AGE (2013) 35:235–249
DOI 10.1007/s11357-011-9340-3

Gene set analysis of GWAS data for human longevity highlights the relevance of the insulin/IGF-1 signaling and telomere maintenance pathways

Joris Deelen · Hae-Won Uh · Ramin Monajemi · Diana van Heemst · Peter E. Thijssen · Stefan Böhringer · Erik B. van den Akker · Anton J. M. de Craen · Fernando Rivadeneira · André G. Uitterlinden · Rudi G. J. Westendorp · Jelle J. Goeman · P. Eline Slagboom · Jeanine J. Houwing-Duistermaat · Marian Beekman

Received: 7 July 2011 / Accepted: 28 October 2011 / Published online: 24 November 2011

Abstract In genome-wide association studies (GWAS) of complex traits, single SNP analysis is still the most applied approach. However, the identified SNPs have small effects and provide limited biological insight. A more appropriate approach to interpret GWAS data of complex traits is to analyze the combined effect of a SNP set grouped per pathway or gene region. We used this approach to study the joint effect on human longevity of genetic variation in two candidate pathways, the insulin/insulin-like growth factor (IGF-1) signaling (IIS) pathway and the telomere maintenance (TM) pathway. For the analyses, we used genotyped

Electronic supplementary material The online version of this article (doi:10.1007/s11357-011-9340-3) contains supplementary material, which is available to authorized users.

J. Deelen (✉) · P. E. Thijssen · E. B. van den Akker · P. E. Slagboom · M. Beekman
Section of Molecular Epidemiology,
Leiden University Medical Center,
Zone S5-P, PO Box 9600, 2300 RC Leiden,
The Netherlands
e-mail: j.deelen@lumc.nl

J. Deelen · H.-W. Uh · A. G. Uitterlinden · R. G. J. Westendorp · P. E. Slagboom · M. Beekman
Netherlands Consortium for Healthy Ageing,
Leiden University Medical Center,
P.O. Box 9600, 2300 RC Leiden, The Netherlands

H.-W. Uh · R. Monajemi · S. Böhringer · J. J. Goeman · J. J. Houwing-Duistermaat
Section of Medical Statistics,
Leiden University Medical Center,
P.O. Box 9600, 2300 RC Leiden, The Netherlands

D. van Heemst · A. J. M. de Craen · R. G. J. Westendorp
Department of Gerontology and Geriatrics,
Leiden University Medical Center,
P.O. Box 9600, 2300 RC Leiden, The Netherlands

P. E. Thijssen
Department of Human Genetics,
Leiden University Medical Center,
P.O. Box 9600, 2300 RC Leiden, The Netherlands

E. B. van den Akker
Department of Mediamatics, Delft Bioinformatics Lab,
Delft University of Technology,
P.O. Box 5031, 2600 GA Delft, The Netherlands

F. Rivadeneira · A. G. Uitterlinden
Department of Epidemiology, Erasmus Medical Center,
P.O. Box 2040, 3015 CE Rotterdam, The Netherlands

F. Rivadeneira · A. G. Uitterlinden
Department of Internal Medicine, Erasmus Medical Center,
P.O. Box 2040, 3015 CE Rotterdam, The Netherlands

GWAS data of 403 unrelated nonagenarians from long-lived sibships collected in the Leiden Longevity Study and 1,670 younger population controls. We analyzed 1,021 SNPs in 68 IIS pathway genes and 88 SNPs in 13 TM pathway genes using four self-contained pathway tests (PLINK set-based test, Global test, GRASS and SNP ratio test). Although we observed small differences between the results of the different pathway tests, they showed consistent significant association of the IIS and TM pathway SNP sets with longevity. Analysis of gene SNP sets from these pathways indicates that the association of the IIS pathway is scattered over several genes (*AKT1*, *AKT3*, *FOXO4*, *IGF2*, *INS*, *PIK3CA*, *SGK*, *SGK2*, and *YWHAG*), while the association of the TM pathway seems to be mainly determined by one gene (*POT1*). In conclusion, this study shows that genetic variation in genes involved in the IIS and TM pathways is associated with human longevity.

Keywords Genetics · Aging · Longevity · Gene set analysis · Insulin/IGF-1 signaling · Telomere maintenance

Introduction

Genome-wide association studies (GWAS) using single SNP analysis have been very successful in identifying loci for various quantitative traits and diseases (Manolio et al. 2008). It became apparent that complex traits are usually determined by many genes with small effects and that results from single SNP analysis provide limited biological insight and only partly explain the genotypic variation of the studied trait. Instead of analyzing single SNPs, the combined effect of a SNP set, grouped per pathway or gene region, can be tested for association with the trait of interest. Such SNP set analysis could be used as an alternative approach for GWAS analysis and, since the composition of SNP sets is often based on pathways, should be able to provide additional biological insight of the studied trait.

Since the amount of tests in SNP set analysis is low compared to single SNP analysis, it requires a lower penalty for multiple testing. Therefore, SNP set analysis is also very suitable in studies with low power for GWAS analysis. The last couple of years, several methods have been developed to perform SNP set analysis on GWAS data (Wang et al. 2010; Fridley and Biernacka 2011; Holmans 2010). There are two main types of methods, the competitive and the self-contained tests. The competitive tests compare the association between a SNP set and trait to a standard defined by the genotyped SNPs outside the SNP set (complement), while the self-contained tests compare the SNP set to a fixed standard that does not depend on the complement (Goeman and Buhlmann 2007).

Human longevity is a complex trait that is assumed to be determined by variation in many genes with small effects. Previous GWA studies, in which single SNP analyses were performed (Newman et al. 2010; Deelen et al. 2011), have identified only one genome-wide significant locus contributing to survival into old age; *APOE*. However, the genetic contribution to human lifespan variation, determined in twin studies, is estimated at 25–30% (Gudmundsson et al. 2000; Hjelmborg et al. 2006; Skytthe et al. 2003) and, although the effect of genetic variation in *APOE* is relatively large, the heritability of longevity is only partially explained by this variation (Deelen et al. 2011). Part of the remaining heritability might be explained by functionally related SNPs with small effects, of which the joint effect could not be detected in a single SNP analysis. Testing of SNP sets of candidate pathways for association with longevity would therefore be valuable.

The insulin/insulin-like growth factor (IGF-1) signaling (IIS) pathway is considered as a candidate pathway for studying human longevity. It is involved in the adaptation of the organism to its (changing) environment (Tatar et al. 2003). When experimentally induced in model organisms like worms, flies, and mice, mutations in genes that play a role in IIS, e.g., homologues of human *IGF1R*, *INSR*, *IRS1*, *PI3K*, and *FOXO*, were shown to have a considerable effect on lifespan (Kenyon et al. 1993; Kimura et al. 1997; Tatar et al. 2001; Holzenberger et al. 2003; Bluher et al. 2003; Morris et al. 1996; Friedman and Johnson 1988; Clancy et al. 2001; Hwangbo et al. 2004; Ogg et al. 1997; Lin et al. 1997; Giannakou et al. 2004; Selman et al. 2011). Although the IIS pathway is evolutionarily conserved, the complexity of the human IIS pathway (Fig. 1) is much larger compared to that of model organisms.

