AGE (2013) 35:1467–1477
DOI 10.1007/s11357-012-9447-1

Linkage and association of successful aging to the 6q25 region in large Amish kindreds

Digna R. Velez Edwards · John R. Gilbert · James E. Hicks · Jamie L. Myers · Lan Jiang · Anna C. Cummings · Shengru Guo · Paul J. Gallins · Ioanna Konidari · Laura Caywood · Lori Reinhart-Mercer · Denise Fuzzell · Claire Knebusch · Renee Laux · Charles E. Jackson · Margaret A. Pericak-Vance · Jonathan L. Haines · William K. Scott

Received: 9 December 2011 / Accepted: 11 June 2012 / Published online: 7 July 2012

Abstract Successful aging (SA) is a multidimensional phenotype involving living to older age with high physical function, preserved cognition, and continued social engagement. Several domains underlying SA are heritable, and identifying health-promoting polymorphisms and their interactions with the environment could provide important information regarding the health of older adults. In the present study, we examined 263 cognitively intact Amish individuals age 80 and older (74 SA and 189 "normally aged") all of whom are part of a single 13-generation pedigree. A genome-wide association study of 630,309 autosomal single nucleotide polymorphisms (SNPs) was performed and analyzed for linkage using multipoint analyses and for association using the modified quasi-likelihood score test. There was evidence for linkage on 6q25-27 near the fragile site *FRA6E* region with a dominant model maximum multipoint heterogeneity LOD score=3.2. The 1-LOD-down support interval for this linkage contained one SNP for which

Electronic supplementary material The online version of this article (doi:10.1007/s11357-012-9447-1) contains supplementary material, which is available to authorized users.

D. R. V. Edwards · J. R. Gilbert · J. E. Hicks · J. L. Myers · S. Guo · P. J. Gallins · I. Konidari · L. Caywood · L. Reinhart-Mercer · M. A. Pericak-Vance · W. K. Scott (✉)
Dr. John T. Macdonald Foundation Department of Human Genetics and John P. Hussman Institute of Human Genomics, Miller School of Medicine, University of Miami, 1501 NW 10th Avenue, Room 414, Miami, FL 33136, USA
e-mail: bscott@med.miami.edu

L. Jiang · A. C. Cummings · D. Fuzzell · C. Knebusch · R. Laux · J. L. Haines
Center for Human Genetics Research, Vanderbilt University, Nashville, TN, USA

C. E. Jackson
Carmel, IN, USA

Present Address:
D. R. V. Edwards
Center for Human Genetics Research, Vanderbilt University Epidemiology Center, Institute of Medicine and Public Health, Department of Obstetrics and Gynecology, Vanderbilt University, Nashville, TN, USA

there was regionally significant evidence of association (rs205990, $p=2.36\times10^{-5}$). This marker survived interval-wide Bonferroni correction for multiple testing and was located between the genes *QKI* and *PDE10A*. Other areas of chromosome 6q25-q27 (including the *FRA6E* region) contained several SNPs associated with SA (minimum $p=2.89\times10^{-6}$). These findings suggest potentially novel genes in the 6q25-q27 region linked and associated with SA in the Amish; however, these findings should be verified in an independent replication cohort.

Keywords Genome-wide association · Longevity · Genetic epidemiology · Family-based study

Introduction

With the rapid expansion of the population over age 60, understanding the components of successful aging (SA) is becoming a critical component of the medical and public health response to this demographic shift. The number of adults over age 60 is predicted to reach two billion worldwide by the year 2050, more than doubling from 2000 (US Census Bureau Population Division and Housing and Household Economic Statistics Division 2001). This growing elderly population could dramatically impact our health care system depending on how successfully this population ages with regard to multiple factors such as physical, mental, and social well-being. Surprisingly, a large proportion of older adults has continued to function at a high level and has "aged successfully" (Depp and Jeste 2006). Understanding why some individuals avoid significant decline in physical function and susceptibility to age-related diseases might help guide clinical care and provide the basis for novel programs to prevent or slow such declines.

Although SA is a complex and multifactorial trait, its sub-domains (chronological age, physical function, cognitive health, and social engagement and well-being) have strong genetic components. Centenarians cluster in families and siblings of centenarians are more likely to live past 85 years than others in their birth cohorts (Perls et al. 2002). Finally, data suggest that several domains underlying SA are heritable including: longevity, grip strength, lower extremity function, and retention of cognitive ability (Carmelli et al. 2000; Farrer et al. 1984; Swedlund et al. 1983). Taken together, these data suggest that SA has a strong genetic component.

Previous genetic studies of SA have focused on three primary categories of genes selected primarily for their relevance to aging-associated pathophysiological processes or processes that are known to change during the aging process (Glatt et al. 2007). These categories include: (1) genes involved in the maintenance of cholesterol, lipid, and/or lipoproteins (Arking et al. 2005; Cellini et al. 2005; Taioli et al. 2001); (2) genes that influence inflammation and immune response, particularly cytokines, many of which also influence cell cycling, growing, motility, and signaling (Capurso et al. 2004; Naumova et al. 2004); and (3) genes involved in drug metabolism (Taioli et al. 2001). To date, few of the candidate genes associated with SA have been successfully replicated, with the exception of a few studies that have associated with longevity but not SA (Atzmon et al. 2009a; Banasik et al. 2011; Chung et al. 2010; Deelen et al. 2011; Kleindorp et al. 2011; Li et al. 2009; Nebel et al. 2011; Snejdrlova et al. 2011), suggesting that this approach may not be the best for a complex trait such as SA.

