Research Paper

Meta-analysis of genetic variants associated with human exceptional longevity

Paola Sebastiani[1], Harold Bae1, Fangui X. Sun[1], Stacy L. Andersen[2], E. Warwick Daw[3], Alberto Malovini[4], Toshio Kojima[5], Nobuyoshi Hirose[6], Nicole Schupf[7], Annibale Puca[8], Thomas T Perls[2]

[1] Department of Biostatistics, Boston University School of Public Health, Boston MA 02118, USA
[2] Section of Geriatrics, Department of Medicine, Boston University School of Medicine and Boston Medical Center, Boston, MA 02118, USA
[3] Division of Statistical Genomics, Washington University School of Medicine, St. Louis, MO 63110, USA
[4] Laboratorio di Informatica Biomedica, Dipartimento di Ingegneria Industriale e dell'Informazione, Università di Pavia, Pavia, Italy
[5] Research Center for Physical Fitness, Sports and Health, Toyohashi University of Technology, Toyohashi , Japan
[6] Division of Geriatric Medicine, Department of Internal Medicine, Keio University School of Medicine, Tokyo, Japan
[7] Taub Institute for Research on Alzheimer's Disease and the Aging Brain, Columbia University Medical Center, New York, NY 10032, USA
[8] Unit of Genetics, Cardiovascular Research Institute Istituto Ricovero Cura Carattere Scientifico Multimedica, Sesto S. Giovanni, Italy; Facolta' di Medicina, Universita' di Salerno, Baronissi, Italy

Key words: *centenarian; exceptional longevity; genetic association study; aging; gene; lifespan; meta-analysis*
Received: *8/2/13; Accepted: 8/22/13; Published: 8/24/13*
Correspondence to: *Paola Sebastiani, PhD;* ***E-mail:*** *sebas@bu.edu*

Abstract: Despite evidence from family studies that there is a strong genetic influence upon exceptional longevity, relatively few genetic variants have been associated with this trait. One reason could be that many genes individually have such weak effects that they cannot meet standard thresholds of genome wide significance, but as a group in specific combinations of genetic variations, they can have a strong influence. Previously we reported that such genetic signatures of 281 genetic markers associated with about 130 genes can do a relatively good job of differentiating centenarians from non-centenarians particularly if the centenarians are 106 years and older. This would support our hypothesis that the genetic influence upon exceptional longevity increases with older and older (and rarer) ages. We investigated this list of markers using similar genetic data from 5 studies of centenarians from the USA, Europe and Japan. The results from the meta-analysis show that many of these variants are associated with survival to these extreme ages in other studies. Since many centenarians compress morbidity and disability towards the end of their lives, these results could point to biological pathways and therefore new therapeutics to increase years of healthy lives in the general population.

INTRODUCTION

In Sebastiani et al "Genetic signatures of exceptional longevity in humans" [1], we presented the results from a genome wide association study of exceptional longevity in 801 centenarians from the New England Centenarian Study (NECS, mean age at death 104 years) and 914 genetically matched controls. The study identified a group of 281 SNPs that, used jointly in a genetic risk model, had 60% sensitivity to discriminate between centenarians and healthy controls. The sensitivity of the model however increased with more extreme ages of the centenarians and reached 85% for subjects age>107 years. The 281 SNPs included rs2075650 in TOMM40/APOE that reached irrefutable genome-wide significance and replicated in an

independent cohort of 253 nonagenarians and centenarians from the Elixir Pharmaceuticals Study of Extreme Longevity and 341 genetically matched controls. The other 280 SNPs were statistically significant with p-values ranging between 10-2 and 10-6 although their level of significance did not meet the stringent criterion for genome-wide significance of 5x10-8, thus raising the possibility that these associations could be false positives. We therefore set out to determine which of these 281 SNPs were associated with longevity in a meta-analysis that included the two original studies, in addition to a case control study of longevity with nonagenarians and centenarians from the Southern Italian Centenarian Study [2], and a case control study of nonagenarians and centenarians from the Long Life Family Study [3]. We also extended the meta-analysis to include genotype data of a subset of SNPs from the Japanese Centenarian Study [4] .

