



Universiteit
Leiden
The Netherlands

Genetic and biomarker studies of human longevity

Deelen, J.

Citation

Deelen, J. (2014, June 25). *Genetic and biomarker studies of human longevity*. Retrieved from <https://hdl.handle.net/1887/26946>

Version: Corrected Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/26946>

Note: To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/26946> holds various files of this Leiden University dissertation

Author: Deelen, Joris

Title: Genetic and biomarker studies of human longevity

Issue Date: 2014-06-25

Chapter 3

Genome-wide association study identifies a single major locus contributing to survival into old age; the APOE locus revisited

Joris Deelen*, Marian Beekman*, Hae-Won Uh, Quinta Helmer, Maris Kuningas, Lene Christiansen, Dennis Kremer, Ruud van der Breggen, H Eka D Suchiman, Nico Lakenberg, Erik B van den Akker, Willemijn M Passtoors, Henning Tiemeier, Diana van Heemst, Anton J M de Craen, Fernando Rivadeneira, Eco J de Geus, Markus Perola, Frans J van der Ouderaa, David A Gunn, Dorret I Boomsma, André G Uitterlinden, Kaare Christensen, Cornelia M van Duijn, Bastiaan T Heijmans, Jeanine J Houwing-Duistermaat, Rudi G J Westendorp and P Eline Slagboom

*These authors contributed equally to this work.

Aging Cell 2011; **10**: 686-98.

Abstract

By studying the loci that contribute to human longevity, we aim to identify mechanisms that contribute to healthy aging. To identify such loci, we performed a genome-wide association study (GWAS) comparing 403 unrelated nonagenarians from long-living families included in the Leiden Longevity Study (LLS) and 1,670 younger population controls. The strongest candidate single nucleotide polymorphisms (SNPs) from this GWAS have been analyzed in a meta-analysis of nonagenarian cases from the Rotterdam Study, Leiden 85-plus study, and Danish 1905 cohort. Only one of the 62 prioritized SNPs from the GWAS analysis ($P < 1 \times 10^{-4}$) showed genome-wide significance with survival into old age in the meta-analysis of 4,149 nonagenarian cases and 7,582 younger controls (OR = 0.71 (95% CI 0.65 – 0.77), $P = 3.39 \times 10^{-17}$). This SNP, rs2075650, is located in *TOMM40* at chromosome 19q13.32 close to the *APOE* gene. Although there was only moderate linkage disequilibrium between rs2075650 and the apolipoprotein E (ApoE) $\epsilon 4$ defining SNP rs429358, we could not find an ApoE-independent effect of rs2075650 on longevity, either in cross-sectional or in longitudinal analyses. As expected, rs429358 associated with metabolic phenotypes in the offspring of the nonagenarian cases from the LLS and their partners. In addition, we observed a novel association between this locus and serum levels of insulin-like growth factor 1 in women ($P = 0.005$). In conclusion, the major locus determining familial longevity up to high age as detected by GWAS was marked by rs2075650, which tags the deleterious effects of the ApoE $\epsilon 4$ allele. No other major longevity locus was found.

Introduction

Worldwide human populations have shown an increase in mean life expectancy in the past two centuries [1]. This is mainly because of environmental factors, such as improved hygiene, nutrition, and health care. The large variation in healthy lifespan among the elderly has prompted research into the determinants of aging and lifespan regulation. The genetic contribution to human lifespan variation was estimated at 25-30% in twin studies [2-4]. The most prominent genetic influence is observed in families in which the capacity to attain a long lifespan clusters [5,6]. Exceptional longevity can be reached with a low degree of age-related disability [7,8], raising the question whether protective mechanisms against disease exist in long-lived subjects.

In most experimentally modified animal model systems, single-gene mutations in many different genes have major life extension effects [9,10]. However, natural human and animal longevity is presumed to be a complex trait [11]. In humans, both candidate gene and genome-wide genetic association approaches have been applied in an attempt to identify longevity loci. The frequency of genetic variants has been typically compared between nonagenarian cases and young controls, revealing loci at which genetic variants may contribute to a higher or lower probability of survival into old age. The initial candidate gene studies aimed at finding human longevity genes were dominated by contradictory results [12]. The more consistent evidence obtained by repeated observation in independent cohort studies for association with longevity

has so far only been observed for three loci, the *APOE* locus [12,13], the *FOXO3A* locus [14-17], and the *AKT1* locus [15]. Thus, despite the expectation that longevity would be influenced by many genetic variants with small effect sizes, the effect of variants has consistently been shown in only three genes.

Hypothesis-free genome-wide approaches have also been undertaken. Genome-wide linkage scans reported evidence for linkage with longevity on chromosome 4q25 [18], 3p24-22, 9q31-34, and 12q24 [19]. However, the evidence for these loci is still very weak, as the results, obtained in centenarians and their families, could not be replicated in nonagenarian sibling pairs [20] or have yet to be tested in other studies. A meta-genome-wide association study (GWAS) for survival to 90 years or older in 1,836 cases and 1,955 controls did not find any significant genome-wide associations [21]. Thus far, hypothesis-free approaches have not identified any loci involved in longevity.

In a few studies, such as the Ashkenazi Jewish Centenarian Study and the Leiden Longevity Study (LLS), different generations of long-lived families are being investigated for parameters and pathways contributing to the longevity phenotype [6,22]. The survival benefit of the LLS families is marked by a 30% decreased mortality risk in the survival analysis of three generations, i.e., the parents of the probands in this study (nonagenarian sibling pairs), their unselected additional siblings, and their offspring [6]. As compared to their partners, the offspring of nonagenarians siblings have a lower prevalence of type 2 diabetes, myocardial infarction, and hypertension

[23], a beneficial glucose, lipid, and thyroid metabolism, and a preservation of insulin sensitivity with age [24-28]. Hence, in middle age, these families display beneficial metabolic profiles.

Because the longevity phenotype is inherited in the LLS families, they offer a route to identify genetic variants that influence human longevity. Previously, we tested whether the absence of GWAS-identified alleles promoting common diseases might explain their familial longevity [29]. Longevity was not easily explained by the absence of disease-susceptibility alleles. More likely therefore, the genome of the long-lived harbors longevity-promoting alleles. To identify such loci, we performed a GWAS comparing nonagenarian siblings from the LLS and younger population

controls. We subsequently investigated emerging candidate single nucleotide polymorphisms (SNPs) in nonagenarian cases from the Rotterdam Study (RS), the Leiden 85-plus study, and the Danish 1905 cohort.

Results

GWAS

A GWAS was performed in nonagenarian participants from the LLS and middle-aged controls from the RS. Genotype data for 516,721 SNPs that passed quality control thresholds were analyzed in a comparison of 403 unrelated nonagenarians (94 years on average) and 1,670 controls (58 years on average). A flow chart of the consecutive

Figure 3.1 Flow chart of experimental work.

