Mechanisms of Ageing and Development 132 (2011) 324–330

Contents lists available at ScienceDirect

Mechanisms of Ageing and Development

journal homepage: www.elsevier.com/locate/mechagedev

A genome-wide association study confirms *APOE* as the major gene influencing survival in long-lived individuals

Almut Nebel [a,1], Rabea Kleindorp [a,1], Amke Caliebe [b,1], Michael Nothnagel [b], Hélène Blanché [c], Olaf Junge [b], Michael Wittig [a], David Ellinghaus [a], Friederike Flachsbart [a], Heinz-Erich Wichmann [d,e,f], Thomas Meitinger [d,g], Susanna Nikolaus [h], Andre Franke [a], Michael Krawczak [b,i], Mark Lathrop [c], Stefan Schreiber [a,h,i,*]

[a] *Institute of Clinical Molecular Biology, at the Christian Albrechts University and the University Hospital Schleswig-Holstein, Kiel, Germany*
[b] *Institute of Medical Informatics and Statistics, at the Christian Albrechts University and the University Hospital Schleswig-Holstein, Kiel, Germany*
[c] *Fondation Jean Dausset-Centre du Polymorphisme Humain (CEPH), Paris, France*
[d] *Institute of Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany*
[e] *Institute of Medical Informatics, Biometry and Epidemiology, Chair of Epidemiology, Ludwig Maximilians University, Munich, Germany*
[f] *Klinikum Grosshadern, Munich, Germany*
[g] *Institute of Human Genetics, Technical University Munich, Munich, Germany*
[h] *Clinic for Internal Medicine I, at the Christian Albrechts University and the University Hospital Schleswig-Holstein, Kiel, Germany*
[i] *Popgen Biobank, at the Christian Albrechts University and the University Hospital Schleswig-Holstein, Kiel, Germany*

ARTICLE INFO

Article history:
Received 3 November 2010
Received in revised form 14 June 2011
Accepted 23 June 2011
Available online 29 June 2011

Keywords:
GWAS
Longevity
APOE

ABSTRACT

We conducted a case–control genome-wide association study (GWAS) of human longevity, comparing 664,472 autosomal SNPs in 763 long-lived individuals (LLI; mean age: 99.7 years) and 1085 controls (mean age: 60.2 years) from Germany. Only one association, namely that of SNP rs4420638 near the *APOC1* gene, achieved genome-wide significance (allele-based $P = 1.8 \times 10^{-10}$). However, logistic regression analysis revealed that this association, which was replicated in an independent German sample, is fully explicable by linkage disequilibrium with the *APOE* allele ε4, the only variant hitherto established as a major genetic determinant of survival into old age. Our GWAS failed to identify any additional autosomal susceptibility genes. One explanation for this lack of success in our study would be that GWAS provide only limited statistical power for a polygenic phenotype with loci of small effect such as human longevity. A recent GWAS in Dutch LLI independently confirmed the *APOE*–longevity association, thus strengthening the conclusion that this locus is a very, if not the most, important genetic factor influencing longevity.

1. Introduction

Approximately 25–30% of the variation in adult lifespan is attributable to genetic factors that become more important with increasing age and exert their strongest effects in nonagenarians and centenarians (Gögele et al., 2010; Hjelmborg et al., 2006). As yet, however, only a few genetic variants have been found consistently to influence longevity. The first to be discovered was the ε4 allele of the apolipoprotein E (*APOE*) gene, a mortality factor that predisposes to both Alzheimer's and cardiovascular diseases (Corder et al., 1993; Panza et al., 2004). *APOE* ε4 is the only variant with a reportedly large adverse effect upon survival at advanced age (Schächter et al., 1994), and this association has been replicated in several populations (Christensen et al., 2006). Variation in the human forkhead box O3A gene (*FOXO3A*), in contrast, has been found to be associated with the ability to live long, an effect corroborated by studies in Japanese, German, Italian, US-American, Jewish, Chinese and Danish populations (Anselmi et al., 2009; Flachsbart et al., 2009; Li et al., 2009; Pawlikowska et al., 2009; Soerensen et al., 2010; Willcox et al., 2008). More recently, we have identified exonuclease 1 (*EXO1*) as a potential novel longevity gene (Nebel et al., 2009). All three genes were detected through candidate-gene approaches.

A few genome-wide association studies (GWAS) for the longevity phenotype have been reported so far. In 2007, a

* Corresponding author at: Institute of Clinical Molecular Biology, Christian-Albrechts-University, Schittenhelmstrasse 12, 24105 Kiel, Germany.
Tel.: +49 431 597 2350; fax: +49 431 597 2196.
E-mail address: s.schreiber@ikmb.uni-kiel.de (S. Schreiber).
[1] These authors contributed equally to this work.

0047-6374/$ – see front matter © 2011 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.mad.2011.06.008

genome-wide screening for longevity-related traits in a community-based population was published, but independent replication of its results is still pending (Lunetta et al., 2007). A meta-analysis of four prospective cohort studies of individuals reaching 90 years and older observed an association of SNP rs9664222 near the gene *MINPP1*, which did not reach genome-wide significance, but was replicated independently (Newman et al., 2010). A case–control GWAS, including a large sample of US-American centenarians, reported 33 replicated genome-wide significant SNPs (Sebastiani et al., in press), but the validity of the results is still critically discussed (Editorial Expression of Concern; http://www.sciencemag.org/content/330/6006/912.2.full). More recently, a GWAS in Dutch nonagenarians from the Leiden Longevity Study identified a SNP, which tagged the deleterious effects of the APOE ε4 allele (Deelen et al., in press).

Here, we report upon a GWAS using a case–control study design in which we compared more than 760 centenarians and nonagenarians, who represent an exceptional longevity phenotype, with more than 1000 younger controls. We replicated and confirmed *APOE* as the major genetic determinant of survival into old age. In addition, when we tested the 33 SNPs of Sebastiani et al. in our samples, only the *APOE*-related marker rs2075650 showed a significant association.