RESULTS

Table 1 lists the studies' characteristics. The ELIX, SICS, LLFS and JCS case-control studies were all smaller than the NECS and cases in the ELIX, SICS and LLFS were younger than the NECS. Controls in the LLFS were males who died by the age of 94 or females who died by the age of 95, and 85% of these controls are relatives of the cases (eg siblings who died at younger ages) from the same family of the cases, so that they provide the strongest type of genetic matching. Some of the controls in the NECS and ELIX studies were chosen from the Illumina repository of controls and their ages are unknown.

Table 1. Description of the Studies

Study Population	Symbol	N Cases	Age of Cases	N Controls	Age of Controls	Genotyping Platform
Elixir Pharmaceutical Longevity Study	ELIX	253	100 (89-114)	341	NA	Illumina 370/550/610
Japanese Centenarian Study	JCS	513	106 (100-114)	561	69 (19-89)	Affymetrix 500KEA/500K/5.0
Long Life Family Study	LLFS	738	98 (95-110)	356	91 (44-95)	Illumina Omni 2.5
New England Centenarian Study	NECS	801	104 (95-119)	914	73 (53-90)(1)	Illumina 370/550/610/1M
Southern Italian Centenarian Study	SICS	410	95 (90-109)	553	NA	Illumina 317/370

Summary characteristics of the studies included in the meta-analysis. Samples genotyped with the Illumina 550 array are from the Illumina iControlDB. Controls in the NECS and ELIX studies were genetically matched as described in [1], controls in the SICS and JCS study were geographically matched, and 86% of controls in LLFS were family matched.
(1):Summary ages for 241 of the 914 controls enrolled in the NECS.

In the meta-analysis of additive genetic associations in NECS, ELIX, SICS and LLFS, 10 SNPs reached statistically significant association after Bonferroni correction (p-value < 0.05/280=0.00018). An additional 4 SNPs reached Bonferroni corrected statistical significance using a dominant model for the top-strand allele, and two SNPs reached Bonferroni corrected statistically significance using a recessive model for the top-strand allele (Table 2). The number of significant associations was much larger when a 5% and 6% false discovery rate corrections were used (Supplement Table 1). The Venn diagram in Figure 1 shows the number of significant SNPs from the meta-analysis of additive, dominant and recessive models, and 128 SNPs reached statistical significance with 6% false discovery rate. Note the substantial overlapping between the results with different genetic models. In fact, the 3 parameters of the additive, dominant and recessive models are functionally related and the tests are not independent. The full list of results for the meta-analysis of the 280 SNPs is in Supplement Table 1.

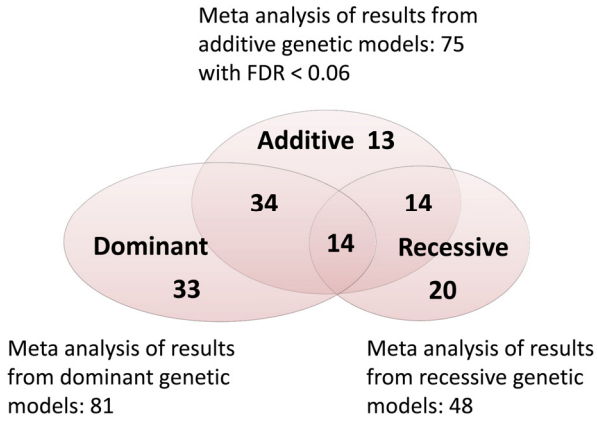

Figure 1. Venn diagram showing the number of significant associations from the meta-analysis of additive, dominant and recessive models when a 6% false discovery rate (FDR) was used. Genotypes were called using the top-strand rule and dominant and recessive models were coded for the top-strand allele A as explained in methods.

