Original Article

Genetic Investigation of an Iranian Supercentenarian by Whole Exome Sequencing

Tara Akhtarkhavari MSc[1], Mohammad Taghi Joghataei PhD[2], Zohreh Fattahi MSc[1], Mohammad Reza Akbari MD PhD[3], Farzaneh Larti MD PhD[1], Hossein Najmabadi PhD[1], Kimia Kahrizi MD•[1]

Abstract

Background: The genetic basis of longevity is an important field of study because the majority of supercentenarian cases experience healthy aging and may only show age-related diseases in their last few years of life. It is clear that genetic factors play an important role in survival beyond 90 years of age, but the exact relationship of genetic variants to this phenomenon remains unknown.

Objective: The aim of this project was to investigate different hypotheses that describe the relationship between genetic variants and human longevity in a living Iranian man by Whole Exome Sequencing (WES).

Methods: Initially, we conducted high quality DNA extraction on a peripheral blood sample. Then, WES was performed on the DNA and different bioinformatic software packages and databases were used to analyze the data. Tertiary analysis was performed on four genetic hypotheses for longevity.

Results: Analysis showed that among 27 metabolic variants which are related to longevity, 18 variants encompassed the exceptional longevity allele. In comparison with the NHGRI GWAS catalog, the case had 58 trait-associated variants of which 11 were homozygous for the risk allele. We also discovered 25 novel variants within candidate genes for aging and longevity and we detected seven longevity-associated variants in the sample.

Conclusion: This study was performed on just one sample and so the results cannot be interpreted as a generalized principle for other elderly societies, but this is the first step towards investigation of the genetic basis of longevity in Iran and provides an insight for further studies in the field of longevity.

Keywords: Aging, computational biology, whole exome sequencing, Iran, longevity

Cite this article as: Akhtarkhavari T, Joghataei MT, Fattahi Z, Akbari MR, Larti F, Najmabadi H, Kahrizi K. Genetic Investigation of an Iranian Supercentenarian by Whole Exome Sequencing. *Arch Iran Med*. 2015; 18(10): 688 – 697.

Introduction

Human longevity is a rare condition which occurs in approximately one individual per seven million.[1,2] This phenotype is one of the most complicated traits[1] and different genetic, epigenetic and environmental factors are involved in it.[3] It has been suggested that genetic factors have an approximately 20%–30% contribution to survival up to 85 years of age,[1,4] but this contribution intensifies above this age.[1,5–8]

In developed countries, it has been shown that the mortality rate among elderly is decreasing and also the average lifespan has risen, and as a result, the number of old people has increased.[9–12] As in developed countries, within the last 50 years, the percentage of individuals over 60 years of age has increased in Iran from 6.24% to 7.27%.[13] Old people often suffer from age-related diseases such as Type 2 diabetes, osteoporosis, and cardiovascular diseases which impose an enormous socioeconomic burden not only on the aged individuals and their families, but also on the society.[9] Therefore, the genetic basis of longevity is an important field of study because the majority of supercentenarians experience healthy aging and may only show age-related diseases in their last few years of life.[14]

Up to now, it is clear that genetic factors play an important role in survival beyond 90 years of age, but the exact correlation between genetic variants and longevity is still unknown; hence, we have studied the different hypotheses that describe the relationship between genetic variants and human longevity in an Iranian supercentenarian.

Authors' affiliations: [1]Genetics Research Center, University of Social Welfare and Rehabilitation Sciences, Tehran, Iran. [2]Division of Neuroscience, Cellular and Molecular Research Center, Iran University of Medical Sciences, Tehran, Iran. [3]Women's College Research Institute, University of Toronto, Toronto, Canada.

•Corresponding author and Reprints: Kimia Kahrizi MD, Genetics Research Center, University of Social Welfare and Rehabilitation Sciences, Tehran, Iran. E-mail: kahrizi@yahoo.com

Accepted for publication: 18 August 2015

Materials and Methods

The study was approved by the Ethics Committee of the University of Social Welfare and Rehabilitation Sciences, Tehran, Iran, and written informed consent was obtained from the case and his legal guardian.