Genome-wide association studies (GWAS) avoid choosing specific candidate genes in favor of broadly surveying common variation across the entire genome. Two GWAS for extreme longevity (living past age 90 (Newman et al. 2010) and living past age 100) have recently been published, and three other GWAS, one for healthy cognitive aging past age 85 (Poduslo et al. 2009), one for disease-free survival after age 65 (Lunetta et al. 2007), and one for survival of 89 years for men and 91 years for women (Deelen et al. 2011) have also been published. There is significant variability across studies with regard to how longevity is defined as surviving beyond a certain age and others are based on being healthy and disease free up to a certain age. This inconsistency in the definition has made it difficult to compare findings across studies. To our knowledge, a GWAS has not been performed for a dichotomous successful aging outcome defined as maintaining function in cognition, physical function, and social engagement at age 80 or older. In this report, we present a genome-wide linkage and association study examining 74 successfully aged and 189 "normally aged" Amish adults over age 80 living in several communities in Indiana and Ohio. Genetic

studies of population isolates such as the Amish have multiple advantages over other study designs because they represent well-ascertained multigenerational pedigrees obtained from relatively few founders and spanning few generations, resulting in both environmental and phenotypic homogeneity (Arcos-Burgos and Muenke 2002). The results of this study suggest that several regions of interest in the genome might harbor SA loci, with the strongest evidence implicating the *FRA6E* region of chromosome 6.

Materials and methods

Study population

This study is a part of the larger ongoing Collaborative Aging and Memory Project (CAMP). CAMP is a multi-institutional prospective population-based study of aging and dementia in the Amish communities of Adams, Elkhart, and LaGrange counties in Indiana and Holmes County in Ohio, conducted from 2002 until the present. These communities were formed in several waves of migration in the eighteenth and nineteenth centuries (van der Walt et al. 2005). The Amish originally emigrated from Europe to Pennsylvania in the 1700s, and a further westward expansion of the Amish population occurred in the early 1800s when a subset of the Pennsylvania population migrated to Ohio and Indiana. A second wave of migration from Switzerland arrived in the nineteenth century, eventually settling in Adams County, Indiana, in the 1850s. Present-day Amish in Holmes County, Ohio and Elkhart and LaGrange counties, Indiana are largely descended from the first wave of westward immigration from Pennsylvania, while the Adams County, Indiana settlements are largely descended from the second wave of immigration that passed through Pennsylvania (van der Walt et al. 2005). Both sets of communities, therefore, have a degree of shared history and ancestry with the Pennsylvania Amish communities. Written and informed consent was obtained for all participants and their legal guardians.

For the SA arm of the CAMP study, individuals over age 80 were identified through public directories published by individual Amish communities and referral from individuals already enrolled in the study. Once individuals were identified, a door-to-door interview was performed for a baseline examination. The only exclusion criterion for the SA arm of the study was cognitive impairment (individuals screening cognitively impaired were referred to the dementia arm of the study). The recruitment and ascertainment methods for CAMP have been previously described (Pericak-Vance et al. 1996; Velez Edwards et al. 2011).

Using an "all common paths" query of the Anabaptist Genealogy Database, all 263 individuals were placed in a 13-generation, 4,998 person pedigree (Agarwala et al. 2003). There is no evidence of a more recent founder effect among the SA individuals. The first individuals in this pedigree born in Indiana or Ohio were born in the 1820s, four to five generations before the oldest sampled individuals in this study.

Definition of successful aging Successful aging was defined according to tests and measurements taken at the time of baseline enrollment in the study. SA was defined as previously described (Velez Edwards et al. 2011), considering functioning in all three domains described by Rowe and Kahn (1997). The specific criteria we used are outlined in Table 1. The first requirement was survival to age 80. All individuals had to be cognitively intact (education-adjusted modified mini-mental state examination (3MS) >86). If someone had an education-adjusted 3MS <87 but was determined not to be cognitively impaired after further neuropsychological testing and evaluation at a consensus conference (Hahs et al. 2006), they would be classified as SA if they met the other criteria. SA individuals did not have significant depressive

Table 1 Criteria to identify "successfully aged" individuals

Criteria	Test/cutoff
Age	>80
Cognitively intact	3MS >86 (education-adjusted)
No significant depressive symptoms	GDS <6
"High function" self-reported measures of physical function	0 or 1 ADL and IADL scales
	0 or 1 Nagi score
	3 or 4 Rosow–Breslau score
Lower extremity function	>8 (12 pt scale) EPESE
Satisfied or very satisfied with life	Questionnaire

symptoms (geriatric depression scale (GDS) score <6). Next, we considered whether individuals met standard cutoffs for "high function" on the self-reported measures of physical function: total scores of 0 or 1 on the activities of daily living (ADL) and instrumental (IADL) scales, indicating no assistance or partial assistance on only a single item needed; Nagi score of 0 or 1 (Nagi 1976), indicating no difficulty or difficulty on only one item; Rosow–Breslau score of 3 or 4, indicating limitation on zero items or one item. Lower extremity function was considered by limiting SA to individuals scoring in the top 1/3 of the sample on the Established Populations for Epidemiologic Studies of the Elderly (EPESE) short physical performance battery summary score (>8 on a 12 point scale). Finally, individuals considered SA self-reported to be satisfied or very satisfied with life.

There were a total of 263 Amish individuals included in this study after performing sample quality control (QC). Among these 263 individuals, 74 were classified as SA and 189 were classified as controls. This study has been reviewed and approved by the Institutional Review Boards at the University of Miami and Vanderbilt University Medical Center.

Collection of clinical and demographic data The baseline examination included collection of demographic characteristics and a screening of cognitive and physical function. Demographic information and family relationships (parents and grandparents) were recorded in standard format to facilitate pedigree construction.

Participants underwent cognitive screening using a revision of the 3MS examination (Teng and Chui 1987) developed for use in the Cache County Memory Study. We established the 3MS cutoff point of <87 (after education adjustment), indicating potential cognitive impairment after examining the receiver operating curves in predicting subsequent dementia and prodromal Alzheimer's disease in the Cache County Memory Study. This cut point was selected as having maximal utility in detection of dementia (Khachaturian et al. 2000), optimizing the sensitivity of case detection (98.4 %). Depression was screened using the 15-item GDS (Yesavage 1988), and life satisfaction was probed using a single question: "Overall how satisfying is your life: very satisfying, satisfying, not so satisfying?"