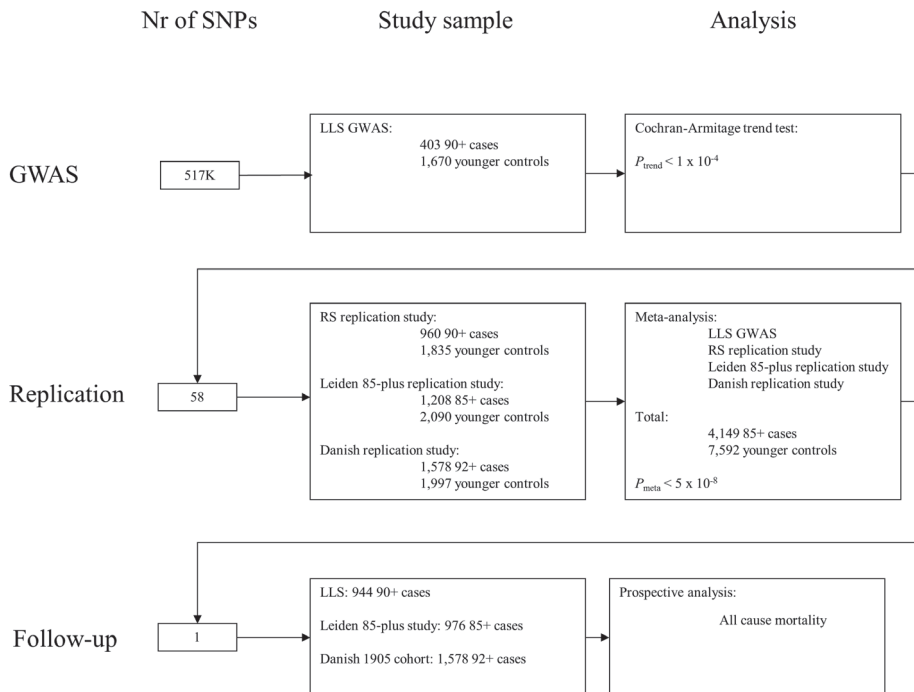


Table 3.1 Characteristics of the genotyped samples used for analysis.

Study	SNPs	Samples	<i>n</i>	Mean age	Age range	Men/women
LLS GWAS	517K	Cases	403	94	89 - 102	137/266
	517K	Controls	1,670	58	55 - 59	745/925
RS replication study	58	Cases	960	94	90 - 106	217/743
	58	Controls	1,835	62	60 - 65	809/1,026
Leiden 85-plus replication study	58	Cases	1,208	92	85 - 109	372/836
	58	Controls	2,090	35	15 - 70	743/1,347
Danish replication study	58	Cases	1,578	93	92 - 93	430/1,148
	58	Controls	1,997	57	46 - 68	900/1,097

SNPs, single nucleotide polymorphisms; LLS, Leiden Longevity Study; GWAS, genome-wide association study; RS, Rotterdam Study.

analysis steps is depicted in Figure 3.1 and a description of the population samples investigated in the GWAS and subsequent replication studies is given in Table 3.1. Results of the association analysis of stage 1 are depicted in Figure S3.1. None of the SNPs reached genome-wide significance ($P < 5 \times 10^{-8}$).

Replication studies

We prioritized the SNPs that had the most significant association with survival into old age according to the analysis of stage 1 ($P < 1 \times 10^{-4}$, Table S3.1). For 58 of the 62 selected SNPs, successful genotyping was obtained in the replication cohorts. In stage 2, these 58 SNPs were tested for association comparing 960 RS replication cases (mean age of 93 years), 1,208 Leiden 85-plus replication cases (mean age of 92 years), and 1,578 Danish replication cases (mean age of 93 years) with appropriate middle-aged population controls (Table 3.1). Meta-analysis for the 58 SNPs, comprising a total of 4,149 nonagenarian cases and 7,582 younger controls (from the LLS GWAS, RS

replication, Leiden 85-plus replication, and Danish replication studies), was performed.

Rs2075650 on chromosome 19 was the only SNP that was associated with survival into old age at the genome-wide significance level ($P = 3.39 \times 10^{-17}$) (Table S3.2A). The minor allele was underrepresented among the older cases as compared to middle-aged controls, hence associated with the decreased probability of carriers surviving into old age, corresponding to an odds ratio (OR) below unity (OR = 0.71 (95% CI 0.65 – 0.77)). This effect is observed in both sexes (Tables S3.2B and S3.2C). The remaining 57 SNPs did not show genome-wide significant effects on longevity either in men or women (Tables S3.2B, for men, and S3.2C, for women). The association of rs2075650 with survival did show some heterogeneity across the four studies ($P = 0.0495$), which is mainly because of the RS replication study.

Rs2075650 and the ApoE ε2/ε3/ε4 polymorphism

Rs2075650 is located in the *TOMM40* gene, next to the *APOE* gene (Figure S3.2). *APOE*

was previously associated with longevity [12,13]. The apolipoprotein E (ApoE) protein has three isoforms (ApoE ϵ 2, ApoE ϵ 3, and ApoE ϵ 4), which are defined by two SNPs, rs7412 (Arg136Cys; ϵ 2) and rs429358 (Cys112Arg; ϵ 4). A meta-analysis of rs7412 and rs429358 in the LLS GWAS, the Leiden 85-plus replication study, and the Danish replication study samples (3,189 cases and 5,757 controls), showed a significant association of rs429358 with longevity (OR = 0.62 (95% CI 0.56 – 0.68), $P = 1.33 \times 10^{-23}$), which was comparable to rs2075650 (OR = 0.67 (95% CI 0.61 – 0.74), $P = 9.15 \times 10^{-17}$). Rs7412 also showed an association with longevity, with a higher prevalence of the minor allele in nonagenarians (OR = 1.31 (95% CI 1.17 – 1.46), $P = 1.35 \times 10^{-6}$).

We observed only moderate linkage disequilibrium (LD) between rs2075650 and rs429358 ($r^2 = 0.553$) and low LD between rs2075650 and rs7412 ($r^2 = 0.014$) when analyzing all samples with genotype data of rs2075650, rs429358, and rs7412 ($n = 8,946$). Nevertheless, in a conditional analysis with rs429358 and rs7412 (Model 1, described in the Materials and methods section), rs2075650 was no longer associated with longevity (OR = 0.93 (95% CI 0.81 – 1.07), $P = 0.337$). The OR increased from 0.67 to 0.93, i.e., the deleterious effect of rs2075650 on longevity diminishes and is statistically non-significant. However, the deleterious effect of rs429358 (OR = 0.64 (95% CI 0.56 – 0.74), $P = 2.68 \times 10^{-9}$) and the protective effect of rs7412 (OR = 1.20 (95% CI 1.07 – 1.36), $P = 0.002$) on longevity remained significant.

To determine whether there was an ApoE-independent effect of rs2075650 on

survival after 90 years, prospective analysis of rs2075650, adjusted for rs429358 and rs7412, was performed. This analysis showed that carriers of the minor allele of rs2075650 displayed no increased mortality, i.e., a significant hazard ratio (HR) above 1, after 90 years of age independently of ApoE in two of the three cohorts analyzed (LLS, HR = 0.99 (95% CI 0.78 – 1.25), $P = 0.914$; Leiden 85-plus study, HR = 1.06 (95% CI 0.89 – 1.27), $P = 0.521$; Danish 1905 cohort, HR = 1.21 (95% CI 1.01 – 1.44), $P = 0.036$, Table S3.3A and Figure S3.3).

Overall, our results suggest that the association of rs2075650 with longevity is most likely a reflection of the effects of rs429358, caused by the moderate LD between the loci.