Case description

The case was a healthy man who was described as a supercentenarian (>110 years old[14] / 110–119 years old[8]) on the basis of his ID card (for reasons of confidentiality, we cannot mention his exact age). He was an Iranian individual who came from a village near Yazd. He had one son and six daughters, and a brother who died at the age of 90 but further information about his mother and father was not available. A physician examined the case and his clinical history is shown in Table 1.

Table 1. Clinical history of our case.

Clinical characteristics	Description
Cognitive decline/forgetfulness/declined consciousness	Negative
Eye examination: refractory errors/cataract/presbyopia/astigmatism	Negative
Skin appearance: pruritus/dryness/eczema/abnormal pigmentation	Negative
Teeth	Dentures
Age-related deafness	Negative
Speech	Normal
Tremor	Negative
Motor movements: walking with a cane	Yes
Musculoskeletal: pain in joint, muscle, bone/muscular atrophy/knee arthritis	Knee arthritis
Blood pressure	Normal
Genitourinary tract (GUI): Urinary infection/urinary incontinence impotency/ benign prostate hyperplasia (BPH)	Negative
History of surgery: cataract/prostate/appendix/glaucoma/cardiac/orthopedic	Negative
Family history	Negative
Nutrition	Natural diet
Medication	Negative
Gastrointestinal (GI): dysphagia/dyspepsia/constipation/hemorrhoid/fissure/abdominal pain/colitis	Negative
Blood biochemistry test FBS 95 mg/dL Blood Urea 25.0 mg/dL Creatinine 0.9 mg/dL Uric Acid 5.5 mg/dL Cholesterol 185.0 mg/dL Triglycerides 120.0 mg/dL HDL 50.0 mg/dL LDL 111 mg/dL S.G.O.T (AST) 35 IU/I S.G.P.T (ALT) 30 IU/I Alkaline phosphatase 220.0 IU/I	Normal

DNA extraction

A peripheral blood sample was obtained and DNA extraction was performed by the salting out protocol. The quality and quantity of the DNA were checked with an Eppendorf BioPhotometer.

Whole exome sequencing (WES) and data analysis

An Illumina Hiseq 2500 was used for whole exome sequencing, and an Agilent SureSelect Human Exome Kit (V4) was used for exome capture and enrichment. The mean depth of coverage for the CCSD exons in the sample was 173x and 95.3% of the CCD exons were covered at 20× or more in that sample. We used Burrows-Wheeler Aligner[15] (http://bio-bwa.sourceforge.net/) for aligning reads and the Genome Analysis Toolkit (GATK)[16,17] (https://www.broadinstitute.org/gatk/) for variant calling. Variant annotation was performed by applying ANNOVAR[18] (http://www.openbioinformatics.org/annovar/) and Variant Effect Predictor (VEP)[19] (http://www.ensembl.org/info/docs/tools/vep/index.html). Furthermore, for visualization of the variants, we used the Integrative Genomics Viewer (IGV) tool[20,21] (http://www.broadinstitute.org/igv/). Tertiary analysis was based on four different hypotheses about the genetic basis of longevity[14] which are described below:

The metabolic variant hypothesis

This hypothesis suggests that metabolic variants could be involved in longevity. We used the same chart as was gathered by Sebastiani, *et al*. in 2012[14] that comprises 27 coding single nucleotide polymorphism (SNPs) in which the exceptional longevity alleles were linked to exceptional longevity. The 27 SNPs are located on *FOXO3A*, *SIRT1*, *SIRT3*, *SIRT5*, *SIRT6*, *IGF1R*, *HSP70*, *CETP*, *PON1*, *MINPP1*, and *KLOTHO* genes.[14]