2. Materials and methods

2.1. Study participants

In stage 1 we conducted a GWAS where a total of 763 unrelated German long-lived individuals (LLI) were studied who were between 94 and 110 years of age (mean age: 99.7 years; Table 1). The 1085 German control individuals in stage 1 (age range: 45–77 years; mean age: 60.2 years) were obtained from the Popgen (Krawczak et al., 2006) and KORA (Wichmann et al., 2005) biobanks. The regional genetic differences between northern and southern Germany are not significant (Lao et al., 2008; Steffens et al., 2006) and therefore the two populations were pooled. The sex ratio was approximately 1:1. The 754 LLI used for validation in stage 2 had an age range of 95–108 years and a mean age of 96.9 years. The stage 2 control sample of 860 individuals (age range: 60–75 years; mean age: 67.3 years) was matched to the LLI as closely as possible in terms of ancestry, gender and geographical origin. The sex ratio in this sample was 12:1 in favor of females. The lack of men in stage 2 implies that any male-specific effects that were observed in the initial screening may not be detectable in the follow-up sample. The description for the stage 1 samples is given in detail elsewhere (Krawczak et al., 2006; Nebel et al., 2005; Wichmann et al., 2005). For stage 2 it is published in Nebel et al. (2005). There are no mortality data available for the controls. However, based on current predictions only 1.5% of all 60-year old and 1.8% of all 75-year old German females will become 100 years (to become 95 years old the respective proportions are 9.5% for 60-year old and 12.3% for 75-year old women). The probability for males is even lower (Human Mortality Database: www.demogr.mpg.de). Hence, we can estimate that only a negligible proportion of younger controls will become long-lived. The French sample used to investigate rs12741354 in *ASTN1* comprised 536 centenarians (mean age: 103.7 years) and 508 younger controls (mean age: 50.8 years). Cases and controls were matched for gender and geographic origin (Blanché et al., 2001). The sex ratio was 5:1 in favor of females.

DNA was isolated from blood samples of all participants using standard methods. All subjects gave written informed consent prior to participation. The protocols were approved by the respective institutional ethics review boards. The study workflow is given in Fig. 1.

2.2. Genotyping

The Affymetrix® Genome-Wide Human SNP Array 6.0 (San Francisco, CA, USA) was used to genotype the individuals in stage 1 (Supplementary methods). For validation, markers were genotyped in the individuals from stage 2 with TaqMan®SNP Genotyping Assays and the SNPlex® Genotyping System (Applied Biosystems, Foster City, USA) on an automated platform (Hampe et al., 2001).

2.3. Quality control and single-marker analysis

Data storage, quality control, and association analysis were performed using a customized database, R v2.7.1 (R Development Team, 2008) (http://www.R-project.org) and PLINK v1.04 (Purcell et al., 2007). Only autosomal SNPs with high-quality genotypes were included in subsequent analyses: SNPs were included if they had (a) a call rate $\geq$ 95% in cases and controls and (b) a minor allele frequency (MAF) $\geq$ 0.02 in controls or if |MAF(controls) − MAF(cases)| > 0.02 (see Fig. 2 for Manhattan-plot and Supplementary Fig. 1 for *Q*–*Q*-plot) and (c) *P* value of Hardy–Weinberg equilibrium (HWE) in controls >0.01. A total of 664,472 SNPs passed quality control. To reduce the rate of false-positive results, the genotyping quality of the 1600 most significantly associated SNPs (with *P* values of up to 9.6×10^{-6}) was investigated by visual inspection of the corresponding cluster plots (Supplementary Figs. 2 and 3). More detailed quality control for GWAS and follow-up is described in detail in Supplementary methods. Genotype–phenotype associations were assessed for statistical significance using a Cochran–Armitage trend (CAT) test and Fisher's exact test, with the latter applied to both an allele- (CCA) and a genotype-based (CCG) case–control comparison. Population stratification was low (genomic inflation (GC) factor (Devlin and Roeder, 1999) for stage 1 λ: 1.064 and stage 2 λ: 1.06). GC-adjusted *P* values were calculated according to Devlin and Roeder (1999). Logistic regression analysis of rs4420638 and *APOE* ε4 status as influence variables was conducted with and without interaction using R v2.7.1 (R Development Team, 2008). To this end, genotypes were encoded by the number of the major allele of rs4420638 or the ε4 allele (*APOE*), respectively. All tests were two-sided.

2.4. Haplotype analysis

Stage 1 data were also subjected to a haplotype-based association analysis, using a two-marker sliding window. To this end, maximum-likelihood estimates of haplotype frequencies were obtained and posterior probabilities of haplotypic genotypes calculated for each individual. The latter were included in a score test for association with a binary phenotype, assuming an additive risk model for haplotypes. SNPs involved in the 5000 most significantly associated haplotypes were subjected to visual inspection of the corresponding cluster plots (see Supplementary methods). Haplotypes for which both SNPs passed the inspection were genotyped and analyzed in stage 2. The analysis was carried out in R v2.7.1 (R Development Team, 2008), using the haplo.score function of R package haplo.stats v1.3.8 (Schaid et al., 2002), which takes phase uncertainty into account.

2.5. Testing for association of 33 SNPs from Sebastiani et al. in German samples

To test the association findings of the study by Sebastiani et al. which were obtained from Illumina 370 CNV chips and Illumina 610 arrays, the German Affymetrix GWAS data had to be imputed, since different genotyping technologies had been used. Genotype imputation was carried out using a hidden Markov model algorithm implemented in the software program BEAGLE v3.1.1 (Browning and Browning, 2009) and the HapMap 3 samples from the CEU, TSI, MEX and GIH populations to infer missing autosomal genotypes *in silico*. To take imputation uncertainty into account, association analysis between the phenotype and the dosage data (expected allele counts) was performed using PLINK's logistic regression framework for dosage data (with flag-dosage).

3. Results

3.1. Stage 1 analyses

We have ranked the SNPs according to their minimum (P_{min}) of P_{CAT}, P_{CCA}, and P_{CCG}. A total of 58,903 markers (8.9%) showed a nominal P_{min} value <0.05. To reduce the rate of false-positive results, cluster plots of the 1600 most significantly associated SNPs (all with P_{min} values $\leq 9.6 \times 10^{-6}$) were subjected to post hoc visual

Table 1
Description of LLI and controls genotyped in the two stages of the study.

Stage	Sample	No.	Mean age (years)	Median age (years)	Age range (years)
1	LLI	763	99.7	101	94–110
1	Controls	1085	60.2	60	45–77
2	LLI	754	96.9	97	95–108
2	Controls	860	67.3	66.5	60–75

LLI: long-lived individuals.

