StataCorp 2009).

Sample and SNP collection for genetic association analysis

The SNP set analysis was performed in the same manner as described in Deelen et al. (42). We used the genotype data of 417 unrelated long-lived individuals and 476 young controls from the LLS, typed with Illumina HumanOmniExpress BeadChips (Supplementary Table S1). We removed SNPs with a SNP call rate < 0.95, MAF < 0.01 or $P_{HWE} < 10^{-4}$ in nonagenarian cases and controls and 1,018 analyzed SNPs within a 10-kb window around the 40 genes from the MTOR pathway (Figure 1) using the PLINK set-based test.

PLINK set-based statistical analysis

In the PLINK set-based test (–set-test, http://pngu.mgh.harvard.edu/purcell/plink; (43)), a SNP analysis of the pathway or gene SNP set is performed. For the SNP set, a mean SNP statistic is calculated from the single SNP statistics of a maximum amount (–setmax) of independent SNPs below a certain P value threshold (–set-p). If SNPs are not independent, i.e., in case linkage disequilibrium (r^2) is above a certain threshold (–set-r^2), the SNP with the lowest P value in the single SNP analysis is selected. The same analysis is performed with a certain amount (–mperm) of simulated SNP sets in which the phenotype status of the individuals is permuted. An empirical P value for the SNP set is computed by calculating the number of times the test statistic of the simulated SNP sets exceeds that of the original SNP set. For the analysis in this study, the parameters were set to –setp 0.05 –set-r^2 0.5, –set-max 99999, and –mperm 10,000. Bonferroni correction for the pathway analysis is not necessary since it contains just one test. When SNP sets were tested per gene (40 SNP sets), the significance level for p values after Bonferroni correction for multiple testing is 0.00125.

Blood cell subtypes

To further investigate the candidate genes, their expression level was again tested for association with familial longevity, but now adjusted for several blood cell counts. In the whole blood samples of the participants the following cell subtypes were counted using the automated Siemens ADVIA 1200 system (SMSD, Tarrytown, NY) in the Leiden Medical Diagnostical Center: leukocytes, trombocytes, neutrophils, lymphocytes, monocytes, basophils and eosinophils. Both longevity-associated genes in middle-age were adjusted for each of these cell counts separately when associating with familial longevity, which analysis is described above.

T cell differentiation measurements

Data on distribution of different T cell subsets were generated in 71 offspring and 73 controls, of whom also gene expression data were available, as described previously (25). Briefly, PBMCs viably cryopreserved with DMSO were thawed and treated with human Ig, GAMUNEX (Bayer, Leverkusen, Germany), and ethidium monoazide (EMA) (Invitrogen, Karlsruhe, Germany). Cells were first stained indirectly for KLRG-1 using a primary antibody kindly provided by Prof. Hans-Peter Pircher, Freiburg, Germany and a Pacific-Orange-conjugated secondary

antibody (Invitrogen). After blocking with mouse serum (Chemicon/Millipore, Schwalbach, Germany), the following directly conjugated monoclonal antibodies were added: CD3-PE (Calltag; Invitrogen), CD4- PerCP, CD8-allophycocyanin-Cy7, CCR7-PE-Cy7 (BD Biosciences, Heidelberg, Germany), CD27-allophycocyanin, CD45RA-Pacific Blue, CD28- Alexa Fluor 700, (BioLegend, San Diego, CA), and CD57-FITC (Immunotools, Freiburg, Germany). After 20 min incubation on ice, cells were washed and analyzed immediately on an LSR II cytometer with FACSDiva software (BD Biosciences). The spectral overlap between all channels was calculated automatically by the BD FACSDiva software, after measuring negative and single-color controls. Data were analyzed using FlowJo software (Tree Star, Portland, OR). For data analysis, EMA-positive dead cells were excluded. T cells were characterized as naïve (CD45RA+CCR7+CD27+CD28+), central memory (CD45RA-CCR7+CD27+CD28+), effector memory (CD45RA-CCR7-CD27-CD28-) and “terminally differentiated” effector memory (TEMRA; CD45RA+CCR7-CD27-CD28-) T cells according to previously published models (35).

Associatin of gene expression with insulin-related phenotypes

Since the mTOR pathway has been implicated in insulin-related phenotypes, their effect on the expression of the longevity-related genes was further investigated. Blood samples were taken at baseline for the determination of non-fasted serum parameters. For the serum measurements, the Hitachi Modular or the Cobas Integra 800, both from Roche, Almere, the Netherlands were applied. CVs of these measurements were all below 5 %. Information on medical history was requested from the participants' general practitioners. As described before, differences in expression level between offspring and controls were assessed using linear regression using the same model, but excluding known Type 2 diabetes patients. The association of glucose and insulin levels with expression levels of the 7 longevity-associated genes was also analyzed using linear regression in controls only. To investigate the influence of glucose and insulin levels on the association with longevity of these 7 genes, linear regression between offspring and controls was performed again, including one of these parameters as covariate.

Acknowledgements

We thank all participants of the Leiden Longevity Study. The research leading to these results has received funding from the European Union's Seventh Framework Programme (FP7/2007-2011) under grant agreement

n° 259679. This study was supported by a grant from the Innovation-Oriented Research Program on Genomics (SenterNovem IGE05007), the Centre for Medical Systems Biology, 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. David A. Gunn is employed by Unilever PLC and was involved in the design, data collection and analysis as well as the decision to publish.

References

(1) Stanfel MN, Shamieh LS, Kaeberlein M, Kennedy BK. The TOR pathway comes of age. Biochim Biophys Acta 2009 Oct;1790(10):1067-74.

(2) Polak P, Hall MN. mTOR and the control of whole body metabolism. Curr Opin Cell Biol 2009 Apr;21(2):209-18.

(3) Laplante M, Sabatini DM. mTOR signaling at a glance. J Cell Sci 2009 Oct 15;122(Pt 20):3589-94.

(4) Kaeberlein M, Powers RW, III, Steffen KK, Westman EA, Hu D, Dang N, et al. Regulation of yeast replicative life span by TOR and Sch9 in response to nutrients. Science 2005 Nov 18;310(5751):1193-6.

(5) Vellai T, Takacs-Vellai K, Zhang Y, Kovacs AL, Orosz L, Muller F. Genetics: influence of TOR kinase on lifespan in C. elegans. Nature 2003 Dec 11;426(6967):620.

(6) Kapahi P, Zid BM, Harper T, Koslover D, Sapin V, Benzer S. Regulation of lifespan in Drosophila by modulation of genes in the TOR signaling pathway. Curr Biol 2004 May 25;14(10):885-90.

(7) Harrison DE, Strong R, Sharp ZD, Nelson JF, Astle CM, Flurkey K, et al. Rapamycin fed late in life extends lifespan in genetically heterogeneous mice. Nature 2009 Jul 16;460(7253):392-5.

(8) Spindler SR. Rapid and reversible induction of the longevity, anticancer and genomic effects of caloric restriction. Mech Ageing Dev 2005 Sep;126(9):960-6.

(9) Newgard CB, An J, Bain JR, Muehlbauer MJ, Stevens RD, Lien LF, et al. A branched-chain amino acid-related metabolic signature that differentiates obese and lean humans and contributes to insulin resistance. Cell Metab 2009 Apr;9(4):311-26.

(10) Guertin DA, Sabatini DM. An expanding role for mTOR in cancer. Trends Mol Med 2005 Aug;11(8):353-61.

(11) Le TC, Faivre S, Serova M, Raymond E. mTORC1 inhibitors: is temsirolimus in renal cancer telling us how they really work? Br J Cancer 2008 Oct 21;99(8):1197-203.

(12) Yap TA, Garrett MD, Walton MI, Raynaud F, de Bono JS, Workman P. Targeting the PI3K-AKT-mTOR pathway: progress, pitfalls, and promises. Curr Opin Pharmacol 2008 Aug;8(4):393-412.

(13) Strimpakos A, Saif MW, Syrigos KN. Pancreatic cancer: from molecular pathogenesis to targeted therapy. Cancer Metastasis Rev 2008 Sep;27(3):495-522.

(14) Krebs M, Brunmair B, Brehm A, Artwohl M, Szendroedi J, Nowotny P, et al. The Mammalian target of rapamycin pathway regulates nutrient-sensitive glucose uptake in man. Diabetes 2007 Jun;56(6):1600-7.

(15) Dowling RJ, Zakikhani M, Fantus IG, Pollak M, Sonenberg N. Metformin inhibits mammalian target of rapamycin-dependent translation initiation in breast cancer cells. Cancer Res 2007 Nov 15;67(22):10804-12.

(16) Holloszy JO, Fontana L. Caloric restriction in humans. Exp Gerontol 2007 Aug;42(8):709-12.

(17) Fontana L, Meyer TE, Klein S, Holloszy JO. Long-term calorie restriction is highly effective in reducing the risk for atherosclerosis in humans. Proc Natl Acad Sci U S A 2004 Apr 27;101(17):6659-63.

(18) Passtoors WM, Boer JM, Goeman JJ, van den Akker EB, Deelen J, Zwaan BJ, et al. Transcriptional Profiling of Human Familial Longevity Indicates a Role for ASF1A and IL7R. PLoS One 2012;7(1):e27759.

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

(20) Westendorp RG, van HD, Rozing MP, Frolich M, Mooijaart SP, Blauw GJ, et al. Nonagenarian siblings and their offspring display lower risk of mortality and morbidity than sporadic nonagenarians: The Leiden Longevity Study. J Am Geriatr Soc 2009 Sep;57(9):1634-7.

(21) Heijmans BT, Beekman M, Houwing-Duistermaat JJ, Cobain MR, Powell J, Blauw GJ, et al. Lipoprotein particle profiles mark familial and sporadic human longevity. PLoS Med 2006 Dec;3(12):e495.

(22) Barzilai N, Gabriely I, Gabriely M, Iankowitz N, Sorkin JD. Offspring of centenarians have a favorable lipid profile. J Am Geriatr Soc 2001 Jan;49(1):76-9.

(23) Rozing MP, Westendorp RG, Frolich M, de Craen AJ, Beekman M, Heijmans BT, et al. Human insulin/IGF-1 and familial longevity at middle age. Aging (Albany NY) 2009 Aug;1(8):714-22.

(24) Wijsman CA, Rozing MP, Streefland TC, le CS, Mooijaart SP, Slagboom PE, et al. Familial longevity is marked by enhanced insulin sensitivity. Aging Cell 2011 Feb;10(1):114-21.

(25) Derhovanessian E, Maier AB, Beck R, Jahn G, Hahnel K, Slagboom PE, et al. Hallmark features of immunosenescence are absent in familial longevity. J Immunol 2010 Oct 15;185(8):4618-24.

(26) Globaltest, R package, version 4.10.0 [computer program]. 2008.

(27) Deelen J, Beekman M, Uh HW, Helmer Q, Kuningas M, Christiansen L, et al. Genome-wide association study identifies a single major locus contributing to survival into old age; the APOE locus revisited. Aging Cell 2011 Aug;10(4):686-98.

(28) Liu Z, Wu Y, Nicklas EW, Jahn LA, Price WJ, Barrett EJ. Unlike insulin, amino acids stimulate p70S6K but not GSK-3 or glycogen synthase in human skeletal muscle. Am J Physiol Endocrinol Metab 2004 Apr;286(4):E523-E528.

(29) Lamming DW, Ye L, Katajisto P, Goncalves MD, Saitoh M, Stevens DM, et al. Rapamycin-induced insulin resistance is mediated by mTORC2 loss and uncoupled from longevity. Science 2012 Mar 30;335(6076):1638-43.

(30) Kumar A, Harris TE, Keller SR, Choi KM, Magnuson MA, Lawrence JC, Jr. Muscle-specific deletion of rictor impairs insulin-stimulated glucose transport and enhances Basal glycogen synthase activity. Mol Cell Biol 2008 Jan;28(1):61-70.

(31) Soukas AA, Kane EA, Carr CE, Melo JA, Ruvkun G. Rictor/TORC2 regulates fat metabolism, feeding, growth, and life span in Caenorhabditis elegans. Genes Dev 2009 Feb 15;23(4):496-511.

(32) Fang Z, Zhang T, Dizeyi N, Chen S, Wang H, Swanson KD, et al. Androgen Receptor Enhances p27 Degradation in Prostate Cancer Cells through Rapid and Selective TORC2 Activation. J Biol Chem 2012 Jan 13;287(3):2090-8.

(33) Brown J, Wang H, Suttles J, Graves DT, Martin M. Mammalian Target of Rapamycin Complex 2 (mTORC2) Negatively Regulates Toll-like Receptor 4-mediated Inflammatory Response via FoxO1. J Biol Chem 2011 Dec 30;286(52):44295-305.

(34) Thedieck K, Polak P, Kim ML, Molle KD, Cohen A, Jeno P, et al. PRAS40 and PRR5-like protein are new mTOR interactors that regulate apoptosis. PLoS One 2007;2(11):e1217.

(35) Koch S, Larbi A, Derhovanessian E, Ozcelik D, Naumova E, Pawelec G. Multiparameter flow cytometric analysis of CD4 and CD8 T cell subsets in young and old people. Immun Ageing 2008;5:6.

(36) Morath C, Arns W, Schwenger V, Mehrabi A, Fonouni H, Schmidt J, et al. Sirolimus in renal transplantation. Nephrol Dial Transplant 2007 Sep;22 Suppl 8:viii61-viii65.

(37) Rozing MP, Westendorp RG, de Craen AJ, Frolich M, de Goeij MC, Heijmans BT, et al. 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 2010 Mar;58(3):564-9.

(38) Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 2001 Dec;25(4):402-8.

(39) Goeman JJ, van de Geer SA, de Kort F, van Houwelingen HC. A global test for groups of genes: testing association with a clinical outcome. Bioinformatics 2004 Jan 1;20(1):93-9.

(40) Goeman JJ, Oosting J, Cleton-Jansen AM, Anninga JK, van Houwelingen HC. Testing association of a pathway with survival using gene expression data. Bioinformatics 2005 May 1;21(9):1950-7.

(41) Globaltest R package, version 5.6.1 [computer program]. 2011.

(42) Deelen J, Uh HW, Monajemi R, van HD, Thijssen PE, Bohringer S, et al. Gene set analysis of GWAS data for human longevity highlights the relevance of the insulin/IGF-1 signaling and telomere maintenance pathways. Age (Dordr) 2011 Nov 24.

(43) Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007 Sep;81(3):559-75.

(44) Sengupta S, Peterson TR, Sabatini DM. Regulation of the mTOR complex 1 pathway by nutrients, growth factors, and stress. Mol Cell 2010 Oct 22;40(2):310-22.

(45) Sarbassov DD, Guertin DA, Ali SM, Sabatini DM. Phosphorylation and regulation of Akt/PKB by the rictor-mTOR complex. Science 2005 Feb 18;307(5712):1098-101.

(46) Ikenoue T, Inoki K, Yang Q, Zhou X, Guan KL. Essential function of TORC2 in PKC and Akt turn motif phosphorylation, maturation and signalling. EMBO J 2008 Jul 23;27(14):1919-31.

(47) Garcia-Martinez JM, Alessi DR. mTOR complex 2 (mTORC2) controls hydrophobic motif phosphorylation and activation of serum- and glucocorticoid-induced protein kinase 1 (SGK1). Biochem J 2008 Dec 15;416(3):375-85.

Supplementary information

Supplementary figure and tables can be found online:
http://onlinelibrary.wiley.com/doi/10.1111/acel.12015/suppinfo

Table S1. *Analyzed samples: genetic variation.*

Table S2. *Measured SNPs in the mTOR pathway.*

Table S3. *Results of gene set analysis of the mTOR pathway gene SNP sets.*

Table S4. *Linear regression results of gene expression of offspring compared to controls, including adjustments for T cell differentiation markers.*

Table S5. *Descriptives of prevalence of disease and metabolic syndrome related parameters in offspring and controls (n = 658).*

Table S6. *Linear regression results of gene expression of offspring compared to controls, excluding diabetics and cancer patients.*

Table S7. *Linear regression results of gene expression of offspring compared to controls (without diabetics), including adjustments for glucose plasma levels.*

Chapter 6

The association of mTOR and familial longevity in the blood transcriptome is not reflected in primary skin fibroblasts

W.M. Passtoors[1], B. Guigas[2,3], P. Dekker[4], A.B. Maier[5], G.C.M. van der Zon[2], R. van der Breggen[1], G.J.B. van Ommen[6,7], P.E. Slagboom[1,8], M. Beekman[1]

1 Section of Molecular Epidemiology, Leiden University Medical Center, Leiden, the Netherlands
2 Department of Molecular Cell Biology, Leiden University Medical Center, Leiden, The Netherlands
3 Department of Parasitology, Leiden University Medical Center, Leiden, The Netherlands
4 Department of Gerontology and Geriatrics, Leiden University Medical Center, Leiden, The Netherlands
5 Section of Gerontology and Geriatrics, Department of Internal Medicine, VU Medical Center, Amsterdam, the Netherlands
6 Center for Human and Clinical Genetics, Leiden University Medical Center, Leiden, The Netherlands
7 Center for Medical Systems Biology, the Netherlands
8 Netherlands Consortium for Healthy Ageing, the Netherlands

Manuscript in preparation

Abstract

The Target of Rapamycin (TOR) pathway is an evolutionary conserved signaling pathway involved in lifespan regulation. In a gene expression study of whole blood, we recently reported that the mammalian TOR pathway was associated with familial longevity in humans. The aim of the present study was to investigate whether such an association can also be found at the cellular level. For this purpose, primary fibroblasts from a subset of our initial discovery cohort consisting of 40 middle aged members of long-lived families and 40 controls of similar age were analyzed. The protein expression, phosphorylation state, and gene expression of key molecules within the mTOR pathway as well as total protein content and markers for autophagy were measured. Although a significant effect on total protein content was observed, that might reflect differences in cell growth and proliferation, no differences were observed on the other parameters. The variation in gene expression levels of mTOR genes in cultured fibroblasts did not reflect at all the variation in these levels in whole blood from the same individuals. We conclude that contribution of the mTOR pathway to familial longevity cannot be studied in fibroblasts of donors which do express marks of such contribution in blood.