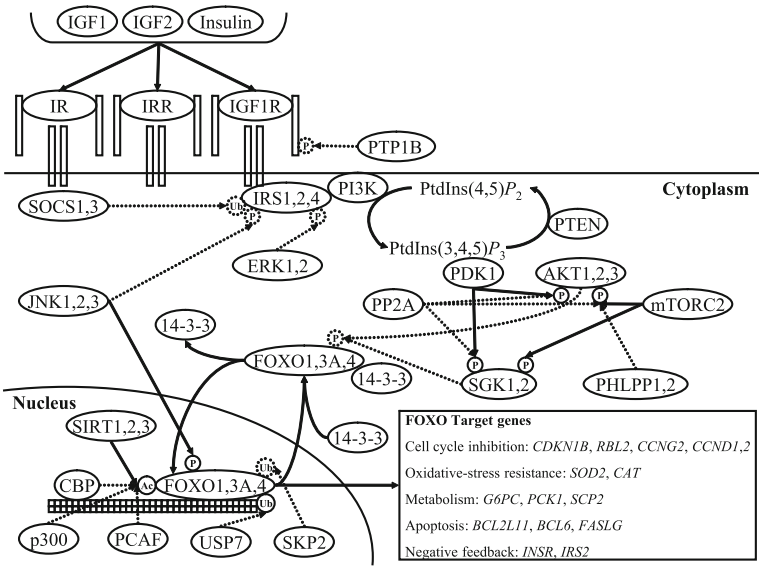

Fig. 1 Insulin/IGF-1 signaling pathway. The insulin/IGF-1 signaling pathway consists of the core components IGF1R/IR/IRR, IRS, PI3K, AKT/SGK, FOXO and SIRT, and proteins that have a direct activating or inhibiting effect on these proteins. The *small closed circles* (containing Ac, P, or Ub) indicate an activating effect of the posttranslational modification on the protein, while the *small dashed circles* indicate an inhibiting effect. The *straight arrows* pointing to these *small circles* indicate an activating effect on the posttranslational modification, while the *dashed arrows* indicate an inhibiting effect. *Ac* acetylation, *P* phosphorylation, *Ub* ubiquitylation

Several studies have investigated associations between single SNPs in genes from the IIS pathway and human longevity. The most prominent results came from *FOXO3* (Willcox et al. 2008; Flachsbart et al. 2009; Anselmi et al. 2009; Pawlikowska et al. 2009; Li et al. 2009; Soerensen et al. 2010) and *AKT1* (Pawlikowska et al. 2009), which showed associations with longevity in several independent cohort studies.

Another candidate pathway for studying human longevity is the mechanism of telomere maintenance (TM). Telomeres are structures at the end of chromosomes, consisting of TTAGGG tandem repeats (Moyzis et al. 1988), which protect chromosomes from degradation or rearrangement (Blackburn 1991). In normal human cells, telomere length declines with every cell division (Harley et al. 1990), and when a critical length is reached, the cell will enter replicative senescence (Allsopp 1996). In human epidemiological studies in blood, increased telomere length has been associated with longevity (Atzmon et al. 2010), while decreased telomere length has been associated with increased mortality (Cawthon et al. 2003; Bakaysa et al. 2007; Kimura et al. 2008), although some studies showed contradictory results (Martin-Ruiz et al. 2005; Bischoff et al. 2006). Telomere integrity is essentially regulated by two protein networks, telomerase and its associated factors, which regulate telomere length, and the shelterin complex, which covers the telomeres (de Lange 2005; Collins and Mitchell 2002) (Fig. 2). Several studies have investigated associations between single SNPs in telomerase and shelterin genes and telomere length. The most promising results came from TERC and TERT (Atzmon et al. 2010; Codd et al. 2010; Levy et al. 2010; Mirabello et al. 2010; Rafnar et al. 2009), of which the latter has also been associated with human longevity (Atzmon et al. 2010).

In this study, we used four self-contained tests (PLINK set-based test, Purcell et al. 2007; GRASS, Chen et al. 2010; Global test, Goeman et al. 2004; and SNP ratio test, O'Dushlaine et al. 2009) and one competitive test (the comparative approach of Global test) to study the joint effect of genetic variation in the IIS and TM pathways on human longevity. For the analyses, we used genotyped GWAS data of nonagenarian siblings from the Leiden Longevity Study (LLS) and younger population controls from the Rotterdam Study (RS) (Deelen et al. 2011).

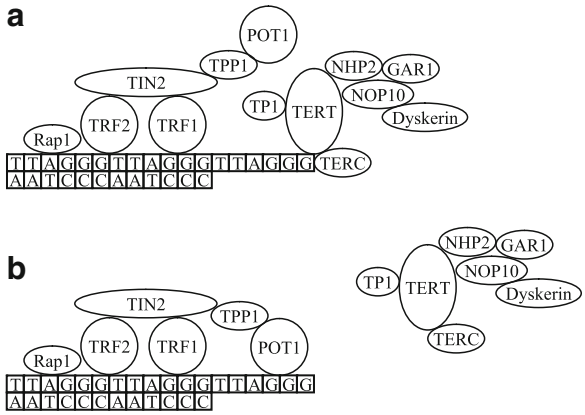

Fig. 2 Telomere maintenance pathway. The telomere maintenance pathway consists of proteins belonging to telomerase and its associated factors or to the shelterin complex. Telomere elongation is performed by telomerase after binding to the telomere (**a**). However, binding of the shelterin protein POT1 to the telomere blocks this process (**b**)

Materials and methods

Study populations

Leiden longevity study

For the LLS, long-lived siblings of European descent were recruited together with their offspring and the partners of the offspring. Families were included if at least two long-lived siblings were alive and fulfilled the age criterion of 89 years or older for males and 91 years or older for females, representing less than 0.5% of the Dutch population in 2001 (Schoenmaker et al. 2006). In total, 944 long-lived proband siblings were included with a mean age of 94 years (range, 89–104), 1,671 offspring (61 years, 39–81), and 744 partners (60 years, 36–79). DNA from the LLS was extracted from samples at baseline using conventional methods (Beekman et al. 2006). For the GWAS, 403 unrelated LLS siblings (one sibling from each sibling pair) were included (LLS GWAS cases) (Deelen et al. 2011).

Rotterdam study

The RS is a prospective population-based study of people aged 55 years and older, which was designed to study neurological, cardiovascular, locomotor, and ophthalmological diseases (Teichert et al. 2009). The study consists of 7,983 participants from the baseline cohort (RS-I) and 3,011 participants from an independent extended cohort formed in 1999 (RS-II) from which DNA was isolated between 1990 and 1993 (RS-I) or between 2000 and 2001 (RS-II). For the GWAS, 1,731 participants from the combined cohort who were below 60 years of age and for whom GWAS data were available were included as controls (LLS GWAS controls) (Deelen et al. 2011).

Population substructure

Multidimensional scaling analysis in PLINK (http://pngu.mgh.harvard.edu/purcell/plink, Purcell et al. 2007) showed that there was no substructure in the GWAS data to an extent that would affect the observations (Deelen et al. 2011).

Genotyping and SNP selection

For the SNP set analyses, we used the genotype data from the GWAS described by Deelen et al. (2011). The LLS GWAS cases were genotyped using Illumina Infinium HD Human660W-Quad BeadChips (Illumina, San Diego, CA, USA). The RS GWAS controls were genotyped using Illumina Infinium II HumanHap 550K Beadchips and Illumina Infinium II HumanHap550-Duo BeadChips (Illumina), respectively (Teichert et al. 2009). Of the 551,606 SNPs measured in both the LLS GWAS cases and RS GWAS controls, 516,712 SNPs passed quality control using the following criteria: SNP call rate ≥0.95 or MAF ≥0.01 in RS GWAS controls and LLS GWAS cases, PHWE $\geq 10^{-4}$ and no between-chip effect in the RS GWAS controls, and good cluster plots in the LLS GWAS cases and RS GWAS controls if $P < 1 \times 10^{-4}$ (Deelen et al. 2011).

We analyzed SNPs within a 10-kb window around genes encoding proteins that belonged to the IIS (Fig. 1) and TM pathway (Fig. 2). A gene was defined as an NCBI Entrez Gene (mRNA or RNA) cluster, corresponding to a set of transcripts (RefSeq) for which the alignments can be obtained from the UCSC genome browser (http://genome.ucsc.edu/), in which all transcripts within a cluster agree on strand and overlap. Due to an overlap of the 10-kb windows

around *IGF2* and *INS*, two SNPs, rs4320932 and rs7924316, were assigned to both genes.