The lack of disease-associated variants hypothesis

This hypothesis suggests that the reduced number or lack of disease-associated variants can play a role in the phenotype of longevity. It is noticeable that this hypothesis has been rejected in several previous studies.[8,14,22] In order to test the second hypothesis, we compared our data with the Catalog of Published Genome-Wide Association Studies (NHGRI GWAS Catalog) 11-26-08 (http://www.genome.gov/gwastudies/), which contains 11,912 SNPs from 1751 publications.[23]

The rare variants hypothesis

This hypothesis suggests that some novel or rare variants within a small number of individuals are important for shaping the phenotype of longevity. To test this hypothesis, we gathered candidate aging and longevity genes from the Genage[24] and Longevity map[25] databases. These databases provide all candidate aging and longevity genes from GWAS studies among different populations. We compared the candidate genes with our data and then isolated novel variants among the genes. Based on the fact that the Iranian population is not included in global population studies such as the 1000 Genomes Project, we decided to compare the novel variants with an in-house exome sequencing database of 285 Iranian individuals to assess the rarity of the variants among our population.

The enrichment of longevity-associated variants hypothesis

This hypothesis suggests that long-lived individuals show this trait because of a high number of longevity-associated variants. For evaluation of this hypotheses, we used longevity-associated variants from all over the world, which were available in the Longevity map[25] database in order to find similar variants.

Results

In total, we found 22,657 coding variants based on RefSeq database Release 63. The results for each hypothesis are as follows:

The metabolic variant hypothesis

Among 27 metabolic variants, the subject had the exceptional longevity allele in 18 variants, of which 14 variants showed a homozygous genotype and 4 variants showed a heterozygous genotype. The details of the variants are shown in Table 2.

The lack of disease-associated variants hypothesis

In comparison with NHGRI GWAS Catalog 11-26-08, we found 58 trait-associated variants for 63 traits and a number of variants were associated to more than one trait. Among these variants, 11 variants were homozygous for the risk allele of the trait and 31 variants were heterozygous, but for 21 variants, the risk allele was not identified. Detailed information about these variants is included in Tables 3, 4, and 5.

The rare variants hypothesis

We found 17 novel variants within aging candidate genes and 13 novel variants within longevity candidate genes. It is noticeable that four variants are common between these two categories. Variants within longevity candidate genes and aging candidate genes are shown in Tables 6 and 7, respectively. To determine whether these variants are rare among the Iranian population, we compared them with our in-house database of 285 Iranian individuals. This comparison is shown in Table 8. We found out that the variant with variant ID 2:227659846 is not rare among our population.

The enrichment of longevity-associated variants hypothesis

In comparison with all of the previous studies about longevity, in our sample, we found just seven longevity-associated variants and two variants that were reported as non-significant variants in previous GWAS studies. These variants are shown in Table 9.

Discussion

The metabolic variant hypothesis

It is believed that variants within different metabolic pathways are involved in developing the phenotype of longevity. In 2012, Sebastiani, *et al.* collected 27 coding SNPs from the published literature which were previously linked to exceptional longevity.[26] Their data on two other supercentenarian cases revealed that 62.96% and 74.74% of these variants encompassed the exceptional longevity allele. Our case had the exceptional longevity allele in 66.66% of these variants; therefore, this study supports the metabolic variant hypothesis.

The lack of disease-associated variants hypothesis

The majority of supercentenarian cases experience healthy aging and they may not show age-related diseases until their last few years of life. This may originate from the reduced number or lack of disease-associated variants in their genome. Although the second hypothesis has been rejected in several previous stud-

Table 2. Metabolic variants.[14]