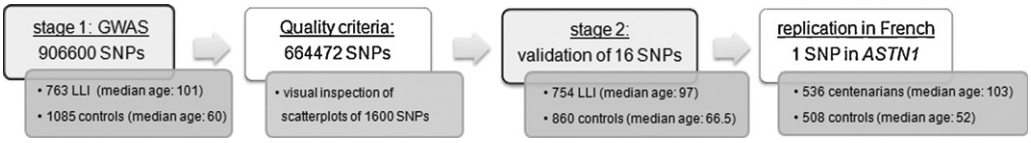

Fig. 1. Workflow diagram summarizing the different stages of the study.

inspection. Of these, only 16 SNPs met our stringent quality criteria (Supplementary methods) and were selected for follow-up (Table 2). Their nominal P_{CCA} values ranged from 1.8×10^{-10} to 9.5×10^{-6}. Due to multiple testing in GWAS usually low initial *P* values in stage 1 are required in order to be able to replicate the association. Our follow-up study has not sufficient power to detect SNP-effects in stage 2 with higher *P* values ($>10^{-5}$). Therefore, additional SNPs were not included in our follow-up. Only the association between longevity and SNP rs4420638 near the *APOC1* gene achieved genome-wide significance (Bonferroni-adjusted $P_{CCA} = 1.2 \times 10^{-4}$, based upon 664,472 tests and assuming no correlation between SNPs; see Fig. 2 and Supplementary Fig. 4). The significance of the other 15 SNPs did not withstand Bonferroni correction, thereby highlighting the need for independent validation.

3.2. Stage 2 analyses

The 16 SNPs selected for follow-up were next genotyped in an additional sample of 754 German LLI (mean age: 96.9 years) and 860 controls (mean age: 67.3 years; stage 2; Table 1), matched for ancestry, gender and geographical origin within Germany. This sample had a nominal power between 83% and 99% to detect the initially observed associations at the 5% significance level (with Bonferroni correction, the power was between 51% and 93%). However, only the association between longevity and rs4420638 was replicated ($P_{CCA} = 1.1 \times 10^{-8}$, Bonferroni adjusted $P_{CCA} = 1.3 \times 10^{-7}$, Table 3), yielding a P_{CCA} value $<2.2 \times 10^{-16}$ in the combined analysis of stage 1 and stage 2.

3.3. ASTN1

Another SNP, rs12741354 in the astrotactin 1 (*ASTN1*) gene, yielded a nominally significant P_{CCA} value of 0.007, but this association did not withstand Bonferroni correction (Bonferroni-adjusted $P_{CCA} = 0.091$ assuming 13 tests; see Supplementary methods). As SNP rs12741354 had the second lowest *P* value, it was also investigated in an independent collection of 536 French centenarians (mean age: 103.7 years) and 508 younger controls (mean age: 50.8 years) (power: 65.5%) where it failed to show any association with longevity ($P_{CCA} = 0.86$); Supplementary Table 1).

3.4. APOE

The array-based genotypes of SNP rs4420638 were confirmed by a TaqMan® SNP Genotyping Assay (99.4% concordance). SNP rs4420638 is located 14 kb downstream of the *APOE* locus (Corder et al., 1993) and has previously been shown to be in strong linkage disequilibrium (LD) with the latter (Coon et al., 2007). Defining *APOE* ε4 by intragenic SNP rs429358, we were able to confirm this LD in the German control sample of stage 2 ($r^2 = 0.72$, $D' = 0.97$). Logistic regression analysis including both rs4420638 genotype and *APOE* ε4 status showed that the longevity association of rs4420638 was explicable by its LD with *APOE* ($P = 0.57$ for the association between longevity and rs4420638, after adjustment for ε4 status). SNP rs4420638 was not further analyzed in the French as this is the sample in which the *APOE*–longevity association was originally detected (Schächter et al., 1994).

3.5. FOXO3A and EXO1

The recently confirmed longevity gene *FOXO3A* (Anselmi et al., 2009; Flachsbart et al., 2009; Li et al., 2009; Pawlikowska et al., 2009; Soerensen et al., 2010; Willcox et al., 2008) and the longevity candidate *EXO1* (Nebel et al., 2009) yielded comparatively high P_{CCA} values of 0.007 and 0.035, respectively, and were therefore far too large to qualify for follow-up in stage 2.

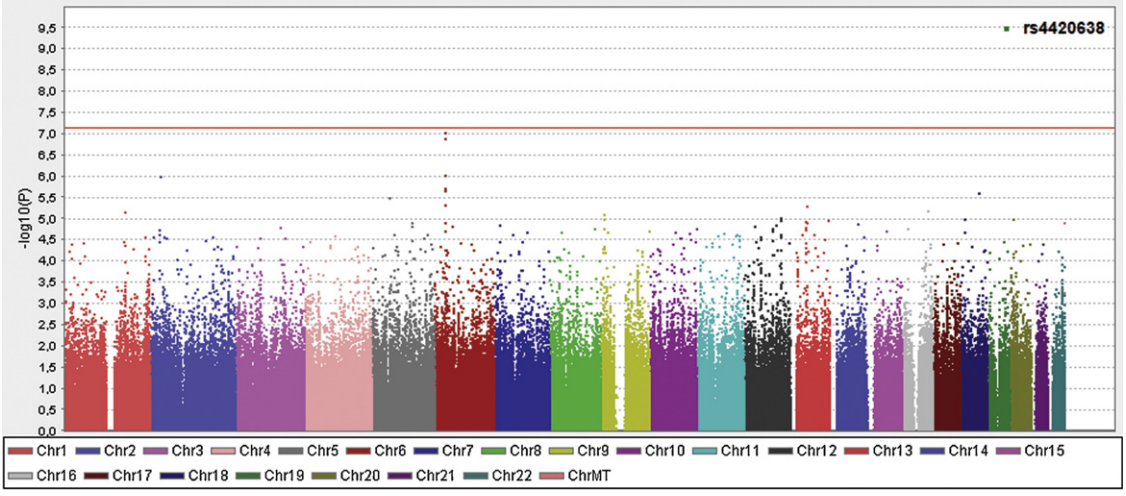

Fig. 2. "Manhattan plot". Calculated for all SNPs that passed quality criteria and showed good cluster plots (cluster plots were only inspected up to a P_{min} value $\leq 9.6 \times 10^{-6}$). Marker positions are in NCBI's build 36. Marker rs4420638 is the only SNP with genome-wide significance (red line). The large peak of significant SNPs on chromosome 6 could not be replicated in stage 2. The plot was created using Haploview 4.2 and allelic *P* values. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

Table 2
Summary statistics for the 16 SNPs most significantly associated with longevity in the GWAS (stage 1).