Statistical analysis

PLINK set-based test

In the PLINK set-based test (–set-test, http://pngu.mgh.harvard.edu/purcell/plink; Purcell et al. 2007), a single SNP analysis (in our case, a trend test) of the original pathway or gene SNP set is performed. For each SNP set, a mean SNP statistic is calculated from the single SNP statistics of a maximum amount (–set-max) 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-r2), 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 –set-p 0.05 –set-r2 0.5, –set-max 99999, and –mperm 10,000.

GRASS

GRASS (http://linchen.fhcrc.org/grass.html; Chen et al. 2010) calculates "eigenSNPs" for each gene in the pathway SNP set by summarizing the variation of a gene using principal component analysis. Subsequently, one or more of these "eigenSNPs" per gene are selected using regularized logistic regression to calculate a test statistic for each pathway SNP set. The same analysis is performed with simulated SNP sets in which the phenotype status of the individuals is permuted. The *P* value per pathway SNP set is calculated by comparing the test statistic of the original pathway SNP set with that of the combined simulated pathway SNP sets. For the analysis in this study, the amount of simulated pathway SNP sets was 10,000.

Global test

In this study, we used a modified version of the Global test (http://www.bioconductor.org/help/bioc-views/release/bioc/html/globaltest.html; Goeman et al. 2004), which is capable and powerful for analyzing GWAS data (Chapman and Whittaker 2008; Pan 2009). This test is based on a multiple logistic regression model that uses the phenotype as the response variable and the SNPs in the SNP set as covariates and which automatically takes the correlations between SNPs into account. The null hypothesis is tested that none of the SNPs in the SNP set are associated with the phenotype. *P* values are calculated using a permutation test based on 10,000 permutations.

For the comparative approach, 10,000 random SNP sets per pathway SNP set were generated and tested to determine the chance to find a similar-sized SNP set with a comparable or lower *P* value as compared to the original pathway SNP set.

SNP ratio test

The SNP ratio test (http://sourceforge.net/projects/snpratiotest/; O'Dushlaine et al. 2009) performs a single SNP analysis (in our case, a trend test) of the original pathway or gene SNP set and of similar-sized SNP sets in which the phenotype status of the individuals is permuted. An empirical *P* value of the SNP set is computed by calculating the ratio between the proportion of SNPs that shows an association below a certain *P* value threshold (*p*) in the original GWAS dataset and in the simulated GWAS datasets. The amount of significant SNPs in the simulated GWAS datasets is defined as the top *n* SNPs with the lowest *P* values, where *n* is the amount of SNPs with an association below *p* in the original GWAS dataset. For the analysis in this study, we made use of the scripts described in "SRT_documentation_090310.pdf" (http://sourceforge.net/projects/snpratiotest/). For the analysis in this study, *p* was set to 0.05, and the amount of simulated datasets used was 10,000.

Statistical significance

To adjust for multiple testing, the significance level was set at the Bonferroni-corrected nominal *P* value (which is 0.05/(number of pathway or gene SNP sets tested)).

Table 1 Characteristics of the insulin/IGF-1 signaling pathway proteins

Protein	Gene	Entrez gene ID	Chr	Start (bp)	End (bp)	Size (kb)	SNPs	Coverage
AKT1	*AKT1*	207	14	104,306,732	104,333,125	26.39	2	25.00%
AKT2	*AKT2*	208	19	45,428,064	45,483,105	55.04	6	45.45%
AKT3	*AKT3*	10000	1	241,718,158	242,073,509	355.35	25	50.00%
BIM	*BCL2L11*	10018	2	111,594,962	111,642,493	47.53	12	47.06%
BCL-6	*BCL6*	604	3	188,921,859	188,946,169	24.31	6	23.81%
CAT	*CAT*	847	11	34,417,048	34,450,182	33.13	18	80.00%
Cyclin D1	*CCND1*	595	11	69,165,054	69,178,423	13.37	3	27.27%
Cyclin D2	*CCND2*	894	12	4,253,163	4,284,782	31.62	20	45.00%
Cyclin G2	*CCNG2*	901	4	78,297,381	78,310,237	12.86	4	33.33%
p27kip	*CDKN1B*	1027	12	12,761,569	12,766,572	5.00	9	75.00%
CBP	*CREBBP*	1387	16	3,715,057	3,870,122	155.07	15	50.00%
Deptor (mTORC2)	*DEPDC6*	64798	8	120,955,081	121,132,338	177.26	47	61.90%
p300	*EP300*	2033	22	39,818,560	39,906,027	87.47	6	50.00%
Fas ligand	*FASLG*	356	1	170,894,808	170,902,635	7.83	7	45.45%
FOXO1	*FOXO1*	2308	13	40,027,801	40,138,734	110.93	19	65.38%
FOXO3A	*FOXO3*	2309	6	108,987,719	109,112,664	124.95	21	68.75%
FOXO4	*FOXO4*	4303	X	70,232,751	70,240,109	7.36	3	NA
G6P	*G6PC*	2538	17	38,306,341	38,318,912	12.57	5	83.33%
IGF1	*IGF1*	3479	12	101,313,775	101,398,508	84.73	20	47.06%
IGF1R	*IGF1R*	3480	15	97,010,284	97,325,282	315.00	102	56.34%
IGF2	*IGF2*	3481	11	2,106,923	2,127,409	20.49	7	63.64%
Insulin	*INS*	3630	11	2,137,585	2,139,015	1.43	4	80.00%
IR	*INSR*	3643	19	7,063,266	7,245,011	181.75	52	50.00%
IRR	*INSRR*	3645	1	155,077,289	155,095,290	18.00	6	37.50%
IRS1	*IRS1*	3667	2	227,304,277	227,371,750	67.47	11	53.33%
IRS2	*IRS2*	8660	13	109,204,185	109,236,915	32.73	15	59.09%
IRS4	*IRS4*	8471	X	107,862,383	107,866,263	3.88	2	NA
PCAF	*KAT2B*	8850	3	20,056,528	20,170,900	114.37	44	62.96%
ERK2	*MAPK1*	5594	22	20,443,947	20,551,970	108.02	12	58.33%
ERK1	*MAPK3*	5595	16	30,032,927	30,042,131	9.20	2	33.33%
JNK1	*MAPK8*	5599	10	49,279,693	49,313,189	33.50	6	35.71%
JNK2	*MAPK9*	5601	5	179,593,203	179,651,677	58.47	17	40.00%
JNK3	*MAPK10*	5602	4	87,155,300	87,593,307	438.01	71	55.81%
mSIN1 (mTORC2)	*MAPKAP1*	79109	9	127,239,494	127,509,334	269.84	26	64.00%
mLST8 (mTORC2)	*MLST8*	64223	16	2,195,451	2,199,419	3.97	4	80.00%
mTOR (mTORC2)	*MTOR*	2475	1	11,089,176	11,245,195	156.02	11	50.00%
PEPCK	*PCK1*	5105	20	55,569,543	55,574,919	5.38	10	33.33%
PDK1	*PDPK1*	5170	16	2,527,971	2,593,190	65.22	0	0.00%
PHLPP1	*PHLPP1*	23239	18	58,533,714	58,798,646	264.93	41	64.29%
PHLPP2	*PHLPP2*	23035	16	70,236,353	70,306,205	69.85	4	15.38%
PI3K	*PIK3CA*	5290	3	180,349,005	180,435,191	86.19	10	44.44%
	PIK3CB	5291	3	139,856,921	139,960,875	103.95	8	42.86%
	PIK3CD	5293	1	9,634,377	9,711,759	77.38	8	42.11%
	PIK3R1	5295	5	67,558,218	67,633,405	75.19	31	65.12%