Gene	SNP	Exceptional longevity allele	Genotype of the case (I = 54462)
SIRT1	rs2273773	T	TT
SIRT3	rs28365927	G	GG
KLOTHO	rs2772364	C	CC
	rs9527026	G	AG
	rs564481	C	CC
	rs648202	C	CC
	rs649964	C	CC
IGF1R	rs35812156	C	CC
SIRT6	rs352493	T	TT
SIRT5	rs3757261	C	CC
PON1	rs854560	A	AA
FOXO3A	rs12206094	T	CC
	rs2764264	C	CC
	rs7762395	A	GG
	rs9400239	T	TT
	rs479744	T	GG
IGF1R	rs2229765	A	AA
	rs34516635	A	GG
	chr15:97068418	A	GG
	chr15:97272104	A	GG
HSP70	rs2227956	A	GG
CETP	rs5882	GG	AA
PON1	rs662	C	CT
MINPP1	rs9664222	C	AA
SIRT1	rs3758391	T	TT
KLOTHO	rs9536314	G	GT
	rs9527025	C	CG

Table 3. Twelve variants with a homozygous genotype for the risk allele.

References	Disease/Trait	Region	Mapped_gene	SNPs	Strongest SNP-Risk Allele	Genotype of the case
[27]	Activated partial thromboplastin time	1q24.2	*F5*	rs6028	rs6028-C	CC
[28]	Cardiovascular disease risk factors	6p22.2	*SLC17A4*	rs11754288	rs11754288-A	AA
[29]	C-reactive protein	1p31.3	*LEPR*	rs1805096	rs1805096-A	AA
[30]	C-reactive protein	6q22.1	*GPRC6A*	rs6901250	rs6901250-A	AA
[31]	Crohn's disease	1p13.2	*PTPN22;LOC101928822*	rs2476601	rs2476601-G	GG
[32]	Inflammatory biomarkers	2q13	*IL1F10*	rs6743376	rs6743376-A	AA
[33]	Lipoprotein-associated phospholipase A2 activity and mass	6p12.3	*PLA2G7*	rs1805017	rs1805017-T	TT
[34]	Lipoprotein-associated phospholipase A2 activity and mass	6p12.3	*PLA2G7*	rs1805017	rs1805017-T	TT
[35]	Obesity (early onset extreme)	9q22.31	*NINJ1*	rs2275848	rs2275848-T	TT
[36]	Obesity-related traits	4p12	*GABRB1*	rs6289	rs6289-G	GG
[37]	Pulmonary function	2q35	*TNS1*	rs2571445	rs2571445-G	GG
[38]	Type 2 diabetes	4p16.1	*WFS1*	rs1801214	rs1801214-T	TT

Table 4. Sixty one variants with a heterozygous genotype for the risk allele.