Association of rs429358 (ϵ 4) and rs2075650 with serum parameters

As previous studies showed that rs429358 was associated with several metabolic phenotypes [30-32], association of this SNP with relevant serum parameters was determined in the offspring of the elderly LLS cases and their partners ($n = 2,324$, Model 2 described in the Materials and methods section). We replicated the previously reported associations of rs429358 with plasma levels of ApoE ($P = 7.42 \times 10^{-28}$), total cholesterol ($P = 0.001$), low-density lipoprotein (LDL) cholesterol ($P = 4.91 \times 10^{-5}$), high-density lipoprotein (HDL) cholesterol ($P = 0.062$), and C-reactive protein (CRP) ($P = 0.028$) and with HDL ($P = 0.061$) and LDL particle size ($P = 0.062$) (Table 3.2). In addition, we detected a minor effect on Insulin-like growth factor 1 (IGF-1) ($P = 0.025$) and insulin-like growth

Table 3.2 Association analysis of serum parameters between carriers and non-carriers of rs429358.

Serum parameter	n	β	95% CI	P	n	β	95% CI	P	n	β	95% CI	P
ApoE (mg/dL)*	2,222	0.83	0.80 - 0.86	7.42×10^{-28}	1,015	0.85	0.80 - 0.89	2.74×10^{-11}	1,207	0.81	0.78 - 0.85	1.13×10^{-22}
Total cholesterol (mmol/L)	2,229	0.18	0.07 - 0.29	0.001	1,019	0.18	0.04 - 0.32	0.011	1,210	0.18	0.02 - 0.33	0.024
HDL cholesterol (mmol/L)	2,228	-0.04	-0.07 - 0.00	0.062	1,018	-0.04	-0.09 - 0.00	0.064	1,210	-0.03	-0.08 - 0.02	0.286
LDL cholesterol (mmol/L)	2,168	0.20	0.10 - 0.29	4.91×10^{-5}	978	0.19	0.07 - 0.31	0.002	1,190	0.20	0.07 - 0.33	0.003
HDL size (nm)	2,219	-0.04	-0.08 - 0.00	0.061	1,011	-0.04	-0.10 - 0.02	0.159	1,208	-0.04	-0.09 - 0.02	0.165
LDL size (nm)	2,219	-0.06	-0.13 - 0.00	0.062	1,011	-0.08	-0.19 - 0.02	0.117	1,208	-0.05	-0.14 - 0.03	0.246
CRP (mg/L)*	2,216	0.90	0.81 - 0.99	0.028	1,014	0.84	0.73 - 0.94	0.005	1,202	0.94	0.83 - 1.08	0.399
IGF-1 (nmol/L)	2,223	-0.49	-0.92 - -0.06	0.025	1,015	-0.10	-0.74 - 0.53	0.748	1,208	-0.80	-1.36 - -0.24	0.005
IGFBP3 (mg/L)	2,223	-0.09	-0.17 - 0.00	0.042	1,015	-0.06	-0.18 - 0.06	0.281	1,208	-0.10	-0.21 - 0.01	0.062
IGF-1/IGFBP3	2,223	-0.03	-0.11 - 0.04	0.384	1,015	0.04	-0.08 - 0.15	0.504	1,208	-0.09	-0.19 - 0.01	0.065

95% CI, 95% confidence interval; *ApoE*, apolipoprotein E; *HDL*, high-density lipoprotein; *LDL*, low-density lipoprotein; *CRP*, C-reactive protein; *IGF-1*, insulin-like growth factor 1; *IGFBP3*, insulin-like growth factor binding protein 3. *Natural log transformed serum parameter was used in the association analysis.

Table 3.3 Association analysis of serum parameters previously associated with familial longevity in middle age in the Leiden Longevity Study families between carriers and non-carriers of rs429358.

Serum parameter	n	β	95% CI	P	n	β	95% CI	P	n	β	95% CI	P
Glucose (mmol/L)	2,234	-0.05	-0.17 - 0.07	0.388	1,021	-0.16	-0.36 - 0.04	0.116	1,213	0.03	-0.11 - 0.18	0.660
Insulin (mU/L)*	2,163	0.95	0.88 - 1.02	0.123	990	0.93	0.84 - 1.03	0.158	1,173	0.96	0.87 - 1.06	0.400
HDL cholesterol (mmol/L)	2,228	-0.04	-0.07 - 0.00	0.062	1,018	-0.04	-0.09 - 0.00	0.064	1,210	-0.03	-0.08 - 0.02	0.286
Triglycerides (mmol/L)*	2,229	1.03	0.98 - 1.08	0.203	1,016	1.07	0.99 - 1.15	0.095	1,208	1.01	0.95 - 1.07	0.834
HDL Size (nm)	2,219	-0.04	-0.08 - 0.00	0.061	1,011	-0.04	-0.10 - 0.02	0.159	1,208	-0.04	-0.09 - 0.02	0.165
LDL Size (nm)	2,219	-0.06	-0.13 - 0.00	0.062	1,011	-0.08	-0.19 - 0.02	0.117	1,208	-0.05	-0.14 - 0.03	0.246
fT3 (pmol/L)	2,223	0.05	-0.02 - 0.12	0.141	1,015	0.07	-0.02 - 0.16	0.127	1,208	0.03	-0.06 - 0.13	0.470

95% CI, 95% confidence interval; *HDL*, high-density lipoprotein; *LDL*, low-density lipoprotein; *fT3*, free triiodothyronine. *Natural log transformed serum parameter was used in the association analysis.

factor binding protein 3 (IGFBP3) levels ($P = 0.042$) (Table 3.2). The effect on IGF-1 seems to be female-specific ($P = 0.005$ and $P = 0.748$, in women and men, respectively) and is still significant after correction for multiple testing. We observed no ApoE-independent effect of rs2075650 on these traits, except for an increase of 0.18 mmol/L total cholesterol ($P = 0.017$) and 0.14 mmol/L LDL cholesterol ($P = 0.014$) with each minor allele of rs2075650 (using Model 3 described in the Materials and methods section).

No significant effects of rs429358 were observed on glucose ($P = 0.388$), insulin ($P = 0.123$), triglyceride ($P = 0.203$), and free triiodothyronine (fT3) ($P = 0.141$) levels (Table 3.3); the phenotypes that have previously been associated, in middle age, with familial longevity in the LLS families [24-28].

Analysis of Alzheimer's disease SNPs

Rs2075650 has consistently been associated with an increased risk of Alzheimer's disease in several independent GWAS [33-35]. Therefore, we studied the effect of SNPs present in the AlzGene database (<http://www.alzgene.org/>) [36], on survival into old age in the LLS GWAS. Apart from rs2075650, none of the 751 measured Alzheimer's disease SNPs showed a significant association after adjustment for multiple testing (Table S3.4).

Analysis of *FOXO3A* and *AKT1* SNPs

Apart from *APOE*, two other genes have shown consistent evidence for association with longevity, *FOXO3A* [14-17] and *AKT1* [15]. For the longevity-promoting *FOXO3A* SNPs previously reported with centenarian longevity, we observed no association with

survival into old age in our nonagenarians (Table S3.5). For *AKT1*, one of the two measured SNPs, rs2498804, showed a significant association with survival into old age (OR = 0.75 (95% CI 0.63 – 0.89), $P = 0.001$) (Table S3.5).

Discussion

To identify common SNPs contributing to longevity, GWAS analysis of 403 nonagenarian cases and 1,670 population controls was performed. Of the 62 top associating SNPs, 58 were tested in a meta-analysis of 4,149 nonagenarian cases and 7,582 younger controls and we identified one SNP, rs2075650, that associated significantly with survival into old age ($P = 3.39 \times 10^{-17}$). Carriers of the minor allele had a 29% decreased probability of reaching 90 years on average. Although cases and controls originate from different generations, we concluded that there was no substructure to an extent that would affect the observations.