Table 1 (continued)

Protein	Gene	Entrez gene ID	Chr	Start (bp)	End (bp)	Size (kb)	SNPs	Coverage
	PIK3R2	5296	19	18,125,016	18,142,343	17.33	5	66.67%
	PIK3R3	8503	1	46,278,399	46,371,295	92.90	10	60.00%
PP2A	*PPP2R5B*	5526	11	64,448,756	64,458,523	9.77	3	37.50%
Protor-1 (mTORC2)	*PRR5*	55615	22	43,443,091	43,512,225	69.13	32	50.00%
PTEN	*PTEN*	5728	10	89,613,175	89,718,512	105.34	8	47.06%
PTP1B	*PTPN1*	5770	20	48,560,298	48,634,493	74.20	17	44.44%
p130Rb2	*RBL2*	5934	16	52,025,852	52,083,061	57.21	3	33.33%
RICTOR (mTORC2)	*RICTOR*	253260	5	38,973,780	39,110,258	136.48	9	30.77%
SCP2	*SCP2*	6342	1	53,165,536	53,289,870	124.33	18	50.00%
SGK1	*SGK1*	6446	6	134,532,077	134,680,889	148.81	38	46.75%
SGK2	*SGK2*	10110	20	41,621,100	41,647,687	26.59	9	34.62%
SIRT1	*SIRT1*	23411	10	69,314,433	69,348,152	33.72	4	33.33%
SIRT2	*SIRT2*	22933	19	44,061,040	44,082,201	21.16	7	38.89%
SIRT3	*SIRT3*	23410	11	205,030	226,362	21.33	17	60.00%
SKP2	*SKP2*	6502	5	36,187,946	36,219,904	31.96	15	51.72%
SOCS1	*SOCS1*	8651	16	11,255,775	11,257,540	1.77	4	50.00%
SOCS3	*SOCS3*	9021	17	73,864,457	73,867,753	3.30	4	50.00%
MnSOD	*SOD2*	6648	6	160,020,139	160,034,343	14.20	4	44.44%
USP7	*USP7*	7874	16	8,893,452	8,964,842	71.39	12	42.11%
14-3-3	*YWHAB*	7529	20	42,947,758	42,970,575	22.82	6	50.00%
	YWHAE	7531	17	1,194,586	1,250,306	55.72	16	70.00%
	YWHAG	7532	7	75,794,044	75,826,278	32.23	5	35.71%
	YWHAH	7533	22	30,670,479	30,683,590	13.11	9	43.75%
	YWHAQ	10971	2	9,641,557	9,688,557	47.00	10	58.33%
	YWHAZ	7534	8	101,999,981	102,034,799	34.82	6	40.00%
Total							1,023	

Chr Chromosome position of the gene according to NCBI Build 36, *Start (bp)* start position of the gene according to NCBI Build 36, *End (bp)* end position of the gene according to NCBI Build 36, *Coverage* coverage of genes based on Phased data HapMap II release 22 CEU, *NA* not available

Results

For the IIS pathway, we selected genes encoding proteins that belong to the well-described core of the pathway, consisting of IGF1R/IR/IRR, IRS, PI3K, AKT/SGK, FOXO, and SIRT, or that had a direct activating or inhibiting effect on these core components (van der Horst and Burgering 2007; Taniguchi et al. 2006). In addition, we selected several FOXO target genes that play a role in cell-cycle inhibition, oxidative-stress resistance, metabolism, and apoptosis (van der Horst and Burgering 2007) (Fig. 1). For the TM pathway, we selected genes encoding proteins that were specifically associated with telomeres and belonged to telomerase and its associated factors or to the shelterin complex (Vulliamy et al. 2008; Harrington et al. 1997; de Lange 2005) (Fig. 2). We analyzed SNPs within a 10-kb window around the selected genes (based on Pawlikowska et al. 2009) from genotyped GWAS data of 403 unrelated nonagenarian participants from the LLS and 1,670 middle-aged controls from the RS (Deelen et al. 2011). A description of the investigated samples is given in Table S1. In total, 1,021 SNPs in 68 IIS pathway genes and 88 SNPs in 13 TM pathway genes were analyzed (Tables 1, 2, S3A, and S3B).

Four methods, PLINK set-based test, Global test, GRASS, and SNP ratio test (Table S2), were used to investigate the association of the SNP sets from the IIS and TM pathways with longevity. As a biological

negative control, we also analyzed a SNP set of 223 SNPs in 9 genes previously associated with eye and hair color (Eriksson et al. 2010) (Tables 3 and S3C). Both candidate pathways were consistently associated with longevity across all four tests (Table 4). We applied Bonferroni correction to adjust for the number of tested pathways (i.e., 2, so for significance $P<0.025$). After Bonferroni correction, the IIS pathway SNP set remained significant in GRASS and Global test, while the TM pathway SNP set remained significant in the PLINK set-based test, GRASS, and Global test. Using the comparative approach in Global test as a competitive test, we also showed that the probability to find a random SNP set with the same amount of genes as the IIS or TM pathway and a comparable or lower *P* value is less than 5% (2.11% for the IIS and 2.95% for the TM pathway).

To determine which genes are mainly responsible for the observed association of the pathway SNP sets from the IIS and TM pathways with longevity, we also investigated the association of gene SNP sets from these pathways. Although the power to detect an association using gene SNP set analysis is lower than for pathway SNP set analysis, due to the larger amount of tests, it provides a ranking of genes based on the contribution to the observed associations of the pathways. To analyze the gene SNP sets, we used the PLINK set-based test, Global test, and SNP ratio test. GRASS was not used, since the underlying statistical method of this test is less suitable for analysis of gene SNP sets. Nine of the 68 IIS pathway gene SNP sets (*AKT1*, *AKT3*, *FOXO4*, *IGF2*, *INS*, *PIK3CA*, *SGK1*, *SGK2*, and *YWHAG*) and 1 of the 13 TM pathway gene SNP sets (*POT1*) showed an association ($P<0.05$) with longevity in at least two tests (Tables 5 and 6).

Discussion

To study the effect of the IIS and TM pathways on longevity, SNP set analysis on GWAS data of 403 nonagenarian cases and 1,670 population controls was performed. Both pathway SNP sets associated significantly with longevity. The gene SNP sets analysis showed that the association of the IIS pathway was scattered over several genes (*AKT1*, *AKT3*, *FOXO4*, *IGF2*, *INS*, *PIK3CA*, *SGK1*, *SGK2*, and *YWHAG*), while the association of the TM pathway seems to be mainly determined by one gene (*POT1*).

The proteins encoded by the IIS gene SNP sets that associate with longevity are involved in several parts of the IIS pathway (Fig. 1). Akt1, Akt3, Foxo4, Igf2, Ins2, Pik3ca, and Sgk1 knockout mice all show

Table 2 Characteristics of the telomere maintenance pathway proteins

Protein	Gene	Entrez gene ID	Chr	Start (bp)	End (bp)	Size (kb)	SNPs	Coverage
TPP1 (shelterin)	*ACD*	65057	16	66,248,916	66,252,219	3.30	2	50.00%
Dyskerin (telomerase)	*DKC1*	1736	X	153,644,225	153,659,158	14.93	1	NA
GAR1 (telomerase)	*GAR1*	54433	4	110,956,115	110,965,342	9.23	1	14.29%
NHP2 (telomerase)	*NHP2*	55651	5	177,509,072	177,513,567	4.50	2	33.33%
NOP10 (telomerase)	*NOP10*	55505	15	32,421,209	32,422,654	1.45	7	45.45%
POT1 (shelterin)	*POT1*	25913	7	124,249,676	124,357,273	107.60	25	55.56%
TP1 (telomerase)	*TEP1*	7011	14	19,903,666	19,951,419	47.75	21	40.00%
TERC (telomerase)	*TERC*	7012	3	170,965,092	170,965,542	0.45	1	25.00%
TRF1 (shelterin)	*TERF1*	7013	8	74,083,651	74,122,541	38.89	10	60.00%
TRF2 (shelterin)	*TERF2*	7014	16	67,946,965	67,977,375	30.41	6	57.14%
RAP1 (shelterin)	*TERF2IP*	54386	16	74,239,136	74,248,842	9.71	4	50.00%
TERT (telomerase)	*TERT*	7015	5	1,306,287	1,348,162	41.88	7	41.18%
TIN2 (shelterin)	*TINF2*	26277	14	23,778,691	23,781,720	3.03	1	14.29%
Total							88	