Introduction

After extensive studies in animal models, mammalian target of rapamycin (mTOR) signaling has recently also been implicated in human ageing and familial longevity (1;2). Gene expression studies in blood showed that expression changes of downstream targets of mTOR complex 1 (mTORC1) are downregulated with aging (3;4). It has been suggested that, in line with what is known in animal models, inhibition of mTOR signaling with advancing age might be a survivor effect, since decreased mTOR signaling is associated with increased survival (3). Our study investigating expression levels of genes belonging to the mTOR pathway showed that familial longevity in middle age associated with downregulation of the expression of genes in the mTOR pathway (4). These results correspond with previous studies showing that increased mTOR signaling is associated with increased prevalence and progression of disease (5;6), and that reduced mTOR signaling extends lifespan in model organisms (7-10) and is associated with better metabolic health (11). Thus, taken together, accumulating evidence suggests that mTOR signaling can play an important role in human metabolic health and longevity.

mTOR is an evolutionarily conserved nutrient-sensing protein kinase (6). The core kinase is present in two distinct multiprotein complexes (mTORC1 and mTORC2), each with different functions (12), although their exact respective roles remain to be elucidated. The main known functions of mTORC1 include translation initiation, protein synthesis and autophagy, while those of mTORC2 include cytoskeletal organization and stress resistance (13). In addition, the two complexes are connected in that mTORC2 regulates Akt phosphorylation, part of an upstream pathway that controls activation of mTORC1 (6). The mTOR pathway has been extensively studied in cell lines, including human and mouse fibroblasts (14-18).

The mTOR pathway was found to be involved in human familial longevity in the Leiden Longevity Study (LLS) in which nonagenarian siblings, their offspring and the partners thereof were studied. Expression levels of mTOR pathway related genes in blood differed between nonagenarian and middle aged individuals and between middle aged offspring and similarly aged partners as controls (4). The middle aged members of the long lived families (offspring) have a decreased prevalence of myocardial infarction and type 2 diabetes (19), and display a healthier metabolism reflected by beneficial lipid profiles (20;21) and preservation of insulin sensitivity (22;23) than the controls. Previously we demonstrated differences in the *in vitro* cellular responses to stress of the offspring's fibroblasts as compared to controls, including a lower increase in senescence, but an increase in apoptosis, which is seen as advantageous to the organism (24;25). To what extent mTOR signaling contributes to human familial longevity is not yet investigated at the protein and signaling level.

In this paper primary skin fibroblasts of members of long-lived families are investigated *in vitro* for differences in mTOR signaling as compared to similarly aged controls. The expression level of the most important proteins in the mTOR pathway were determined, as well as the phosphorylation of these proteins, total protein content and autophagy. We also inspected the correlation between the transcription levels of mTOR related genes in cultured skin fibroblasts and in blood from the same donors.

Results

The characteristics of the randomly selected fibroblast strains subset of subjects from the Leiden Longevity Study are shown in Table 1. Offspring of long-lived individuals and their partners, which will act as controls, were of similar age. Experiments were performed in fibroblasts obtained from 37

offspring and 36 controls, with a similar proportion of male and female donors.

Table 1 Analyzed samples in this study

Samples	Number	Mean Age	Age Range	Women (%)
Offspring	37	63.0	45-78	20 (54.1%)
Controls	36	62.5	44-78	19 (52.8%)

Protein expression and phosphorylation in skin fibroblasts

To investigate whether gene expression differences in blood represent lower mTOR signaling in peripheral tissues like the skin, we measured protein expression and phosphorylation of the most important proteins in this pathway (Figure 1) in skin fibroblasts using; mTOR, P70S6K, 4EBP1, Raptor, Rictor and Akt. mTOR itself is part of both mTORC1 and mTORC2 with complex-specific proteins Raptor and Rictor respectively. P70S6K and 4EBP1 are downstream targets of mTORC1 and Akt is a downstream target of mTORC2 as well as an upstream effector of mTORC1.

The comparison of offspring of long-lived individuals and controls for association with familial longevity showed no significant difference in expression or phosphorylation of any of the proteins after multiple testing correction (Bonferroni p=0.005, Table 2).

Protein synthesis and autophagy

Biological processes regulated by the mTOR pathway include autophagy and protein synthesis. Lower mTOR signaling, which is hypothesized to be the case in familial longevity, results in higher autophagy levels. An increase in autophagy levels could be indicated by a decrease in Light Chain3-I-on-Light Chain3-II isoform (LC3-I-on-LC3-II) ratio. However, the members of long-lived families show the same ratio as controls.

Although total protein content also reflects cell growth and differentiation, we could speculate that decreased total protein levels might indicate a decreased protein synthesis or an increased protein turnover to which a decreased mTORC1 activity may have contributed. Furthermore, since a lower mTOR signaling reduces protein synthesis, we speculate that total protein yield from 10^6 cells, as a surrogate marker for protein synthesis, might be decreased in members of long-lived families as compared to controls. The skin fibroblasts of members of long-lived families indeed had a lower protein content than controls (Table 2, p=0.007).

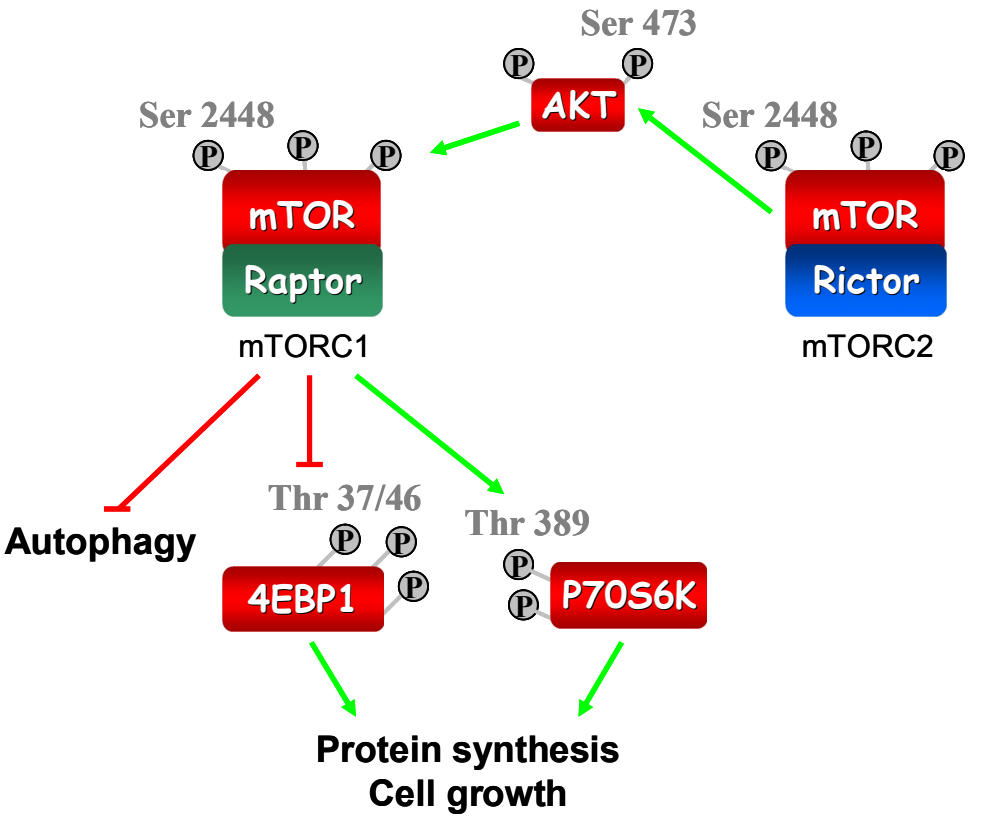

Figure 1. Core of mTOR signaling pathway of which protein expression and phosphorylation are investigated for their association with human longevity. Phosphorylation at Ser2448 activates mTOR in both complexes mTORC1 and mTORC2. mTORC1 inhibits 4EBP1 by phosphorylation at Thr37/46 and activates P70S6K by phosphorylating Thr389. mTORC2 activates Akt by phosphorylation at Ser473 and Akt indirectly activates mTORC1.

Gene expression and protein expression in skin fibroblasts

In blood we have previously established a difference between the offspring and controls in mTOR related gene expression levels (4). To investigate whether gene expression levels in skin fibroblasts correlate with protein levels we additionally measured transcription levels of *MTOR*, *RPTOR* and *RPS6KB1* in 66 fibroblast samples using RT-qPCR. Pearson pairwise correlation analysis of protein and gene expression levels in skin fibroblasts of the same individuals showed unexpectedly no significant correlation except for 4EBP1 ($r=0.324$, $p=0.008$ (Supplementary Table S1).

Gene expression in blood and gene expression in skin fibroblasts

The observed gene expression differences in blood between members of long-lived families and controls was apparently not reflected by any marked difference in mTOR signaling in the skin fibroblasts of these groups. The question raises therefore to what extent mTOR gene expression in blood and fibroblasts of the same donor correlates. We tested gene expression levels of the 40 mTOR genes in 72 individuals for such correlation by Pearson correlation analysis. We found no significant correlation of transcription levels for any of the genes in blood and skin fibroblasts of the same individuals (Supplementary Table S2).

Table 2. Linear regression results of protein expression in skin fibroblasts of offspring compared to controls

	Controls vs Offspring	
	Coef	p
Proteins		
mTOR	0.054	0.277
p-mTOR (S2448)	0.113	0.024
P70S6K	-0.003	0.938
p-P70S6K (T389)	0.003	0.967
p-4EBP1 (T37/46)	0.013	0.812
Raptor	0.039	0.504
Rictor	-0.038	0.606
p-Akt (S473)	0.117	0.270
Protein synthesis		
Total protein	-0.065	0.007
Autophagy		
LC3-I/LC3-II ratio	-0.005	0.962

Coef: coefficient from linear regression of protein expression (arbitrary units), positive indicated higher expression in offspring. p: raw p value from linear regression.

Discussion

The aim of this study was to further investigate our previous finding that expression levels of genes from the mTOR pathway in blood are associated with human familial longevity. In the skin fibroblasts of members of long-lived families and controls, however, we found no differences in the levels of proteins from the mTOR pathway nor their phosphorylation state. Hence we conclude that we are not able to detect an effect of mTOR signaling on human familial longevity in cultured skin fibroblasts. This is also reflected by the similarity between members of long-lived families and

controls in levels of autophagy as downstream effect of the mTOR pathway. The total protein content in members of long-lived families as compared to controls was slightly lower, potentially indicating downregulation of mTOR signaling. But since gene expression and protein levels of mTOR related genes in cultured skin fibroblasts and blood from the same donors did not correlate, we conclude that cultured skin fibroblasts may not be suitable to investigate the basis and consequences of natural variation of mTOR related gene expression in blood that marked familial longevity at a functional level.

Besides the biological explanation that mTOR signaling itself might differ between blood and skin of the same donor, our lack of clear validation might be due to the technical variations, of which the major one is that the skin fibroblasts were cultured several passages *in vitro* before RNA extraction, while blood was drawn and it's RNA immediately stabilized and frozen, representing the *in vivo* status. Thus, our current findings may stress that mTOR signaling in cultured primary skin fibroblasts is different from that in whole blood and therefore cannot provide sufficient insights about the role of mTOR signaling in human ageing and familial longevity.

Another explanation for the lack of clear validation might be that investigating natural occurring genetic variation in a primary cell system does not result in clearly measurable changes in protein expression and phosphorylation state. Most successful studies on mTOR signaling in skin fibroblasts investigate the effects of the knockdown of genes (14;18), the use of inhibiting compounds (17), mutant genes, overexpression of genes (14), amino acid deprivation or induction of oxidative stress (15). Because the effect of naturally occurring genetic variation is expected to be much smaller than the effect of such challenges, it may be that that the culturing of cells introduces too much background variation in mTOR signaling to detect the subtle differences between members of long-lived families and controls. Hence we may need to challenge the mTOR pathway to be able to detect the differences associated with human familial longevity in fibroblasts like was applied in the study of MTOR signaling in fibroblasts of long lived mouse strains (15).

In summary, the gene expression differences in mTOR related genes in blood associated with human familial longevity was not observed for primary skin fibroblast. Moreover, we were not able to detect differences in levels or phosphorylation status of mTOR proteins in skin fibroblast in relation to human familial longevity. We conclude that under unchallenged conditions cultured skin fibroblasts may not be suitable to investigate the

basis and consequences of natural variation in mTOR signaling in humans as determinant of human longevity. Potentially, cultured PBMCs better resemble the effects found in blood, or the exploration of differential response to a challenged mTOR pathway may provide functional insight in the association of mTOR related gene expression in blood with human familial longevity.

Materials and methods

Study design

The individuals investigated in this study are participants of the Leiden Longevity Study. The families participating in this study have at least two siblings with a minimum age for men of 89 years and for women of 91 years (28). The offspring of these long-lived individuals, who have an increased potential to become long-lived individuals (30% reduced standardized mortality rate), were also included. In addition, the partners of the offspring were included as population controls of similar age and environmental exposures as the offspring, and as a young control group for the nonagenarian siblings. Blood samples were taken from all the participants. During the period November 2006 and May 2008, a biobank was established from fibroblasts cultivated skin biopsies from 150 offspring-control pairs. Because it was expected that the difference in biological age between the offspring and control groups would be small, a relatively large sample size of 80 strains from 40 couples was chosen, consisting of 20 fibroblast strains from the 65 female offspring, 20 strains from the 86 male offspring, as well as 20 strains from the 86 female controls and 20 strains from the 63 male controls (Table 1).

Fibroblast cultures

Four mm skin biopsies were taken from the sun unexposed medial side of the upper arm. Fibroblasts were grown in D-MEM:F-12 (1:1) medium supplemented with 10% fetal calf serum (FCS), 1 mM MEM sodium pyruvate, 10 mM HEPES, 2 mM glutamax I and antibiotics (100 U mL^{-1} penicillin, 100 µg mL^{-1} streptomycin and 0.25-2.5 µg mL^{-1} amphotericin B), all obtained from Gibco (Breda, the Netherlands). One specific FCS batch was used during the whole experiment (Bodinco, Alkmaar, the Netherlands, batch no. 162229). Fibroblasts were incubated at 37 °C with 5% CO_2 and 100% humidity. All cultures that are used in the present study were grown under predefined, highly standardized conditions as published earlier (29) and frozen at low passage. Of the 73 individuals from which the fibroblasts

are tested, gene expression was measured in whole blood for 72 of them (4).

Experiments were set up in batches of maximally eight fibroblasts strains which were thawed from frozen stocks on day zero. On day one, the medium was changed and on day 4 fibroblasts were passaged 1:4 and passed further in equal numbers to have similar confluences for experiments. On day 25, fibroblasts were seeded in 9 cm dishes and grown for three days to similar confluences before harvesting cells for RNA or protein isolation.

Western Blot preparation and analysis

Fibroblast cells were washed with PBS, lysed in EBSB buffer (4 mM glycerol, 3% SDS, 100 mM TRIS-PO4 pH 6.8), boiled at 100 °C for 5 minutes and stored at -20 °C.

Whole cell lysates were prepared as described (30), and cleared by centrifugation. Protein content was determined using a bicinchoninic acid protein assay kit (Pierce, Rockford, IL). Expression and phosphorylation of proteins were determined by SDS-PAGE Western Blotting. Antibodies against total mTOR (#2972), P70S6K (#9202), LC3-I-on-LC3-II (#4108), and phospho-specific mTOR-Ser2448 (#2971S), P70S6K-Thr389 (#9205S), 4EBP1-Thr37/46 (#9459), and Akt-Ser473 (#9271) were from Cell Signalling Technology (Danvers, MA, USA). Antibodies against total Raptor (sc-27744) and Rictor (sc-50676) were from Santa Cruz Biotechnology (Santa Cruz, CA, USA). Bound antibodies were visualized by enhanced chemiluminescence and quantified by densitometry analysis of scanned films (Image J 1.34S; National Institutes of Health, Bethesda, MD, USA).

Statistical analysis was conducted with the software package STATA/SE 11.0 (DPC Software, StataCorp 2009). Differences in protein expression or phosphorylation (in arbitrary units) between the offspring and controls were determined with linear regression with age, gender and their interactions as covariates and adjusted for cell batch, cell confluence and where applicable western blot batch.

RNA sample preparation

From the 40 offspring and 40 controls RNA was extracted from both the cultured fibroblasts and whole blood. Harvesting of fibroblast cells included

washing with PBS and the cells were lysed in RLT buffer (Qiagen, Venlo, the Netherlands) and stored at -80 °C. RNA from fibroblasts was extracted using the RNeasy mini kit, RNA from whole blood (PAXgene tubes) was extracted using PAXgene Blood RNA kit (both Qiagen, Venlo, the Netherlands) following the manufacturer's recommended protocols. The quality and integrity of the total RNA was evaluated on the 2100 Bioanalyzer (Agilent Technologies, Amstelveen, The Netherlands) and the concentration was measured using a NanoDrop spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA).

RT-qPCR and correlation analysis of gene expression data

For all 40 mTOR signaling genes the suggested Taqman® assay (Applied Biosystems) was selected. Reverse transcription was performed with total RNA from fibroblasts and blood of in total 80 individuals, using the First Strand cDNA Synthesis Kit according to the manufacturer's protocol (Roche Applied Science). cDNA was amplified using the DNA Engine Tetrad® 2 Peltier Thermal Cycler (Bio-Rad). qPCR was then performed with the Taqman® method using the BiomarkTM 48.48 and 96.96 Dynamic Arrays (Fluidigm). Relative gene expression of the BioMark™ Array data were calculated by using the 2-ΔΔCt method, in which Ct indicates cycle threshold, the fractional cycle number where the fluorescent signal reaches the detection threshold (31). YKT6 was used as housekeeping gene and commercially available human total reference RNA (Clontech Laboratories, Mountain View, CA, USA) as reference RNA.

Gene expression and protein expression levels of 5 genes/proteins were correlated between the two tissues, skin fibroblasts and whole blood, within the same individuals using pairwise Pearson correlation in STATA adjusted for age, gender, their interaction, cell batch and cell confluence (Table 3). Relative gene expression levels of the 40 genes were also correlated between the two tissues within the same individuals using pairwise Pearson correlation in STATA adjusted for age, gender, their interaction, cell batch and cell confluence (Table 4).

References

(1) Laplante M, Sabatini DM. mTOR signaling in growth control and disease. Cell 2012 April 13;149(2):274-93.

(2) Remondini D, Salvioli S, Francesconi M, Pierini M, Mazzatti DJ, Powell JR et al. Complex patterns of gene expression in human T cells during in vivo aging. Mol Biosyst 2010 October 1;6(10):1983-92.

(3) Harries LW, Fellows AD, Pilling LC, Hernandez D, Singleton A, Bandinelli S et al. Advancing age is associated with gene expression changes resembling mTOR inhibition: evidence from two human populations. Mech Ageing Dev 2012 August;133(8):556-62.

(4) Passtoors WM, Beekman M, Deelen J, van der BR, Maier AB, Guigas B et al. Gene expression analysis of mTOR pathway: association with human longevity. Aging Cell 2012 October 13.

(5) Guertin DA, Sabatini DM. An expanding role for mTOR in cancer. Trends Mol Med 2005 August;11(8):353-61.

(6) Stanfel MN, Shamieh LS, Kaeberlein M, Kennedy BK. The TOR pathway comes of age. Biochim Biophys Acta 2009 October;1790(10):1067-74.