Reference	Disease/Trait	Region	Mapped_gene	SNPs	Strongest SNP-Risk Allele	Genotype of the case
[27]	Activated partial thromboplastin time	3q27.3	*KNG1*	rs710446	rs710446-C	CT
[39]	Ankylosing spondylitis	5q15	*ERAP1*	rs30187	rs30187-T	CT
[40]	Ankylosing spondylitis	5q15	*ERAP1*	rs27434	rs27434-A	AG
[41]	Bipolar disorder	3p21.1	*ITIH1*	rs1042779	rs1042779-A	AG
[42]	Blood pressure	3p22.1	*ULK4*	rs3774372	rs3774372-T	CT
[43]	Butyrylcholinesterase levels	3q26.1	*BCHE*	rs1803274	rs1803274-T	CT
[28]	Cardiovascular disease risk factors	3q26.1	*BCHE*	rs1803274	rs1803274-T	CT
[28]	Cardiovascular disease risk factors	2p23.3	*GCKR*	rs1260326	rs1260326-T	CT
[44]	Cholesterol, total	2p23.3	*GCKR*	rs1260326	rs1260326-T	CT
[45]	Chronic kidney disease	2p23.3	*GCKR*	rs1260326	rs1260326-T	CT
[30]	C-reactive protein	2p23.3	*GCKR*	rs1260326	rs1260326-T	CT
[31]	Crohn's disease	9q32	*TNFSF15*	rs3810936	rs3810936-C	CT
[46]	Electrocardiographic conduction measures	3p22.2	*SCN10A*	rs6795970	rs6795970-A	AG
[47]	Electrocardiographic traits	3p22.2	*SCN10A*	rs6795970	rs6795970-A	AG
[48]	Electrocardiographic traits	3p22.2	*SCN10A*	rs6795970	rs6795970-A	AG
[49]	Endometriosis	2p25.1	*GREB1*	rs13394619	rs13394619-G	AG
[50]	Glycemic traits (pregnancy)	2p23.3	*GCKR*	rs1260326	rs1260326-T	CT
[44]	HDL cholesterol	2q34	*CPS1*	rs1047891	rs1047891-A	AC
[51]	HDL cholesterol	8p21.3	*LPL*	rs328	rs328-G	CG
[52]	Height	1q21.2	*MTMR11*	rs11205303	rs11205303-C	CT
[53]	Hematological and biochemical traits	2p23.3	*GCKR*	rs1260326	rs1260326-C	CT
[54]	Hypertriglyceridemia	2p23.3	*GCKR*	rs1260326	rs1260326-T	CT
[55]	Interstitial lung disease	3q26.2	*LRRC34*	rs6793295	rs6793295-C	CT
[56]	Lipoprotein-associated phospholipase A2 activity and mass	2p23.3	*GCKR*	rs1260326	rs1260326-T	CT
[57]	Liver enzyme levels (gamma-glutamyl transferase)	2p23.3	*GCKR*	rs1260326	rs1260326-T	CT
[58]	Magnesium levels	1q22	*MUC1*	rs4072037	rs4072037-C	CT
[59]	Metabolite levels	2q34	*ACADL*	rs2286963	rs2286963-T	GT
[60]	Monocyte count	3p22.1	*ACKR2*	rs2228468	rs2228468-C	AC
[36]	Obesity-related traits	1q23.1	*KIRREL*	rs6427419	rs6427419-C	AC
[36]	Obesity-related traits	4q35.2	*KLKB1*	rs3733402	rs3733402-G	AG
[36]	Osteoarthritis	3p21.1	*GNL3;PBRM1;SNORD19*	rs11177	rs11177-A	AG
[61]	Periodontal microbiota	5q32	*FBXO38*	rs10043775	rs10043775-T	CT
[62]	Platelet counts	2p23.3	*GCKR*	rs1260326	rs1260326-T	CT
[63]	Primary tooth development (number of teeth)	7q32.1	*OPN1SW*	rs1799922	rs1799922-T	GT
[63]	Primary tooth development (time to first tooth eruption)	7q32.1	*OPN1SW*	rs1799922	rs1799922-T	GT
[64]	QT interval	1p36.31	*RNF207*	rs846111	rs846111-C	CG
[65]	QT interval	1p36.31	*RNF207*	rs846111	rs846111-C	CG

66	Schizophrenia	6p22.1	*POM121L2*	rs16897515	rs16897515-C	AC
67	Serum albumin level	2p23.3	*GCKR*	rs1260326	rs1260326-T	CT
67	Serum total protein level	2p23.3	*GCKR*	rs1260326	rs1260326-T	CT
68	Systemic lupus erythematosus	6q23.3	*TNFAIP3*	rs2230926	rs2230926-G	GT
44	Triglycerides	2p23.3	*GCKR*	rs1260326	rs1260326-T	CT
69	Triglycerides	2p23.3	*GCKR*	rs1260326	rs1260326-T	CT
53	Triglycerides	2p23.3	*GCKR*	rs1260326	rs1260326-C	CT
70	Triglycerides	2p23.3	*GCKR*	rs1260326	rs1260326-T	CT
51	Triglycerides	8p21.3	*LPL*	rs328	rs328-G	CG
71	Two-hour glucose challenge	2p23.3	*GCKR*	rs1260326	rs1260326-T	CT
72	Type 2 diabetes	8q24.11	*SLC30A8*	rs13266634	rs13266634-C	CT
73	Type 2 diabetes	6q13	*C6orf57*	rs1048886	rs1048886-G	AG
74	Type 2 diabetes	8q24.11	*SLC30A8*	rs13266634	rs13266634-C	CT
75	Type 2 diabetes	8q24.11	*SLC30A8*	rs13266634	rs13266634-C	CT
76	Type 2 diabetes	8q24.11	*SLC30A8*	rs13266634	rs13266634-C	CT
77	Type 2 diabetes	8q24.11	*SLC30A8*	rs13266634	rs13266634-C	CT
78	Type 2 diabetes	8q24.11	*SLC30A8*	rs13266634	rs13266634-C	CT
79	Type 2 diabetes	8q24.11	*SLC30A8*	rs13266634	rs13266634-C	CT
80	Type 2 diabetes and other traits	8q24.11	*SLC30A8*	rs13266634	rs13266634-C	CT
81	Upper aerodigestive tract cancers	4q23	*ADH1B*	rs1229984	rs1229984-T	CT
81	Upper aerodigestive tract cancers	4q21.23	*HELQ*	rs1494961	rs1494961-C	CT
82	Urate levels	2p23.3	*GCKR*	rs1260326	rs1260326-T	CT
83	Ventricular conduction	1p13.1	*CASQ2*	rs4074536	rs4074536-C	CT
84	Waist Circumference - Triglycerides (WC-TG)	2p23.3	*C2orf16*	rs1919128	rs1919128-A	AG