Chr.	Position B36	ID	MAF (LLI)	AF (controls)	P_{CAT}	P_{CCA}	P_{CCG}	OR	95% C.I.	Bonferroni adjusted P_{CCA}	GC *P* value	SNP call rate (LLI)	SNP call rate (controls)
1	175309154	rs12741354	0.451	0.527	7.4×10^{-6}	8.0×10^{-6}	4.3×10^{-5}	0.74	0.65–0.84	1	1.31×10^{-5}	0.99	1
2	28515768	rs2338013	0.194	0.264	7.4×10^{-7}	8.9×10^{-7}	3.4×10^{-6}	0.67	0.58–0.79	0.591	2.07×10^{-6}	0.98	1
5	52022995	rs350450	0.219	0.287	4.2×10^{-6}	3.2×10^{-6}	2.2×10^{-5}	0.69	0.59–0.80	1	6.34×10^{-6}	1	1
6	29731718	rs29228	0.176	0.238	4.7×10^{-6}	4.5×10^{-6}	2.5×10^{-5}	0.68	0.58–0.80	1	9.12×10^{-6}	0.99	0.99
6	29753592	rs3129063	0.175	0.240	2.2×10^{-6}	2.1×10^{-6}	1.1×10^{-5}	0.67	0.57–0.79	1	4.33×10^{-6}	0.99	1
6	29778631	rs3129046	0.200	0.276	1.2×10^{-7}	8.0×10^{-8}	6.4×10^{-7}	0.66	0.56–0.77	0.053	2.21×10^{-7}	1	1
6	29785931	rs1610742	0.200	0.276	1.6×10^{-7}	1.2×10^{-7}	8.7×10^{-7}	0.66	0.56–0.77	0.079	3.02×10^{-7}	1	1
6	29808162	rs1610601	0.200	0.270	1.1×10^{-6}	1.0×10^{-6}	6.0×10^{-6}	0.68	0.58–0.79	0.664	1.98×10^{-6}	1	1
6	29834025	rs1633063	0.202	0.270	2.3×10^{-6}	1.8×10^{-6}	1.1×10^{-5}	0.68	0.58–0.80	1	3.82×10^{-6}	0.99	1
9	10715294	rs11790055	0.375	0.305	7.1×10^{-6}	9.4×10^{-6}	2.1×10^{-5}	1.38	1.20–1.58	1	1.49×10^{-5}	1	1
9	10726580	rs10959258	0.374	0.304	9.1×10^{-6}	1.3×10^{-5}	3.1×10^{-5}	1.36	1.19–1.57	1	1.86×10^{-5}	1	1
13	46905819	rs9595687	0.036	0.069	8.0×10^{-6}	9.5×10^{-6}	1.1×10^{-5}	0.49	0.36–0.68	1	2.09×10^{-5}	0.99	0:99
13	47001519	rs1575892	0.032	0.066	4.5×10^{-6}	3.7×10^{-6}	1.1×10^{-6}	0.45	0.32–0.63	1	9.32×10^{-6}	1	1
16	73418057	rs16947526	0.056	0.097	6.3×10^{-6}	5.0×10^{-6}	1.8×10^{-5}	0.56	0.43–0.72	1	1.23×10^{-5}	1	1
18	54052953	rs158869	0.484	0.406	2.3×10^{-6}	3.0×10^{-6}	1.8×10^{-6}	1.37	1.20–1.56	1	4.98×10^{-6}	0.99	1
19	50114786	rs4420638	0.109	0.185	3.7×10^{-10}	1.8×10^{-10}	1.6×10^{-10}	0.53	0.44–0.65	1.2×10^{-4}	9.55×10^{-10}	1	1

Chr.: chromosome; LLI: long-lived individuals; MAF (LLI): minor allele frequency in LLI; AF (controls): frequency in controls of the minor allele in LLI; P_{CAT}: *P* value for a Cochran–Armitage trend test for association; P_{CCA}: *P* value for an allele-based case–control comparison, using Fisher's exact test; P_{CCG}: *P* value for a genotype-based case–control comparison using Fisher's exact test; OR: allelic odds ratio for attaining old age, calculated for the minor allele using the major allele in LLI as reference; 95% C.I.: confidence interval (95%) of the OR; Bonferroni adjustment was made for 664,472 tests; GC *P* value: P_{CCA} adjusted for genomic control (λ: 1.06).

Table 3
Stage 2 validation of the 16 most significantly associated SNPs from the GWAS (stage 1).

Chr.	Position B36	ID	Min AF (LLI)	AF (controls)	P_{CAT}	P_{CCA}	P_{CCG}	OR	95% C.I.	Bonferroni adjusted P_{CCA}	GC *P* value	SNP call rate (LLI)	SNP call rate (controls)
1	175309154	rs12741354	0.483	0.532	0.006	0.007	0.021	0.82	0.71–0.94	0.084	0.020	0.95	0.98
2	28515768	rs2338013	0.252	0.241	0.496	0.527	0.469	1.06	0.90–1.25	1	0.320	0.96	0.98
5	52022995	rs350450	0.264	0.26	0.775	0.802	0.948	1.02	0.87–1.21	1	0.711	0.97	0.96
6	29731718	rs29228	0.199	0.211	0.446	0.473	0.682	0.93	0.78–1.11	1	0.425	0.97	0.99
6	29753592	rs3129063	0.202	0.216	0.353	0.367	0.528	0.92	0.77–1.10	1	0.354	0.97	0.97
6	29778631	rs3129046	0.237	0.247	0.523	0.525	0.744	0.95	0.80–1.12	1	0.432	0.96	0.98
6	29785931	rs1610742	0.234	0.246	0.454	0.470	0.742	0.94	0.79–1.11	1	0.378	0.95	0.98
6	29808162	rs1610601	0.225	0.234	0.548	0.575	0.734	0.95	0.80–1.13	1	0.549	0.97	0.98
6	29834025	rs1633063	0.224	0.240	0.289	0.303	0.464	0.91	0.77–1.08	1	0.255	0.96	0.99
9	10715294	rs11790055	0.328	0.327	0.942	0.969	0.876	1.01	0.86–1.17	1	0.929	0.96	0.98
9	10726580	rs10959258	0.331	0.329	0.897	0.895	0.831	1.01	0.87–1.17	1	0.897	0.98	0.99
13	46905819	rs9595687	0.053	0.048	0.470	0.507	0.957	1.13	0.81–1.56	1	0.488	0.97	0.99
13	47001519	rs1575892	0.055	0.049	0.442	0.461	0.75	1.13	0.82–1.56	1	0.504	0.97	0.99
16	73418057	rs16947526	0.077	0.087	0.314	0.354	0.628	0.88	0.68–1.14	1	0.338	0.95	0.98
18	54052953	rs158869	0.450	0.425	0.169	0.175	0.308	1.11	0.96–1.28	1	0.226	0.95	0.98
19	50114786	rs4420638	0.104	0.174	1.9×10^{-08}	1.1×10^{-08}	8.8×10^{-08}	0.55	0.44–0.68	1.3×10^{-07}	3.53×10^{-08}	0.98	0.97

Bonferroni adjustment was made for 13 tests (see Supplementary methods). For details see legend to Table 2.