Chr Chromosome position of the gene according to NCBI Build 36, *Start (bp)* start position of the gene according to NCBI Build 36, *End (bp)* end position of the gene according to NCBI Build 36, *Coverage* coverage of genes based on Phased data HapMap II release 22 CEU, *NA* not available

Table 3 Characteristics of the eye and hair color pathway proteins

Protein	Gene	Entrez gene ID	Chr	Start (bp)	End (bp)	Size (kb)	SNPs	Coverage
ASIP	*ASIP*	434	20	32,311,832	32,320,809	8.98	5	50.00%
HERC2	*HERC2*	8924	15	26,029,783	26,240,890	211.11	9	41.67%
IRF4	*IRF4*	3662	6	336,739	356,443	19.70	14	65.00%
MC1R	*MC1R*	4157	16	88,511,788	88,514,886	3.10	3	33.33%
OCA2	*OCA2*	4948	15	25,673,616	26,018,053	344.44	82	58.00%
SLC24A4	*SLC24A4*	123041	14	91,858,678	92,037,578	178.90	62	53.68%
SLC45A2	*SLC45A2*	51151	5	33,980,478	34,020,537	40.06	15	44.83%
TYR	*TYR*	7299	11	88,550,688	88,668,575	117.89	22	56.00%
TYRP1	*TYRP1*	7306	9	12,683,386	12,700,266	16.88	11	50.00%
Total							223	

Chr Chromosome position of the gene according to NCBI Build 36, *Start (bp)* start position of the gene according to NCBI Build 36, *End (bp)* end position of the gene according to NCBI Build 36, *Coverage* coverage of genes based on Phased data HapMap II release 22 CEU

abnormalities in growth and/or increased mortality (www.informatics.jax.org; Blake et al. 2011), which indicates that these genes are indeed responsible for the growth- and lifespan-regulating effects of the IIS pathway. Previously, SNPs in several of the significant IIS pathway genes (*AKT1*, *FOXO4*, *INS*, and *PIK3CA*) were studied by single SNP analysis, and only one SNP, rs3803304 in *AKT1*, which was not measured in our study, showed an association with longevity (Pawlikowska et al. 2009). However, gene set testing, which could have detected association of additional genes containing SNPs with many small effects, was not applied in that study. Most signaling cascades require cooperation of several genes in multiple branches of the cascade. This indicates that, for signaling pathways, mutations in different genes could result in similar downstream effects, which would explain the scattered association in the IIS pathway.

Table 4 Results of gene set analysis of insulin/IGF-1 signaling, telomere maintenance, and eye and hair color pathway SNP sets

Pathway test	Insulin/IGF-1 signaling	Telomere maintenance	Eye and hair color
PLINK set-based test[a]	0.064	0.019	0.340
GRASS[a]	0.010	0.023	0.540
Global test[a]	0.011	0.023	0.362
SNP ratio test[a]	0.044	0.034	0.337

[a] Permutation (n=10,000)
P value

Although SNPs in *FOXO3A* have previously been associated with longevity in several independent studies (Willcox et al. 2008; Flachsbart et al. 2009; Anselmi et al. 2009; Pawlikowska et al. 2009; Li et al. 2009; Soerensen et al. 2010), the gene SNP set showed no effect in our study in the PLINK set-based test, Global test, and SNP ratio test (P=0.181, P=0.138, and P=0.180, respectively) (Table 5). This might be due to the fact that the effects of *FOXO3A* on longevity are most prominent in centenarians. As was previously reported by Flachsbart et al., centenarians represent a highly selected phenotype even among nonagenarians (Flachsbart et al. 2009). In addition, the genetic contribution to longevity in general is increased at higher ages (Hjelmborg et al. 2006), and the small effects of longevity-promoting gene variants, relative to other factors, may be larger in centenarians (Perls et al. 2002) and not detectable in nonagenarians. The cases in our study, which are from long-lived families, have a mean age of 94 years, yet we had only 11 individuals >100 years, which may explain the absence of significance of the *FOXO3A* association in our population.

POT1 is part of the shelterin complex and is responsible for the binding of this complex to the TTAGGG repeats of telomeres. Binding of POT1 to the telomere leads to decreased elongation by telomerase (de Lange 2005). Reduction of POT1 in human fibroblasts by RNAi leads to induction of

Table 5 Results of gene set analysis of insulin/IGF-1 signaling pathway gene SNP sets

Gene	PLINK set-based test[a]	Global test[a]	SNP ratio test[a]
AKT1	0.003	0.002	0.099
AKT2	0.193	0.461	0.197
AKT3	0.101	0.023	0.043
BCL2L11	1	0.678	1
BCL6	1	0.539	1
CAT	1	0.661	1
CCND1	1	0.471	1
CCND2	0.248	0.073	0.073
CCNG2	1	0.528	1
CDKN1B	1	0.675	1
CREBBP	1	0.495	1
DEPDC6	1	0.378	1
EP300	1	0.823	1
FASLG	1	0.219	1
FOXO1	1	0.688	1
FOXO3	0.181	0.138	0.180
FOXO4	0.023	0.023	0.055
G6PC	0.156	0.172	0.173
IGF1	0.342	0.042	0.148
IGF1R	0.054	0.373	0.491
IGF2	0.028	0.019	0.084
INS	0.022	0.049	0.188
INSR	0.154	0.217	0.286
INSRR	0.139	0.247	0.224
IRS1	1	0.873	1
IRS2	1	0.569	1
IRS4	1	0.605	1
KAT2B	1	0.905	1
MAPK1	1	0.248	1
MAPK3	1	0.132	1
MAPK8	0.185	0.531	0.215
MAPK9	1	0.198	1
MAPK10	0.191	0.885	0.068
MAPKAP1	1	0.372	1
MLST8	1	0.593	1
MTOR	1	0.722	1
PCK1	1	0.547	1
PHLPP1	0.113	0.398	0.200
PHLPP2	1	0.364	1
PIK3CA	0.003	9.36×10^{-4}	0.022
PIK3CB	1	0.726	1
PIK3CD	1	0.828	1
PIK3R1	1	0.666	1
PIK3R2	1	0.722	1
PIK3R3	1	0.263	1
PPP2R5B	1	0.363	1
PRR5	0.355	0.163	0.257
PTEN	1	0.855	1
PTPN1	1	0.982	1
RBL2	1	0.061	1
RICTOR	1	0.343	1
SCP2	1	0.729	1
SGK1	0.091	0.007	0.016
SGK2	0.027	0.042	0.349
SIRT1	1	0.941	1
SIRT2	1	0.282	1
SIRT3	0.241	0.232	0.326
SKP2	1	0.898	1
SOCS1	1	0.349	1
SOCS3	1	0.996	1
SOD2	1	0.692	1
USP7	0.025	0.101	0.103
YWHAB	1	0.223	1
YWHAE	0.067	0.124	0.196
YWHAG	0.090	0.032	0.018
YWHAH	1	0.236	1
YWHAQ	0.228	0.175	0.293
YWHAZ	1	0.756	1

[a] Permutation (n=10,000) P value

apoptosis, chromosomal instability, and senescence (Yang et al. 2005). The same effects are observed in Pot1b knockout mice (He et al. 2009; Hockemeyer et al. 2008). In addition, telomerase-deficient Pot1b knockout mice show a reduction in lifespan compared to "normal" telomerase-deficient mice (Hockemeyer et al. 2008), which stresses the importance of TM in lifespan regulation. Most protein complexes contain one or several proteins essential for specific functions of the complex, e.g., binding, transport, or activation/repression activity. This indicates that, for pathways containing a protein complex, mutations in a single gene, encoding such an essential protein, could be sufficient to alter the function of the complex, which would explain the single-gene association in the TM pathway.