Rs2075650 is located in the *TOMM40* gene at chromosome 19q13.32 close to and centromeric of the *APOE* gene (Figure S3.2), which has shown consistent evidence for association with longevity [12,13]. The ApoE protein has three isoforms (ApoE $\epsilon 2$, ApoE $\epsilon 3$, and ApoE $\epsilon 4$) that are defined by two SNPs, rs7412 (Arg136Cys; $\epsilon 2$) and rs429358 (Cys112Arg; $\epsilon 4$). ApoE $\epsilon 4$ carriers have an increased risk of cardiovascular disease and Alzheimer's disease, while ApoE $\epsilon 2$ carriers are protected from these diseases [12,37,38]. Although we detected only moderate LD ($r^2 = 0.553$) between rs2075650 and the ApoE $\epsilon 4$ -defining SNP rs429358, we could

not detect a significant effect of rs2075650 on longevity independent of rs429358. Several prospective studies, including one with the Danish 1905 cohort [39], reported increased mortality for ApoE ϵ 4 carriers, even though there is still much debate about *APOE* being a "frailty gene" or a "longevity gene" [12,39-41]. The prospective data in the LLS and Leiden 85-plus study support the "frailty gene" hypothesis, as rs429358 affects mortality after 85 years and continues the effect after 90 years (HR = 1.08 (95% CI 1.03 – 1.13), P = 0.001 and HR = 1.08 (95% CI 1.03 – 1.12), P = 0.001, respectively, Table S3.3B and Figure S3.4). In these prospective studies, carriers of the minor allele of rs2075650 showed no increased mortality independent of rs429358, which indicates that the association of rs2075650 with longevity is most likely due to variation in the *APOE* gene. Although GWAS have reported significant associations between rs2075650 and Alzheimer's disease, brain imaging, total cholesterol, and CRP plasma levels [35,42-44], no analyses were performed to determine whether these associations are ApoE-independent. We observed no ApoE-independent effect on the phenotypes investigated in the LLS offspring and partners except for total and LDL cholesterol.

Previously, rs429358 had been associated with several metabolic phenotypes, such as ApoE, total cholesterol, HDL cholesterol, LDL cholesterol, and CRP levels, as well as HDL and LDL particle size [30-32] and, here, we have confirmed these findings using serum measurements of the offspring and partners from the LLS. Because the insulin/IGF-1 signaling

(IIS) pathway has a lifespan regulating effect in several model organisms [9,10] and humans [45], we also investigated the effect of rs429358 on serum levels of IGF-1 and IGFBP3, which both play a role in this pathway. Both proteins are involved in the etiology of several age-related diseases. However, up till now, it is not clear whether higher or lower serum levels are beneficial for longevity. Low IGF-1 serum levels associate with a decreased risk of cancer, but an increased risk of cardiovascular disease and neurodegenerative disease [46]. Previously, we showed in the Leiden 85-plus study cohort that genetic variants known to associate with lower IIS activity and IGF-1 serum levels at younger age associated with better survival at ages above 85 years [47]. However, the effect of these genetic variants on IGF-1 serum levels was not tested in the Leiden 85-plus study cohort. In addition, we showed previously that neither IGF-1 and IGFBP3 levels nor their ratio differed between partners and offspring from the LLS [24], which indicates that IGF-1 serum levels are, in middle age, not a marker for longevity, whereas a decreased risk of metabolic diseases was evident at that age in long-lived families [23]. In the current study, we found that the minor allele of rs429358 associates with lower IGF-1 levels in middle-aged women, which, to our knowledge, has not previously been reported. Like low IGF-1 levels, ApoE ϵ 4 was previously associated with an increased risk of developing cardiovascular disease and neurodegenerative disease [12,37,38]. Thus, the mechanism behind the increased risk of female ApoE ϵ 4 carriers of developing cardiovascular and/or neurodegenerative

diseases might involve serum levels of IGF-1 or other aspects of IIS activity reflected by these levels. Apart from lipid metabolism, the parameters determining the longevity phenotype in middle age in the LLS, such as glucose metabolism, insulin sensitivity, and thyroid hormone metabolism [24-28], were not influenced by the presence of the minor allele of rs429358. This indicates that it is likely that other loci could explain the differences in these phenotypes between LLS offspring and partners.

The strength of this study is that, by using a GWAS, we were able to replicate the previously reported association of the *APOE* locus with longevity [12,13] as the major locus. This was not observed in the previously published meta-GWAS of Newman *et al.* [21], possibly because of differences in the study design and population control selection between the studies. While Newman *et al.* used nonagenarian cases in a population-based design, we made use of a family-based design in which the families are genetically enriched for longevity. In addition, Newman *et al.* used population controls from the same cohort which had died before the age of 80. Between 60 and 80 years, however, there might already have been a selection on survival, decreasing the frequency of ApoE ϵ 4 carriers in the control group. In contrast, we made comparisons to a younger population group (55-60 years) from a different cohort (RS).

As we previously reported that long-lived individuals carry the same number of disease risk alleles for cardiovascular disease, cancer, and type 2 diabetes as young controls [29], we expected to primarily find longevity-promoting alleles. However, although most

of the 58 prioritized SNPs ($n = 43$) from the LLS GWAS showed a longevity-promoting effect ranging from 36 to 168%, none of them could be replicated in additional study populations of nonagenarian singletons. The only replicated locus is *APOE*, which is a mortality locus that has previously been reported to be the major locus responsible for Alzheimer's disease [33-35], a well-known age-related disease. Nevertheless, none of the other Alzheimer's disease loci showed an association with survival to 90 years, which indicates that the remaining genetic variation in longevity in the LLS could not be explained by the genetic variation which contributes to Alzheimer's disease. In addition to *APOE*, we also observed evidence for association at the previously reported *AKT1* locus [15] with survival into old age in the LLS GWAS, although the effect of this SNP is relatively small (25% decreased probability of becoming 90 years) compared to the effect of rs429358 (51%). The previously reported longevity-promoting effect of the *FOXO3A* locus could not be replicated in this study. This is probably due to the relatively low number of centenarians in the LLS GWAS case group, in which the effect of SNPs in *FOXO3A* on longevity seems to be most prominent. The still unexplained genetic variation in longevity might be attributable to rare variants or variants with small effects, which has previously been reported for other complex traits, such as Alzheimer's disease. These loci could not be identified in this study because of the relatively small number of cases in the LLS GWAS, the heterogeneity of factors influencing lifespan within populations, and the difference in the design of the studies used for replication.

One way to identify variants with small effects would be to increase the initial sample size of the GWAS and perform replication in other studies of nonagenarians. Given the higher heritability of longevity at older ages [48], one may also limit the study population to centenarians or supercentenarians. In addition to common variants with small effects, rare variants with large effects might play a role in longevity. By whole-genome/exome sequencing of long-lived subjects and their families, rare variants can be identified and associated with human longevity.

In conclusion, we have shown that the deleterious effect of the ApoE ϵ 4 allele, tagged by rs2075650, is the single major hit in our GWAS for longevity, indicating that no other major longevity locus was present among these nonagenarians. We confirmed the previously reported associations of the ApoE ϵ 4 allele with lipid metabolism parameters and report an additional effect on IGF-1 signaling in women. To identify genetic variants with smaller and protective effects on human lifespan, a meta-GWAS for longevity with a larger sample size is merited.