Table 4
Association with longevity of 13 two-SNP haplotypes in stages 1 (detection) and 2 (validation).

Chromosome	1st SNP	2nd SNP	P_{score} (stage 1)	P_{score} (stage 2)
1	rs11577254	rs11205832	4.5×10^{-9}	0.767
1	rs705536	rs7540568	2.8×10^{-8}	0.466
2	rs11123531	rs13030366	2.4×10^{-7}	0.209
2	rs4673981	rs4465766	4.4×10^{-7}	0.205
4	rs4613559	rs16878067	1.3×10^{-11}	0.411
4	rs6448493	rs4613559	7.1×10^{-8}	0.233
4	rs17600829	rs10025735	2.3×10^{-6}	0.922
6	rs16872733	rs6458227	$<1.00 \times 10^{-16}$	1.000
8	rs1389109	rs13254091	4.8×10^{-8}	0.185
10	rs11597067	rs576289	6.1×10^{-9}	0.621
10	rs11009149	rs12251479	1.3×10^{-7}	0.382
11	rs4469890	rs10833433	7.6×10^{-12}	0.280
14	rs17127261	rs747053	2.0×10^{-6}	0.910

3.6. Haplotype analysis

We also conducted a genome-wide haplotype-based association analysis, using a sliding window comprising pairs of consecutive SNPs. Out of the 677,291 two-marker haplotypes considered, 51,912 (7.7%) showed a nominally significant association with longevity in a score test assuming an additive risk model for haplotypes ($P_{score} < 0.05$). Visual inspection of the scatter plots for SNPs included in the 5000 most significantly associated haplotypes yielded only 19 marker pairs with high-quality genotypes for both SNPs. Of these, two haplotypes including the *APOE* locus and four redundant haplotypes from a region of strong LD on chromosome 6 were not considered for follow-up. Thus, 13 haplotype pairs were chosen for validation in the stage 2 samples. However, none of them showed a significant haplotype association with longevity, neither in the stage 2 sample as a whole (Table 4) nor in the subset of centenarians (data not shown).

3.7. Testing for association of 33 SNPs from Sebastiani et al. in German samples

A recent GWAS identified 70 genome-wide significant SNPs, 33 of which were subsequently replicated in an additional sample (Sebastiani et al., in press). Here, we tested these 33 SNPs for association with longevity in our large German collection. Except for SNP rs2075650 (P value = 7.5×10^{-6}, Table 5), which is in high LD with the known *APOE* alleles (Sebastiani et al., in press; Yu et al., 2007), none of the other markers showed a significant association in our study.

4. Discussion

Our study confirmed the known association between *APOE* genotype and survival at advanced age, but failed to identify any

Table 5
Association results of SNPs identified in the GWAS of Sebastiani et al. in the German GWAS data set. Markers not genotyped on the Affymetrix 6.0 Chip were imputed.

Chr.	Position B36	ID	MAF (controls)	AF (LLI)	P value	INFO score
1	243871287	rs10924270	0.461	0.467	0.738	0.795
1	246927264	**rs4916176**	0.125	0.122	0.820	–
2	5210976	rs1377638	0.132	0.123	0.380	0.979
2	171530712	rs4668356	0.054	0.065	0.185	0.991
2	173328711	rs6433379	0.123	0.131	0.527	0.914
2	234432806	**rs579327**	0.177	0.159	0.152	–
2	235521853	rs2042831	0.076	0.080	0.488	0.940
3	14874782	rs294636	0.183	0.189	0.566	0.963
3	97267562	rs4393926 (in LD with rs4390941,[a] $r^2 = 0.99$)	0.341	0.330	0.543	0.662
4	80183611	rs1455311	0.177	0.172	0.767	0.991
5	55450713	rs415407	0.467	0.471	0.834	0.926
5	65819465	rs10069397	0.484	0.448	0.055	0.993
5	83892667	rs7717527	0.182	0.196	0.443	0.967
6	31237686	**rs2073724**	0.086	0.085	0.861	–
6	33198814	rs4713607	0.495	0.519	0.197	1.000
8	58468572	rs954295	0.212	0.207	0.764	0.992
8	58940672	rs1436013	0.214	0.219	0.714	0.885
8	79864431	rs2717536	NA	NA	NA	–
8	110758554	rs3133926	0.215	0.202	0.444	0.973
8	135681127	rs1036819	0.112	0.106	0.727	0.969
10	67755359	rs16922827	0.093	0.109	0.139	0.942
10	99744963	rs508001	0.102	0.117	0.096	1.014
11	60282362	rs7930940	NA	NA	NA	–
11	63782795	rs2244621	0.149	0.120	0.014	0.933
11	125565159	rs522486 (in LD with rs1695739,[a] $r^2 = 0.83$)	0.220	0.233	0.462	1.057
15	36552135	**rs11073328**	0.109	0.100	0.372	–
17	9277095	**rs10521157**	0.223	0.227	0.758	–
17	44426482	**rs1390154 (in LD with rs9899404,[a] $r^2 = 0.99$)**	0.448	0.416	0.055	–
17	52793555	**rs792376**	0.178	0.176	0.859	–
19	50087459	rs2075650	0.143	0.109	7.5×10^{-6}	0.433
20	2602925	**rs1810636**	0.351	0.354	0.854	–
20	3775059	**rs4815617**	0.097	0.102	0.630	–
22	45173805	**rs9615362**	0.098	0.103	0.632	–

LLI: long-lived individuals; MAF (LLI): minor allele frequency in LLI; AF (controls): frequency in controls of the minor allele in LLI; P value: P value from logistic regression dosage test, P values are adjusted for uncertainty in imputation; INFO score: r^2 quality metric, quality measure for the imputed SNPs, the info metric is calculated based on the entire file, based on the ratio of empirical and expected variance in dosage. Values closer to 1 indicate better expected quality of imputation, http://pngu.mgh.harvard.edu/~purcell/plink/dosage.shtml; LD: linkage disequilibrium; NA: no information about this SNP was available. Marker was not in the imputed data set nor could a SNP in LD ($r^2 > 0.8$) be found. SNPs in bold were genotyped on Affymetrix 6.0, the P value is for an allele-based case–control comparison, using Fisher's exact test; data from all other SNPs were derived from imputation.