Self-rated physical function was assessed by several well-established instruments. Basic self-care was assessed using a modified Katz ADL scale consisting of eight items: bathing, dressing, grooming, toileting, continence, transferring, walking, and eating (Katz 1983). Difficulty with musculoskeletal function was assessed by a five-item scale based on the work of Nagi (1976) used in the EPESE (Seeman et al. 1994): pushing large objects, stooping or kneeling, carrying weights over 10 lb, extending arms above shoulder level, and writing. Advanced daily living tasks were evaluated by a three-item modified Rosow–Breslau scale (Rosow and Breslau 1966) assessing difficulty performing heavy work around the house or farm, walking up and down a flight of stairs, and walking a half mile without help. A final scale assessing more integrative functions IADLs consisted of six of the eight items suggested by Lawton and Brody (1969): traveling, shopping, preparing meals, doing housework, managing medications, and handling money.

In addition to these self-reported measures of functioning, a short physical performance test developed for use in the EPESE was used to assess lower extremity function (Seeman et al. 1994). The tasks include three measures of balance (side by side, semi-tandem, and tandem) for 10 s, the time required to sit and stand from a chair five times without using the hands and the time required to walk 10 ft.

Genotyping

The Affymetrix Human SNP Array 6.0 (Affymetrix®, Inc Santa Clara, CA) was used for genome-wide SNP genotyping. DNA was obtained from the Hussman Institute of Human Genomics at the University of Miami and the Center for Human Genetics Research at Vanderbilt University DNA Banks. Genomic DNA was quantitated via the ND-8000 spectrophotometer, and DNA quality was evaluated via gel electrophoresis. The genomic DNA (250 ng/5 μl) samples were processed according to standard Affymetrix procedures for processing of the assay. The arrays were then scanned using the GeneChip Scanner 3000 7 G operated by the Affymetrix® GeneChip® Command Console® software. The data were processed for genotype calling using the Affymetrix® Power Tools (APT) software using the birdseed calling algorithm version 2.0 Affymetrix®, Inc Santa Clara, CA (Korn et al. 2008).

Samples with call rates less than 95 % were reexamined by re-clustering analysis. The following sample QC procedures were applied: (1) one CEPH sample was placed on each plate to ensure reproducibility of results across the plates; (2) all samples with call rates <95 % were rechecked individually to ensure quality of genotypes; the DNA sample was examined using gel electrophoresis to determine the quality of the actual DNA sample before making a call to exclude it from analysis; (3) A sample call rate of 95 % was applied to all samples, if the call rate fell below this average after several chips were run, genotyping was stopped, and molecular procedures and equipment were tested for potential problems.

Statistical analysis

Quality control Additional sample QC was performed using the APT software release 1.12.0 available from Affymetrix®, Inc Santa Clara, CA. Twenty-nine samples were dropped due to low contrast QC. All individuals with inconsistent genders were dropped if the inconsistency could not be resolved; this removed 20 samples. RELPAIR software (Epstein et al. 2000) was used to test for misspecification of pedigree relationships and to obtain a list of problematic individuals to be followed up and resolved; all misspecifications were resolved. Seven samples were dropped due to sample genotyping efficiency (<95 %).

PLINK (Purcell et al. 2007) statistical software was used both for marker QC and to format data for analysis. Markers included in analyses had a minor allele frequency (MAF) >0.05 and a genotyping efficiency >95 %; this resulted in the removal of 221,982 markers for MAF and 30,303 markers for genotyping efficiency. Markers that did not map to a chromosome, mitochondrial SNPs, and X chromosome SNPs were also excluded (23,997). A goodness of fit test was used to examine deviations from Hardy–Weinberg equilibrium (HWE) in the most unrelated individuals (166 individuals determined by pairwise kinship coefficients obtained from the KinInbcoef program (Thornton and McPeek 2007)). Markers with strong deviations from HWE ($p<1.0\times10^{-7}$) were removed from the analysis; this removed seven markers. Genotypes on 551 individuals generated by the CAMP dementia study were also available for these same markers and were used for SNP allele frequency estimation, determining deviations from HWE, and setting linkage phase in the pedigrees.

Linkage analysis We used the software program PedCut (Liu et al. 2008) to subdivide the 13 generation pedigree into computationally tractable sub-pedigrees. The 74 SA individuals and their 7 full siblings were grouped into 16 sub-pedigrees informative for linkage (e.g., at least two SA individuals in each pedigree) that were small enough (bit size ≤24) to generate multipoint LOD scores using the Lander–Green algorithm implemented in Merlin (Abecasis et al. 2002). PLINK was used to select a set of 4,941 markers in low linkage disequilibrium (pairwise $r^2<0.16$ for all markers) for multipoint linkage analysis. Inter-marker map distances were obtained from the Rutgers Combined Physical–Linkage Map (Matise et al. 2007). Multipoint LOD scores allowing for heterogeneity (multipoint heterogeneity LOD score (MHLOD)) were calculated using affected-only dominant (trait allele frequency=0.001), affected-only recessive (trait allele frequency=0.20), and nonparametric allele-sharing (multipoint nonparametric allele-sharing LOD score (MLOD*)) models.

Case–control association The case–control analysis of SA was performed using the modified quasi-likelihood score (MQLS) method, a quasi-likelihood score test that accounts for correlations among individuals to maintain correct type I error rates and tests for allelic association (Thornton and McPeek 2007). MQLS uses a kinship variance/covariance matrix of inbreeding coefficients and kinship coefficients to account for inbreeding and relatedness between individuals using a specified pedigree structure. Population prevalence for SA of 30 % was specified, consistent with the frequency of SA in the current study. The method is computationally tractable for GWAS. Only analyses of autosomal SNPs were performed.

STATA 10.0 statistical software (STATA 10, College Station, TX, 2008) was used to create summary statistics for demographic data. Generalized estimating equations with an independence correlation matrix and robust standard errors were used to compare SA cases to controls for demographic data using nuclear families as the correlated grouping variable. p values are reported for demographic data analyses.