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 men and 91 years or older for women, representing <

0.5% of the Dutch population in 2001 [6]. 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 [20]. For the GWAS, 403 unrelated LLS siblings (one sibling from each sibling pair) were included (LLS GWAS cases).

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 [49]. 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 (RS GWAS controls). For the replication study, 960 cases above 90 years at time of recruitment (RS replication cases) and 1,825 controls between 60 and 65 years at baseline (RS replication controls) from the combined cohorts, for whom GWAS data were also available, were included.

Leiden 85-plus study

In the Leiden 85-plus study, two prospective population-based cohorts were recruited from inhabitants of Leiden [50,51]. Between 1987 and 1989, 673 subjects aged 85 years and older were enrolled in a prospective

study (Cohort 1). Between 1997 and 1999, 563 subjects were enrolled in the month of their 85th birthday with follow-up (Cohort 2). Subjects were visited at their home and there were no exclusion criteria related to health. DNA was available from the combined cohorts consisting of 1,208 subjects aged 85 years and older (Leiden 85-plus replication cases).

Netherlands Twin Registry

From the Netherlands Twin Registry (NTR), 2,090 unrelated participants of European descent for whom DNA was available were selected as control samples [52] (Leiden 85-plus replication controls). The substructure in the NTR has been reported before [53], and in this study, we included samples aged between 15 and 70 years at the time of blood sampling, without known family relations (i.e., those without any substructure).

Danish 1905 cohort

The participants in this study are from the Danish 1905 birth cohort recruited in 1998 [54] when they were aged 92-93 years. From this cohort, 3,600 subjects were still alive, of whom 2,262 participated in the study. Participants were subjected to a home-based interview on health and lifestyle parameters, physical and cognitive function tests, and the collection of biological material. The current genetic study comprises a total of 1,578 of these individuals (Danish replication cases). Survival was followed up until January 2010. Ninety-nine percent (1,561 subjects) of subjects died in the 12 years of follow-up. Control samples were 1997 twins (one twin for each pair) between 46 and 68 years of

age collected from all over Denmark (Danish replication controls).

The cases in all three replication cohorts originate from population-based cohort studies from a genetic background similar to the LLS [55]. All the participants in these studies have signed an informed consent.

Genotyping

Genome-wide association study

LLS GWAS cases were genotyped using Illumina Infinium HD Human660W-Quad BeadChips (Illumina, San Diego, CA, USA). The RS-I and RS-II cohorts were genotyped using Illumina Infinium II HumanHap 550K Beadchips and Illumina Infinium II HumanHap550-Duo BeadChips (Illumina), respectively [49].

For the GWAS, we selected 551,606 SNPs for analysis because these were genotyped in both the LLS GWAS cases and (some of) the RS GWAS controls. Of these 551,606 SNPs, 34,885 SNPs were excluded on the basis of the following criteria: SNP call rate < 0.95 or minor allele frequency < 0.01 in RS GWAS controls or LLS GWAS cases ($n = 8,908$ and $n = 24,586$, respectively), and $P_{\text{HWE}} < 10^{-4}$ in RS GWAS controls ($n = 1,355$). In addition, SNPs with a between-chip effect in the RS GWAS controls were removed using a genotype trend test comparing the RS GWAS controls from RS-I with RS-II ($n = 36$), leaving 516,721 SNPs for statistical analysis. The Illumina clusterplots of the SNPs with $P < 1 \times 10^{-4}$ ($n = 71$) were visually inspected to confirm high-quality genotyping and 9 SNPs were excluded on the basis of bad clustering in the LLS GWAS cases or RS GWAS controls.

Genotype data were used to confirm gender and family relationships. Two RS GWAS control samples were excluded because of abnormalities in the sex chromosomes (both samples had Triple X Syndrome). Latent clustering of genotypes because of population substructure was assessed by pairwise identity-by-state (IBS) distance using Graphical Relationship Representation (<http://bioinformatics.well.ox.ac.uk/GRR>) [56]. LLS GWAS cases showed no relationship errors. From the RS GWAS controls, 59 samples were excluded because of high IBS. In total, 403 LLS GWAS cases and 1,670 RS GWAS control samples with a sample call rate > 0.95 were analyzed. Because cases and controls originate from different generations, we investigated whether substructure in these cohorts could have influenced the observed associations. IBS estimates for all pairs of subjects in the data set were computed on a randomly selected set of 10% of the SNPs that passed quality control thresholds, using the `--genome`, `--cluster`, and `--mds-plot 4` commands in PLINK (<http://pngu.mgh.harvard.edu/purcell/plink>) [57]. The first two resulting principal components (C1 and C2) were plotted against each other, which gives a representation of the data in two dimensions. In the resulting scatter plot, each point represents an individual (green = LLS GWAS case and blue = RS GWAS control, Figure S3.5). If there had been substructure, one would see multiple clusters in one plot. However, because all our samples seem to be in one cluster, we concluded that there was no substructure to an extent that would affect the observations.

Replication studies

For the RS replication study, we used the existing GWAS data in the Rotterdam Study after the quality control screening described by Teichert *et al.* [49]. For the Leiden 85-plus and Danish replication studies, genotyping was performed using the Sequenom MassARRAY iPLEX Gold and TaqMan SNP Genotyping assays. Of the 62 prioritized SNPs, 58 could be designed for replication studies using Sequenom, of which 56 were successfully genotyped in > 95% of the samples displayed in Table 3.1. The average genotype call rate for SNPs genotyped with Sequenom was 98.40%, and the average concordance rate with GWAS data among the LLS GWAS cases was 99.97%. For 2 of the 6 SNPs that could not be genotyped with Sequenom, rs2075650 and rs642990, pre-designed TaqMan SNP genotyping assays (C___3084828_20 and C___2206314_20, respectively) were used for genotyping, following the manufacturer's instructions. The average genotype call rate for the SNPs genotyped with TaqMan was 99.04%, and the average concordance rate with GWAS data among the LLS GWAS cases was 100%.

ApoE $\epsilon 2/\epsilon 3/\epsilon 4$ polymorphism

The ApoE $\epsilon 2/\epsilon 3/\epsilon 4$ defining SNPs, rs429358 (Cys112Arg; $\epsilon 4$) and rs7412 (Arg136Cys; $\epsilon 2$), were genotyped in the LLS GWAS cases, Leiden 85-plus replication study, and Danish replication study controls using pre-designed TaqMan SNP genotyping assays (C___3084793_20 and C___904973_10, respectively). For the RS GWAS controls and Danish replication study cases, previously measured data were used [39,58].

Measurement of serum parameters

All standard serum measurements were performed using fully automated equipment.

Glucose, total cholesterol, HDL cholesterol, and triglyceride levels were measured using the Hitachi Modular P 800 (Roche, Almere, the Netherlands) [24]. LDL cholesterol was calculated using the Friedewald formula [59].

LDL and HDL particle sizes were measured using proton NMR spectroscopy (LipoScience Inc, Raleigh, NY, USA) [60].

IGF-1, IGFBP3, and insulin levels were measured using the Immulite 2500 (DPC, Los Angeles, CA, USA) [24].

ft3 was measured using the Modular E170 and high-sensitivity CRP was measured using Cobas Integra 800 (both from Roche) [26].

The level of ApoE was determined in serum samples using a human ApoE-specific sandwich ELISA [61,62].