(7) Harrison DE, Strong R, Sharp ZD, Nelson JF, Astle CM, Flurkey K et al. Rapamycin fed late in life extends lifespan in genetically heterogeneous mice. Nature 2009 July 16;460(7253):392-5.

(8) Kaeberlein M, Powers RW, III, Steffen KK, Westman EA, Hu D, Dang N et al. Regulation of yeast replicative life span by TOR and Sch9 in response to nutrients. Science 2005 November 18;310(5751):1193-6.

(9) Kapahi P, Zid BM, Harper T, Koslover D, Sapin V, Benzer S. Regulation of lifespan in Drosophila by modulation of genes in the TOR signaling pathway. Curr Biol 2004 May 25;14(10):885-90.

(10) Vellai T, Takacs-Vellai K, Zhang Y, Kovacs AL, Orosz L, Muller F. Genetics: influence of TOR kinase on lifespan in C. elegans. Nature 2003 December 11;426(6967):620.

(11) Krebs M, Brunmair B, Brehm A, Artwohl M, Szendroedi J, Nowotny P et al. The Mammalian target of rapamycin pathway regulates nutrient-sensitive glucose uptake in man. Diabetes 2007 June;56(6):1600-7.

(12) Polak P, Hall MN. mTOR and the control of whole body metabolism. Curr Opin Cell Biol 2009 April;21(2):209-18.

(13) Laplante M, Sabatini DM. mTOR signaling at a glance. J Cell Sci 2009 October 15;122(Pt 20):3589-94.

(14) Astle MV, Hannan KM, Ng PY, Lee RS, George AJ, Hsu AK et al. AKT induces senescence in human cells via mTORC1 and p53 in the absence of DNA damage: implications for targeting mTOR during malignancy. Oncogene 2012 April 12;31(15):1949-62.

(15) Wang M, Miller RA. Fibroblasts from long-lived mutant mice exhibit increased autophagy and lower TOR activity after nutrient deprivation or oxidative stress. Aging Cell 2012 August;11(4):668-74.

(16) Jozwiak J, Kotulska K, Lojek M, Galus R, Jozwiak S, Polnik D et al. Fibroblasts from normal skin of a tuberous sclerosis patient show upregulation of mTOR pathway. Am J Dermatopathol 2009 February;31(1):68-70.

(17) Liu M, Liu F. Resveratrol inhibits mTOR signaling by targeting DEPTOR. Commun Integr Biol 2011 July;4(4):382-4.

(18) Kang HT, Lee KB, Kim SY, Choi HR, Park SC. Autophagy impairment induces premature senescence in primary human fibroblasts. PLoS One 2011;6(8):e23367.

(19) Westendorp RG, van HD, Rozing MP, Frolich M, Mooijaart SP, Blauw GJ et al. Nonagenarian siblings and their offspring display lower risk of mortality and morbidity than sporadic nonagenarians: The Leiden Longevity Study. J Am Geriatr Soc 2009 September;57(9):1634-7.

(20) Heijmans BT, Beekman M, Houwing-Duistermaat JJ, Cobain MR, Powell J, Blauw GJ et al. Lipoprotein particle profiles mark familial and sporadic human longevity. PLoS Med 2006 December;3(12):e495.

(21) Barzilai N, Gabriely I, Gabriely M, Iankowitz N, Sorkin JD. Offspring of centenarians have a favorable lipid profile. J Am Geriatr Soc 2001 January;49(1):76-9.

(22) Rozing MP, Westendorp RG, Frolich M, de Craen AJ, Beekman M, Heijmans BT et al. Human insulin/IGF-1 and familial longevity at middle age. Aging (Albany NY) 2009 August;1(8):714-22.

(23) Wijsman CA, Rozing MP, Streefland TC, le CS, Mooijaart SP, Slagboom PE et al. Familial longevity is marked by enhanced insulin sensitivity. Aging Cell 2011 February;10(1):114-21.

(24) Dekker P, Maier AB, van HD, de Koning-Treurniet C, Blom J, Dirks RW et al. Stress-induced responses of human skin fibroblasts in vitro reflect human longevity. Aging Cell 2009 September;8(5):595-603.

(25) Dekker P, de Lange MJ, Dirks RW, van HD, Tanke HJ, Westendorp RG et al. Relation between maximum replicative capacity and oxidative stress-induced responses in human skin fibroblasts in vitro. J Gerontol A Biol Sci Med Sci 2011 January;66(1):45-50.

(26) Tanida I, Ueno T, Kominami E. LC3 and Autophagy. Methods Mol Biol 2008;445:77-88.

(27) Kabeya Y, Mizushima N, Yamamoto A, Oshitani-Okamoto S, Ohsumi Y, Yoshimori T. LC3, GABARAP and GATE16 localize to autophagosomal membrane depending on form-II formation. J Cell Sci 2004 June 1;117(Pt 13):2805-12.

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

(29) Maier AB, le CS, de Koning-Treurniet C, Blom J, Westendorp RG, van HD. Persistence of high-replicative capacity in cultured fibroblasts from nonagenarians. Aging Cell 2007 February;6(1):27-33.

(30) Linssen MM, van Raalte DH, Toonen EJ, Alkema W, van der Zon GC, Dokter WH et al. Prednisolone-induced beta cell dysfunction is associated with impaired endoplasmic reticulum homeostasis in INS-1E cells. Cell Signal 2011 November;23(11):1708-15.

(31) Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 2001 December;25(4):402-8.

Supplementary information

Table S1 Pearson correlation between gene expression and protein expression in fibroblasts of the same donors (n = 66)

Gene	Protein	Correlation	P
MTOR	mTOR	-0.058	0.646
RPTOR	Raptor	-0.132	0.292
RPS6KB1	P70S6K	-0.198	0.112

p: raw p value from 2-tailed Pearson correlation.

Table S2 Pearson correlation between gene expression in fibroblasts and gene expression in whole blood of the same donors (n = 72)

Gene	Protein	Correlation	P
EIF4EBP1	4EBP1	0.060	0.639
AKT1	AKT	0.070	0.585
C12orf44	ATG101	0.067	0.575
ATG13	ATG13	-0.130	0.276
DEPTOR	Deptor	-0.139	0.245
EIF4B	eIF4B	0.074	0.538
EIF4E	eIF4E	0.011	0.930
RB1CC1	FIP200	0.031	0.793
FOXO1	FOXO1	-0.072	0.550
FOXO3	FOXO3	0.030	0.803
HIF1A	HIF1a	-0.056	0.638
MLST8	mLST8	0.127	0.289
MAPKSP1	MP1	-0.107	0.370
MAPKAP1	mSIN1	0.140	0.249
MTOR	mTOR	0.063	0.621
LAMTOR2	p14	0.075	0.530
CDKN2C	p18	0.218	0.066
PPARGC1A	PGC1a	-0.056	0.651
PRKCA	PKCa	0.105	0.381
PPARG	PPARg	-0.057	0.632
AKT1S1	PRAS40	0.046	0.701
PRR5	Protor1	-0.145	0.224
PRR5L	Protor2	0.117	0.327
RRAGA	RagA	-0.088	0.467
RRAGB	RagB	0.038	0.752
RRAGC	RagC	-0.110	0.356
RRAGD	RagD	-0.139	0.245
RPTOR	Raptor	-0.153	0.221
RHEB	Rheb	-0.057	0.632

RHOA	RhoA	0.032	0.787
RICTOR	Rictor	0.113	0.344
RPS6	S6	0.034	0.777
RPS6KB1	S6K1	-0.070	0.581
SGK1	SGK	-0.160	0.179
POLDIP3	SKAR	-0.131	0.315
SREBF1	SREBP	0.149	0.210
TSC1	TSC1	-0.109	0.361
TSC2	TSC2	0.107	0.372
ULK1	ULK1	-0.074	0.559
YY1	YY1	0.033	0.786

p: raw p value from 2-tailed Pearson correlation

Chapter 7

Meta-analysis on blood transcriptomic studies identifies consistently co-expressed PPI modules as robust markers of human ageing

E.B. van den Akker[1,2,§] , W.M. Passtoors[1,§], R. Jansen[3], E.W. van Zwet[4], J.J. Goeman[4], M. Hulsman[2], V. Emilsson[5], M. Perola[6], A.H.M. Willemsen[7], B.W.J.H. Penninx[3], B.T. Heijmans[1], A.B. Maier[8],D.I. Boomsma[3,7], J.N. Kok[1,9], P.E. Slagboom[1,10], M.J.T. Reinders[2,], M. Beekman[1,10,]**

1) Molecular Epidemiology, Leiden University Medical Center, Leiden, The Netherlands
2) The Delft Bioinformatics Lab, Delft University of Technology, Delft, The Netherlands
3) Department of Psychiatry, VU University Medical Center, Amsterdam, The Netherlands
4) Medical Statistics, Leiden University Medical Center, Leiden, The Netherlands
5) Icelandic Heart Association, Kópavogur, Iceland
6) National Institute for Health and Welfare, Helsinki, Finland
7) Department of Biological Psychology, VU University, Amsterdam, The Netherlands
8) Gerontology and Geriatrics, Leiden University Medical Center, Leiden, The Netherlands
9) Algorithms, Leiden Institute of Advanced Computer Science, University of Leiden, Leiden, The Netherlands
10) Netherlands Consortium for Healthy Ageing, Leiden, The Netherlands

§ Both authors contributed equally to the studies

Aging Cell. 2013 Oct 14 published online.

Abstract

The bodily decline that occurs with advancing age strongly impacts on the prospects for future health and life expectancy. Despite the profound role of age in disease etiology, knowledge about the molecular mechanisms driving the process of ageing in humans is limited. Here we used an integrative network-based approach for combining multiple large-scale expression studies in blood (2,539 individuals) with Protein-Protein Interaction (PPI) data for the detection of consistently co-expressed PPI modules that may reflect key processes that change throughout the course of normative ageing. Module detection followed by a meta-analysis on chronological age identified fifteen consistently co-expressed PPI modules associated with chronological age, including a highly significant module (p=3.5×10^{-38}) enriched for "*T-cell activation*" marking age-associated shifts in lymphocyte blood cell counts (R^2=0.603; p=1.9×10^{-10}). Adjusting the analysis in the compendium for the "*T-cell activation*" module showed five consistently co-expressed PPI modules that robustly associated with chronological age and included modules enriched for "*Translational elongation*", "*Cytolysis*" and "*DNA metabolic process*". In an independent study of 3,535 individuals, four out of five modules consistently associated with chronological age, underpinning the robustness of the approach. We found three out of five modules to be significantly enriched with ageing-related genes, as defined by the GenAge database, and association with prospective survival at high ages for one of the modules including *ASF1A*. The hereby-detected age associated and consistently co-expressed PPI modules therefore may provide a molecular basis for future research into mechanisms underlying human ageing.

Introduction

A steadily growing life expectancy of the general western population throughout the past two centuries (1) has imposed the urgency for understanding the adverse effects of ageing for public health and its relation to the observed large variation in healthy lifespan (2). Age-dependent detrimental processes strongly attenuate prospects for future health, with chronological age being the major risk factor for mortality and virtually all common diseases in the western world (3). Ageing is a systemic ailment marked by a gradual metabolic decline eventually leading to a state of senescence on both the cellular and organismal level that seems to be caused by the accumulation of damage over time (4). Despite their profound role for disease etiology, the existing knowledge concerning the

molecular mechanisms driving biological ageing processes in humans is limited.

Construction of consistent age-associated signatures has proven to be challenging as a multitude of gene expression studies have identified age associated genes so far, though with limited mutual overlap (5;6). This inconsistency is most likely due to variable technical circumstances, small study sizes and low signal-to-noise ratios, typically observed when analysing the ageing transcriptome. More similarity was observed at the pathway level, across tissues and even species (7;8) suggesting that the analysis of the ageing transcriptome by functionally grouped gene sets is a promising alternative for the classical individual-gene analyses.

Rather than employing literature-based sets of genes sharing similar biological functions, so-called network approaches are increasingly used, which infer functional clusters of genes from the expression data itself by exploiting gene co-expression patterns hidden within the data (9). Alternatively, changes in these gene co-expression patterns that occur with age might be used for inferring a functional grouping from the data (10). However, co-expression patterns may contain spurious gene-gene correlations (11), which make the use of multiple data-sources simultaneously or the integration with other additional information sources on functional relationships between genes desirable.

Established modulators of ageing processes in model organisms were reported to spatially cluster within networks constructed of Protein-Protein Interaction (PPI) data (12;13). Hence, PPI networks can be exploited for prioritizing new ageing-associated genes (14;15) or for refining modules of co-expressed genes that are correlated during the course of ageing (16). We previously demonstrated that the inference of these so-called co-expressed PPI modules has a high reproducibility across multiple expression datasets in breast-cancer (17), and here we extend this algorithm to combine multiple gene expression datasets on ageing.

Though many algorithms for network inference exist (18), relatively little attention has gone to the problem of network inference and subsequent associations with a phenotype using multiple heterogeneous expression data sources simultaneously. Merging the expression data into a single set and using this for network inference clearly surpasses the differences in correlation structures present within each dataset. Irrespective of the type of network inference chosen, we propose to handle such heterogeneity by integrating the gene-gene similarity measures obtained across expression

datasets using a suitable meta-analysis setting. Thus, in the approach described in this paper we employ a meta-analysis for inferring a consistent gene-gene network that serves as a basis for identifying consistently co-expressed PPI modules, which are subsequently analysed with respect to chronological age across datasets using again a meta-analysis.

To robustly characterize the changes of the blood transcriptome associated with chronological age, we have build a compendium using three large-scale transcriptomic studies (19-21) generated in blood comprising 2,539 individuals on which we applied our integrative network approach. For comparison, two types of individual-gene meta-analyses were performed as well, which in combination with an enrichment analysis yielded only broad terms for age-associated cellular processes. Application of our integrative network-based approach yielded five consistently co-expressed PPI modules showing robust age associations and functional enrichments for *"Translational elongation"*, *"Cytolysis"* and *"DNA metabolic process"*, that seem to reflect downstream mTOR signaling events or cell-cycle checkpoints. Finally, we show that four out of five modules replicate in an independent cohort, that they are enriched for known longevity- and ageing-related genes and that the expression of one module associates with prospective survival at old age.

Results

The largest transcriptome compendium for normative ageing

To robustly characterize the changes of the blood transcriptome throughout the course of normative ageing in the range of 15-94 years, we built a gene expression compendium using three large-scale transcriptomic studies performed in blood: the San Antonio Family Heart Study (SAFHS) (21), the Icelandic Family Blood (IFB) cohort (19) and the Dietary, Lifestyle, and Genetic determinants of Obesity and Metabolic syndrome (DILGOM) study (20). Data of IFB was measured in two roughly equally sized batches, from this point on referred to as IFB_A and IFB_B, and was treated as two separate datasets in the downstream analysis. Data quality was critically re-assessed and re-annotated yielding a compendium of 9,047 unique genes expressed in 2,539 individuals divided over four datasets (SAFHS: 1,240, IFB_A: 411, IFB_B: 435, DILGOM: 454, Table 1).

Table 1: Descriptives of the datasets composing the compendium

Study	Tissue	Cohort	Ethnicity	#start total[††]	#end total[†]	# males (%)[†]	mean age[†]	min age[†]	max age[†]
SAFHS *	Lymphocytes	San Antonio Family Heart Study	Mexican Americans (USA)	1240	1240	506 (40.8%)	39.3	15	94
IFB_A	Peripheral blood	Icelandic Family Blood (IFB) cohort	Caucasian (Icelandic)	904[A]	411	198 (48.2%)	48.8	19	84
IFB_B	Peripheral blood	Icelandic Family Blood (IFB) cohort	Caucasian (Icelandic)	904[A]	434	180 (41.5%)	46.2	20	76
DILGOM	Peripheral blood	DILGOM (Dietary, Lifestyle, and Genetic determinants of Obesity and Metabolic syndrome)	Caucasian (Finnish)	518[B]	454	195 (43.0%)	51.6	30	70

[*] Expression and phenotypic data were obtained from ArrayExpress under accessions: E-TABM-305, [††] Number of individuals with matching phenotypic data per study when obtained. (A) Data of IFB was measured in two batches. This figure indicates the total number of individuals before preprocessing or removal of duplicates across batches. (B) A small batch was detected and all samples belonging to it were removed. [†] Statistics computed after preprocessing.

Limited overlap of age-associated genes between studies within the compendium

The most straightforward method for an integrative analysis across datasets is to first compute the age-association genes per dataset and subsequently inspect the overlap of significant results. A linear model adjusted for gender yielded between 111 (1.2%) and 1,103 (12.2%) significantly age-associated genes per dataset (Bonferroni correction, $\alpha \leq 0.05$), of which 26 genes were significantly associated with age in all four datasets (Figure 1 and Table S2). These results confirmed the high discrepancy between lists of age-associated genes previously reported in literature, though now observed in equal or similar tissues (5;6).

Rank-based integration of age-associated genes improves consistency between studies

As repeatedly applied cutoffs across multiple heterogeneous datasets may lead to high false exclusion rates of age-associated genes, we investigated whether age-association rankings were consistently high across datasets by applying a rank integration approach (5;22). From the 9,047 genes present in the compendium, 247 consistently showed highly ranked differential expressions with age across the four datasets, of which 195 remained significant after permutation tests (both at FDR<=0.05). Out of these 195 genes, 128 (65.6%) showed decreased and 67 (34.4%) showed

increased expression levels with age. The top 25 genes with increased and decreased expression are displayed in Table 2 and Table 3 respectively and include many of the age associated genes previously identified, like *LRNN3*, *LEF1*, *and SYT11* (23-25). Results for all 9,047 in the compendium are provided in Table S3.