Results

A description of the criteria used to define SA and normally aged control individuals is provided in Table 1, and a description of the study sample is provided in Table 2. SA cases were slightly younger and had a higher proportion of males than the controls. A larger proportion of SA cases were from Ohio than Indiana, although a majority of participants were from Indiana. SA cases had slightly lower but not statistically different mean body mass index than the control. By design, measures used to define SA (tests of cognitive impairment, depression, and physical impairment tests) were lower in SA cases than the control.

Multipoint linkage analysis

The results of multipoint genome-wide linkage analysis for all chromosomes using parametric affected-only dominant, parametric affected-only recessive, and nonparametric linkage models are summarized in Fig. 1. Five regions of the genome generated multipoint LOD scores (MHLOD or MLOD*) exceeding three under at least one of the models tested. The peak MHLOD of 3.24 on chromosome 6 occurred under a dominant model at 166.19 Mb, with a 1-LOD-down support interval extending from 164.38 to 169.62 Mb. The peak MHLOD of 3.30 on chromosome 16 was obtained at 23.97 Mb under a recessive model (1-LOD-down support interval, 18.97 to 26.57 Mb). Three peaks were detected under the nonparametric model on chromosomes 10 (peak MLOD* 4.30 at 4.26 Mb, 1-LOD-down support interval 3.28 to 5.59 Mb), 17 (peak MLOD* 3.50 at 10.09 Mb, 1-LOD-down support interval 9.82 to 11.49 Mb), and 20 (peak MLOD* 3.3 at 0.93 Mb, 1-LOD-down support interval 0.58 to 4.07 Mb).

Association analysis

To assess association under linkage peaks, markers were examined for regionally significant association within the 1-LOD-down support intervals for the five multipoint linkage peaks. Only the association at intergenic SNP rs205990 (chromosome 6, 164.44 Mb (GRCh37/hg19 build) between quaking homolog KH domain RNA binding (*QKI*) and phosphodiesterase 10A (*PDE10A*), $p=2.36\times10^{-5}$) survived a Bonferroni correction for the number of SNPs located in each 1-LOD-down confidence interval ($p<2.8\times10^{-5}$ on

Table 2 Demographic and clinical characteristics of study population

Variable	Total (N=263)	Successfully aged (N=74)	Controls (N=189)	p value
Gender (% male)	42	65	33	$<1.0\times10^{-3}$
Age (mean)	83.20	81.82	83.74	$<1.0\times10^{-3}$
		Center (%)		
Ohio	45	66	37	$<1.0\times10^{-3}$
Indiana	55	34	63	
		County (%)		
Adams	9	1	12	$<1.0\times10^{-3}$
Elkhart/LaGrange	46	33	51	
Holmes	45	66	37	
BMI (mean)	27.74	26.70	28.24	0.024
3MS (median [range])	92 [58–103]	94 [87–103]	92 [58–101]	$<1.0\times10^{-3}$
GDS (median [range])	2 [0–14]	1 [0–4]	2 [0–14]	$<1.0\times10^{-3}$
ADL (median [range])	0 [0–9]	0 [0–1]	1 [0–9]	$<1.0\times10^{-3}$
IADL (Median [range])	0 [0–12]	0 [0–1]	1 [0–12]	$<1.0\times10^{-3}$
Nagi (median [range])	1 [0–5]	0 [0–1]	2 [0–5]	$<1.0\times10^{-3}$
Rosow–Breslau (median [range])	3 [3–6]	3 [3–4]	4 [3–6]	$<1.0\times10^{-3}$
EPESE (median [range])	8 [0–12]	10 [9–12]	7 [0–12]	$<1.0\times10^{-3}$

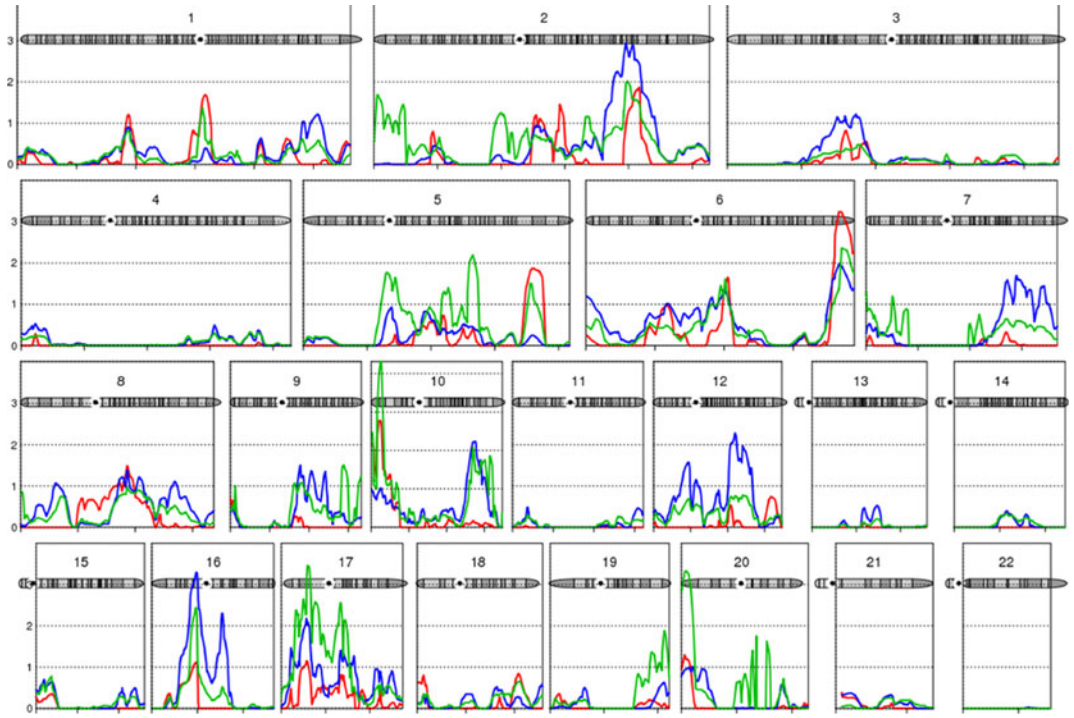

Fig. 1 Multipoint linkage results for all chromosomes. MHLOD for the dominant multipoint model (*red or moderate gray line*), recessive multipoint model (*blue or darkest gray line*), and MLOD* for the nonparametric multipoint model (*green or lightest gray line*) are presented in this figure