Statistical analysis

GWAS and replication studies

For the association analysis of the GWAS data, we applied a Cochran-Armitage trend test [63,64]. For X-linked SNPs, the genotypes of the men were considered as homozygous genotypes. SNPs with a P -value $< 1 \times 10^{-4}$ ($n = 62$) were selected for replication. Odds ratios were estimated and the corresponding 95% confidence intervals were computed. For meta-analyses, a fixed effect approach was used. Scores and their variances were computed within each study and combined across the four studies to obtain a single meta-statistic. P -values below 5×10^{-8} were considered as genome-wide significant [65]. The between-study variance

was calculated to determine heterogeneity across the four studies. All these analysis were performed using Bioconductor R (<http://www.bioconductor.org>) [66].

The quantile–quantile plot (Figure S3.6), constructed using Bioconductor R (<http://www.bioconductor.org>) [66], showed that the P -value distribution of stage 1 conformed to a null distribution at all but the extreme tail. The genomic inflation factor (λ), which measures over-dispersion of test statistics from association tests indicating population stratification, was 1.027 and we therefore decided not to adjust for population stratification.

Linkage Disequilibrium between rs2075650 and the ApoE $\epsilon 2/\epsilon 3/\epsilon 4$ polymorphism

Pairwise linkage disequilibrium (LD) between rs2075650 and the ApoE $\epsilon 2/\epsilon 3/\epsilon 4$ polymorphism determining SNPs rs7412 and rs429358 was calculated in 8,946 individuals using the `--ld` command in PLINK (<http://pengu.mgh.harvard.edu/purcell/plink>) [57].

ApoE-independent association of rs2075650 with longevity

To determine whether the association of rs2075650 with longevity was independent of the ApoE $\epsilon 2/\epsilon 3/\epsilon 4$ polymorphism, a logistic regression model with adjustment for rs429358, rs7412 and an interaction term for $\epsilon 2/\epsilon 3$ with $\epsilon 3/\epsilon 4$ was tested [67]:

$$\text{Logit}(P_{\text{status}} = 1) = \beta_0 + \beta_1 * \text{rs2075650} + \beta_2 * \text{rs429358} + \beta_3 * \text{rs7412} + \beta_4 * (\text{rs429358} * \text{rs7412}) + \beta_5 * \text{study (Model 1)}$$

status was coded as 0 (control) or 1 (long-lived case), *study* was coded as 0 (LLS

GWAS), 1 (Leiden 85-plus replication study) or 2 (Danish replication study) and the genotypes of *rs2075650*, *rs429358* and *rs7412* were coded as 0 (homozygous for the common allele), 1 (heterozygous) or 2 (homozygous for the rare allele). STATA/SE 11.1 (StataCorp LP, TX, USA) was used for this analysis.

Prospective analysis

Prospective analysis of *rs2075650* and *rs429358* was performed with 944 nonagenarian siblings from the LLS, 976 octogenarians and nonagenarians from the Leiden 85-plus study, and 1,578 nonagenarians from the Danish 1905 cohort.

After a mean follow-up time of 5.7 years (LLS), 14.8 years (Leiden 85-plus study), and 11.4 years (Danish 1905 cohort), 73.2% ($n = 691$) (LLS), 84.8% ($n = 828$) (Leiden 85-plus study), and 98.9% ($n = 1,561$) (Danish 1905 cohort) of the individuals had died.

Mortality analyses were performed with STATA/SE 11.1 (StataCorp LP) using a gender-adjusted, left-truncated Cox proportional hazards model to adjust for late entry into the data set according to age.

*Association of *rs429358* ($\epsilon 4$) and *rs2075650* with serum parameters*

To determine the association of *rs429358* and the ApoE-independent association of *rs2075650* with serum parameters in the offspring and their partners from the LLS the following regression models were tested:

$$\text{Serum parameter} = \beta_0 + \beta_1 * \text{age} + \beta_2 * \text{gender} + \beta_3 * (\text{age} * \text{gender}) + \beta_4 * \text{group} + \beta_5 * \text{rs429358} \text{ (Model 2)}$$

$$\text{Serum parameter} = \beta_0 + \beta_1 * \text{age} + \beta_2 * \text{gender} + \beta_3 * (\text{age} * \text{gender}) + \beta_4 * \text{group} + \beta_5 * \text{rs2075650} + \beta_6 * \text{rs429358} + \beta_7 * \text{rs7412} + \beta_8 * (\text{rs429358} * \text{rs7412}) \text{ (Model 3)}$$

age was coded in years, *gender* was coded as 1 (male) or 2 (female), *group* was coded as 0 (partner) or 1 (offspring) and the genotypes of *rs2075650*, *rs429358* and *rs7412* were coded as 0 (homozygous for the common allele), 1 (heterozygous) or 2 (homozygous for the rare allele). Robust standard errors were used to account for sibship relations. STATA/SE 11.1 (StataCorp LP) was used for these analyses.

Acknowledgments

We thank all participants of the Leiden Longevity Study, Rotterdam Study, Leiden 85-plus study, Netherlands Twin Register, and the Danish 1905 cohort. 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 is 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. The data collection of the Danish 1905 cohort is supported by the US National Institute on Aging, research grant NIA-P01-AG08761, and the Danish Aging Research Center is supported by a grant from the VELUX foundation. The data collection of the Netherlands Twin Register is supported by NWO (904-61-090; 904-61-193; 480-

04-004; 400-05-717; SPI 56-464-14192), Center for Medical Systems Biology (NWO Genomics; Centre for Neurogenomics and Cognitive Research (CNCR-VU); the EU (EU/QLRT-2001-01254); Genome scan for neuroticism (NIMH R01 MH059160); the Geestkracht program of ZonMW (10-000-1002), and institutes involved in NESDA (VU University Medical Center, Leiden University Medical Center, GGZ Buitendamstel-Geestgronden, Rivierduinen, University Medical Center Groningen, GGZ Lentis, GGZ Friesland, GGZ Drenthe). The genotyping of the samples was provided through the Genetic Association Information Network (GAIN).