Table 2. SNPs that reached Bonferroni corrected significance in meta-analysis of results using additive, dominant and recessive models of Caucasian studies.

Row	SNP	Gene	Alleles	A.AF[1]	CA[2]	NECS. OR	SICS. OR	ELIX. OR	LLFS. OR	MetaOR (95% CI)	pval
1	rs2075650	TOMM40/APOE	A/G	0.925	G	0.492	0.726	0.499	0.507	0.527 (0.452;0.616)	4.44E-16
2	rs1525501	NA	A/G	0.128	G	0.761	0.836	0.510	0.719	0.724 (0.630;0.831)	4.94E-06
3	rs3803833	NA	A/C	0.875	C	0.712	0.974	0.856	0.667	0.765 (0.675;0.867)	2.72E-05
4	rs1016013	NA	A/G	0.373	G	1.300	1.168	1.186	1.089	1.206 (1.105;1.317)	2.87E-05
5	rs216148	CSF1R	A/G	0.124	G	0.692	1.011	0.550	0.798	0.753 (0.655;0.865)	6.36E-05
6	rs1867102	C9orf3	A/G	0.477	G	1.295	1.144	1.107	1.117	1.195 (1.094;1.304)	7.28E-05
7	rs1822590	NA	A/C	0.277	C	1.309	1.120	1.109	1.190	1.208 (1.100;1.327)	7.57E-05
8	rs4918255	SORCS1	A/G	0.665	G	1.265	1.103	1.180	1.238	1.212 (1.101;1.333)	8.81E-05
9	rs1456669	NA	A/C	0.160	C	0.638	0.835	1.118	0.817	0.777 (0.685;0.882)	9.89E-05
10	rs915179	LMNA	A/G	0.541	G	1.342	1.112	1.198	0.999	1.191 (1.090;1.301)	0.00010
11	rs2738679	WWOX	A/G	0.700	GG	1.840	1.213	1.802	1.587	1.600 (1.270;2.017)	6.85E-05
12	rs17702471	GPC6	A/G	0.785	GG	2.512	1.673	1.613	1.142	1.810 (1.347;2.431)	8.33E-05
13	rs1042663	C2	A/G	0.109	GG	0.617	0.808	0.629	0.988	0.720 (0.612;0.848)	8.386E-05
14	rs651922	DCPS	A/G	0.724	GG	2.312	1.035	2.040	1.079	1.650 (1.270;2.145)	0.00018
15	rs11218921	NA	A/G	0.920	AG/GG	0.502	0.655	0.810	0.523	0.590 (0.457;0.762)	5.47E-05
16	rs2738173	DEFB1	A/G	0.842	AG/GG	0.652	0.793	0.973	0.945	0.778 (0.683;0.887)	0.00016

Sixteen SNPs that reached Bonferroni corrected statistical significance (0.05/281=0.00018) in the meta-analysis of additive models (rows 1—10); dominant models for the A allele (rows 11—16) and recessive model for the A allele (row 15-16).

[1] A.AF= frequency of A allele in NECS controls

[2] CA= coded allele in genetic models.

Table 3. SNPs that reached Bonferroni corrected significance in meta-analysis of results using additive, dominant and recessive models of Caucasian and Japanese studies.

Row	SNP	Gene	Allele	A.AF[1]	CA[2]	NECS.OR	SICS.OR	ELIX.OR	JCS.OR	LLFS.OR	MetaOR	pval
1	rs1525501	NA	A/G	0.128	G	0.761	0.836	0.510	0.949	0.719	0.807 (0.725;0.898)	8.97E-05
2	rs1456669	NA	A/C	0.160	C	0.638	0.835	1.118	0.937	0.817	0.825 (0.743;0.916)	0.000312
3	rs4729049	CDK6	A/G	0.889	G	1.357	1.170	1.132	1.201	1.085	1.212 (1.078;1.362)	0.001311
4	rs11954180	SLC6A7	A/G	0.058	G	1.540	1.214	1.401	0.915	1.022	1.302 (1.105;1.535)	0.001639
5	rs2596230	RYR3	A/G	0.876	GG	4.175	2.755	3.099	1.099	0.893	2.607 (1.536;4.423)	0.000383
6	rs1800392	WRN	A/C	0.481	AC/CC	0.636	0.815	0.923	0.898	0.906	0.787 (0.685;0.904)	0.000708