Table 6 Results of gene set analysis of telomere maintenance pathway gene SNP sets

Gene	PLINK set-based test[a]	Global test[a]	SNP ratio test[a]
ACD	1	0.491	1
DKC1	1	0.642	1
GAR1	1	0.281	1
NHP2	1	0.759	1
NOP10	1	0.208	1
POT1	0.007	0.014	0.019
TEP1	1	0.525	1
TERC	1	0.202	1
TERF1	1	0.821	1
TERF2	0.018	0.160	0.164
TERF2IP	1	0.825	1
TERT	1	0.471	1
TINF2	1	0.587	1

[a] Permutation (n=10,000)
P value

There are two main kinds of pathway analyses, explorative and candidate based. Since we want to focus on two pathways, the IIS and TM pathways, we performed candidate-based pathway analysis. The advantage of testing candidate pathways instead of explorative testing is the decreased penalty for multiple testing, due to the limited amount of tests performed. For information about pathways, several databases are available, e.g., Gene Ontology (Ashburner et al. 2000) and Kyoto Encyclopedia of Genes and Genomes (KEGG) (Kanehisa and Goto 2000), which are particularly useful for explorative studies (Wang et al. 2010). However, to our knowledge, the IIS and TM pathways are not described in sufficient detail in these databases, and we therefore assembled these pathways based on literature. Although the IIS pathway is available in KEGG (hsa04910; insulin-signaling pathway), only four of the nine IIS pathway genes that were associated with longevity, *AKT1*, *AKT3*, *INS*, and *PIK3CA*, were part of this pathway, which indicates that the pathway definition used in this study could have had a large influence on the results of the analysis.

Different pathway tests could show contradictory results, even when analyzing the same GWAS data (Wang et al. 2010). These discrepancies are caused by differences in, for example, the underlying statistical methods of the tests. Therefore, we used several pathway tests in parallel for our analysis. Some of the available pathway tests require SNP *P* values as input data, while others require raw genotypes (Wang et al. 2010). Given that we have GWAS data available, we selected pathway tests that make use of raw genotypes. All four selected pathway tests are self-contained tests which deal with the complexity of SNP set testing by permuting the case-control status. While, the PLINK set-based test, Global test, and SNP ratio test do not completely incorporate LD information, GRASS employs PCA to deal with correlations within each gene. A simulation study showed that in general, GRASS was more powerful than the PLINK set-based test (Chen et al. 2010). Simulation studies for Global test or SNP ratio test are not yet available. However, despite the differences between the methods, they all showed similar results for the IIS and TM pathways in this study.

SNP set analysis could have power to detect significant association, even if the power to detect associations in single SNP analysis is low (Fridley and Biernacka 2011), as was previously shown in the Welcome Trust Case Control Consortium (Torkamani et al. 2008). Our study has a power <1% to detect single SNP associations of the tested SNPs with an OR of 1.2 and a minor allele frequency of 0.25 (the mean frequency of the tested SNPs). However, because the small (non-significant) effects of the

SNPs are jointly tested, the pathway SNP set analysis is able to detect a significant association of the IIS and TM pathway. This indicates that SNP set analysis could be a useful approach for studies which showed no significant associations in single SNP analysis.

There is still much debate about the optimal size of the window used in SNP set analysis (Holmans 2010; Fridley and Biernacka 2011; Wang et al. 2010), and we choose a fixed window of 10 kb to take into account effects of SNPs in regulatory regions surrounding the genes. The same window was also used in a previous study of the IIS pathway (Pawlikowska et al. 2009). Although there is a chance that we will miss some functional SNPs, increasing the window would increase the chance that SNPs are included with no functional relationship to the tested gene.

The amount and diversity of SNPs measured per gene/pathway is highly variable between genotyping platforms used for GWAS. In addition, there is a large variety in allele frequencies and presence of SNPs between populations. For single SNP analysis, one is dependent on association of the same SNP (or a SNP in high LD) for replication. However, when due to varying allele frequencies, different SNPs associate in different populations, SNP set analysis determines the combined effect of SNPs within a gene and is able to overcome this problem. Therefore, replication of SNP set analysis is assumed to be more reproducible between genotyping platforms and populations (Luo et al. 2010; Wang et al. 2010). To support these assumptions, our findings should be replicated in other cohorts.

In conclusion, we have shown that genetic variation in genes involved in the IIS and TM pathways is associated with human longevity. In addition, we provide evidence that different self-contained tests show similar results when applied to candidate-based pathway analysis.

Acknowledgments We thank all participants of the Leiden Longevity Study and Rotterdam Study. The research leading to these results has received funding from the European Union's Seventh Framework Programme (FP7/2007-2011) under grant agreement no. 259679. This study was supported by a grant from the Innovation-Oriented Research Program on Genomics (SenterNovem IGE05007), the Centre for Medical Systems Biology, and the Netherlands Consortium for Healthy Ageing (grant 050-060-810), all in the framework of the Netherlands Genomics Initiative, Netherlands Organization for Scientific Research (NWO), and by Unilever Colworth. The generation and management of GWAS genotype data for the Rotterdam study are supported by the Netherlands Organisation for Scientific Research NWO Investments (nr. 175.010.2005.011, 911-03-012). This study is funded by the Research Institute for Diseases in the Elderly (014-93-015; RIDE2) and the Netherlands Genomics Initiative (NGI)/Netherlands Organisation for Scientific Research (NWO) project nr. 050-060-810. The Rotterdam Study is funded by the Erasmus Medical Center and Erasmus University, Rotterdam; the Netherlands Organization for the Health Research and Development (ZonMw); the Research Institute for Diseases in the Elderly (RIDE); the Ministry of Education, Culture and Science; the Ministry for Health, Welfare and Sports; the European Commission (DG XII); and the Municipality of Rotterdam.

References

Allsopp RC (1996) Models of initiation of replicative senescence by loss of telomeric DNA. Exp Gerontol 31:235–243

Anselmi CV, Malovini A, Roncarati R, Novelli V, Villa F, Condorelli G, Bellazzi R, Puca AA (2009) Association of the FOXO3A locus with extreme longevity in a southern Italian centenarian study. Rejuv Res 12:95–104

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet 25:25–29

Atzmon G, Cho M, Cawthon RM, Budagov T, Katz M, Yang X, Siegel G, Bergman A, Huffman DM, Schechter CB, Wright WE, Shay JW, Barzilai N, Govindaraju DR, Suh Y (2010) Evolution in health and medicine Sackler colloquium: genetic variation in human telomerase is associated with telomere length in Ashkenazi centenarians. Proc Natl Acad Sci USA 107(Suppl 1):1710–1717

Bakaysa SL, Mucci LA, Slagboom PE, Boomsma DI, McClearn GE, Johansson B, Pedersen NL (2007) Telomere length predicts survival independent of genetic influences. Aging Cell 6:769–774

Beekman M, Blauw GJ, Houwing-Duistermaat JJ, Brandt BW, Westendorp RG, Slagboom PE (2006) Chromosome 4q25, microsomal transfer protein gene, and human longevity: novel data and a meta-analysis of association studies. J Gerontol A-Biol 61:355–362

Bischoff C, Petersen HC, Graakjaer J, ndersen-Ranberg K, Vaupel JW, Bohr VA, Kolvraa S, Christensen K (2006) No association between telomere length and survival among the elderly and oldest old. Epidemiology 17:190–194