[a] SNP associated in the study of Sebastiani et al. (in press).

additional susceptibility loci for longevity. This result is in marked contrast to other complex phenotypes, for which GWAS have successfully uncovered multiple genetic risk factors at unprecedented speed and with high levels of reproducibility (Esparza-Gordillo et al., 2009; Franke et al., 2008; Hofmann et al., 2008). One explanation for this lack of success in our study would be that GWAS provide only limited statistical power for a polygenic phenotype with loci of small effect such as human longevity. GWAS are more suitable for the detection of genetic variants with large effects, which presumably explain only a small fraction of the total variation (Barrett et al., 2008). It is assumed that many causal gene variants, each with a weak to moderate contribution, influence our ability to become long-lived (Christensen et al., 2006). A recent study supports this notion and concludes that the alleles with small individual effects may jointly influence life span so that the resulting impact can be both substantial and significant (Yashin et al., 2010). Another case in point is the fact that the recently detected longevity gene *FOXO3A* (Anselmi et al., 2009; Flachsbart et al., 2009; Li et al., 2009; Pawlikowska et al., 2009; Soerensen et al., 2010; Willcox et al., 2008) and the candidate *EXO1* (Nebel et al., 2009) yielded P_{CCA} values in our GWAS of 0.007 and 0.035, respectively. These *P* values were far too large to qualify the corresponding SNPs for follow-up in stage 2, suggesting that most genetic factors of similarly small effect (i.e. OR $\leq$ 1.4) were unfortunately likely to be missed by our study. Furthermore, especially male-specific effects could not be detected in our sample due to the small number of men in the replication study. The German and French centenarian/LLI samples used here are among the largest to-date and yet, they may not have been sufficiently powerful to detect small-effect variants in a longevity GWAS. The German sample used in stage 1, for example, had a power of 3.2% to detect an association of an OR of 1.4 (minor allele frequency of 50% and applying a significance level of 7.5×10^{-8} (Supplementary Fig. 5)).

A previous meta-analysis of four prospective cohort studies revealed a statistically significant and replicated association of rs9664222 with longevity (defined as reaching 90 years and older) (Newman et al., 2010). This SNP did not show any association in our large case–control samples (P_{CCA} = 0.611). This result could be due to the different designs and/or levels of statistical power of the two studies. Moreover, we also tested the 33 polymorphisms from Sebastiani et al. (in press) in our imputed GWAS data set, of which SNP rs2075650 that is in LD with the known *APOE* alleles (Sebastiani et al., in press) showed a significant association in our German sample. Also the Dutch GWAS identified an *APOE*-related SNP as the most relevant marker with a deleterious effect on survival into old age. No other major longevity locus was found (Deelen et al., in press). Thus, our and the independent Dutch study (Deelen et al., in press) both strengthen the conclusion that *APOE* is a very important genetic factor and it may very well be the most relevant one influencing human longevity.

In summary, the GWAS, as it was carried out in our study, revealed apart from *APOE* no novel candidate genes for longevity in humans. Larger samples, a more specific and genetically enriched longevity phenotype (so-called 'super-centenarians') or a meta-analysis involving several thousands of participants may be required to successfully and reliably identify additional genes through a GWAS. In addition, some genetic factors involved in longevity could not be detected in the present study on principal ground, including low-frequency polymorphisms, copy-number variations or epigenetic changes. Furthermore, it is noteworthy that although individual-specific mechanisms may be important for longevity, our study only considers probabilities on a population level. In the future, a genome-wide pathway-based approach (Chen et al., 2009; Luo et al., 2010) could be a useful means to explore the joint effects of different gene variants in longevity-related pathways like, for example, IGF1-insulin-signaling or repair.

Acknowledgements

We wish to acknowledge the excellent technical support by all laboratory staff at the Institute of Clinical Molecular Biology, the Popgen biobank and the Biological Resource Center of the Fondation Jean Dausset-CEPH. This study was supported by the German Federal Ministry of Education and Research (BMBF) through the German National Genome Research Network (NGFN-2 and NGFN-plus), the DFG Excellence Cluster 'Inflammation at Interfaces', the Helmholtz Center Munich – German Research Center for Environmental Health, the Munich Center of Health Sciences (MC Health) as part of LMUinnovativ and the French Ministère de l'Enseignement Supérieur et de la Recherche.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.mad.2011.06.008.

References

Anselmi, C.V., Malovini, A., Roncarati, R., Novelli, V., Villa, F., Condorelli, G., Bellazzi, R., Puca, A.A., 2009. Association of the FOXO3A locus with extreme longevity in a Southern Italian Centenarian Study. Rejuvenation Res. 12, 95–104.

Barrett, J.C., Hansoul, S., Nicolae, D.L., Cho, J.H., Duerr, R.H., Rioux, J.D., Brant, S.R., Silverberg, M.S., Taylor, K.D., Barmada, M.M., Bitton, A., Dassopoulos, T., Datta, L.W., Green, T., Griffiths, A.M., Kistner, E.O., Murtha, M.T., Regueiro, M.D., Rotter, J.I., Schumm, L.P., Steinhart, A.H., Targan, S.R., Xavier, R.J., Libioulle, C., Sandor, C., Lathrop, M., Belaiche, J., Dewit, O., Gut, I., Heath, S., Laukens, D., Mni, M., Rutgeerts, P., Van Gossum, A., Zelenika, D., Franchimont, D., Hugot, J.P., de Vos, M., Vermeire, S., Louis, E., Cardon, L.R., Anderson, C.A., Drummond, H., Nimmo, E., Ahmad, T., Prescott, N.J., Onnie, C.M., Fisher, S.A., Marchini, J., Ghori, J., Bumpstead, S., Gwilliam, R., Tremelling, M., Deloukas, P., Mansfield, J., Jewell, D., Satsangi, J., Mathew, C.G., Parkes, M., Georges, M., Daly, M.J., 2008. Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. Nat. Genet. 40, 955–962.