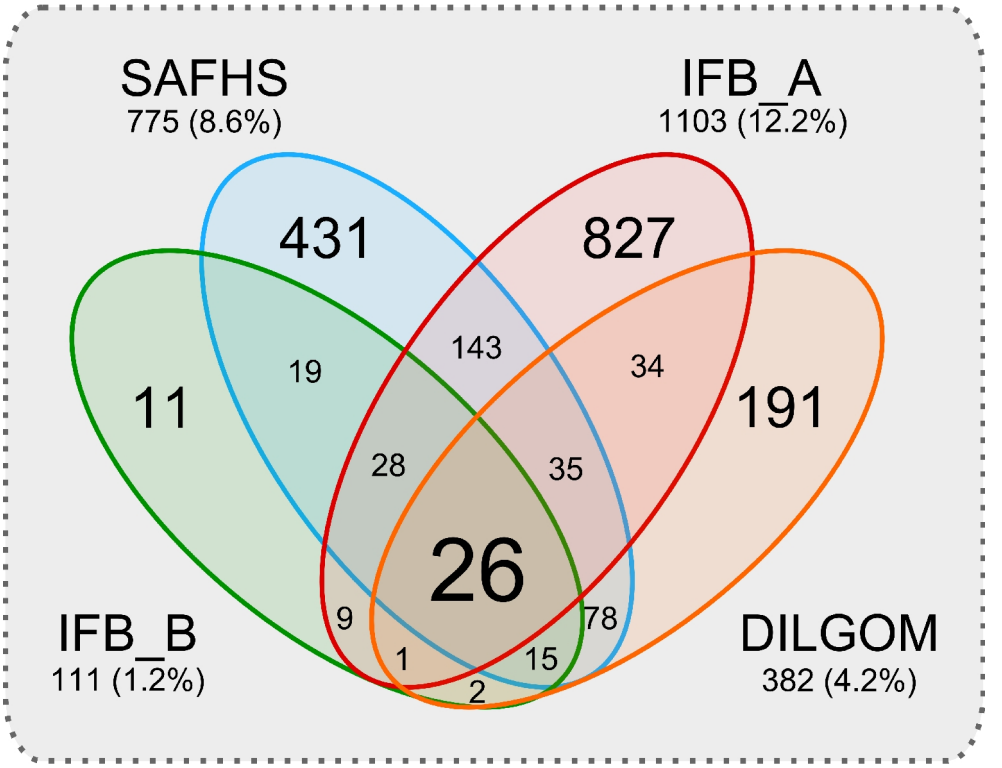

Figure 1: Significantly age-associated genes in studies of the blood compendium
A Venn analysis was performed for inspecting the overlap of the significantly age-associated genes found within different studies. The majority of the consistently detected age-associated genes (24 out of 26) show a decreased expression with advancing age and include: *ARH*, *BACH2*, *CCR7*, *ECRG4*, *EDAR*, *EPHA1*, *EPHX2*, *FAM102A*, *FAM134B*, *FBLN2*, *FCGBP*, *FLNB*, *IL24*, *LRRN3*, *NELL2*, *NMT2*, *NRCAM*, *OXNAD1*, *PDE9A*, *PHGDH*, *PIK3IP1*, *SIRPB2*, *SUSD3* and *TSGA14*. The remaining 2 consistently age-associated genes showing increased expressions are *ARP10* and *SYT11*. See Table S2 for more details.

Functional enrichments of individual-gene analysis are not informative for normative ageing

We next identified enriched functional groupings amongst genes significantly associated with normative ageing using DAVID focusing on

GO_FAT terms. Whereas the 26 genes from the overlap did not yield any significantly enriched terms, the 195 significant genes obtained with the rank integration approach yielded 11 significant enriched groupings when run at default settings (Table S4 and S5 respectively). Interestingly, enriched terms include *"Glycosylation site:N-linked"* (p=6.1×10^{-5}, Benjamini corrected), previously linked to the inflamm-ageing theory (26). However, as most of the 11 identified terms are rather broadly defined, like *"disulfide bond"* or *"signal peptide",* little detailed knowledge is gained on potential molecular mechanisms underlying normative ageing following the individual-gene analysis approach.

Table 2. Top 25 genes according to the gene statistic (U_i) having increased expression with age

Symbol	GeneID	p-value[1]	q-value[1]	p-value[2]	q-value[2]
GPR56	9289	5.3 × 10^{-09}	4.8 × 10^{-05}	1.0 × 10^{-06}	0.0018
HF1	3075	2.3 × 10^{-08}	8.1 × 10^{-05}	1.0 × 10^{-06}	0.0018
SYT11	23208	2.7 × 10^{-08}	8.1 × 10^{-05}	≤ 5.0 × 10^{-7}	0.0018
ARP10	164668	7.3 × 10^{-08}	1.7 × 10^{-04}	1.0 × 10^{-06}	0.0018
B3GAT1 (CD57)	27087	1.1 × 10^{-07}	2.0 × 10^{-04}	3.0 × 10^{-06}	0.0021
SLC1A7	6512	1.8 × 10^{-07}	2.6 × 10^{-04}	3.2 × 10^{-05}	0.0110
IFNG	3458	5.0 × 10^{-07}	6.4 × 10^{-04}	1.1 × 10^{-05}	0.0065
DSCR1L1	10231	6.1 × 10^{-07}	6.8 × 10^{-04}	2.0 × 10^{-06}	0.0021
ARK5	9891	7.9 × 10^{-07}	7.9 × 10^{-04}	3.0 × 10^{-06}	0.0021
PIG13	81563	9.3 × 10^{-07}	8.8 × 10^{-04}	1.0 × 10^{-06}	0.0018
SPUVE	11098	1.1 × 10^{-06}	8.8 × 10^{-04}	1.2 × 10^{-05}	0.0067
PDGFRB	5159	1.2 × 10^{-06}	8.8 × 10^{-04}	1.5 × 10^{-06}	0.0021
EDG8	53637	1.4 × 10^{-06}	9.4 × 10^{-04}	7.8 × 10^{-05}	0.015
MARLIN1	152789	1.5 × 10^{-06}	9.4 × 10^{-04}	5.0 × 10^{-06}	0.0032
TGFBR3	7049	2.0 × 10^{-06}	0.0012	2.8 × 10^{-05}	0.011
GZMB	3002	2.4 × 10^{-06}	0.0013	5.0 × 10^{-04}	0.050
CX3CR1	1524	2.9 × 10^{-06}	0.0014	2.9 × 10^{-05}	0.011
STYK1	55359	3.3 × 10^{-06}	0.0015	4.8 × 10^{-05}	0.013
ADRB2	154	3.7 × 10^{-06}	0.0016	3.0 × 10^{-06}	0.0021
GAF1	26056	7.1 × 10^{-06}	0.0029	7.2 × 10^{-05}	0.015
CTSL	1514	7.7 × 10^{-06}	0.0030	3.2 × 10^{-04}	0.040
GFI1	2672	1.1 × 10^{-05}	0.0040	3.0 × 10^{-06}	0.0021
TTC38	55020	1.1 × 10^{-05}	0.0040	7.6 × 10^{-05}	0.015
AGPAT4	56895	1.2 × 10^{-05}	0.0041	2.5 × 10^{-06}	0.0021
GZMA	3001	1.4 × 10^{-05}	0.0045	3.3 × 10^{-04}	0.040

[1]*p*- and *q*-values determined using the gamma-distribution of the gene statistic, U_i , [2]*p*- and *q*-values determined using permutation of the gene statistic, U_i

Table 3. Top 25 genes according to the gene statistic (U_i) having decreased expression with age

Symbol	GeneID	p-value[1]	q-value[1]	p-value[2]	q-value[2]
LRRN3	54674	1.3 × 10-12	1.2 × 10-8	≤ 5.0 × 10-7	3.2 × 10-4
FCGBP	8857	3.2 × 10-10	1.5 × 10-6	≤ 5.0 × 10-7	3.2 × 10-4
CCR7	1236	1.1 × 10-9	3.2 × 10-6	≤ 5.0 × 10-7	3.2 × 10-4
NELL2	4753	2.0 × 10-8	4.5 × 10-5	1.0 × 10-6	3.8 × 10-4
NRCAM	4897	3.1 × 10-8	5.6 × 10-5	≤ 5.0 × 10-7	3.2 × 10-4
IGJ	3512	1.5 × 10-7	2.3 × 10-4	2.6 × 10-4	0.019
LEF1	51176	1.9 × 10-7	2.5 × 10-4	≤ 5.0 × 10-7	3.2 × 10-4
FAM134B	54463	2.2 × 10-7	2.5 × 10-4	≤ 5.0 × 10-7	3.2 × 10-4
PACAP	51237	2.5 × 10-7	2.5 × 10-4	1.5 × 10-6	4.8 × 10-4
ITM2C	81618	2.8 × 10-7	2.5 × 10-4	3.5 × 10-6	8.1 × 10-4
PIK3IP1	113791	3.0 × 10-7	2.5 × 10-4	1.0 × 10-6	3.8 × 10-4
PDE9A	5152	5.1 × 10-7	3.8 × 10-4	1.0 × 10-6	3.8 × 10-4
BACH2	60468	6.9 × 10-7	4.8 × 10-4	1.0 × 10-6	3.8 × 10-4
FLJ12895	65982	9.5 × 10-7	6.0 × 10-4	1.5 × 10-6	4.8 × 10-4
FAM102A	399665	1.1 × 10-6	6.0 × 10-4	≤ 5.0 × 10-7	3.2 × 10-4
FBLN2	2199	1.1 × 10-6	6.0 × 10-4	≤ 5.0 × 10-7	3.2 × 10-4
FLNB	2317	1.2 × 10-6	6.0 × 10-4	≤ 5.0 × 10-7	3.2 × 10-4
APEG1	10290	1.2 × 10-6	6.0 × 10-4	1.0 × 10-6	3.8 × 10-4
EPHX2	2053	1.3 × 10-6	6.0 × 10-4	1.5 × 10-6	4.8 × 10-4
TNFRSF17	608	1.3 × 10-6	6.1 × 10-4	1.2 × 10-4	0.011
MYC	4609	1.6 × 10-6	6.6 × 10-4	3.5 × 10-6	8.1 × 10-4
NT5E	4907	1.7 × 10-6	6.6 × 10-4	1.0 × 10-6	3.8 × 10-4
TOSO	9214	1.7 × 10-6	6.6 × 10-4	1.0 × 10-6	3.8 × 10-4
ARH	26119	3.2 × 10-6	0.0012	2.0 × 10-6	6.2 × 10-4
OXNAD1	92106	3.3 × 10-6	0.0012	≤ 5.0 × 10-7	3.2 × 10-4

[1]*p*- and *q*-values determined using the gamma-distribution of the gene statistic, U_i , [2]*p*- and *q*-values determined using permutation of the gene statistic, U_i

A novel integrative network approach for detecting consistent co-expressed PPI modules

To improve robustness against noise and increase power, we used a novel integrative network-based approach to explore functional age-associated groupings of genes. The proposed approach detects consistently co-expressed PPI modules across multiple datasets (for details see Materials & Methods, and Supplemental Methods). Using the four transcriptomic datasets mapped onto the PPI network, we detected a total of 162 consistently co-expressed PPI modules ranging in size from 2 to 37 genes (see Figure S6 for a complete overview). The following steps in our

analysis were limited to the subset of 27 co-expressed PPI modules counting at least 5 genes. Application of DAVID yielded significant functional enrichments for 19 out of the 27 identified co-expressed PPI modules (Table S7), suggesting that the applied approach grouped genes according to plausible biological functions.

Age-associated co-expressed PPI modules point towards T-cell activation

To test whether transcriptional changes of the 27 identified modules associate with chronological age, an expression profile for each module was constructed by determining the mean expression of the genes within a detected co-expressed PPI module per individual. As with the individual-gene analysis, we proceeded by computing the associations of the module expressions with age while adjusting for gender for each dataset separately. Only one module (Figure 2A), enriched for "*T-cell activation*", was significantly associated with age in each of the four datasets of the compendium. This module A contains genes commonly employed as markers for assessing the differentiation status of T-cell lineages, such as *CCR7*, *CD28* and *TNFRSF7* (*CD27*). A fixed-effect meta-analysis on the expression of the different modules across the datasets showed again that the "*T-cell activation*" module was most significantly associated with age (Bonferroni corrected $p=3.5\times10^{-38}$). The consistent age association of the "*T-cell activation*" module, however, raises the concern that the identified modules reflect age-related changes in the proportions of cell populations in blood, as previously reported (27), rather than changes in gene expression.

T-cell activation module expression marks blood lymphocyte counts

To investigate the relation between the expression of the "*T-cell activation*" module and the proportions of blood cell populations, for which we have no data in the compendium, we revisited a transcriptomic dataset on peripheral blood measured in the Leiden Longevity Study (LLS) (25) (Supplemental Methods). Using the expression data of 50 middle-aged and 50 ninety-year-old individuals, we first confirmed the association with age of the expression of the "*T-cell activation*" module ($p=3.7\times10^{-5}$), and subsequently observed a significant correlation between the expression of the "*T-cell activation*" module and lymphocyte counts ($R^2=0.603$, $p=1.9\times10^{-10}$). These findings suggest that the previously observed age associations in the blood compendium are most probably confounded by the age-associated decline in lymphocyte counts. We also conclude that the

expression of the "*T-cell activation*" module could serve as a proxy for the age-associated decline in lymphocyte counts in the compendium.

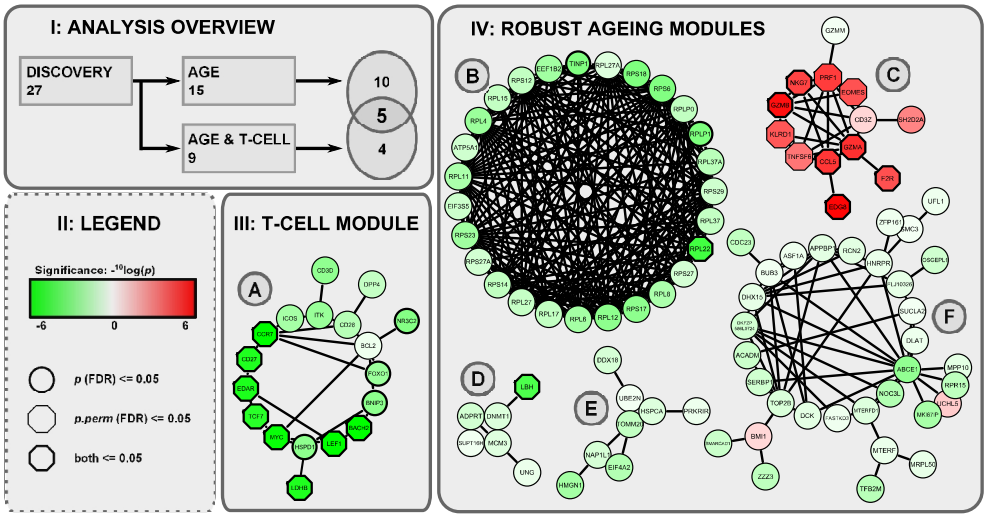

Figure 2: Overview main results of the integrative network-based approach
Panel 1: Overlap of the PPI network and cluster analysis of the transcriptomic data reveals 27 modules, 15 are significantly associated with age, 9 are significantly associated with age when corrected for the "*T-cell activation*" module expression and the 5 most robust findings are found in the overlap. Panel 2: The co-expressed PPI module that is enriched for "*T-cell activation*". Panel 3: B-F: 5 co-expressed PPI modules with expressions robustly associated with age. B, C and F: modules enriched for "*Translational elongation*", "*Cytolysis*" and "*DNA metabolic process*", respectively. Node's shape and color reflect the results of the individual gene statistic (U_i). The red and green colors denote a correlating or anti-correlating relationship of gene expression with age, respectively. The intensity of the coloring indicates the significance of the gamma distributed transformed rank product statistics. Nodes marked by a thick bordering or a hexagon shaped bordering represent genes with FDR adjusted *p*-values ≤ 0.05 for respectively the analytical and permutation-based approach.

Five co-expressed PPI modules associate with age independent of T-cell activation

Based on these findings we adapted the fixed-effect meta-analysis to re-analyse the 27 modules in the compendium while adjusting for gender as well as the expression of the "*T-cell activation*" module. This revealed nine modules significantly associated with chronological age, of which five also showed a significant association without adjusting for "*T-cell activation*" (Figure 2 B-F). These five modules thus exhibit the most robust expression changes with age and include: 1) a large consistently down-regulated ribosomal module (p=9.4×10^{-19}), enriched for "*Translational elongation*" (p=4.5×10^{-46}); 2) an up-regulated module containing amongst others

several granzymes and the perforin gene (p=2.9×10^{-24}), enriched for "*Cytolysis*" (p=9.4×10^{-05}); and, 3) a down-regulated module containing the *PARP1 (ADPRT)* gene (p=3.1×10^{-39}) enriched for "*DNA metabolic process*" (p=0.0036). The two remaining modules were both down-regulated with advancing age and lacked any significant functional enrichments (Figure 2E and 2F; p=3.9×10^{-11} and p=2.5×10^{-18}, respectively).

Replication of Co-Expressed PPI Modules as Robust Markers for Ageing

We conducted an independent replication study of the identified network modules as robust markers for chronological age using gene expression data from the Netherlands Twin Register and Netherlands Study of Depression and Anxiety (NTR & NESDA) consortium (N=3,535) (28) assayed on individuals within age range 17-79 years (Supplemental Methods). An association analysis between the mean expression of a module and chronological age, adjusted for sex and the mean expression of the "*T-cell activation*" module, yielded significant results for four out of the five identified modules, all with directions corresponding to those found in the compendium (Table S8). These results emphasize the robustness of the findings produced by our approach and confirm that the mean module expression in whole blood of module B, C, E and F may be considered as robust markers of chronological age.

Co-expressed PPI modules are enriched for GenAge longevity and ageing genes

As a validation of the identified modules, we computed whether ageing-related genes stored by GenAge (13), a database providing a comprehensive overview of ageing-related genes in humans and model systems, were enriched within modules A-F (Figure 2). Whereas module A was supported by human derived annotations only (OR=12.1, 95% CI 2.88-39.2, p=6.95×10^{-4}), module B was solely based on knowledge derived from model organisms (OR=16.9, 95% CI 7.26-39.1, p=2.52×10^{-10}, Table S9). Modules D, E and F had annotations balanced over both sources and therefore the significance of the joint enrichment was assessed by using a resampling approach (Supplemental Methods), which yielded significant enrichments for modules E (p=0.016) and F (p=0.0029). These findings provide additional evidence that the joint expression of these modules may play a relevant role in human ageing.

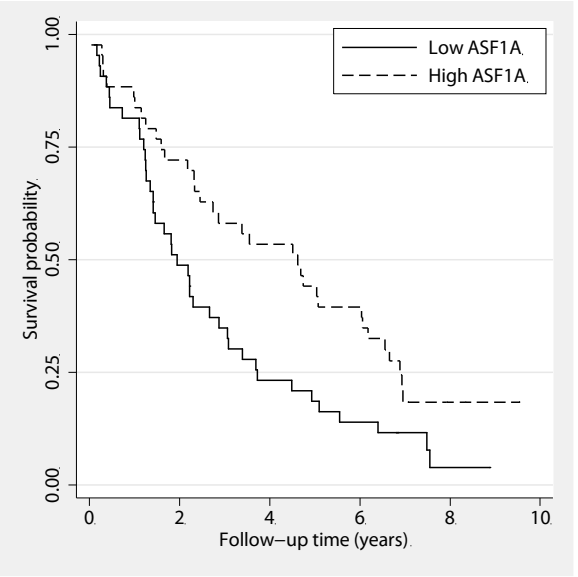

Figure 3: Expression of ASF1A associates with prospective survival in nonagenarians
High expression of *ASF1A* confers a prospective survival benefit at old age.