chromosome 6, $p<5.6\times10^{-5}$ on chromosome 10, $p<3.1\times10^{-5}$ on chromosome 16, $p<1.0\times10^{-4}$ on chromosome 17, and $p<4.4\times10^{-5}$ on chromosome 20).

In addition to the association mapping under linkage peaks, all genotyped SNPs were analyzed to detect any genome-wide significant associations (Bonferroni corrected $p<8\times10^{-8}$). Quantile–quantile (q–q) plots (Online Resource Fig. 1) demonstrated little deviation from expectations. As summarized in the Manhattan plot (Online Resource Fig. 2), no associations exceeded this threshold, but the strongest results genome-wide ($p<1\times10^{-5}$) are summarized in Table 3. Ten of these SNPs are in the fragile site 6 (*FRA6E*) region on chromosome 6 and are less than 3 Mb from the regionally significant SNP on chromosome 6 (rs205590). Eight additional SNPs were further than 3 Mb. Stratification by state of residence indicated that associations were concentrated in Amish from Ohio, consistent with linkage results in the 6q region.

Three SNPs in the *FRA6E* region were in a gene, Parkinson juvenile disease protein 2 (*PARK2*), and are in high linkage disequilibrium (LD) (Online Resource Fig. 3) with each other with r^2 ranging from 0.86 to 1.00. Upstream of the *PARK2* association, seven *FRA6E* markers in high LD ($r^2=1.00$) were moderately associated with SA. These markers were located ~60 Mb upstream of mitogen-activated protein kinase, kinase 4 (*MAP3K4*) and downstream of plasminogen (*PLG*). However, there is low pairwise LD between these three associations ($r^2<0.01$ between rs205990 and the other two blocks, $r^2=0.08$ between the *FRA6E* blocks), indicating that the signals are independent.

Discussion

We performed a GWAS examining linkage to and association with SA in the Amish and observed suggestive evidence for both at intergenic SNP rs205990 on chromosome 6. This marker survived interval-wide Bonferroni correction and is located between the genes *QKI* and *PDE10A*. Neither gene has been previously associated with risk for SA or any other longevity or aging trait. *PDE10A* is involved in neurologic cascades and has been associated with schizophrenia and Huntington's disease (Arnaud-

Table 3 Strongest MQLS results ($p<1\times10^{-5}$) from binary trait analyses

SNP	Chromosome	Position (bp)	MAF		HWE *p*	MQLS *p*	Nearest
			SA	Control			Genes
rs11585386	1	227085771	0.19	0.05	0.635	4.50×10^{-6}	*RHOU\RAB4A*
rs294588	5	162931093	0.49	0.26	0.400	1.05×10^{-6}	*MAT2B\LOC391844*
rs12518984	5	162940090	0.45	0.25	0.724	7.80×10^{-6}	*MAT2B\LOC391844*
rs1572438	6	803970	0.49	0.27	0.502	9.22×10^{-6}	*HUS1B\FOXQ1*
rs6942089	6	80491598	0.30	0.47	0.112	7.30×10^{-6}	*SH3BGRL2\ELOVL4*
rs9359390	6	80493147	0.30	0.49	0.269	3.38×10^{-6}	*SH3BGRL2\ELOVL4*
rs2038292	6	80509889	0.39	0.40	0.349	6.08×10^{-6}	*SH3BGRL2\ELOVL4*
rs2092745	6	80511550	0.40	0.39	0.276	7.33×10^{-6}	*SH3BGRL2\ELOVL4*
rs6925255	6	80522843	0.29	0.49	0.274	1.18×10^{-6}	*SH3BGRL2\ELOVL4*
rs932492	6	80531291	0.37	0.46	0.086	6.49×10^{-6}	*SH3BGRL2\ELOVL4*
rs9443716	6	80543068	0.37	0.46	0.087	6.29×10^{-6}	*SH3BGRL2\ELOVL4*
rs1247322	6	161250642	0.24	0.10	0.771	3.56×10^{-6}	*PLG\MAP3K4*
rs1247320	6	161253771	0.24	0.10	0.771	3.56×10^{-6}	*PLG\MAP3K4*
rs1247319	6	161253850	0.24	0.10	0.771	3.56×10^{-6}	*PLG\MAP3K4*
rs1247318	6	161253927	0.24	0.10	0.771	2.89×10^{-6}	*PLG\MAP3K4*
rs1247317	6	161254404	0.24	0.10	0.771	3.56×10^{-6}	*PLG\MAP3K4*
rs1247316	6	161255135	0.24	0.10	0.771	3.56×10^{-6}	*PLG\MAP3K4*
rs1247363	6	161262680	0.24	0.10	0.771	7.61×10^{-6}	*PLG\MAP3K4*
rs16892673	6	161853925	0.21	0.06	0.651	4.69×10^{-6}	*PARK2*
rs16892698	6	161863355	0.19	0.05	0.319	6.24×10^{-6}	*PARK2*
rs16892700	6	161864569	0.19	0.05	0.319	6.24×10^{-6}	*PARK2*
rs9918668	7	93529680	0.19	0.43	0.403	5.84×10^{-7}	*BET1\COL1A2*
rs643473	12	128041892	0.19	0.40	0.863	9.42×10^{-6}	*LOC100129252\NLRP9P*