Blanché, H., Cabanne, L., Sahbatou, M., Thomas, G., 2001. A study of French centenarians: are ACE and APOE associated with longevity? CR Acad. Sci. III 324, 129–135.

Browning, B.L., Browning, S.R., 2009. A unified approach to genotype imputation and haplotype-phase inference for large data sets of trios and unrelated individuals. Am. J. Hum. Genet. 84, 210–223.

Chen, Y., Xiong, D.H., Guo, Y.F., Pan, F., Zhou, Q., Zhang, F., Deng, H.W., 2009. Pathway-based genome-wide association analysis identified the importance of EphrinA-EphR pathway for femoral neck bone geometry. Bone 46, 129–136.

Christensen, K., Johnson, T.E., Vaupel, J.W., 2006. The quest for genetic determinants of human longevity: challenges and insights. Nat. Rev. Genet. 7, 436–448.

Coon, K.D., Myers, A.J., Craig, D.W., Webster, J.A., Pearson, J.V., Lince, D.H., Zismann, V.L., Beach, T.G., Leung, D., Bryden, L., Halperin, R.F., Marlowe, L., Kaleem, M., Walker, D.G., Ravid, R., Heward, C.B., Rogers, J., Papassotiropoulos, A., Reiman, E.M., Hardy, J., Stephan, D.A., 2007. A high-density whole-genome association study reveals that APOE is the major susceptibility gene for sporadic late-onset Alzheimer's disease. J. Clin. Psychiatry 68, 613–618.

Corder, E.H., Saunders, A.M., Strittmatter, W.J., Schmechel, D.E., Gaskell, P.C., Small, G.W., Roses, A.D., Haines, J.L., Pericak-Vance, M.A., 1993. Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. Science 261, 921–923.

Deelen, J., Beekman, M., Uh, H.W., Helmer, Q., Kuningas, M., Christiansen, L., Kremer, D., van der Breggen, R., Suchiman, H.E., Lakenberg, N., van den Akker, E.B., Passtoors, W.M., Tiemeier, H., van Heemst, D., de Craen, A.J., Rivadeneira, F., de Geus, E.J., Perola, M., van der Ouderaa, F.J., Gunn, D.A., Boomsma, D.I., Uitterlinden, A.G., Christensen, K., van Duijn, C.M., Heijmans, B.T., Houwing-Duistermaat, J.J., Westendorp, R.G., Slagboom, P.E. Genome-wide association study identifies a single major locus contributing to survival into old age; the APOE locus revisited. Aging Cell, in press.

Devlin, B., Roeder, K., 1999. Genomic control for association studies. Biometrics 55, 997–1004.

Esparza-Gordillo, J., Weidinger, S., Folster-Holst, R., Bauerfeind, A., Ruschendorf, F., Patone, G., Rohde, K., Marenholz, I., Schulz, F., Kerscher, T., Hubner, N., Wahn, U., Schreiber, S., Franke, A., Vogler, R., Heath, S., Baurecht, H., Novak, N., Rodriguez, E., Illig, T., Lee-Kirsch, M.A., Ciechanowicz, A., Kurek, M., Piskackova, T., Macek, M., Lee, Y.A., Ruether, A., 2009. A common variant on chromosome 11q13 is associated with atopic dermatitis. Nat. Genet. 41, 596–601.