Module F associates with prospective survival at old age

To investigate whether the identified modules could potentially serve as biomarkers, we studied the microarray data assayed on 50 nonagenarian individuals from the Leiden Longevity Study (25). A left truncated Cox proportional hazard model adjusted for sex and cell counts indicates that the mean expression of module F associates with prospective survival beyond the age of 90 years (N=50, N_{death}=45, HR=0.265, 95% CI 0.12-0.57, *p*=0.001). By showing that module F associates with prospective survival at old age, we illustrate its potential biological relevance.
Interestingly, the *ASF1A* gene is part of module F and has previously been identified by our group as one of the genes that was differentially expressed in blood of members of long-lived families as compared to similarly aged controls at middle age (25). To confirm that the expression of the *ASF1A* gene in module F also associates with prospective survival at old age, we analysed the gene expression of *ASF1A* measured with RT-qPCR in 74 nonagenarians from the Leiden Longevity Study (of which 24 overlapped with the micro-array experiment) for association with prospective survival. Since we observe a similar association (N_{death}=64, HR=0.54, 95% CI 0.34-0.85, *p*=0.008, Figure 3), these results indicate that modules, of which the expression in blood is consistently associated with chronological age across various datasets, may associate with variation in lifespan, and

therefore provide valid gene targets for studying relevant biological endpoints in human ageing.

Discussion

Age associated changes in gene expression may provide meaningful leads to pathways affected by and involved in ageing, however, are in general difficult to detect consistently (5). Therefore, we constructed the largest compendium of human whole blood expression studies (19-21) comprising 2,539 individuals on which we performed a novel integrative network-based analysis. This yielded fifteen consistently age-associated co-expressed PPI modules. Because the most significant age-associated module appeared to correlate with lymphocyte cell counts in an independent gene expression dataset, the expression of this module, enriched for "*T-cell activation*", was subsequently used as a proxy for possible confounding shifts in the distribution of lymphocyte subsets. This enabled the identification of five age-associated modules (Figure 2 Panel I and IV), including three modules enriched for "*Translational elongation*", "*Cytolysis*" and "*DNA metabolic process*" (Figure 2B-D). Replication in an independent cohort confirmed these findings for four out of five modules (Figure 2B, C, E and F), underpinning the robustness of the proposed approach. The enrichments against a database for ageing-related genes (Figure 2B, E & F) emphasize the relevance of these biological findings for ageing research, which is even further substantiated by the fact that the mean expression of module F associates with prospective survival at old age.

Mitochondrion-related ageing

Two of the identified modules are down-regulated with age and seem to be related to the mitochondrion, though lacking any significant functional enrichment (Figure 2E-F). Despite the absence of functional enrichments, both modules were significantly enriched for ageing-related genes, as defined by GenAge, implying that known age-related single genes can be put into a novel biological perspective by our network approach.

Module (Figure 2F) contains several mitochondrial factors and enzymes, like, for instance, the mitochondrial transcription termination factor *MTERF,* the *ACADM* enzyme used for fatty acid metabolism, or the mitochondrial tRNA synthetase *IARS2*, whose homolog was shown to increase lifespan upon disruption in worms (29). This module also includes several genes previously associated with age or age-associated diseases such as the mitotic checkpoint protein *BUB3*, previously associated with accelerated

ageing in mice (30), and the cell-cycle checkpoint protein *APPBP1* found in increased quantities in the brain affected by Alzheimer's disease (31). This broad range of gene characteristics composing the module could be explained by the fact that the functionality of mitochondria is not confined to cellular energy metabolism alone, but also seems to make up an integral part of multiple cell signaling cascades including cell-cycle control and cell death (32).

Interestingly, module F also includes the *ASF1A* histone chaperone of which we previously have shown that its expression associates with familial longevity in the Leiden Longevity Study (25). We revisited the RT-qPCR data assayed on 74 nonagenarians and now show that the expression of *ASF1A* also associates with prospective survival. This result illustrates that modules, of which the expression in blood is consistently associated with chronological age across various datasets, may associate with variation in lifespan, and therefore provide valid gene targets for studying relevant biological endpoints in human ageing.

The other mitochondrion-related module (Figure 2E) contains the heat shock protein *HSPCA* (*HSP90)* and the mitochondrial receptor *TOMM20*, which jointly play a central role in translocating pre-proteins into the mitochondria (33). They seem to be consistently co-expressed in blood with *EIF4A2,* a eukaryotic translation initiation factor and *DDX18*, an ATP dependent RNA helicase, of which the worm homologs were shown to extent lifespan upon disruption (29;34). To summarize, this module seems to relate to ageing by influencing protein translation and mitochondrial translocation efficiency.

Age-associated limitation of protein synthesis

One of the identified modules predominantly consisted of ribosomal proteins and translation elongation factors comprising part of the ribosomal complex (Figure 2B). The module was significantly enriched for "*Translational elongation*" and for previous findings in model organisms with respect to ageing and longevity. In addition, the module was down-regulated with advancing age fitting previous observations of the ageing blood transcriptome (23-25), which could be interpreted as an attempt of the cell to limit global protein synthesis in response to stress arising from damage accumulating throughout lifespan (35). Whether caused by response to stress or other factors, the change in protein translation may be ascribed to the mTORC1 complex (36). This complex modulates cellular growth and metabolisms by determining the balance between protein

synthesis and degradation in response to nutrient availability. Inhibition of mTOR signaling through the mTORC1 complex not only inhibits protein synthesis, but also has been shown to positively affect the lifespan in various invertebrates and mammals (37). Moreover, human blood transcriptome studies showed that the gene expression of mTOR pathway is down-regulated with chronological age (23;38) and is even associated with human familial longevity (25). Hence, a consistently down-regulated ribosomal module with advancing age corresponds with the age-associated demise of mTOR signaling. Although it is well established that mTOR signaling links to both lifespan regulation and "*Translational elongation*", it remains to be determined whether down-regulation of "*Translational elongation*" is causal for human ageing.

WRN related cell-cycle checkpoint on DNA integrity

A module down-regulated with age and enriched for "*DNA metabolic process*" identified in the compendium could not be replicated in the NTR & NESDA cohort (Figure 2D). Interestingly, this module contains the *PARP1* (*ADPRT*) gene, which directly binds to *WRN* to induce apoptosis upon oxidative stress induced DNA damage and is as such a prime suspect for Werner syndrome (39), a premature ageing disease. Furthermore, the activity of the Parp1 protein in mononuclear cells has previously been shown to positively correlate with the species-specific lifespan across 13 mammalian species (40). Taken together, findings in the compendium suggest that the lowered transcription rate of *PARP1* negatively affects DNA integrity and thus life span, though more experiments are required to investigate this hypothesis.

Age-associated shifts in T-cell composition

Another identified module is up-regulated with age and enriched for "*Cytolysis*" (Figure 2C). It contains several genes used to dispatch virus-infected cells and may reflect the decreased competence for fighting infections in an early stage, caused by an age-related deterioration of the immune system, known as immuno-senescence (41). We can, however, not rule out that the age-associated expression of *GZMA*, *GZMB* and *PRF1* that are part of this module point to an age-associated shift in T-cytotoxic cells (27;42).

Though identified co-regulated PPI modules may show extensive correlation with confounding factors, we should be careful to dismiss modules as such only. For instance the "*T-cell activation*" module (Figure

2A), which is down-regulated with age, also contained *BNIP3*, an inhibitor of the mTORC1 complex shown to modulate lifespan in worms, flies and mice (37); and *FOXO1*, also displaying an intricate interplay with both complexes of *mTOR* (36), and shown to extent lifespan in various invertebrates (43). Additionally, human *mTOR* signaling may play a central role in orchestrating T-cell maturation and T-cell fate decisions (44), and could thereby also explain the age-associated decline in lymphocytes as marked by the "*T-cell activation*" module. Taken together, these examples illustrate that what is confounding the analysis of the blood transcriptome for molecular mechanisms associated to ageing is subjective to debate and might even not be possible to determine given the complex interplay between the different biological levels on which ageing acts.

The proposed network approach into perspective

Network analyses have clear advantages over individual gene analyses, as they enable the incorporation of useful prior knowledge, which can be exploited for improving the robustness of the analysis and the subsequent interpretation of the results. The improved robustness of the network approach over the individual-gene analyses was reflected by the low mutual overlap between the individual gene results (Figure 1) as opposed to the high concordance between the results obtained in the compendium and replication cohort. The advantages for the interpretation were clearly illustrated by the modest insights gained from the two different strategies for individual gene analysis (*"Glycosylation site:N-linked"*), as opposed to the detailed gene modules produced by our approach that can serve as a novel basis for further investigation into the molecular mechanisms underlying normative ageing. Moreover, our approach is capable of inferring biological coherence from the data, without the explicit need of predefined functional groupings, as was shown by the enrichments of the identified modules found for genes within the GenAge database.
Though the analysis benefits from incorporating Protein-Protein Interaction data, the type and source clearly affect the results. To be as inclusive as possible for types and sources of PPI data, we have chosen to employ data obtained from the STRING database, which systematically collects and integrates interaction data derived from various sources for predicting functional relations between gene pairs. This choice results in a vast and comprehensive source of data. However, STRING data is not confined to physical interactions, as is the case with for instance IntAct (http://www.ebi.ac.uk/intact/) and unlike KEGG (http://www.genome.jp/kegg/), STRING data is not manually curated. For network inference, a trade-off exists between the sparsity and the quality of

the employed gene-gene interactions. We made use of a threshold on the quality of reported interactions that are created by STRING by benchmarking the different interaction data sources to KEGG. Varying this threshold would affect the size and nature of the obtained co-expressed PPI modules. As the threshold determines the scale of the analysis, an interesting observation is that the results can be confounded to parts of the global network that do not necessarily overlap with the predefined known biological pathways. The latter is illustrated by the fact that some of our modules are not enriched for biological pathways and could basically be valued as a strong point of our data-driven approach.

Conclusion

By applying a network approach to multiple blood transcriptomics datasets we have identified five co-expression PPI modules that associate with chronological age in humans. The confirmation of most of our findings in an independent dataset underpins the robustness of our approach. The modules are significantly enriched for ageing-related genes as curated by the GenAge database. This implies that these age-related single genes, in the absence of a clear understanding of their joint functioning belong to a network that finds its basis in protein-protein interactions and will serve as novel input for ageing research. We reinforced the biological relevance of one of the modules by showing that it associates with prospective survival beyond 90 years in humans as was observed also for a single known age-related gene in this module (*ASF1A*). These findings collectively warrant further investigations into the biological function of module F and its potential as a biomarker for healthy ageing and human longevity.

Materials & Methods

Creating the blood expression compendium

Analyses were based on gene expression data derived from individuals enrolled in three large cohort studies for which details on sample inclusion and employed expression protocols are provided in depth in the original publications (19-21). Gene expression and accompanying phenotypic data was obtained from either the original authors or from the public data repository ArrayExpress. Data quality was stringently re-examined per dataset for the presence of outlier samples or outlier measurements and annotated to a common annotation standard (EntrezGeneID). A detailed description of the data processing and an overview on the resulting sample statistics is given in the Supplemental Methods and Table 1 respectively.

Rank integration approach

A rank integration approach (5;22) was used to identify genes consistently up- or down-regulated with age across multiple heterogeneous datasets. This type of meta-analysis integrates individual-gene statistics across datasets, by ranking the statistics per dataset and assessing the significance of the observed combined ranking using a Gamma distribution (45), or through permutation. Gender adjusted linear fits between expression and age were used as gene statistics that were obtained by fitting the following multivariate linear regression model:

$$E_{ijk} = \beta_{0ik} + \beta_{1ik} G_{jk} + \beta_{2ik} A_{jk} + \varepsilon_{ijk} \quad \text{(model 1)}$$

where E_{ijk} is the gene expression of gene i for individual j in the k^{th} dataset, with $1 \le i \le M$, $1 \le j \le N$ and $1 \le k \le K$, where G_{jk} and A_{jk} are the gender and age of individual j in the k^{th} dataset respectively and where ε_{ijk} is the residual error of gene i for individual j in the k^{th} dataset. Genes were ranked on the regression coefficients between age and expression, $\beta_{2ik.}$ The rank position of gene i in dataset k is denoted by R_{ik}. Ranks across the datasets were integrated per gene by computing rank product statistics as previously defined by Koziol (45):

$$RP_i = \sum_{k=1}^{K} \log(R_{ik}) \quad \text{(model 2)}$$

The significance of the observed rank products was assessed in two ways. Following Koziol, rank products RP_i were transformed using:

$$U_i = -RP_i + K \times \log(M+1) \quad \text{(model 3)}$$

The significance of the U-statistics could be assessed by employing the Gamma distribution (45) or through permutation as described in the Supplemental Methods.

Extracting co-expressed PPI modules

Genes were mapped to the Protein-Protein Interaction network (STRING v9.0), which yielded a compendium of about 81.3% of the initial set of genes (N = 7,353) in the compendium. Ranked co-expression matrices were computed for each dataset separately by computing a correlation matrix composed of first order partial correlations between all pairs of genes adjusted for sex and subsequently assigned a rank to each of them. A higher positive correlation resulted in a higher ranking. The ranked co-expression matrices were integrated by computing rank products as in the section on individual-gene analysis. The resulting gene-gene rank product matrix together with the PPI network matrix was subsequently used as

input for the method that identifies co-expressed PPI sub-networks as described in Van den Akker et al. (17), see also Supplemental Methods. In short, a cluster analysis on the gene-gene rank product matrix yielded co-expressed modules of genes. High confidence co-expressed genes were obtained by applying a threshold on the gene-gene rank product matrix. We obtained co-expressed PPI modules by intersecting the co-expressed gene-modules with the PPI network matrix. Co-expressed PPI modules were subsequently visualized using Cytoscape (Supplemental Methods).

Fixed effect meta-analysis on module expressions across the blood compendium

Gene expression data was summarized per co-expressed PPI module for each dataset separately by taking the mean expression per individual over all genes in the module, resulting in a module expression for each dataset. Associations with age were tested for each co-expressed PPI module, by performing a fixed effect meta-analysis across the four datasets using a first order partial correlation between age and the module expression, computed with the controlling variable gender to adjust for sex differences. Per dataset *k,* we thus computed:

$$\rho_{a_k m_k \bullet g_k} = \frac{\rho_{a_k m_k} - \rho_{a_k g_k}\rho_{m_k g_k}}{\sqrt{1-(\rho_{a_k g_k})^2}\sqrt{1-(\rho_{m_k g_k})^2}} \qquad \text{(model 4)}$$

where $\rho_{a_k m_k}$ is the correlation between age and the expression of the n^{th} module across individuals of the k^{th} dataset; $\rho_{a_k g_k}$ is the correlation between age and gender across individuals of the k^{th} dataset and $\rho_{m_k g_k}$ is the correlation between expression of the m^{th} module and gender of individuals in the k^{th} dataset. To correct for multiple controlling variables, higher order partial correlations were computed by repeatedly computing first order partial correlations as described above. The function *metacor* of R package *meta* was used for integrating and testing the meta correlation statistic between age and module expression across the four datasets using default settings. Modules with significant correlations (bonferroni corrected *p*-value ≤ 0.05) were considered age dependent.

Acknowledgements

The research leading to these results has received funding from the Medical Delta (COMO) and the European Union's Seventh Framework Programme (FP7/2007-2011) IDEAL-ageing under grant agreement n°

259679. This study was supported by a grant from the Innovation-Oriented Research Program on Genomics (SenterNovem IGE05007), the Centre for Medical Systems Biology, 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. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.
The authors thank the participants of the SAFHS, IFB, DILGOM, LLS, NTR and NESDA studies for their contributions. Furthermore, we would like to thank Michael Inouye for helpful discussions.

References

(1) Oeppen J, Vaupel JW. Demography. Broken limits to life expectancy. Science 2002 May 10;296(5570):1029-31.

(2) Hitt R, Young-Xu Y, Silver M, Perls T. Centenarians: the older you get, the healthier you have been. Lancet 1999 August 21;354(9179):652.

(3) Wilson PW, D'Agostino RB, Levy D, Belanger AM, Silbershatz H, Kannel WB. Prediction of coronary heart disease using risk factor categories. Circulation 1998 May 12;97(18):1837-47.

(4) Kirkwood TB. Evolution of ageing. Nature 1977 November 24;270(5635):301-4.

(5) de Magalhaes JP, Curado J, Church GM. Meta-analysis of age-related gene expression profiles identifies common signatures of aging. Bioinformatics 2009 April 1;25(7):875-81.

(6) Passtoors WM, Beekman M, Gunn D, Boer JM, Heijmans BT, Westendorp RG et al. Genomic studies in ageing research: the need to integrate genetic and gene expression approaches. J Intern Med 2008 February;263(2):153-66.

(7) Partridge L, Gems D. Mechanisms of ageing: public or private? Nat Rev Genet 2002 March;3(3):165-75.

(8) Zahn JM, Sonu R, Vogel H, Crane E, Mazan-Mamczarz K, Rabkin R et al. Transcriptional profiling of aging in human muscle reveals a common aging signature. PLoS Genet 2006 July;2(7):e115.

(9) Zhang B, Horvath S. A general framework for weighted gene co-expression network analysis. Stat Appl Genet Mol Biol 2005;4:Article17.

(10) Southworth LK, Owen AB, Kim SK. Aging mice show a decreasing correlation of gene expression within genetic modules. PLoS Genet 2009 December;5(12):e1000776.

(11) Stuart JM, Segal E, Koller D, Kim SK. A gene-coexpression network for global discovery of conserved genetic modules. Science 2003 October 10;302(5643):249-55.

(12) Bell R, Hubbard A, Chettier R, Chen D, Miller JP, Kapahi P et al. A human protein interaction network shows conservation of aging processes between human and invertebrate species. PLoS Genet 2009 March;5(3):e1000414.

(13) de Magalhaes JP, Toussaint O. GenAge: a genomic and proteomic network map of human ageing. FEBS Lett 2004 July 30;571(1-3):243-7.

(14) Tacutu R, Shore DE, Budovsky A, de Magalhaes JP, Ruvkun G, Fraifeld VE et al. Prediction of C. elegans longevity genes by human and worm longevity networks. PLoS One 2012;7(10):e48282.

(15) Witten TM, Bonchev D. Predicting aging/longevity-related genes in the nematode Caenorhabditis elegans. Chem Biodivers 2007 November;4(11):2639-55.

(16) Xue H, Xian B, Dong D, Xia K, Zhu S, Zhang Z et al. A modular network model of aging. Mol Syst Biol 2007;3:147.

(17) van den Akker EB, Verbruggen B, Heijmans BT, Beekman M, Kok JN, Slagboom PE et al. Integrating protein-protein interaction networks with gene-gene co-expression networks improves gene signatures for classifying breast cancer metastasis. J Integr Bioinform 2011;8(2):188.