Blackburn EH (1991) Structure and function of telomeres. Nature 350:569–573

Blake JA, Bult CJ, Kadin JA, Richardson JE, Eppig JT (2011) The Mouse Genome Database (MGD): premier model organism resource for mammalian genomics and genetics. Nucleic Acids Res 39:D842–D848

Bluher M, Kahn BB, Kahn CR (2003) Extended longevity in mice lacking the insulin receptor in adipose tissue. Science 299:572–574

Cawthon RM, Smith KR, O'Brien E, Sivatchenko A, Kerber RA (2003) Association between telomere length in blood and mortality in people aged 60 years or older. Lancet 361:393–395

Chapman J, Whittaker J (2008) Analysis of multiple SNPs in a candidate gene or region. Genet Epidemiol 32:560–566

Chen LS, Hutter CM, Potter JD, Liu Y, Prentice RL, Peters U, Hsu L (2010) Insights into colon cancer etiology via a regularized approach to gene set analysis of GWAS data. Am J Hum Genet 86:860–871

Clancy DJ, Gems D, Harshman LG, Oldham S, Stocker H, Hafen E, Leevers SJ, Partridge L (2001) Extension of life-span by loss of CHICO, a *Drosophila* insulin receptor substrate protein. Science 292:104–106

Codd V, Mangino M, van der Harst P, Braund PS, Kaiser M, Beveridge AJ, Rafelt S, Moore J, Nelson C, Soranzo N, Zhai G, Valdes AM, Blackburn H, Mateo Leach I, de Boer RA, Goodall AH, Ouwehand W, van Veldhuisen DJ, van Gilst WH, Navis G, Burton PR, Tobin MD, Hall AS, Thompson JR, Spector T, Samani NJ (2010) Common variants near TERC are associated with mean telomere length. Nat Genet 42:197–199

Collins K, Mitchell JR (2002) Telomerase in the human organism. Oncogene 21:564–579

de Lange T (2005) Shelterin: the protein complex that shapes and safeguards human telomeres. Genes Dev 19:2100–2110

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 (2011) Genome-wide association study identifies a single major locus contributing to survival into old age; the APOE locus revisited. Aging Cell 10:686–698

Eriksson N, Macpherson JM, Tung JY, Hon LS, Naughton B, Saxonov S, Avey L, Wojcicki A, Pe'er I, Mountain J (2010) Web-based, participant-driven studies yield novel genetic associations for common traits. PLoS Genet 6: e1000993

Flachsbart F, Caliebe A, Kleindorp R, Blanche H, von Eller-Eberstein H, Nikolaus S, Schreiber S, Nebel A (2009) Association of FOXO3A variation with human longevity confirmed in German centenarians. Proc Natl Acad Sci USA 106:2700–2705

Fridley BL, Biernacka JM (2011) Gene set analysis of SNP data: benefits, challenges, and future directions. Eur J Hum Genet 19:837–843

Friedman DB, Johnson TE (1988) A mutation in the age-1 gene in *Caenorhabditis elegans* lengthens life and reduces hermaphrodite fertility. Genetics 118:75–86

Giannakou ME, Goss M, Junger MA, Hafen E, Leevers SJ, Partridge L (2004) Long-lived *Drosophila* with overexpressed dFOXO in adult fat body. Science 305:361

Goeman JJ, Buhlmann P (2007) Analyzing gene expression data in terms of gene sets: methodological issues. Bioinformatics 23:980–987

Goeman JJ, van de Geer SA, de Kort F, van Houwelingen HC (2004) A global test for groups of genes: testing association with a clinical outcome. Bioinformatics 20:93–99

Gudmundsson H, Gudbjartsson DF, Frigge M, Gulcher JR, Stefansson K (2000) Inheritance of human longevity in Iceland. Eur J Hum Genet 8:743–749

Harley CB, Futcher AB, Greider CW (1990) Telomeres shorten during ageing of human fibroblasts. Nature 345:458–460

Harrington L, McPhail T, Mar V, Zhou W, Oulton R, Bass MB, Arruda I, Robinson MO (1997) A mammalian telomerase-associated protein. Science 275:973–977

He H, Wang Y, Guo X, Ramchandani S, Ma J, Shen MF, Garcia DA, Deng Y, Multani AS, You MJ, Chang S (2009) Pot1b deletion and telomerase haploinsufficiency in mice initiate an ATR-dependent DNA damage response and elicit phenotypes resembling dyskeratosis congenita. Mol Cell Biol 29:229–240

Hjelmborg JV, Iachine I, Skytthe A, Vaupel JW, McGue M, Koskenvuo M, Kaprio J, Pedersen NL, Christensen K (2006) Genetic influence on human lifespan and longevity. Hum Genet 119:312–321

Hockemeyer D, Palm W, Wang RC, Couto SS, de Lange T (2008) Engineered telomere degradation models dyskeratosis congenita. Genes Dev 22:1773–1785

Holmans P (2010) Statistical methods for pathway analysis of genome-wide data for association with complex genetic traits. Adv Genet 72:141–179

Holzenberger M, Dupont J, Ducos B, Leneuve P, Geloen A, Even PC, Cervera P, Le BY (2003) IGF-1 receptor regulates lifespan and resistance to oxidative stress in mice. Nature 421:182–187

Hwangbo DS, Gershman B, Tu MP, Palmer M, Tatar M (2004) *Drosophila* dFOXO controls lifespan and regulates insulin signalling in brain and fat body. Nature 429:562–566

Kanehisa M, Goto S (2000) KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res 28:27–30

Kenyon C, Chang J, Gensch E, Rudner A, Tabtiang R (1993) A *C. elegans* mutant that lives twice as long as wild type. Nature 366:461–464

Kimura KD, Tissenbaum HA, Liu Y, Ruvkun G (1997) Daf-2, an insulin receptor-like gene that regulates longevity and diapause in *Caenorhabditis elegans*. Science 277:942–946

Kimura M, Hjelmborg JV, Gardner JP, Bathum L, Brimacombe M, Lu X, Christiansen L, Vaupel JW, Aviv A, Christensen K (2008) Telomere length and mortality: a study of leukocytes in elderly Danish twins. Am J Epidemiol 167:799–806

Levy D, Neuhausen SL, Hunt SC, Kimura M, Hwang SJ, Chen W, Bis JC, Fitzpatrick AL, Smith E, Johnson AD, Gardner JP, Srinivasan SR, Schork N, Rotter JI, Herbig U, Psaty BM, Sastrasinh M, Murray SS, Vasan RS, Province MA, Glazer NL, Lu X, Cao X, Kronmal R, Mangino M, Soranzo N, Spector TD, Berenson GS, Aviv A (2010)

Genome-wide association identifies OBFC1 as a locus involved in human leukocyte telomere biology. Proc Natl Acad Sci USA 107:9293–9298

Li Y, Wang WJ, Cao H, Lu J, Wu C, Hu FY, Guo J, Zhao L, Yang F, Zhang YX, Li W, Zheng GY, Cui H, Chen X, Zhu Z, He H, Dong B, Mo X, Zeng Y, Tian XL (2009) Genetic association of FOXO1A and FOXO3A with longevity trait in Han Chinese populations. Hum Mol Genet 18:4897–4904

Lin K, Dorman JB, Rodan A, Kenyon C (1997) Daf-16: An HNF-3/forkhead family member that can function to double the life-span of *Caenorhabditis elegans*. Science 278:1319–1322

Luo L, Peng G, Zhu Y, Dong H, Amos CI, Xiong M (2010) Genome-wide gene and pathway analysis. Eur J Hum Genet 18:1045–1053

Manolio TA, Brooks LD, Collins FS (2008) A HapMap harvest of insights into the genetics of common disease. J Clin Invest 118:1590–1605