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. U. S. A. 106, 2700–2705.
Franke, A., Fischer, A., Nothnagel, M., Becker, C., Grabe, N., Till, A., Lu, T., Muller-Quernheim, J., Wittig, M., Hermann, A., Balschun, T., Hofmann, S., Niemiec, R., Schulz, S., Hampe, J., Nikolaus, S., Nurnberg, P., Krawczak, M., Schurmann, M., Rosenstiel, P., Nebel, A., Schreiber, S., 2008. Genome-wide association analysis in sarcoidosis and Crohn's disease unravels a common susceptibility locus on 10p12.2. Gastroenterology 135, 1207–1215.
Gögele, M., Pattaro, C., Fuchsberger, C., Minelli, C., Pramstaller, P.P., Wjst, M., 2010. Heritability analysis of life span in a semi-isolated population followed across four centuries reveals the presence of pleiotropy between life span and reproduction. J. Gerontol. A Biol. Sci. Med. Sci. 66, 26–37.
Hampe, J., Wollstein, A., Lu, T., Frevel, H.J., Will, M., Manaster, C., Schreiber, S., 2001. An integrated system for high throughput TaqMan based SNP genotyping. Bioinformatics 17, 654–655.
Hjelmborg, J.v.B., Iachine, I., Skytthe, A., Vaupel, J.W., McGue, M., Koskenvuo, M., Kaprio, J., Pedersen, N.L., Christensen, K., 2006. Genetic influence on human lifespan and longevity. Hum. Genet. 119, 312–321.
Hofmann, S., Franke, A., Fischer, A., Jacobs, G., Nothnagel, M., Gaede, K.I., Schurmann, M., Muller-Quernheim, J., Krawczak, M., Rosenstiel, P., Schreiber, S., 2008. Genome-wide association study identifies ANXA11 as a new susceptibility locus for sarcoidosis. Nat. Genet. 40, 1103–1106.
Krawczak, M., Nikolaus, S., von Eberstein, H., Croucher, P.J., El Mokhtari, N.E., Schreiber, S., 2006. PopGen: population-based recruitment of patients and controls for the analysis of complex genotype–phenotype relationships. Community Genet. 9, 55–61.
Lao, O., Lu, T.T., Nothnagel, M., Junge, O., Freitag-Wolf, S., Caliebe, A., Balascakova, M., Bertranpetit, J., Bindoff, L.A., Comas, D., Holmlund, G., Kouvatsi, A., Macek, M., Mollet, I., Parson, W., Palo, J., Ploski, R., Sajantila, A., Tagliabracci, A., Gether, U., Werge, T., Rivadeneira, F., Hofman, A., Uitterlinden, A.G., Gieger, C., Wichmann, H.E., Ruther, A., Schreiber, S., Becker, C., Nurnberg, P., Nelson, M.R., Krawczak, M., Kayser, M., 2008. Correlation between genetic and geographic structure in Europe. Curr. Biol. 18, 1241–1248.
Li, Y., Wang, W.J., Cao, H., Lu, J., Wu, C., Hu, F.Y., Guo, J., Zhao, L., Yang, F., Zhang, Y.X., Li, W., Zheng, G.Y., Cui, H., Chen, X., Zhu, Z., He, H., Dong, B., Mo, X., Zeng, Y., Tian, X.L., 2009. Genetic association of FOXO1A and FOXO3A with longevity trait in Han Chinese populations. Hum. Mol. Genet. 18, 4897–4904.
Lunetta, K.L., D'Agostino Sr., R.B., Karasik, D., Benjamin, E.J., Guo, C.Y., Govindaraju, R., Kiel, D.P., Kelly-Hayes, M., Massaro, J.M., Pencina, M.J., Seshadri, S., Murabito, J.M., 2007. Genetic correlates of longevity and selected age-related phenotypes: a genome-wide association study in the Framingham Study. BMC Med. Genet. 8 (Suppl. 1), S13.
Luo, L., Peng, G., Zhu, Y., Dong, H., Amos, C.I., Xiong, M., 2010. Genome-wide gene and pathway analysis. Eur. J. Hum. Genet. 18, 1045–1053.
Nebel, A., Croucher, P.J., Stiegeler, R., Nikolaus, S., Krawczak, M., Schreiber, S., 2005. No association between microsomal triglyceride transfer protein (MTP) haplotype and longevity in humans. Proc. Natl. Acad. Sci. U. S. A. 102, 7906–7909.
Nebel, A., Flachsbart, F., Till, A., Caliebe, A., Blanche, H., Arlt, A., Hasler, R., Jacobs, G., Kleindorp, R., Franke, A., Shen, B., Nikolaus, S., Krawczak, M., Rosenstiel, P., Schreiber, S., 2009. A functional EXO1 promoter variant is associated with prolonged life expectancy in centenarians. Mech. Ageing Dev. 130, 691–699.
Newman, A.B., Walter, S., Lunetta, K.L., Garcia, M.E., Slagboom, P.E., Christensen, K., Arnold, A.M., Aspelund, T., Aulchenko, Y.S., Benjamin, E.J., Christiansen, L., D'Agostino Sr., R.B., Fitzpatrick, A.L., Franceschini, N., Glazer, N.L., Gudnason, V., Hofman, A., Kaplan, R., Karasik, D., Kelly-Hayes, M., Kiel, D.P., Launer, L.J., Marciante, K.D., Massaro, J.M., Miljkovic, I., Nalls, M.A., Hernandez, D., Psaty, B.M., Rivadeneira, F., Rotter, J., Seshadri, S., Smith, A.V., Taylor, K.D., Tiemeier, H., Uh, H.W., Uitterlinden, A.G., Vaupel, J.W., Walston, J., Westendorp, R.G., Harris, T.B., Lumley, T., van Duijn, C.M., Murabito, J.M., 2010. A meta-analysis of four genome-wide association studies of survival to age 90 years or older: the Cohorts for Heart and Aging Research in Genomic Epidemiology Consortium. J. Gerontol. A Biol. Sci. Med. Sci. 65, 478–487.
Panza, F., D'Introno, A., Colacicco, A.M., Capurso, C., Capurso, S., Kehoe, P.G., Capurso, A., Solfrizzi, V., 2004. Vascular genetic factors and human longevity. Mech. Ageing Dev. 125, 169–178.
Pawlikowska, L., Hu, D., Huntsman, S., Sung, A., Chu, C., Chen, J., Joyner, A.H., Schork, N.J., Hsueh, W.C., Reiner, A.P., Psaty, B.M., Atzmon, G., Barzilai, N., Cummings, S.R., Browner, W.S., Kwok, P.Y., Ziv, E., 2009. Association of common genetic variation in the insulin/IGF1 signaling pathway with human longevity. Aging Cell 8, 460–472.
Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., de Bakker, P.I., Daly, M.J., Sham, P.C., 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. 81, 559–575.
R Development Core Team, 2008. A language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
Schächter, F., Faure-Delanef, L., Guenot, F., Rouger, H., Froguel, P., Lesueur-Ginot, L., Cohen, D., 1994. Genetic associations with human longevity at the APOE and ACE loci. Nat. Genet. 6, 29–32.
Schaid, D.J., Rowland, C.M., Tines, D.E., Jacobson, R.M., Poland, G.A., 2002. Score tests for association between traits and haplotypes when linkage phase is ambiguous. Am. J. Hum. Genet. 70, 425–434.
Sebastiani, P., Solovieff, N., Puca, A., Hartley, S.W., Melista, E., Andersen, S., Dworkis, D.A., Wilk, J.B., Myers, R.H., Steinberg, M.H., Montano, M., Baldwin, C.T., Perls, T.T. Genetic signatures of exceptional longevity in humans. Science, in press (Editorial Expression of Concern; http://www.sciencemag.org/content/330/6006/912.2.full).
Soerensen, M., Dato, S., Christensen, K., McGue, M., Stevnsner, T., Bohr, V.A., 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.
Steffens, M., Lamina, C., Illig, T., Bettecken, T., Vogler, R., Entz, P., Suk, E.K., Toliat, M.R., Klopp, N., Caliebe, A., Konig, I.R., Kohler, K., Ludemann, J., Diaz Lacava, A., Fimmers, R., Lichtner, P., Ziegler, A., Wolf, A., Krawczak, M., Nurnberg, P., Hampe, J., Schreiber, S., Meitinger, T., Wichmann, H.E., Roeder, K., Wienker, T.F., Baur, M.P., 2006. SNP-based analysis of genetic substructure in the German population. Hum. Hered. 62, 20–29.
Wichmann, H.E., Gieger, C., Illig, T., 2005. KORA-gen-resource for population genetics, controls and a broad spectrum of disease phenotypes. Gesundheitswesen 67 (Suppl. 1), S26–S30.
Willcox, B.J., Donlon, T.A., He, Q., Chen, R., Grove, J.S., Yano, K., Masaki, K.H., Willcox, D.C., Rodriguez, B., Curb, J.D., 2008. FOXO3A genotype is strongly associated with human longevity. Proc. Natl. Acad. Sci. U. S. A. 105, 13987–13992.
Yashin, A.I., Wu, D., Arbeev, K.G., Ukraintseva, S.V., 2010. Joint influence of small-effect genetic variants on human longevity. Aging 2, 612–620.
Yu, C.E., Seltman, H., Peskind, E.R., Galloway, N., Zhou, P.X., Rosenthal, E., Wijsman, E.M., Tsuang, D.W., Devlin, B., Schellenberg, G.D., 2007. Comprehensive analysis of APOE and selected proximate markers for late-onset Alzheimer's disease: patterns of linkage disequilibrium and disease/marker association. Genomics 89, 655–665.