(18) Marbach D, Costello JC, Kuffner R, Vega NM, Prill RJ, Camacho DM et al. Wisdom of crowds for robust gene network inference. Nat Methods 2012 August;9(8):796-804.

(19) Emilsson V, Thorleifsson G, Zhang B, Leonardson AS, Zink F, Zhu J et al. Genetics of gene expression and its effect on disease. Nature 2008 March 27;452(7186):423-8.

(20) Inouye M, Silander K, Hamalainen E, Salomaa V, Harald K, Jousilahti P et al. An immune response network associated with blood lipid levels. PLoS Genet 2010 September;6(9).

(21) Goring HH, Curran JE, Johnson MP, Dyer TD, Charlesworth J, Cole SA et al. Discovery of expression QTLs using large-scale transcriptional profiling in human lymphocytes. Nat Genet 2007 October;39(10):1208-16.

(22) Breitling R, Armengaud P, Amtmann A, Herzyk P. Rank products: a simple, yet powerful, new method to detect differentially regulated genes in replicated microarray experiments. FEBS Lett 2004 August 27;573(1-3):83-92.

(23) Harries LW, Hernandez D, Henley W, Wood AR, Holly AC, Bradley-Smith RM et al. Human aging is characterized by focused changes in gene expression and deregulation of alternative splicing. Aging Cell 2011 October;10(5):868-78.

(24) Hong MG, Myers AJ, Magnusson PK, Prince JA. Transcriptome-wide assessment of human brain and lymphocyte senescence. PLoS One 2008;3(8):e3024.

(25) Passtoors WM, Boer JM, Goeman JJ, van den Akker EB, Deelen J, Zwaan BJ et al. Transcriptional Profiling of Human Familial Longevity Indicates a Role for ASF1A and IL7R. PLoS One 2012;7(1):e27759.

(26) Dall'olio F, Vanhooren V, Chen CC, Slagboom PE, Wuhrer M, Franceschi C. N-glycomic biomarkers of biological aging and longevity: A link with inflammaging. Ageing Res Rev 2012 February 14.

(27) Derhovanessian E, Maier AB, Beck R, Jahn G, Hahnel K, Slagboom PE et al. Hallmark features of immunosenescence are absent in familial longevity. J Immunol 2010 October 15;185(8):4618-24.

(28) Boomsma DI, Willemsen G, Sullivan PF, Heutink P, Meijer P, Sondervan D et al. Genome-wide association of major depression: description of samples for the GAIN

Major Depressive Disorder Study: NTR and NESDA biobank projects. Eur J Hum Genet 2008 March;16(3):335-42.

(29) Smith ED, Tsuchiya M, Fox LA, Dang N, Hu D, Kerr EO et al. Quantitative evidence for conserved longevity pathways between divergent eukaryotic species. Genome Res 2008 April;18(4):564-70.

(30) Baker DJ, Jeganathan KB, Malureanu L, Perez-Terzic C, Terzic A, van Deursen JM. Early aging-associated phenotypes in Bub3/Rae1 haploinsufficient mice. J Cell Biol 2006 February 13;172(4):529-40.

(31) Chen Y, Liu W, McPhie DL, Hassinger L, Neve RL. APP-BP1 mediates APP-induced apoptosis and DNA synthesis and is increased in Alzheimer's disease brain. J Cell Biol 2003 October 13;163(1):27-33.

(32) McBride HM, Neuspiel M, Wasiak S. Mitochondria: more than just a powerhouse. Curr Biol 2006 July 25;16(14):R551-R560.

(33) Fan AC, Kozlov G, Hoegl A, Marcellus RC, Wong MJ, Gehring K et al. Interaction between the human mitochondrial import receptors Tom20 and Tom70 in vitro suggests a chaperone displacement mechanism. J Biol Chem 2011 September 16;286(37):32208-19.

(34) Curran SP, Ruvkun G. Lifespan regulation by evolutionarily conserved genes essential for viability. PLoS Genet 2007 April 6;3(4):e56.

(35) Clemens MJ. Initiation factor eIF2 alpha phosphorylation in stress responses and apoptosis. Prog Mol Subcell Biol 2001;27:57-89.

(36) Laplante M, Sabatini DM. mTOR signaling at a glance. J Cell Sci 2009 October 15;122(Pt 20):3589-94.

(37) Johnson SC, Rabinovitch PS, Kaeberlein M. mTOR is a key modulator of ageing and age-related disease. Nature 2013 January 17;493(7432):338-45.

(38) Passtoors WM, Beekman M, Deelen J, van der Breggen R, Maier AB, Guigas B et al. Gene expression analysis of mTOR pathway: association with human longevity. Aging Cell 2013 February;12(1):24-31.

(39) von KC, Harrigan JA, May A, Opresko PL, Dawut L, Cheng WH et al. Central role for the Werner syndrome protein/poly(ADP-ribose) polymerase 1 complex in the poly(ADP-ribosyl)ation pathway after DNA damage. Mol Cell Biol 2003 December;23(23):8601-13.

(40) Grube K, Burkle A. Poly(ADP-ribose) polymerase activity in mononuclear leukocytes of 13 mammalian species correlates with species-specific life span. Proc Natl Acad Sci U S A 1992 December 15;89(24):11759-63.

(41) Pawelec G, Solana R. Immunosenescence. Immunol Today 1997 November;18(11):514-6.

(42) Napoli AM, Fast LD, Gardiner F, Nevola M, Machan JT. Increased granzyme levels in cytotoxic T lymphocytes are associated with disease severity in emergency department patients with severe sepsis. Shock 2012 March;37(3):257-62.

(43) Calnan DR, Brunet A. The FoxO code. Oncogene 2008 April 7;27(16):2276-88.

(44) Chi H. Regulation and function of mTOR signalling in T cell fate decisions. Nat Rev Immunol 2012 May;12(5):325-38.

(45) Koziol JA. Comments on the rank product method for analyzing replicated experiments. FEBS Lett 2010 March 5;584(5):941-4.

Supplementary information

Supplementary figure and tables can be found online:
http://onlinelibrary.wiley.com/resolve/doi?DOI=10.1111/acel.12160

Text S1	*Supplemental Methods*
Table S2	*Results of 26 significantly age-associated genes in all four datasets*
Table S3	*Results of the individual-gene rank product test*
Table S4	*Enrichment analyses on the 26 significantly age-associated genes in all four datasets*
Table S5	*Enrichment analyses on 195 significant genes obtained with an individual-gene rank product test*
Figure S6	*Overview of detected consistently co-expressed PPI modules.*
Table S7	*Gene enrichment analyses using DAVID on the 27 consistently co-expressed PPI modules counting at least 5 genes*
Table S8	*Replication of the identified modules as robust markers of chronological age in the NTR & NESDA cohort.*
Table S9	*GenAge enrichment analyses of identified co-expressed PPI modules.*

Chapter 8

General discussion

The aim of this thesis was to identify in the human blood transcriptome, relevant pathways and potential biomarker profiles that associate with chronological age and discriminate between 'healthy agers' from long-lived families and normative ageing controls. Such profiles may harbour determinants of the biological ageing rate.

We studied genome-wide gene expression profiles in blood of members of the Leiden Longevity Study (LLS) and replicated our findings by extended sampling within the unique LLS cohort. The findings of the exploratory analysis prompted us to investigate multiple genes in the *IL7R* and *MTOR* pathways for association with familial longevity. The results obtained by examining mRNA from blood samples brought us to study mTOR protein levels and signalling in primary skin fibroblasts from the corresponding donors in the LLS. Finally, to discover robust, biologically relevant gene networks as markers of chronological ageing in larger sample sizes, we performed an explorative network-based meta-analysis on large publicly available transcriptomic datasets.

Main findings

Chapter 3 describes the genome-wide explorative transcriptomic study performed within the Leiden Longevity Study cohort. The explorative study resulted in 360 probes (259 genes) to be associated with familial longevity, out of 2,953 probes (1,853 genes) differentially expressed in nonagenarians and middle-aged controls. We selected 22 genes from the 259 for follow up studies by using a text mining tool applying criteria that genes have known connections with cellular ageing. In Chapter 3 therefore we only investigated a fraction of the total number of potential transcriptomics biomarkers associated with healthy ageing, which will be described later in this discussion.

The microarray study showed that reduced expression of *ASF1A* and *IL7R* marked both chronological ageing and familial longevity in middle-age. The follow up experiments for *ASF1A* included RNAi experiments in *c. elegans* (David Gems, personal communication); no homologue gene is present in worms for *IL7R*. Two homologues of the *ASF1A* gene, *asfl-1* and *unc-85* were knocked down using RNAi in young adult worms and survival was measured in both the RNAi worms and the controls. Though downregulation of *asfl-1* resulted in a small significant increase in lifespan, the combination of knocking down both *asfl-1* and *unc-85* did not result in any change in lifespan. These results most likely illustrate random variation of nematode lifespan.

To then expand the study of the *IL7R* pathway in humans, we measured expression levels of six genes interacting with *IL7R* according to the STRING software and tested this network for association with familial longevity in the LLS (**Chapter 4**). We found several genes in the network to be associated with familial longevity in middle-age indeed, independent of the difference in blood cell proportion and disease prevalence between offspring of nonagenarians and their partners. Hence we conclude that expression levels of *IL7R* in blood may represent a marker of biological age (detecting differences between healthy and normative ageing). Subsequent measurement in the future on the complete LLS and additional cohort studies will enable an extensive survival analysis to be performed to indicate whether expression levels of these genes mark mortality risks and to test for the association with morbidity. The same analysis steps have revealed for example the leucocyte telomere length as a biomarker of ageing (1;2).

MTOR signalling has been implicated extensively in lifespan regulation in studies of animal models and in disease prevalence and progression in both animal models and humans. In our study, *MTOR* gene expression was part of the “normative ageing-signature” resulting from the explorative analysis (**Chapter 3**). Therefore we further investigated mTOR signalling in the Leiden Longevity Study (**Chapter 5**). Expression of genes belonging to both complexes mTORC1 and mTORC2 or downstream thereof showed association with chronological age in the LLS. The observed relation of decreased transcriptional activity of mTORC1 and lifespan extension described in the literature on animal models was in agreement with the low expression level in the human longevity families. The increased transcriptional activity of mTORC2 we observed with chronological ageing was not described previously. For two of the genes, *RPTOR* and *PRR5L*, belonging to respectively the mTORC1 and mTORC2 part of the pathway, we found decreased expression to be associated with familial longevity in middle-age, which association only for *RPTOR* was independent of differences between offspring and partners in cell proportions, glucose levels and prevalence of type 2 diabetes.

The mTOR pathway has been extensively studied in cell lines before, and previous studies in the LLS cohort demonstrated different cellular responses to stress in skin fibroblasts of the offspring and control groups. We questioned whether primary fibroblasts from LLS donors would also reflect reduced transcriptional activity of the mTOR pathway, altered activity at the level of protein expression and whether such a cellular model would allow further studies to link mTOR pathway genes to the metabolic

characteristics of longevity families. Thus, as a functional follow-up of the conclusion that gene expression of mTOR pathway genes in blood seem to mark familial longevity, gene and protein expression of mTOR-related genes were investigated in cultured primary skin fibroblasts of donors from the LLS (**Chapter 6**). Though fibroblasts of LLS offspring of nonagenarians and controls did show differences in cellular senescence and apoptosis previously, we were not able to detect differential mTOR gene or protein expression levels between these groups, neither did we observe any correspondence of variation in expression of mTOR genes in blood and fibroblasts of the same subjects in a group. Our main conclusion was that mTOR signalling in blood and skin fibroblasts appear to be differentially regulated. We included an indirect oxidative stress challenge by adding rotenone for 3 days in the fibroblast experiments. This challenge did not enhance any potential mTOR gene or protein expression difference in the cells from offspring and controls. Such challenges may be tested more extensively using different concentration and incubation time of rotenone, as described in previous senescence and apoptosis experiments within the LLS study (3).

Potentially suitable next approaches for functional studies include the investigation of protein expression or downstream processes like autophagy directly in Peripheral Blood Mononuclear Cells. Challenge tests such as nutrient deprivation or stress induction in the fibroblast cell system might otherwise enhance differences between the groups. Another option would be to investigate the mTOR pathway expression and activity in other tissues such as fat and muscle in which mTOR signalling is active. In these tissues nutrient intake, glucose and insulin levels and exercise influence protein synthesis, energy metabolism and glucose uptake (4), all processes in which the mTOR pathway is involved and which play a role in biological ageing and health.

In **Chapter 7** we describe an integrative network-based approach to discover robust markers for normative ageing combining multiple large-scale expression studies in blood. The combination of co-expression of genes and protein-protein interaction networks resulted in five robust networks enriched for "*Translational elongation*", "*Cytolysis*" and "*DNA metabolic process*" associated with normative ageing. One of the identified modules predominantly consisted of ribosomal proteins and translation elongation factors. This module was found to be down-regulated with chronological age and suggested age-associated limitation of protein synthesis. Another major player in protein synthesis and ageing is mTOR signalling. Not only does mTORC1 regulates protein synthesis, mTORC2

activity is influenced by ribosomes (5), although the exact way in which this happens is unclear at this moment. The next step would be to establish whether the five co-expressed PPI networks that may be considered as robust markers for normative ageing, are also markers for biological age. In this regard the network including *ASF1A* already showed an association with survival at old age, suggesting their gene expression to be indeed markers for biological ageing.

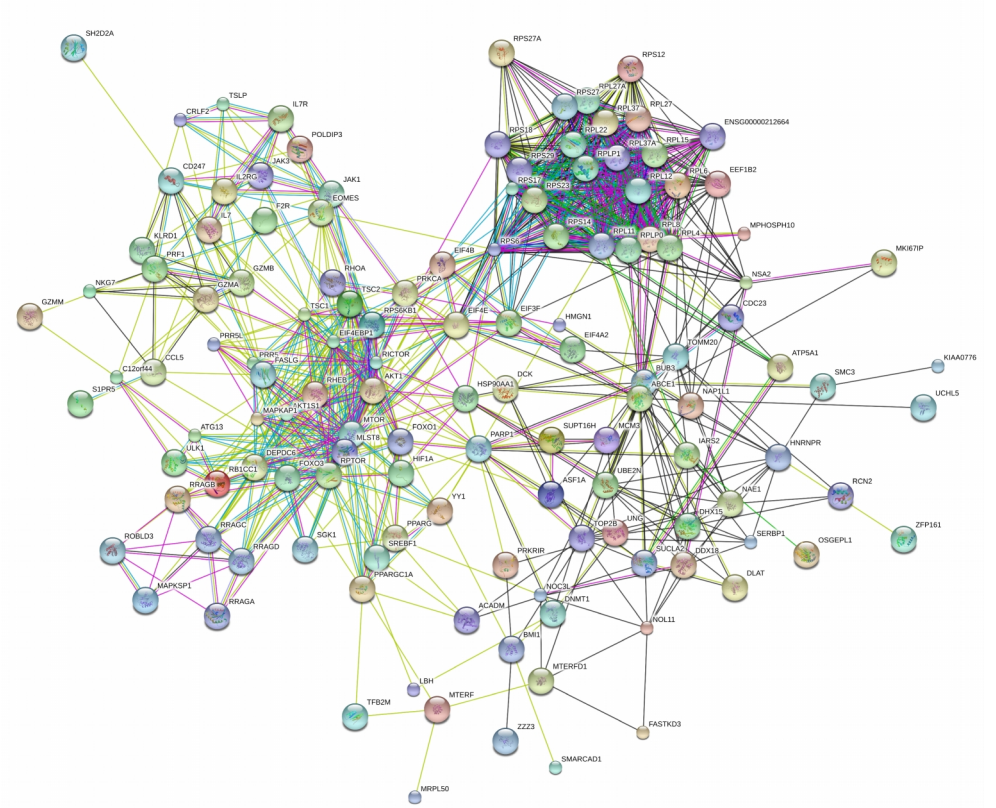

Figure 1. Total PPI network for the genes contributing to the mTOR pathway, IL7R network and the five robust co-expressed PPI modules

Connecting our findings and conclusions

We have identified five networks of which the gene expression levels in blood are robustly associated with chronological ageing, including genes involved in "*Translational elongation*", "*Cytolysis*" and "*DNA metabolic process*". In addition three major candidate genes were found for which the

transcription in blood differed between longevity families and controls, thus potentially marking the biological age and the rate of ageing: *ASF1A*, *IL7R* and *RPTOR*. Prospective studies relating the expression levels to morbidity and mortality will reveal the value of these levels as biomarkers of ageing.

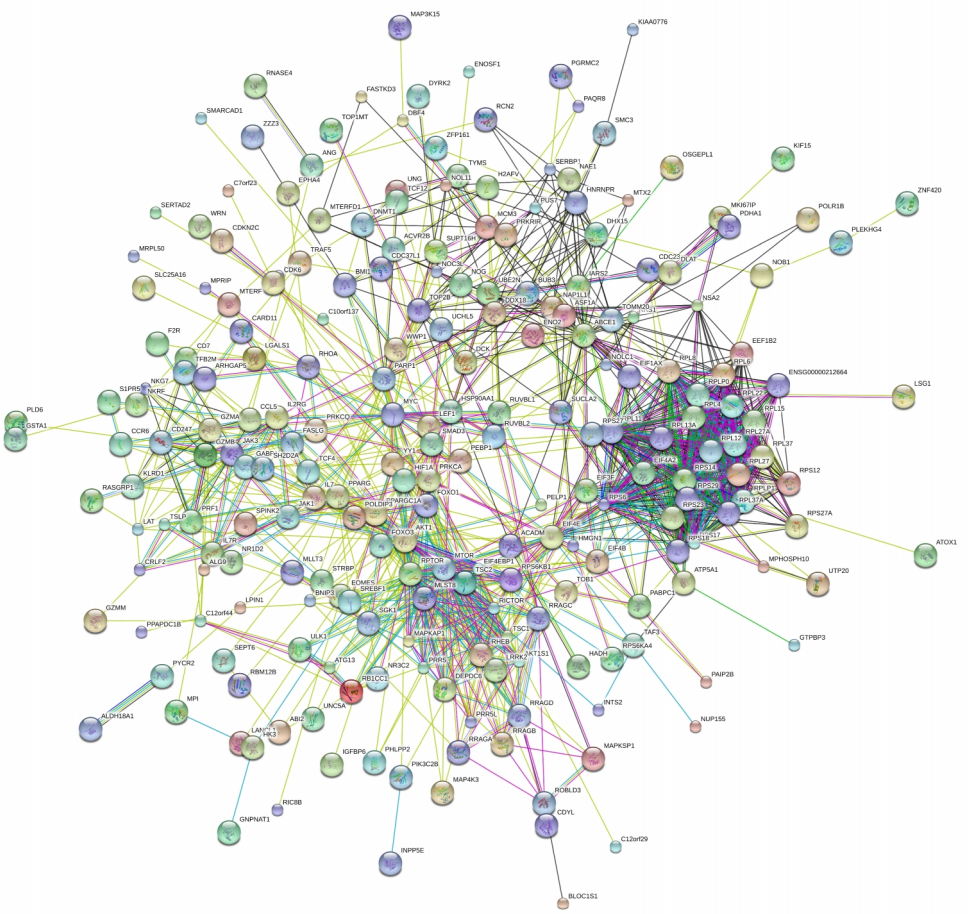

Figure 2. Total PPI network for the genes contributing to the mTOR pathway, IL7R network and the five robust co-expressed PPI modules and 99 potential marker genes

To prioritize the 236 uninvestigated potential marker genes that resulted from the explorative analyses in Chapter 3, we explored their relation with the interactive networks identified in Chapter 4, 5 and 7. By using STRING software we created a protein-protein interaction network for the genes contributing to the mTOR pathway, IL7R network and the five robust co-expressed PPI modules (Figure 1). Next we explored whether the 236

uninvestigated potential marker genes connected to the normative ageing networks. We found that 230 genes were recognized by STRING, 8 of which were already part of the mTOR or IL7R network or were part of a network found in the meta-analysis in chapter 7 (namely *MTOR*, *IL7R*, *ABCE1*, *ASF1A*, *DCK*, *LBH*, *MPHOSPH10* and *ZZZ3*) and 99 showed direct interactions with networks of normative ageing (Figure 2). Thus almost half of the genes that resulted as potential marker genes from chapter 3 showing association with chronological age and familial longevity connected to robust networks of normative chronological ageing resulting from the meta-analysis, plus the mTOR and IL7R networks. The majority of the connected genes showed at least two interactions enhancing the evidence that they connect to these networks. Table 1 shows the 99 genes prioritized for future replication studies. The other 123 genes do not show any interaction with the normative ageing network.