References

- Oeppen J, Vaupel JW. Demography. Broken limits to life expectancy. *Science* 2002; **296**: 1029-31.
- Gudmundsson H, Gudbjartsson DF, Frigge M, Gulcher JR, Stefansson K. Inheritance of human longevity in Iceland. *Eur J Hum Genet* 2000; **8**: 743-9.
- Hjelmberg JV, Iachine I, Skyttthe A, Vaupel JW, McGue M, Koskenvuo M, *et al.* Genetic influence on human lifespan and longevity. *Hum Genet* 2006; **119**: 312-21.
- Skyttthe A, Pedersen NL, Kaprio J, Stazi MA, Hjelmberg JV, Iachine I, *et al.* Longevity studies in GenomeUtwinn. *Twin Res* 2003; **6**: 448-54.
- Perls T, Shea-Drinkwater M, Bowen-Flynn J, Ridge SB, Kang S, Joyce E, *et al.* Exceptional familial clustering for extreme longevity in humans. *J Am Geriatr Soc* 2000; **48**: 1483-5.
- 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; **14**: 79-84.
- Christensen K, McGue M, Petersen I, Jeune B, Vaupel JW. Exceptional longevity does not result in excessive levels of disability. *Proc Natl Acad Sci U S A* 2008; **105**: 13274-9.
- Terry DF, Sebastiani P, Andersen SL, Perls TT. Disentangling the roles of disability and morbidity in survival to exceptional old age. *Arch Intern Med* 2008; **168**: 277-83.
- Fontana L, Partridge L, Longo VD. Extending healthy life span--from yeast to humans. *Science* 2010; **328**: 321-6.
- Kenyon CJ. The genetics of ageing. *Nature* 2010; **464**: 504-12.
- Finch CE, Tanzi RE. Genetics of aging. *Science* 1997; **278**: 407-11.
- Christensen K, Johnson TE, Vaupel JW. The quest for genetic determinants of human longevity: challenges and insights. *Nat Rev Genet* 2006; **7**: 436-48.
- 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; **6**: 29-32.
- 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; **106**: 2700-5.
- Pawlikowska L, Hu D, Huntsman S, Sung A, Chu C, Chen J, *et al.* Association of common genetic variation in the insulin/IGF1 signaling pathway with human longevity. *Aging Cell* 2009; **8**: 460-72.
- Soerensen M, Dato S, Christensen K, McGue M, Stevnsner T, Bohr VA, *et al.* Replication of an Association of Variation in the FOXO3A Gene with Human Longevity Using Both Case-control and Longitudinal Data. *Aging Cell* 2010; **9**: 1010-7.
- 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; **105**: 13987-92.
- Puca AA, Daly MJ, Brewster SJ, Matise TC, Barrett J, Shea-Drinkwater M, *et al.* A genome-wide scan for linkage to human exceptional longevity identifies a locus on chromosome 4. *Proc Natl Acad Sci U S A* 2001; **98**: 10505-8.
- Boyden SE, Kunkel LM. High-density genomewide linkage analysis of exceptional human longevity identifies multiple novel Loci. *PLoS One* 2010; **5**: e12432.
- Beekman M, Blauw GJ, Houwing-Duistermaat JJ, Brandt BW, Westendorp RG, Slagboom PE. Chromosome 4q25, microsomal transfer protein gene, and human longevity: novel data and a meta-analysis of association studies. *J Gerontol A Biol Sci Med Sci* 2006; **61**: 355-62.
- Newman AB, Walter S, Lunetta KL, Garcia ME, Slagboom PE, Christensen K, *et al.* A meta-analysis of four genome-wide association studies of survival to age 90 years or older: the Cohorts for Heart and Aging Research in Genomic Epidemiology Consortium. *J Gerontol A Biol Sci Med Sci* 2010; **65**: 478-87.
- Atzmon G, Schechter C, Greiner W, Davidson D, Rennett G, Barzilai N. Clinical

- phenotype of families with longevity. *J Am Geriatr Soc* 2004; **52**: 274-7.
23. Westendorp RG, van Heemst D, 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; **57**: 1634-7.
 24. 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; **1**: 714-22.
 25. 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; **58**: 564-9.
 26. 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; **65**: 365-8.
 27. Vaarhorst AA, Beekman M, Suchiman HE, van Heemst D, Houwing-Duistermaat JJ, Westendorp RG, *et al.* Lipid metabolism in long-lived families: the Leiden Longevity Study. *Age (Dordr)* 2011; **33**: 219-27.
 28. Wijsman CA, Rozing MP, Streefland TC, le Cessie S, Mooijaart SP, Slagboom PE, *et al.* Familial longevity is marked by enhanced insulin sensitivity. *Aging Cell* 2011; **10**: 114-21.
 29. Beekman M, Nederstigt C, Suchiman HE, Kremer D, van der Breggen R, Lakenberg N, *et al.* Genome-wide association study (GWAS)-identified disease risk alleles do not compromise human longevity. *Proc Natl Acad Sci USA* 2010; **107**: 18046-9.
 30. Boerwinkle E, Utermann G. Simultaneous effects of the apolipoprotein E polymorphism on apolipoprotein E, apolipoprotein B, and cholesterol metabolism. *Am J Hum Genet* 1988; **42**: 104-12.
 31. Hubacek JA, Peasey A, Pikhart H, Stavek P, Kubinova R, Marmot M, *et al.* APOE polymorphism and its effect on plasma C-reactive protein levels in a large general population sample. *Hum Immunol* 2010; **71**: 304-8.
 32. Topic A, Spasojevic K, V, Zeljkovic A, Vekic J, Jelic IZ. Gender-related effect of apo E polymorphism on lipoprotein particle sizes in the middle-aged subjects. *Clin Biochem* 2008; **41**: 361-7.
 33. Harold D, Abraham R, Hollingworth P, Sims R, Gerrish A, Hamshere ML, *et al.* Genome-wide association study identifies variants at CLU and PICALM associated with Alzheimer's disease. *Nat Genet* 2009; **41**: 1088-93.
 34. Lambert JC, Heath S, Even G, Campion D, Sleegers K, Hiltunen M, *et al.* Genome-wide association study identifies variants at CLU and CR1 associated with Alzheimer's disease. *Nat Genet* 2009; **41**: 1094-9.
 35. Seshadri S, Fitzpatrick AL, Ikram MA, DeStefano AL, Gudnason V, Boada M, *et al.* Genome-wide analysis of genetic loci associated with Alzheimer disease. *JAMA* 2010; **303**: 1832-40.
 36. Bertram L, McQueen MB, Mullin K, Blacker D, Tanzi RE. Systematic meta-analyses of Alzheimer disease genetic association studies: the AlzGene database. *Nat Genet* 2007; **39**: 17-23.
 37. Corder EH, Saunders AM, Strittmatter WJ, Schmechel DE, Gaskell PC, Small GW, *et al.* Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. *Science* 1993; **261**: 921-3.
 38. Eichner JE, Kuller LH, Orchard TJ, Grandits GA, McCallum LM, Ferrell RE, *et al.* Relation of apolipoprotein E phenotype to myocardial infarction and mortality from coronary artery disease. *Am J Cardiol* 1993; **71**: 160-5.
 39. Jacobsen R, Martinussen T, Christiansen L, Jeune B, Andersen-Ranberg K, Vaupel JW, *et al.* Increased effect of the ApoE gene on survival at advanced age in healthy and long-lived Danes: two nationwide cohort studies. *Aging Cell* 2010; **9**: 1004-9.
 40. Ewbank DC. Differences in the association between apolipoprotein E genotype and mortality across populations. *J Gerontol A Biol Sci Med Sci* 2007; **62**: 899-907.