Bold indicates a deviation from HWE

bp base pair position from NCBI build 37.1, *MAF* minor allele frequency

Lopez et al. 2008; Kehler 2011; Kleiman et al. 2011). A previously published genome-wide scan associating thyroid stimulating hormone (TSH) and thyroid levels with gene variants observed a strong association with *PDE8A* and *PDE10A* in age diverse populations (Sardinians, Tuscans, and Old Order Amish) (Arnaud-Lopez et al. 2008). The strongest association in this previously published study was in *PDE8A*; however, *PDE10A* is in the same enzyme family and is involved in the cyclic AMP (cAMP) and cyclic GMP signaling pathways. The PDE10A gene protein product has been found to be involved in the pathophysiology of human disease. TSH controls thyroid function through binding of G protein coupled receptor, TSH receptor (TSHR), and cAMP production. TSH plays a role in muscle, bone, central nervous system, and heart physiology, and therefore may contribute to several complex disorders. In addition, the TSHR expression has previously associated with exceptional longevity in a previously published study of Ashkenazi Jewish centenarians, which may support a role for *PDE10A* in SA (Atzmon et al. 2009b). Although we observed strong evidence for linkage at three additional regions on chromosomes 10, 17, and 20, markers under these linkage peaks did not show individual evidence for association. This may be due to either lack of power to find association or it may be possibly due to segregation of alleles in sub-pedigrees and not population-level association.

The SNPs with the strongest, but not genome-wide significant, evidence for association ($p<1.00\times10^{-5}$) were not located in genes and did not show evidence for linkage. Several of these markers are on chromosome 6 within the *FRA6E* region, which is less than

3 Mb from the top linkage result. However, there is no extensive LD between SNPs in the linkage region and the *FRA6E* region, indicating that these association signals might be independent. *FRA6E* has not been previously implicated with SA, although some studies have proposed *PARK2*, located with the *FRA6E* region, as a putative candidate for longevity and aging (Lunetta et al. 2007). The nearest neighboring genes to these markers in the *FRA6E* region were *PLG* and *MAP3K4*. PLG is a circulating zymogen with the primary function of dissolving fibrin (OMIM #173350). *PLG* is not an established candidate gene for SA but has been suggested to be involved with wasting (weight loss and poor wound healing) and decreased longevity due to disseminated fibrin deposition in animal model studies (Cheng et al. 2005). There is no established biological evidence implicating *MAP3K4* with longevity or aging.

We observed a higher rate of SA in males relative to females in our cohort, results that have not been observed consistently in previous literature. Studies in the MacArthur Studies of Successful Aging observed that males were more likely to have better physical performance at baseline in a cohort of 70–79-year olds with high function, consistent with our observations (Seeman et al. 1994). The studies by McLaughlin et al. of adults aged 65 and older, however observed the opposite association between SA and gender (McLaughlin et al. 2010). A possible explanation for these differences is the definition of SA used. We attempted to use a definition as comprehensive as possible, integrating factors pertaining to physical and cognitive function, social engagement, and health, whereas other studies often used only one or two of these components. Another possible explanation for the larger number of SA males relative to females could be the effect of marital status on successful aging which previously has been indicated (Li et al. 2006). Our sample also had a greater proportion of successfully aged men (52 % of men SA vs. 17 % of women SA) who were currently married (78 % of men currently married vs. 45 % of women currently married). The relationship between gender, marriage, and SA has not been consistent across studies, with the majority finding no association; however, it merits further follow-up (Depp and Jeste 2006). Further work is necessary to fully understand the significance of these findings.

A potential limitation to our study was the inability of MQLS to allow for covariate adjustment, and as a result we were unable to adjust for heterogeneity in study findings across ascertainment center (Ohio and Indiana). However, our linkage analyses should implicitly take care of confounding through pedigree structure. Since we looked at a consensus of results across linkage and association analyses, we believe that the results are not likely confounded by ascertainment center and merit further follow-up.

We previously reported evidence for linkage of SA to chromosomes 6, 7, and 14 in a smaller subset of the sample (48 SA cases) used here (Velez Edwards et al. 2011). In the current sample of 74 SA cases, genotyped for a much denser map of markers, the strongest evidence for linkage was obtained in a different area of chromosome 6. This region, near *FRA6E*, did provide modest evidence for linkage (MHLOD=1.10–2.09) in our original study, indicating that the expanded sample improved power to detect linkage (Velez Edwards et al. 2011). The difference in evidence for linkage across the three models considered (dominant, recessive, and nonparametric allele-sharing) may reflect differences in the goodness of fit for each model given different modes of inheritance. While evidence for linkage maximizes under different models for chromosome 6 (dominant), 16 (recessive), and 10, 17 and 20 (nonparametric), the linkage curves for the other models follow a similar pattern (nonparametric model reaches 2.5 on chromosomes 6 and 16, the dominant model peaks just under 3 on chromosome 10, and the recessive model just over 2 on chromosome 17), suggesting that these findings are fairly robust to the specification of a genetic model. A notable exception is chromosome 20, which generates much stronger linkage under nonparametric analysis than the other two models (maximum MHLOD ~1 under a dominant model). We note, however, that given relatively small sample size from an isolated population and differences in results from our previous study in this population, further replication of these results in an independent cohort is desirable. The correlation between linkage and association on chromosome 6 establishes a priority area for

further study in this sample to identify functional variations in genes such as *QKI*, *PDE10A*, and other genes in the region that might be responsible for promoting SA in these individuals.

Acknowledgments We acknowledge the of the Vanderbilt CHGR DNA Resources core for processing some of the samples. Some of the samples used in this study were collected while WKS, JRG, and MAP-V were faculty members at Duke University. This work supported by NIH grants AG019085 to Jonathan L. Haines and AG019726 to William K. Scott.