Remarkably, five of the interacting genes that have not yet been tested for replication in Chapter 3, *ABCE1*, *DCK*, *ASF1A*, *MPHOSPH10* (*MPP10*) and *ZZZ3*, are part of the robust age-associated co-expressed PPI module F in the meta-analysis (Chapter 7, Figure 2, module F). This module was not enriched for any biological process but seems to be related to the mitochondrion and cell cycle. Since these five genes have not only been associated with normative chronological ageing in Chapter 3 and the robust meta-analysis of Chapter 7, but also with familial longevity (Chapter 3), these genes may be the best next genes to be studied as potential markers of biological ageing and longevity.

Table 1. The 99 genes prioritized for future replication studies

Gene name							
ABI2	CCR6	GSTA1	LPIN1	NOLC1	PLD6	RRS1	TOP1MT
ACVR2B	CD7	GTPBP3	LRRK2	NR1D2	PLEKHG4	RUVBL1	TRAF5
ALDH18A1	CDC37L1	H2AFV	LSG1	NR3C2	POLR1B	RUVBL2	TYMS
ALG9	CDK6	HADH	MAP3K15	NUP155	PPAPDC1B	SEPT6	UNC5A
ANG	CDYL	HK3	MAP4K3	PABPC1	PRKCQ	SERTAD2	UTP20
ARHGAP5	DBF4	IGFBP6	MLLT3	PAIP2B	PUS7	SLC25A16	WRN
ATOX1	DYRK2	INPP5E	MPI	PAQR8	PYCR2	SMAD3	WWP1
BLOC1S1	EIF1AX	INTS2	MPRIP	PDHA1	RASGRP1	SPINK2	ZNF420
BNIP3	ENO2	KIF15	MTX2	PEBP1	RBM12B	STRBP	
C10orf137	ENOSF1	LANCL1	MYC	PELP1	RIC8B	TAF3	
C12orf29	EPHA4	LAT	NKRF	PGRMC2	RNASE4	TCF12	
C7orf23	GABPA	LEF1	NOB1	PHLPP2	RPL13A	TCF4	
CARD11	GNPNAT1	LGALS1	NOG	PIK3C2B	RPS6KA4	TOB1	

Implications and future research

One of the main pathways resulting from animal research and from this thesis to mark normative ageing is the mTOR signaling pathway, for which we identified the gene expression of *RPTOR* in blood to be a potential biomarker for biological ageing in middle-age. We observed that the decreased transcriptional activity of the MTOR pathway in members of long-lived families corresponds to the lower prevalence of cardiovascular disease and type 2 diabetes (6), and increased sensitivity to insulin (7). As compared to their partners, matched for age and social and lifestyle factors, the longevity families further display a beneficial lipid profile (8) and their cells are better resistant to stress (3). This pattern of phenotypic characteristics resembles the characteristics of caloric restricted mice (9), including the lower mTOR signaling and the longer lifespan. The interesting difference is that in the human studies the longevity family members and partners do not differ in BMI or in calorie intake. We hypothesized that members of long-lived families have an endogenously beneficial setting of their metabolism and hence protection from diseases. The question rises whether the controls (representing normative ageing subjects) are able to adjust the settings of their nutrient sensing pathways by behavioral changes.

Growing Old Together

Based on the finding that nutrient sensing and mTOR signaling as in animal models is relevant for human ageing, we initiated intervention studies within the LLS cohort. To investigate whether the partners of the offspring may be able to improve their metabolic health by diet intervention and physical activity, we founded the “Growing Old Together” Study. Middle-aged couples of the Leiden Longevity Study; offspring of nonagenarian siblings and their partners, are subjected to an intervention of dietary restriction and physical exercise for 13 weeks, a 25% lowered energy expenditure in total. By examining the potential biomarkers of biological ageing identified in this thesis before and after the intervention, we may be able to determine whether improvement of metabolism is acquired by the intervention, for which participants this is the case and whether the genes identified in this thesis contribute to such improvement.
Since blood may not be the best tissue to investigate mTOR signaling, in the “Growing Old Together” study also abdominal fat and skeletal muscle tissue are being collected. Since skeletal muscle is the major site of glucose disposal in response to food intake and insulin, and since an impairment of glucose uptake in this tissue contributes to type 2 diabetes,

we will investigate whether expression differences in mTOR genes between offspring and controls are more pronounced in muscle tissue. Also, we will investigate which candidate genes for biomarkers for the rate of ageing are the best in marking metabolic health.

Future indications

In general, altered gene expression levels in the longevity family members may be the consequence of ageing, or merely a trait shared by the long-lived families or may truly contribute to human longevity. To investigate whether changes in potential biomarkers causally contribute to ageing, gene expression of specific genes should be measured quantitatively in blood of a large, long-term population-based longitudinal study from the age of 40 years onwards. Now that we have a number of candidate genes such studies may indeed be performed to establish the association with prospective morbidity and mortality. Ideally, longitudinal measurements should be performed for example every five years, including collection of information about cell type proportions in the samples and information regarding the subjects' health. Such analyses could be done in large ongoing cohort studies such as the Rotterdam Study cohort (10) and, since the LLS study was initiated over 10 years ago, includes highly and middle-aged subjects, also in the LLS study. The future longitudinal study of genomic biomarkers may include also quantitative measures of epigenetic variation (such as DNA methylation) and should go hand in hand with measures of novel physiological parameters of biological ageing, as reflected by proteomics and metabolomic measures (11). Novel markers of biological ageing will stimulate studies into identification of determinants and mechanistic factors causally involved in biological ageing. Potentially causal factors will need to be investigated in animal models in which all tissues can be studied and not as in healthy humans, just blood, skin, muscle and fat tissue. In future studies of human longevity families, next generation sequencing of both DNA and RNA can provide further insight into the genetic variation underlying the gene expression differences we observed and test whether this could form the basis of the familial longevity.

References

(1) Cawthon RM, Smith KR, O'Brien E, Sivatchenko A, Kerber RA. Association between telomere length in blood and mortality in people aged 60 years or older. Lancet 2003 Feb 1;361(9355):393-5.

(2) Fitzpatrick AL, Kronmal RA, Kimura M, Gardner JP, Psaty BM, Jenny NS, et al. Leukocyte telomere length and mortality in the Cardiovascular Health Study. J Gerontol A Biol Sci Med Sci 2011 Apr;66(4):421-9.

(3) Dekker P, Maier AB, van HD, de Koning-Treurniet C, Blom J, Dirks RW, et al. Stress-induced responses of human skin fibroblasts in vitro reflect human longevity. Aging Cell 2009 Sep;8(5):595-603.

(4) Laplante M, Sabatini DM. mTOR signaling in growth control and disease. Cell 2012 Apr 13;149(2):274-93.

(5) Zinzalla V, Stracka D, Oppliger W, Hall MN. Activation of mTORC2 by association with the ribosome. Cell 2011 Mar 4;144(5):757-68.

(6) Westendorp RG, van HD, Rozing MP, Frolich M, Mooijaart SP, Blauw GJ, et al. Nonagenarian siblings and their offspring display lower risk of mortality and morbidity than sporadic nonagenarians: The Leiden Longevity Study. J Am Geriatr Soc 2009 Sep;57(9):1634-7.

(7) Wijsman CA, Rozing MP, Streefland TC, le CS, Mooijaart SP, Slagboom PE, et al. Familial longevity is marked by enhanced insulin sensitivity. Aging Cell 2011 Feb;10(1):114-21.

(8) Vaarhorst AA, Beekman M, Suchiman EH, van HD, Houwing-Duistermaat JJ, Westendorp RG, et al. Lipid metabolism in long-lived families: the Leiden Longevity Study. Age (Dordr) 2010 Sep 3.

(9) Blagosklonny MV. Calorie restriction: decelerating mTOR-driven aging from cells to organisms (including humans). Cell Cycle 2010 Feb 15;9(4):683-8.

(10) Hofman A, Grobbee DE, de Jong PT, van den Ouweland FA. Determinants of disease and disability in the elderly: the Rotterdam Elderly Study. Eur J Epidemiol 1991 Jul;7(4):403-22.

(11) Barzilai N, Guarente L, Kirkwood TB, Partridge L, Rando TA, Slagboom PE. The place of genetics in ageing research. Nat Rev Genet 2012 Aug;13(8):589-94.

Chapter 9

Summary

The most prevalent diseases in the Western world have in common that age is the major risk factor. Life expectancy has increased dramatically in the last two centuries and will continue to rise. The healthy lifespan in western societies however, is usually not more than 75% of a lifetime. Many molecular changes occur as a function of chronological age in all tissues. Such changes could represent a marker of the biological ageing rate. The most valuable biomarkers of ageing rate would associate not only with chronological age and health status, but also with morbidity and mortality in prospective studies. To find these biomarkers we investigated chronological and biological ageing in long-lived families, collected in the Leiden Longevity Study (LLS). The LLS consists of nonagenarian sibling pairs, their middle-aged offspring and the partners thereof which act as controls. As compared to controls, the offspring has a lower prevalence of myocardial infarction, hypertension and type 2 diabetes and a lower mortality risk and are therefore regarded as healthy agers.

Transcriptomic studies for identification of human pathways and biomarkers of ageing have so far mainly focused on chronological ageing in several tissues (the largest studies investigated blood) and reported on enrichment for pathways like immune response, energy metabolism, mitochondrial processes and post-transcriptional regulation which may harbor potential biomarkers for chronological ageing. These studies have been reviewed in chapter 2 of this thesis. In this thesis we aimed for identification of relevant pathways and potential biomarker profiles that associate with chronological age in human populations and we tested which of these discriminate between healthy ageing members from long-lived families and normative ageing controls by studying the blood transcriptome using various approaches. Such pathways and biomarkers may harbor determinants of the biological ageing rate.

In chapter 3 we describe a genome-wide gene expression microarray study of 47,209 probes (representing 17,896 known genes) within 50 families of the LLS to screen for genes of which expression changes with age and associated with biological ageing. We identified 2,953 probes that associated with chronological age when comparing gene expression profiles from whole blood samples between 50 nonagenarians and 50 middle-aged controls, the latter representing the general population. From these, the expression of 360 probes was found to change differentially with age in offspring of the nonagenarians as compared to controls. In a RT-qPCR replication experiment utilizing 312 controls, 332 offspring and 79 nonagenarians, we confirmed a ‘‘chronological aging-signature’’ formed of 21 genes, of which reduced expression of *ASF1A* and *IL7R* marked familial

longevity already in middle-age. This indicates that expression changes of *ASF1A* and *IL7R*, which are involved in chromatin modeling and the immune system respectively, represent early features of familial longevity and healthy ageing.

The *IL7R* gene encodes the interleukin 7 receptor, which is important in immune response processes and critical for T cell development. Increased expression of IL7R and IL7 has been associated with prevalence of immune-related diseases. In chapter 4 we explore whether the expression of the *IL7R* gene and six interaction partners in blood associated with familial longevity, and whether the association is dependent on whole blood cell proportions and immune-related diseases including type 2 diabetes, COPD and rheumatoid arthritis. Investigating the expression of these seven genes in blood of the LLS using RT-qPCR resulted in a significant association of the network in a set-based test with chronological age ($p=4.6 \times 10^{-4}$). *IL7R* expression was found to be lower in the offspring as compared to the controls, which is in concordance with the observation that higher *IL7R* gene expression in blood is associated with higher prevalence of immune-related diseases (p-value=0.001). The association of lower IL7R gene expression with familial longevity was not explained by variation between the groups in blood cell counts and in the prevalence of the tested immune-related diseases. Paradoxically, *IL7R* gene expression decreases with age and higher *IL7R* gene expression levels associate with better survival in old and middle age. Taken together, the *IL7R* network reflected by gene expression levels in blood may mark the biological age and health status of elderly individuals, although the complexity of the mechanism still requires elucidation.

In addition to the explorative analysis, we also investigated candidate pathways for human longevity. The pathway that extensively has been implicated in lifespan extension but also development of disease in model organisms is mTOR signaling. This pathway has also been associated with diseases like diabetes and cancer in humans, but its impact on familial longevity or healthy ageing has not been determined yet. In chapter 5 we therefore investigated the gene expression level of the mTOR signaling pathway in blood within the LLS study. Forty genes were measured using RT-qPCR and as a set it showed significant association with chronological age when comparing mRNA levels of nonagenarians and middle-aged controls ($p=4.6 \times 10^{-7}$). Seven out of the 40 mTOR pathway genes had a significant differential expression of at least 5%. Of these, *RPTOR* showed a decreased expression in members of the long-lived families as compared to controls, already in middle-age, independent of prevalence of type 2

diabetes and cancer or variation in glucose levels. At the level of suggestive evidence this was also the case for *PRRL5*. Taken together, the mTOR pathway is not only involved in regulation of disease and lifespan in animal models, but also in humans. The gene expression in blood of *RPTOR* is a potential biomarker for biological ageing.

The mTOR pathway has been broadly studied in cell lines before, and cellular responses to stress in skin fibroblasts of offspring and controls from the LLS were shown to be associated with familial longevity. Therefore we questioned whether the mTOR pathway in this cellular model might reflect the metabolic characteristics of familial longevity. In chapter 6 we described this functional follow-up of the mTOR signaling results in blood, which was performed in skin fibroblasts from middle-aged LLS offspring with their partners as controls. Besides gene expression levels of all 40 mTOR signaling genes, expression level and phosphorylation of the most important proteins in the mTOR pathway were measured and analyzed, as well as total protein content and autophagy markers. However, only decreased total protein content was observed in offspring of long-lived individuals compared to controls. Besides not finding the same difference as was found in blood before, gene expression levels in fibroblasts did not reflect the corresponding protein levels in cultured fibroblasts or gene expression levels in whole blood in the same donors. Hence, the effect of mTOR signaling on human familial longevity may not be present in all tissues.

In chapter 7 we performed a meta-analysis combining multiple large-scale expression studies in blood to discover robust markers for chronological ageing in an integrative network-based approach. The combination of co-expression of genes and protein-protein interaction networks resulted in five robust networks significantly associated with normative ageing and included modules enriched for "*Translational elongation*", "*Cytolysis*" and "*DNA metabolic process*". An independent study in a Dutch cohort resulted in the validation of four out of five networks, indicating the robustness of our results. Whether these networks may be considered as markers for biological ageing as well has yet to be explored.

Striking is that five genes (*ABCE1*, *ASF1A*, *DCK, MPHOSPH10* and *ZZZ3*) of one of the robust networks are also among the 259 possible candidate markers for biological ageing as described in Chapter 3. An independent study also shows this network to be associated to survival at old age. As a relatively unknown group of genes it is important to further investigate this network in regards to function and relationship with health and ageing.

In this thesis we identified several pathways that change their gene expression as a function of chronological age. These may represent potential biomarkers in blood that associate with chronological age in human populations, including genes involved in translational elongation, regulation of transcription and lymphocyte development. We found indication that genes in these pathways may be relevant for ageing or could represent biomarkers for biological ageing, including expression levels of *ASF1A*, *IL7R*, *IL7* and *RPTOR*, which discriminate healthy agers (from longevity families) from normally ageing controls. Our explorative analysis resulted in 259 possible candidate markers for biological ageing of which 23 have been tested for replication in a larger part of the LLS. Of the yet untested genes, 99 were part of a large ageing-associated interaction network based on genes contributing to the mTOR pathway, IL7R network and the five robust co-expressed PPI modules. These 99, but also 4 other genes that formed a centrosome network separate from the large network, need to be validated in a larger part of the study population as possible markers for biological ageing.

A first step to investigate the biological relevance of the observed associations of the nutrient sensing mTOR pathway with familial longevity is initiated by the follow-up of these candidate genes in an intervention study called “Growing Old Together”. Offspring of nonagenarian siblings and their partners are subjected to an intervention of dietary restriction and physical exercise for 13 weeks. By examining the (candidate) biomarkers for the rate of ageing in blood, but also in muscle and fat, before and after the intervention, we aim to determine to what extent the controls change their molecular profile towards that of healthy agers and whether variation in mTOR expression associates to variation in response to the intervention. Such studies are ultimately aimed to stimulate middle aged individuals in the general population to improve their metabolic health and slow down their biological ageing rate and the gene expression profiles studied in this thesis are tested as monitoring tools for such improvements in health.

Chapter 10

Nederlandse samenvatting

Mensen verouderen met een verschillende snelheid. Je kunt het proces aan de buitenkant zien; ouderen krijgen rimpels, worden grijs, bewegen stijver en krijgen te maken met diverse leeftijdsgerelateerde ziekten. Ook aan de binnenkant, in moleculen en cellen van mensen kun je het verouderingsproces aflezen. Zo verschilt ook de activiteit van genen in het erfelijk materiaal tussen ouderen en jonge mensen. Door de verschillen tussen ouderen is de biologische leeftijd niet altijd dezelfde als de kalenderleeftijd. In een groep mensen van 70 jaar zijn grote verschillen aanwezig met betrekking tot de gezondheid, het functioneren en welbevinden, maar ook op het niveau van de cellen en moleculen. De levensverwachting is zeer sterk gestegen in de laatste tweehonderd jaar en veel ouderen van nu maken een jongere indruk dan ouderen in vorige generaties. De stijging in levensverwchting werd veroorzaakt door omgevinggerelateerde oorzaken zoals verbetering van hygiëne, voeding en medische zorg. Men kan zich afvragen of erfelijkheid een rol speelt bij het bepalen van de levensverwachting.