41. Gerdes LU, Jeune B, Ranberg KA, Nybo H, Vaupel JW. Estimation of apolipoprotein E genotype-specific relative mortality risks from the distribution of genotypes in centenarians and middle-aged men: apolipoprotein E gene is a “frailty gene,” not a “longevity gene”. *Genet Epidemiol* 2000; **19**: 202-10.
42. Aulchenko YS, Ripatti S, Lindqvist I, Boomsma D, Heid IM, Pramstaller PP, *et al.* Loci influencing lipid levels and coronary heart disease risk in 16 European population cohorts. *Nat Genet* 2009; **41**: 47-55.
43. Reiner AP, Barber MJ, Guan Y, Ridker PM, Lange LA, Chasman DI, *et al.* Polymorphisms of the HNF1A gene encoding hepatocyte nuclear factor-1 alpha are associated with C-reactive protein. *Am J Hum Genet* 2008; **82**: 1193-201.
44. Shen L, Kim S, Risacher SL, Nho K, Swaminathan S, West JD, *et al.* Whole genome association study of brain-wide imaging phenotypes for identifying quantitative trait loci in MCI and AD: A study of the ADNI cohort. *Neuroimage* 2010; **53**: 1051-63.
45. Suh Y, Atzmon G, Cho MO, Hwang D, Liu B, Leahy DJ, *et al.* Functionally significant insulin-like growth factor I receptor mutations in centenarians. *Proc Natl Acad Sci U S A* 2008; **105**: 3438-42.
46. Juul A. Serum levels of insulin-like growth factor I and its binding proteins in health and disease. *Growth Horm IGF Res* 2003; **13**: 113-70.
47. van Heemst D, Beekman M, Mooijaart SP, Heijmans BT, Brandt BW, Zwaan BJ, *et al.* Reduced insulin/IGF-1 signalling and human longevity. *Aging Cell* 2005; **4**: 79-85.
48. Tan Q, Zhao JH, Iachine I, Hjelmberg J, Vach W, Vaupel JW, *et al.* Power of non-parametric linkage analysis in mapping genes contributing to human longevity in long-lived sib-pairs. *Genet Epidemiol* 2004; **26**: 245-53.
49. Teichert M, Eijgelsheim M, Rivadeneira F, Uitterlinden AG, van Schaik RH, Hofman A, *et al.* A genome-wide association study of acenocoumarol maintenance dosage. *Hum Mol Genet* 2009; **18**: 3758-68.
50. Bootsma-van der Wiel A, van Exel E, de Craen AJ, Gussekloo J, Lagaay AM, Knook DL, *et al.* A high response is not essential to prevent selection bias: results from the Leiden 85-plus study. *J Clin Epidemiol* 2002; **55**: 1119-25.
51. Weverling-Rijnsburger AW, Blauw GJ, Lagaay AM, Knook DL, Meinders AE, Westendorp RG. Total cholesterol and risk of mortality in the oldest old. *Lancet* 1997; **350**: 1119-23.
52. 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; **16**: 335-42.
53. Sullivan PF, de Geus EJ, Willemsen G, James MR, Smit JH, Zandbelt T, *et al.* Genome-wide association for major depressive disorder: a possible role for the presynaptic protein piccolo. *Mol Psychiatry* 2009; **14**: 359-75.
54. Nybo H, Petersen HC, Gaist D, Jeune B, Andersen K, McGue M, *et al.* Predictors of mortality in 2,249 nonagenarians--the Danish 1905-Cohort Survey. *J Am Geriatr Soc* 2003; **51**: 1365-73.
55. Heath SC, Gut IG, Brennan P, McKay JD, Bencko V, Fabianova E, *et al.* Investigation of the fine structure of European populations with applications to disease association studies. *Eur J Hum Genet* 2008; **16**: 1413-29.
56. Abecasis GR, Cherny SS, Cookson WO, Cardon LR. GRR: graphical representation of relationship errors. *Bioinformatics* 2001; **17**: 742-3.
57. 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; **81**: 559-75.
58. Slooter AJ, Cruts M, Kalmijn S, Hofman A, Breteler MM, Van Broeckhoven C, *et al.* Risk estimates of dementia by apolipoprotein E genotypes from a population-based incidence study: the Rotterdam Study. *Arch Neurol* 1998; **55**: 964-8.

59. Friedewald WT, Levy RI, Fredrickson DS. Estimation of the concentration of low-density lipoprotein cholesterol in plasma, without use of the preparative ultracentrifuge. *Clin Chem* 1972; **18**: 499-502.
60. 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; **3**: e495.
61. Mooijaart SP, Berbee JF, van Heemst D, Havekes LM, de Craen AJ, Slagboom PE, *et al.* ApoE plasma levels and risk of cardiovascular mortality in old age. *PLoS Med* 2006; **3**: e176.
62. van Vlijmen BJ, van den Maagdenberg AM, Gijbels MJ, van der Boom H, HogenEsch H, Frants RR, *et al.* Diet-induced hyperlipoproteinemia and atherosclerosis in apolipoprotein E3-Leiden transgenic mice. *J Clin Invest* 1994; **93**: 1403-10.
63. Armitage P. Tests for linear trends in proportions and frequencies. *Biometrics* 1955; **11**: 375-86.
64. Cochran WG. Some methods for strengthening the common X^2 tests. *Biometrics* 1954; **10**: 417-51.
65. Pe'er I, Yelensky R, Altshuler D, Daly MJ. Estimation of the multiple testing burden for genomewide association studies of nearly all common variants. *Genet Epidemiol* 2008; **32**: 381-5.
66. 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**: R80.
67. Ronald J, Rajagopalan R, Ranchalis JE, Marshall JK, Hatsukami TS, Heagerty PJ, *et al.* Analysis of recently identified dyslipidemia alleles reveals two loci that contribute to risk for carotid artery disease. *Lipids Health Dis* 2009; **8**: 52.

Supplementary Information

The supplementary information belonging to this chapter can be found at: <http://onlinelibrary.wiley.com/doi/10.1111/j.1474-9726.2011.00705.x/supinfo>.

Figure S3.1 Manhattan plot presenting the $-\log_{10} P$ -value from the Cochran-Armitage trend test for the 516,721 SNPs that passed the quality control thresholds in the LLS GWAS.

Figure S3.2 Genomic region surrounding rs2075650 (obtained from the UCSC genome browser (<http://genome.ucsc.edu/>)). The physical distances between rs2075650 and rs429358 and between rs2075650 and rs7412 are 16.32 kb and 16.46 kb, respectively.

Figure S3.3 Kaplan-Meier curves showing the survival rate over years to follow-up of ApoE $\epsilon 3\epsilon 3$ carriers with zero (solid line) or one (large dashed line) minor allele(s) of rs2075650 in the LLS, Leiden 85-plus study and Danish 1905 cohort.

Figure S3.4 Kaplan-Meier curves showing the survival rate over years to follow-up of carriers of zero (solid line), one (large dashed line), or two (small dashed line) $\epsilon 4$ allele(s) of rs429358 in the LLS and Leiden 85-plus study.

Figure S3.5 C1 values plotted against the C2 values, both resulting from the multidimensional scaling analysis, of the 403 LLS GWAS cases (green) and the 1,670 RS GWAS controls (blue).

Figure S3.6 Quantile–quantile plot of expected vs. observed chi-square values for the test statistic from the Cochran-Armitage trend test for 516,721 SNPs that passed the quality control thresholds in the LLS GWAS. The slope of the dashed line represents the genomic inflation factor ($\lambda = 1.027$). The shaded region represents the 95% confidence band.

Table S3.1 SNPs ($n = 62$) selected for replication analysis, associating at $P < 1 \times 10^{-4}$ with survival into old age in the analysis of the LLS GWAS.

Table S3.2A Results of the association analysis with survival into old age of the 58 prioritized SNPs from the LLS GWAS in the RS replication study, Leiden 85-plus replication study, Danish replication study, and the meta-analysis.

Table S3.2B Results of the meta-association analysis with survival into old age of the 58 prioritized SNPs in male cases compared to all controls.

Table S3.2C Results of the meta-association analysis with survival into old age of the 58 prioritized SNPs in female cases compared to all controls.

Table S3.3A Results of the prospective analysis of rs2075650 adjusted for rs429358 ($\epsilon 4$) and rs7412 ($\epsilon 2$).

Table S3.3B Results of the prospective analysis of rs429358 ($\epsilon 4$).

Table S3.4 Association of LLS GWAS SNPs selected from the AlzGene database (<http://www.alzgene.org/>) [36] with survival into old age.

Table S3.5 Association of LLS GWAS SNPs within a 10-kb window around *FOXO3A* and *AKT1* with survival into old age.