References

Abecasis GR, Cherny SS, Cookson WO, Cardon LR (2002) Merlin—rapid analysis of dense genetic maps using sparse gene flow trees. Nat Genet 30:97–101

Agarwala R, Biesecker LG, Schaffer AA (2003) Anabaptist genealogy database. Am J Med Genet C Semin Med Genet 121C:32–37

Arcos-Burgos M, Muenke M (2002) Genetics of population isolates. Clin Genet 61:233–247

Arking DE, Atzmon G, Arking A, Barzilai N, Dietz HC (2005) Association between a functional variant of the KLOTHO gene and high-density lipoprotein cholesterol, blood pressure, stroke, and longevity. Circ Res 96:412–418

Arnaud-Lopez L, Usala G, Ceresini G, Mitchell BD, Pilia MG, Piras MG, Sestu N, Maschio A, Busonero F, Albai G, Dei M, Lai S, Mulas A, Crisponi L, Tanaka T, Bandinelli S, Guralnik JM, Loi A, Balaci L, Sole G, Prinzis A, Mariotti S, Shuldiner AR, Cao A, Schlessinger D, Uda M, Abecasis GR, Nagaraja R, Sanna S, Naitza S (2008) Phosphodiesterase 8B gene variants are associated with serum TSH levels and thyroid function. Am J Hum Genet 82:1270–1280

Atzmon G, Barzilai N, Hollowell JG, Surks MI, Gabriely I (2009a) Extreme longevity is associated with increased serum thyrotropin. J Clin Endocrinol Metab 94:1251–1254

Atzmon G, Barzilai N, Surks MI, Gabriely I (2009b) Genetic predisposition to elevated serum thyrotropin is associated with exceptional longevity. J Clin Endocrinol Metab 94:4768–4775

Banasik K, Ribel-Madsen R, Gjesing AP, Wegner L, Andersson A, Poulsen P, Borglykke A, Witte DR, Pedersen O, Hansen T, Vaag A (2011) The FOXO3A rs2802292 G-allele associates with improved peripheral and hepatic insulin sensitivity and increased skeletal muscle-FOXO3A mRNA expression in twins. J Clin Endocrinol Metab 96:E119–E124

Capurso C, Solfrizzi V, D'Introno A, Colacicco AM, Capurso SA, Semeraro C, Capurso A, Panza F (2004) Interleukin 6-174G/C promoter gene polymorphism in centenarians: no evidence of association with human longevity or interaction with apolipoprotein E alleles. Exp Gerontol 39:1109–1114

Carmelli D, Kelly-Hayes M, Wolf PA, Swan GE, Jack LM, Reed T, Guralnik JM (2000) The contribution of genetic influences to measures of lower-extremity function in older male twins. J Gerontol A Biol Sci Med Sci 55:B49–B53

Cellini E, Nacmias B, Olivieri F, Ortenzi L, Tedde A, Bagnoli S, Petruzzi C, Franceschi C, Sorbi S (2005) Cholesteryl ester transfer protein (CETP) I405V polymorphism and longevity in Italian centenarians. Mech Ageing Dev 126:826–828

Cheng Q, Zhao Y, Lawson WE, Polosukhin VV, Johnson JE, Blackwell TS, Gailani D (2005) The effects of intrinsic pathway protease deficiencies on plasminogen-deficient mice. Blood 106:3055–3057

Chung WH, Dao RL, Chen LK, Hung SI (2010) The role of genetic variants in human longevity. Ageing Res Rev 9 (Suppl 1):S67–S78

Deelen J, Beekman M, Uh HW, Helmer Q, Kuningas M, Christiansen L, Kremer D, van der Breggen R, Suchiman HE, Lakenberg N, van den Akker EB, Passtoors WM, Tiemeier H, van Heemst D, de Craen AJ, Rivadeneira F, de Geus EJ, Perola M, van der Ouderaa FJ, Gunn DA, Boomsma DI, Uitterlinden AG, Christensen K, van Duijn CM, Heijmans BT, Houwing-Duistermaat JJ, Westendorp RG, Slagboom PE (2011) Genome-wide association study identifies a single major locus contributing to survival into old age; the APOE locus revisited. Aging Cell 10:686–698

Depp CA, Jeste DV (2006) Definitions and predictors of successful aging: a comprehensive review of larger quantitative studies. Am J Geriatr Psychiatry 14:6–20

Epstein MP, Duren WL, Boehnke M (2000) Improved inference of relationship for pairs of individuals. Am J Hum Genet 67:1219–1231

Farrer LA, Conneally PM, Yu PL (1984) The natural history of Huntington's disease: possible role of "aging genes". Am J Med Genet 18:115–123

Glatt SJ, Chayavichitsilp P, Depp C, Schork NJ, Jeste DV (2007) Successful aging: from phenotype to genotype. Biol Psychiatry 62:282–293

Hahs DW, McCauley JL, Crunk AE, McFarland LL, Gaskell PC, Jiang L, Slifer SH, Vance JM, Scott WK, Welsh-Bohmer KA, Johnson SR, Jackson CE, Pericak-Vance MA, Haines JL (2006) A genome-wide linkage analysis of dementia in the Amish. Am J Med Genet B Neuropsychiatr Genet 141B:160–166

Katz S (1983) Assessing self-maintenance: activities of daily living, mobility, and instrumental activities of daily living. J Am Geriatr Soc 31:721–727

Kehler J (2011) PDE10A inhibitors: novel therapeutic drugs for schizophrenia. Curr Pharm Des 17:137–150

Khachaturian AS, Gallo JJ, Breitner JC (2000) Performance characteristics of a two-stage dementia screen in a population sample. J Clin Epidemiol 53:531–540

Kleiman RJ, Kimmel LH, Bove SE, Lanz TA, Harms JF, Romegialli A, Miller KS, Willis A, des Estages S, Kuhn M, Schmidt CJ (2011) Chronic suppression of phosphodiesterase 10A alters striatal expression of genes responsible for neurotransmitter synthesis, neurotransmission, and signaling pathways implicated in Huntington's disease. J Pharmacol Exp Ther 336:64–76