SNPs that reached Bonferroni corrected statistical significance (0.05/19=0.0026) in the meta-analysis of additive models (rows 1—4); dominant models for the A allele (row 5), and recessive models for the A allele (row 6).

[1] A.AF= frequency of A allele in NECS controls

[2] CA= coded allele in genetic models.

Only 19 SNPs in the set of 28that reached FDR corrected significance had genotype data available in the JCS set, and the meta-analysis of these 19 SNPs was extended to include the results for the JCS set Six SNPs (see Table 3) reached Bonferroni corrected significance (p-value< 0.0026=0.05/19) in the meta-analysis of results from additive models (4 SNPs),dominant models (1 SNP), and recessive models (1 SNP) and 4 of these 6 SNPs were not included in the list of 16 that reached Bonferroni corrected significance in the meta-analysis of NECS, ELIX, SICS and LLFS. Fourteen SNPs in 19 reached 6% FDR corrected statistical significance. Full details of the analysis that included the JCS set is in Supplement Table 2.

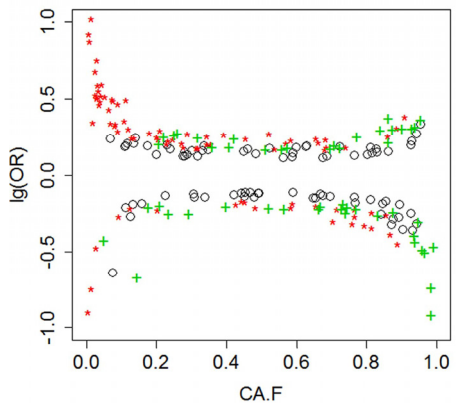

Figure 2. Genetic effects versus allele frequency. Black circles= additive effects; Red asterisks: dominant effects; Green crosses= recessive effects.

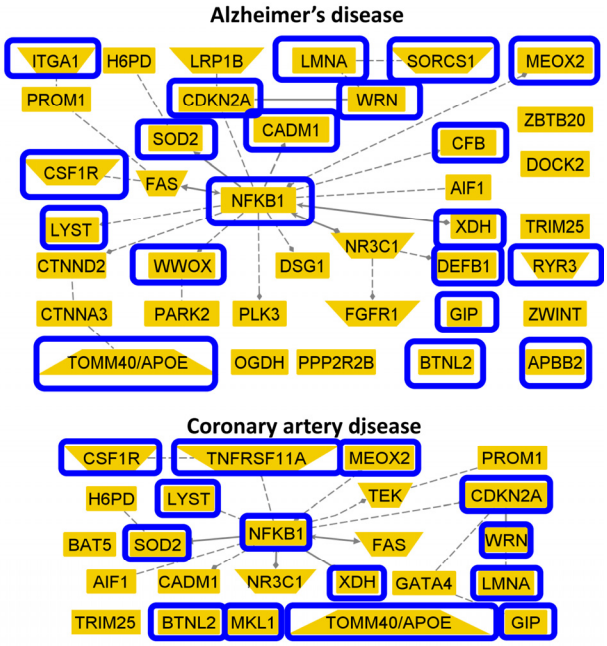

Figure 3. Genes with SNPs that reach statistical significance with meta-analysis and were implicated in Alzheimer's and coronary artery disease. The two networks display 38 genes linked to Alzheimer's disease (top) and 24 genes linked to coronary artery disease (bottom) that included SNPs in the list of 281 in [1]. Genes circled in blue include SNPs that reached statistical significance in the meta-analysis.