Uit de Leiden LangLeven Studie is gebleken dat kinderen, maar ook de ouders, van negentigjarige broer-zuster paren 30% minder kans hebben om te overlijden in vergelijking met mensen van de algemene bevolking uit hetzelfde geboortejaar. Onder de kinderen (ongeveer 60 jaar oud) van de negentigjarige broers en zussen uit de langlevende families komen hartinfarcten, hoge bloeddruk en suikerziekte minder voor dan bij de partners van deze kinderen. Anders gezegd, de leden van langlevende families lijken trager te verouderen en biologisch jonger te zijn dan hun eigen partners die je kunt zien als controlepersonen: leeftijdsgenoten uit andere families.

Het doel van dit proefschrift was om indicatoren, ook wel biomarkers genoemd, voor biologische veroudering te vinden in het bloed van langlevende families en controles van de Leiden LangLeven Studie aan de hand van genexpressiemetingen waarin de activiteit van genen wordt bepaald. Deze biomarkers zouden ons kunnen wijzen op de processen die te maken hebben met de snelheid van veroudering in de mens. Hoofdstuk 2 geeft een overzicht van het onderzoek dat bij aanvang van het promotie onderzoek is gedaan naar de relatie tussen genexpressieprofielen en kalenderleeftijd. Echter, in verschillende studies die kalenderleeftijd bestuderen werden steeds weer andere groepen genen gevonden. Dit kan komen door verschillen in onderzoekspopulaties, de grootte van de studie, de onderzochte weefsels en testmethodes. Een aanpak die niet elk gen apart onderzoekt, maar kijkt naar groepen genen die betrokken zijn bij veroudering, laat in meerdere studies wel zien welke processen consistent

lijken te associëren met de kalenderleeftijd (hoofdstuk 2). In dit proefschrift hebben we verschillende methoden gebruikt om de biomarkers voor biologische veroudering te vinden. Deze worden in de verschillende hoofdstukken van dit proefschrift beschreven.

In hoofdstuk 3 beschrijven we een zogenaamde hypothesis-vrije aanpak waarin we expressieniveaus van bijna 18.000 bekende genen hebben kunnen meten. Hiervoor onderzochten we bloed van deelnemers van de Leiden LangLeven Studie: vijftig negentigjarigen, vijftig van hun kinderen en vijftig partners van de kinderen die als controles fungeren. In deze studieopzet kunnen we verschillende vergelijkingen maken.
Als eerste die tussen negentigjarigen en mensen van middelbare leeftijd. We kozen daarvoor de partners van de kinderen (die zijn geen familie van de negentigjarigen). In deze vergelijking bleken er 1853 genen in hun expressie te verschillen tussen de twee groepen. Deze genen vertegenwoordigen óf een kenmerk van kalenderleeftijd (90 versus ± 60 jaar oud), óf een kenmerk van de langlevende families, die je ook in andere familieleden kunt aantreffen. Wat uniek is in vergelijking met andere studies is dat we ook de kinderen (middelbare leeftijd) van deze negentigjarigen konden meten. In de vergelijking van die kinderen en hun partners bleken 259 van de 1853 genen verschillend tot expressie te komen. Aangezien deze groepen van dezelfde leeftijd zijn, zijn dit waarschijnlijk kenmerken van langlevende families. De expressie van deze genen heeft mogelijk een relatie met de biologische leeftijd die in de kinderen van lang levende families lager is dan die van hun echtgenoten. Deze categorie genen zou een aanwijzing kunnen geven over de processen die mede bepalen hoe snel de biologische veroudering verloopt.
Van de 259 genen hebben we er 22 opnieuw gemeten in een groter deel van de Leiden LangLeven Studie; 79 negentigjarigen, 332 kinderen en 312 partners. Hieruit bleek dat een lagere expressie van de ASF1 anti-silencing function 1 homolog A (*ASF1A)* en het interleukine 7 receptor (*IL7R)* gen kenmerkend is voor de langlevende families. Dit duidt erop dat de genexpressie van deze genen, betrokken bij modellering van chromosomen en immuunsysteem functie, samen gaat met een jongere biologische leeftijd.

Een juiste werking van de IL7 receptor waar het *IL7R* gen voor codeert is erg belangrijk voor ons immuunsysteem. De ontwikkeling van T cellen die nodig zijn voor een optimale immuun reactie is afhankelijk van de IL7 signalering waarvan de IL7 receptor aan de basis ligt. Om te zien of genen die met het *IL7R* gen te maken hebben ook veranderen met biologische leeftijd, hebben we een *IL7R* netwerk getest dat bestaat uit *IL7R* zelf en zes

andere genen die een directe connectie hebben met *IL7R*. De expressie van deze genen hebben we getest in de grote groep deelnemers van de Leiden LangLeven Studie (hoofdstuk 4). Het netwerk als geheel bleek nu anders tot expressie te komen in leden van langlevende families, in het bijzonder *IL7R* zelf. De langelevende families worden gekarakteriseerd door een relatieve lage genexpressie, onafhankelijk van de samenstelling van celtypen in het bloed en de aanwezigheid van immuun-gerelateerde ziekten. We hebben tevens gekeken naar de relatie tussen de *IL7R* expressie en overleving gedurende 7 jaar. Nu bleek dat een hoger expressieniveau gerelateerd is aan een betere overleving op middelbare en hoge leeftijd. Hieruit blijkt dat de relatie van *IL7R* genexpressie met biologische leeftijd enerzijds en gezondheidsstatus anderzijds op het eerste gezicht tegenstrijdig is. Deze complexe relatie wordt ook bij andere immuun parameters gevonden in ouderen en moet verder onderzocht worden.

Een andere aanpak die we hebben gebruikt om de relatie tussen langlevendheid en genexpressie te onderzoeken is de hypothesis-gebaseerde aanpak. In dieren is het bekend dat mammalian target of rapamycin (mTOR) signalering de levensduur beïnvloedt. Bij mensen is dit proces betrokken bij het optreden van ziekten zoals diabetes en kanker. De basis van mTOR signalering ligt in twee eiwitcomplexen, mTOR-complex 1 en mTOR-complex 2. Deze reageren op onder andere voeding, energieverbruik (sporten) en het glucoseniveau in het bloed. Tot nu toe is vooral bewezen dat verlaagde activiteit van het mTOR-complex 1 de levensduur verlengt (in diermodellen) en dat verhoogde activiteit van datzelfde complex samen gaat met de ernst van ziekten bij de mens.
De rol van mTOR-signalering bij het bepalen van levensduur bij de mens was nog niet eerder onderzocht. Dat hebben we in dit proefschrift in de Leiden LangLeven Studie gedaan (hoofdstuk 5). Expressie van veertig genen uit de mTOR signalering werd gemeten in 87 negentigjarigen, 337 van hun kinderen en 321 partners daarvan. Hiervan lieten zeven genen een duidelijk significant verschil zien tussen de negentigjarigen en de partners van de kinderen. In de vergelijking van kinderen en hun partners vonden we nu dat op middelbare leeftijd een lagere expressie voor *RPTOR* en *PRR5L* kenmerkend was voor de langlevende families. Als genexpressie daadwerkelijk een reflectie is van mTOR signalering dan zou een lagere MTORC1 en MTORC2 expressie bevestigen dat ook in mensen een lage mTOR signalering samengaat met familiaire langlevendheid.

In hoofdstuk 6 hadden we als doel de mTOR-bevindingen uit hoofdstuk 5, die gedaan zijn in bloed, te bevestigen op eiwitniveau in cellen uit de huid (fibroblasten) die bij een huidbiopt van een deel van de studiegroep uit

hoofdstuk 5 waren verzameld. In de fibroblasten werd niet alleen de genexpressie gemeten, maar ook de expressie en activiteit van eiwitten die bij de mTOR signalering betrokken zijn. Behalve een verschil in totale eiwit hoeveelheid, zagen we in de fibroblasten geen verschil tussen de kinderen van langlevenden en hun echtgenoten. Het verschil tussen de groepen dat we aantroffen in het bloed was dus niet aanwezig in huidcellen van dezelfde mensen. De expressie van genen reflecteerde niet het corresponderende eiwitniveau in dezelfde cellen. Tevens bleek dat genexpressie in huidcellen niet correspondeerde met de expressie van diezelfde genen in bloed. Of dit komt door de gekozen groep mensen, het kleine aantal mensen dat in deze stap kon worden onderzocht (37 kinderen en 36 partners), het onderzochte weefsel of de gebruikte technieken, is niet duidelijk. Het lijkt erop dat de gekweekte huidfibroblasten wat betreft het samengaan van lage mTOR signalering en familiaire langlevendheid niet hetzelfde laten zien als bloed.

Bij het vergelijken van verschillende studies naar genexpressie in bloed is er weinig overeenkomst tussen studies in de gevonden individuele genen, waarvan de expressie verandert met kalenderleeftijd. Er is meer overeenkomst tussen de biologische processen waarvan de expressie verandert met leeftijd. Om te weten te komen welke processen een consistente rol spelen in veroudering kan onderzocht worden van welke groepen genen in gezamenlijkheid de expressie verandert bij het ouder worden, zodat daarna onderzocht kan worden binnen welk proces deze groep genen een functie vervult. We hebben daarom in hoofdstuk 7 de datasets van drie grote studies uit de literatuur verzameld waarin genexpressie werd gemeten in bloed van mensen met verschillende leeftijden. We hebben gezocht naar die genen die als groep (netwerk) een verandering in expressie met leeftijd ondergaan in meerdere studies (zogenoemde meta-analyse). De netwerken van genen die we in de data vonden zijn gebaseerd op een combinatie van de vergelijkbare genexpressieniveaus in de op bloed gebaseerde datasets de tot nu toe bekende eiwit-eiwitverbindingen van de eiwitten waar de genen voor coderen. Deze combinatie levert netwerken op van genen waarvan we veronderstellen dat ze in een gezamenlijk biologische proces zijn betrokken. Vervolgens zijn deze netwerken getest voor differentiële expressie met kalenderleeftijd, wat resulteerde in identificatie van vijf netwerken waarvan de expressie consistent over studies verandert met veroudering. Een onafhankelijke herhalingsstudie in een Nederlandse groep mensen liet zien dat vier van de vijf netwerken ook in deze studie veranderde met veroudering. Om te onderzoeken of deze netwerken ook echt potentiële biomarkers zijn is er een overlevingsanalyse gedaan in de

LLS welke liet zien dat een van de netwerken, degene die *ASF1A* bevat, daadwerkelijk samengaat met overleving op hoge leeftijd.
Opvallend is dat vijf genen (*ABCE1*, *ASF1A*, *DCK, MPHOSPH10* en *ZZZ3*) van het netwerk dat indicatief is voor betere overleving ook voorkomen op de lijst van genen die gerelateerd zijn aan biologische leeftijd, zoals beschreven in hoofdstuk 3. Dit netwerk heeft geen duidelijk bekende biologische functie en is nu een kandidaatbiomarker voor gezonde veroudering. Het is van belang dit netwerk te onderzoeken in relatie met gezondheid en veroudering. Tevens is het van belang om de biologische functie van dit netwerk te achterhalen.

Als we alle resultaten combineren, komen we uit op één groot netwerk van genen dat zowel de mTOR signalering, het IL7R netwerk, als de vijf netwerken van de meta-analyse omvat. Dit proefschrift laat zien dat dit grote netwerk (figuur 1 van hoofdstuk 8) met al zijn genen gerelateerd is aan veroudering. De hypothese-vrije aanpak beschreven in hoofdstuk 3 heeft in totaal 259 kandidaatgenen opgeleverd als mogelijke marker voor biologische leeftijd. Voor enkele van deze genen, *ASF1A*, *IL7R* en *RPTOR,* hebben we laten zien dat verlaagde genexpressie samengaat met een jongere biologische leeftijd. Er zijn dus veel genen over die nog onderzocht moeten worden. Om de prioriteit van deze nog te testen kandidaatgenen aan te geven, hebben we gekeken of ze ook onderdeel zijn van het bovengenoemde grote verouderingsnetwerk (de combinatie van de resultaten uit hoofdstukken 4, 5 en 7). Van deze kandidaatgenen hebben er 99 een directe connectie met een of meerdere van de netwerkgenen. Deze 99 zullen dan ook verder onderzocht moeten worden op hun eigenschappen als mogelijke biomarker voor biologische veroudering.

Een van de belangrijkste processen die aan veroudering wordt gerelateerd in dieronderzoek en ook in dit proefschrift is de mTOR signalering. We hebben laten zien dat verlaagde expressie van deze genen onder andere samen gaat met minder ziektes en een gezonder suiker metabolisme in de langlevende families. In muizen observeert men een vergelijkbaar beeld als de voedselinname wordt beperkt; verlaagde mTOR signalering en een verlengde levensduur. Er is geen verschil in voedselinname tussen leden van de lang-levende families en hun partners, dus wij concluderen dat deze families een betere ‘afstelling’ van nature hebben. De vraag is nu, kunnen de partners iets doen om net zo gezond te worden?

Iemands leefstijl kan de expressie beïnvloeden, mTOR signalering is bijvoorbeeld zogenaamd ‘nutrient sensing’ en wordt beïnvloed door voeding en lichaamsbeweging. Naar aanleiding van de bevindingen uit dit

proefschrift is de Samen Oud Samen Thuis (Growing Old Together) Studie opgezet waarin de kinderen van langlevende mensen en hun partners een dieet volgen met verminderde hoeveelheid caloriën en een bewegingsprogramma. De effecten van deze interventie op gen- en eiwitexpressie wordt gevolgd in bloed, spier en vet van de personen. Door middel van het veranderen van eet- en bewegingspatronen proberen we erachter te komen of de partners zonder de gunstige genen toch net zo gezond oud kunnen worden als de kinderen van de negentigjarigen.

Curriculum vitae

List of publications

Curriculum vitae

Willemijn Marie-Antoinette Passtoors was born on the 28th of August 1980 in Haarlem. In 1998 she finished her VWO at the Veurs Lyceum in Leidschendam and after a one year break she started Biomedical Science at the University of Amsterdam in 1999. Her first research internship was at the department of Molecular Carcinogenesis at the Dutch Cancer Institute in Amsterdam. The goal of this internship was to study the structure of mutants of the mismatch repair protein MutS by protein crystallography. Her main internship as an undergraduate was at the department of Biological Lead Optimization at Solvay Pharmaceuticals in Weesp. The topic was the development of an in vitro screening method to select compounds viable to induce antioxidative enzymes in vivo in the search for a protective compound in Parkinson's disease. After this internship she obtained a job as technician at the same department where she worked until 2006 while graduating in 2005. In 2006 she started her PhD at the Molecular Epidemiology section of the Leiden University Medical Center under supervision of Professor P. E. Slagboom and Dr. M. Beekman. Her main aim was to identify relevant pathways and potential biomarker profiles in human populations that associate with chronological age and familial longevity. The results of this research are outlined in this thesis. Currently she is employed as a postdoctoral researcher at the Shanghai Center for Systems Biomedicine at the Shanghai Jiao Tong University, performing drug discovery studies for several diseases and studying possible candidate pathways for Alzheimer's disease.

List of publications

Passtoors WM, Maier AB, van der Breggen R, Deelen J, Trompet S, Jukema JW, Derhovanessian E, Pawelec G, van Ommen GJB, Slagboom PE*, Beekman M*. *IL7R* gene expression network associates with human familial longevity. Manuscript in preparation.

Passtoors WM, Guigas B, Dekker P, Maier AB, van der Zon GCM, van der Breggen R, van Ommen GJB, Slagboom PE, Beekman M. The association of mTOR and familial longevity in the blood transcriptome is not reflected in primary skin fibroblasts. Manuscript in preparation.

van den Akker EB*, **Passtoors WM***, Jansen R, van Zwet EW, Goeman JJ, Hulsman M, Emilsson V, Perola M, Willemsen AHM, Penninx BWJH, Heijmans BT, Maier AB, Boomsma DI, Kok JN, Slagboom PE, Reinders MJT*, Beekman M*. Meta-analysis on blood transcriptomic studies identifies consistently co-expressed PPI modules as robust markers of human ageing. Aging Cell. 2013 Oct 14 published online.

Passtoors WM, Beekman M, Deelen J, van der Breggen R, Maier AB, Guigas B, Derhovanessian E, van Heemst D, de Craen AJ, Gunn DA, Pawelec G, Slagboom PE. Gene expression analysis of mTOR pathway: association with human longevity. Aging Cell. 2013 Feb;12(1):24-31.

Passtoors WM, Boer JM, Goeman JJ, van den Akker EB, Deelen J, Zwaan BJ, Scarborough A, van der Breggen R, Vossen RH, Houwing-Duistermaat JJ, Ommen GJ, Westendorp RG, van Heemst D, de Craen AJ, White AJ, Gunn DA, Beekman M, Slagboom PE. Transcriptional profiling of human familial longevity indicates a role for ASF1A and IL7R. PLoS One. 2012;7(1):e27759.

Deelen J, Beekman M, Uh HW, Helmer Q, Kuningas M, Christiansen L, Kremer D, van der Breggen R, Suchiman HE, Lakenberg N, van den Akker EB, **Passtoors WM**, Tiemeier H, van Heemst D, de Craen AJ, Rivadeneira F, de Geus EJ, Perola M, van der Ouderaa FJ, Gunn DA, Boomsma DI, Uitterlinden AG, Christensen K, van Duijn CM, Heijmans BT, Houwing-Duistermaat JJ, Westendorp RG, Slagboom PE. Genome-wide association study identifies a single major locus contributing to survival into old age; the APOE locus revisited. Aging Cell. 2011 Aug;10(4):686-98.

Slagboom PE, Beekman M, **Passtoors WM**, Deelen J, Vaarhorst AA, Boer JM, van den Akker EB, van Heemst D, de Craen AJ, Maier AB, Rozing M,

Mooijaart SP, Heijmans BT, Westendorp RG. Genomics of human longevity. Philos Trans R Soc Lond B Biol Sci. 2011 Jan 12;366(1561):35-42.

Passtoors WM, Beekman M, Gunn D, Boer JM, Heijmans BT, Westendorp RG, Zwaan BJ, Slagboom PE. Genomic studies in ageing research: the need to integrate genetic and gene expression approaches. J Intern Med. 2008 Feb;263(2):153-66.

*Both authors contributed equally to the studies