Kleindorp R, Flachsbart F, Puca AA, Malovini A, Schreiber S, Nebel A (2011) Candidate gene study of FOXO1, FOXO4, and FOXO6 reveals no association with human longevity in Germans. Aging Cell 10:622–628

Korn JM, Kuruvilla FG, McCarroll SA, Wysoker A, Nemesh J, Cawley S, Hubbell E, Veitch J, Collins PJ, Darvishi K, Lee C, Nizzari MM, Gabriel SB, Purcell S, Daly MJ, Altshuler D (2008) Integrated genotype calling and association analysis of SNPs, common copy number polymorphisms and rare CNVs. Nat Genet 40:1253–1260

Lawton MP, Brody EM (1969) Assessment of older people: self-maintaining and instrumental activities of daily living. Gerontologist 9:179–186

Li C, Wu W, Jin H, Zhang X, Xue H, He Y, Xiao S, Jeste DV, Zhang M (2006) Successful aging in Shanghai, China: definition, distribution, and related factors. Int Psychogeriatr 18:551–563

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

Liu F, Kirichenko A, Axenovich TI, van Duijn CM, Aulchenko YS (2008) An approach for cutting large and complex pedigrees for linkage analysis. Eur J Hum Genet 16:854–860

Lunetta KL, D'Agostino RB Sr, Karasik D, Benjamin EJ, Guo CY, Govindaraju R, Kiel DP, Kelly-Hayes M, Massaro JM, Pencina MJ, Seshadri S, Murabito JM (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

Matise TC, Chen F, Chen W, De La Vega FM, Hansen M, He C, Hyland FC, Kennedy GC, Kong X, Murray SS, Ziegle JS, Stewart WC, Buyske S (2007) A second-generation combined linkage physical map of the human genome. Genome Res 17:1783–1786

McLaughlin SJ, Connell CM, Heeringa SG, Li LW, Roberts JS (2010) Successful aging in the United States: prevalence estimates from a national sample of older adults. J Gerontol B Psychol Sci Soc Sci 65B:216–226

Nagi SZ (1976) An epidemiology of disability among adults in the United States. Milbank Mem Fund Q Health Soc 54:439–467

Naumova E, Mihaylova A, Ivanova M, Michailova S, Penkova K, Baltadjieva D (2004) Immunological markers contributing to successful aging in Bulgarians. Exp Gerontol 39:637–644

Nebel A, Kleindorp R, Caliebe A, Nothnagel M, Blanche H, Junge O, Wittig M, Ellinghaus D, Flachsbart F, Wichmann HE, Meitinger T, Nikolaus S, Franke A, Krawczak M, Lathrop M, Schreiber S (2011) A genome-wide association study confirms APOE as the major gene influencing survival in long-lived individuals. Mech Ageing Dev 132:324–330

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

Pericak-Vance MA, Johnson CC, Rimmler JB, Saunders AM, Robinson LC, D'Hondt EG, Jackson CE, Haines JL (1996) Alzheimer's disease and apolipoprotein E-4 allele in an Amish population. Ann Neurol 39:700–704

Perls TT, Wilmoth J, Levenson R, Drinkwater M, Cohen M, Bogan H, Joyce E, Brewster S, Kunkel L, Puca A (2002) Life-long sustained mortality advantage of siblings of centenarians. Proc Natl Acad Sci U S A 99:8442–8447

Poduslo SE, Huang R, Spiro A III (2009) A genome screen of successful aging without cognitive decline identifies LRP1B by haplotype analysis. Am J Med Genet B Neuropsychiatr Genet 153B:114–119

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

Rosow I, Breslau N (1966) A Guttman health scale for the aged. J Gerontol 21:556–559

Rowe JW, Kahn RL (1997) Successful aging. Gerontologist 37:433–440

Seeman TE, Charpentier PA, Berkman LF, Tinetti ME, Guralnik JM, Albert M, Blazer D, Rowe JW (1994) Predicting changes in physical performance in a high-functioning elderly cohort: MacArthur studies of successful aging. J Gerontol 49:M97–108

Snejdrlova M, Kalvach Z, Topinkova E, Vrablik M, Prochazkova R, Kvasilova M, Lanska V, Zlatohlavek L, Prusikova M, Ceska R (2011) APOE polymorphism as a potential determinant of functional fitness in the elderly regardless of nutritional status. Neuro Endocrinol Lett 32:51–54

Swedlund AC, Meindl RS, Nydon J, Gradie M (1983) Family patterns in longevity and longevity patterns of the family. Hum Biol 55:115–129

Taioli E, Mari D, Franceschi C, Bonafe M, Monti D, Bertolini S, Marinelli D, Garte S (2001) Polymorphisms of drug-metabolizing enzymes in healthy nonagenarians and centenarians: difference at GSTT1 locus. Biochem Biophys Res Commun 280:1389–1392

Teng EL, Chui HC (1987) The Modified Mini-Mental State (3MS) examination. J Clin Psychiatry 48:314–318

Thornton T, McPeek MS (2007) Case–control association testing with related individuals: a more powerful quasi-likelihood score test. Am J Hum Genet 81:321–337

US Census Bureau Population Division and Housing and Household Economic Statistics Division. National Population Projections. 2001. Ref Type: Report

van der Walt JM, Scott WK, Slifer S, Gaskell PC, Martin ER, Welsh-Bohmer K, Creason M, Crunk A, Fuzzell D, McFarland L, Kroner CC, Jackson CE, Haines JL, Pericak-Vance MA (2005) Maternal lineages and Alzheimer disease risk in the old order Amish. Hum Genet 118:115–122

Velez Edwards D, Gilbert J, Jiang L, Gallins P, Caywood L, Creason M, Fuzzell D, Knebusch C, Jackson C, Pericak-Vance M, Haines J, Scott W (2011) Successful aging shows linkage to chromosome 6, 7, and 14 in the Amish. Ann Hum Genet 75(4):516–528

Yesavage JA (1988) Geriatric depression scale. Psychopharmacol Bull 24:709–711