Martin-Ruiz CM, Gussekloo J, van Heemst D, von Zglinicki T, Westendorp RG (2005) Telomere length in white blood cells is not associated with morbidity or mortality in the oldest old: a population-based study. Aging Cell 4:287–290

Mirabello L, Yu K, Kraft P, De Vivo I, Hunter DJ, Prescott J, Wong JY, Chatterjee N, Hayes RB, Savage SA (2010) The association of telomere length and genetic variation in telomere biology genes. Hum Mutat 31:1050–1058

Morris JZ, Tissenbaum HA, Ruvkun G (1996) A phosphatidylinositol-3-OH kinase family member regulating longevity and diapause in *Caenorhabditis elegans*. Nature 382:536–539

Moyzis RK, Buckingham JM, Cram LS, Dani M, Deaven LL, Jones MD, Meyne J, Ratliff RL, Wu JR (1988) A highly conserved repetitive DNA sequence, (TTAGGG) n, present at the telomeres of human chromosomes. Proc Natl Acad Sci USA 85:6622–6626

Newman AB, Walter S, Lunetta KL, Garcia ME, Slagboom PE, Christensen K, Arnold AM, Aspelund T, Aulchenko YS, Benjamin EJ, Christiansen L, D'Agostino RB Sr, Fitzpatrick AL, Franceschini N, Glazer NL, Gudnason V, Hofman A, Kaplan R, Karasik D, Kelly-Hayes M, Kiel DP, Launer LJ, Marciante KD, Massaro JM, Miljkovic I, Nalls MA, Hernandez D, Psaty BM, Rivadeneira F, Rotter J, Seshadri S, Smith AV, Taylor KD, Tiemeier H, Uh HW, Uitterlinden AG, Vaupel JW, Walston J, Westendorp RG, Harris TB, Lumley T, van Duijn CM, Murabito JM (2010) A meta-analysis of four genome-wide association studies of survival to age 90 years or older: the cohorts for heart and aging research in genomic epidemiology consortium. J Gerontol A-Biol 65:478–487

O'Dushlaine C, Kenny E, Heron EA, Segurado R, Gill M, Morris DW, Corvin A (2009) The SNP ratio test: pathway analysis of genome-wide association datasets. Bioinformatics 25:2762–2763

Ogg S, Paradis S, Gottlieb S, Patterson GI, Lee L, Tissenbaum HA, Ruvkun G (1997) The Fork head transcription factor DAF-16 transduces insulin-like metabolic and longevity signals in *C. elegans*. Nature 389:994–999

Pan W (2009) Asymptotic tests of association with multiple SNPs in linkage disequilibrium. Genet Epidemiol 33:497–507

Pawlikowska L, Hu D, Huntsman S, Sung A, Chu C, Chen J, Joyner AH, Schork NJ, Hsueh WC, Reiner AP, Psaty BM, Atzmon G, Barzilai N, Cummings SR, Browner WS, Kwok PY, Ziv E (2009) Association of common genetic variation in the insulin/IGF1 signaling pathway with human longevity. Aging Cell 8:460–472

Perls T, Kunkel LM, Puca AA (2002) The genetics of exceptional human longevity. J Mol Neurosci 19:233–238

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 81:559–575

Rafnar T, Sulem P, Stacey SN, Geller F, Gudmundsson J, Sigurdsson A, Jakobsdottir M, Helgadottir H, Thorlacius S, Aben KK, Blondal T, Thorgeirsson TE, Thorleifsson G, Kristjansson K, Thorisdottir K, Ragnarsson R, Sigurgeirsson B, Skuladottir H, Gudbjartsson T, Isaksson HJ, Einarsson GV, Benediktsdottir KR, Agnarsson BA, Olafsson K, Salvarsdottir A, Bjarnason H, Asgeirsdottir M, Kristinsson KT, Matthiasdottir S, Sveinsdottir SG, Polidoro S, Hoiom V, Botella-Estrada R, Hemminki K, Rudnai P, Bishop DT, Campagna M, Kellen E, Zeegers MP, de Verdier P, Ferrer A, Isla D, Vidal MJ, Andres R, Saez B, Juberias P, Banzo J, Navarrete S, Tres A, Kan D, Lindblom A, Gurzau E, Koppova K, de Vegt F, Schalken JA, van der Heijden HF, Smit HJ, Termeer RA, Oosterwijk E, van Hooij O, Nagore E, Porru S, Steineck G, Hansson J, Buntinx F, Catalona WJ, Matullo G, Vineis P, Kiltie AE, Mayordomo JI, Kumar R, Kiemeney LA, Frigge ML, Jonsson T, Saemundsson H, Barkardottir RB, Jonsson E, Jonsson S, Olafsson JH, Gulcher JR, Masson G, Gudbjartsson DF, Kong A, Thorsteinsdottir U, Stefansson K (2009) Sequence variants at the TERT-CLPTM1L locus associate with many cancer types. Nat Genet 41:221–227

Schoenmaker M, de Craen AJ, de Meijer PH, Beekman M, Blauw GJ, Slagboom PE, Westendorp RG (2006) Evidence of genetic enrichment for exceptional survival using a family approach: the Leiden longevity study. Eur J Hum Genet 14:79–84

Selman C, Partridge L, Withers DJ (2011) Replication of extended lifespan phenotype in mice with deletion of insulin receptor substrate 1. PLoS One 6:e16144

Skytthe A, Pedersen NL, Kaprio J, Stazi MA, Hjelmborg JV, Iachine I, Vaupel JW, Christensen K (2003) Longevity studies in GenomEUtwin. Twin Res 6:448–454

Soerensen M, Dato S, Christensen K, McGue M, Stevnsner T, Bohr VA, Christiansen L (2010) Replication of an association of variation in the FOXO3A gene with human longevity using both case-control and longitudinal data. Aging Cell 9:1010–1017

Taniguchi CM, Emanuelli B, Kahn CR (2006) Critical nodes in signalling pathways: insights into insulin action. Nat Rev Mol Cell Biol 7:85–96

Tatar M, Kopelman A, Epstein D, Tu MP, Yin CM, Garofalo RS (2001) A mutant *Drosophila* insulin receptor homolog that extends life-span and impairs neuroendocrine function. Science 292:107–110

Tatar M, Bartke A, Antebi A (2003) The endocrine regulation of aging by insulin-like signals. Science 299:1346–1351

Teichert M, Eijgelsheim M, Rivadeneira F, Uitterlinden AG, van Schaik RH, Hofman A, De Smet PA, van Gelder T, Visser LE, Stricker BH (2009) A genome-wide association study of acenocoumarol maintenance dosage. Hum Mol Genet 18:3758–3768

Torkamani A, Topol EJ, Schork NJ (2008) Pathway analysis of seven common diseases assessed by genome-wide association. Genomics 92:265–272

van der Horst A, Burgering BM (2007) Stressing the role of FoxO proteins in lifespan and disease. Nat Rev Mol Cell Biol 8:440–450

Vulliamy T, Beswick R, Kirwan M, Marrone A, Digweed M, Walne A, Dokal I (2008) Mutations in the telomerase component NHP2 cause the premature ageing syndrome dyskeratosis congenita. Proc Natl Acad Sci USA 105:8073–8078

Wang K, Li M, Hakonarson H (2010) Analysing biological pathways in genome-wide association studies. Nat Rev Genet 11:843–854

Willcox BJ, Donlon TA, He Q, Chen R, Grove JS, Yano K, Masaki KH, Willcox DC, Rodriguez B, Curb JD (2008) FOXO3A genotype is strongly associated with human longevity. Proc Natl Acad Sci USA 105:13987–13992

Yang Q, Zheng YL, Harris CC (2005) POT1 and TRF2 cooperate to maintain telomeric integrity. Mol Cell Biol 25:1070–1080