Table 5. Twenty Seven variants within which the risk allele was not recognized.

Reference	Disease/Trait	Region	Mapped_gene	SNPs	Strongest SNP-Risk Allele	Genotype of the case
85	Alcohol consumption	4q23	*ADH1B*	rs1229984	rs1229984-?	CT
86	Alcohol dependence	4q23	*ADH1B*	rs1229984	rs1229984-?	CT
87	Celiac disease	1p36.32	*MMEL1*	rs3748816	rs3748816-?	AG
88	Chemerin levels	6p21.2	*PI16*	rs1405069	rs1405069-?	AC
89	Cognitive performance	1q41	*FAM177B*	rs6683071	rs6683071-?	GG
89	Cognitive performance	2q31.1	*GORASP2*	rs4668356	rs4668356-?	TT
90	Coronary heart disease	2q32.1	*TFPI*	rs7586970	rs7586970-?	CT
91	Coronary heart disease	6q25.2	*OPRM1*	rs675026	rs675026-?	AG
92	Height	3q26.31	*FNDC3B*	rs7652177	rs7652177-?	GG
93	Hypertriglyceridemia	2p23.3	*GCKR*	rs1260326	rs1260326-?	CT
94	Iron status biomarkers	3q22.1	*TF*	rs1799852	rs1799852-?	CT
95	Lipid metabolism phenotypes	2p24.1	*APOB*	rs676210	rs676210-?	AA
95	Lipid metabolism phenotypes	2p23.3	*GCKR*	rs1260326	rs1260326-?	CT
96	Major mood disorders	3p21.1	*PBRM1*	rs2251219	rs2251219-?	CT
97	Metabolite levels	2p23.3	*GCKR*	rs1260326	rs1260326-?	CT
98	Metabolite levels	2p23.3	*GCKR*	rs1260326	rs1260326-?	CT
99	Myopia (pathological)	4p15.2	*DHX15*	rs6841898	rs6841898-?	CT
100	Panic disorder	3p26.1	*GRM7*	rs3749380	rs3749380-?	CT
101	Protein quantitative trait loci	2q13	*IL1F10*	rs6761276	rs6761276-?	CC
102	Pulmonary function (interaction)	2q35	*TNS1*	rs2571445	rs2571445-?	GG
103	Reasoning	3p22.1	*LYZL4*	rs2286720	rs2286720-?	AG
104	Rheumatoid arthritis	1p36.13	*PADI4*	rs2240335	rs2240335-?	AC
105	Rheumatoid arthritis	1p36.13	*PADI4*	rs2240335	rs2240335-?	AC
106	Stroke	6q21	*AIM1*	rs783396	rs783396-?	AC
107	Triglycerides	8p21.3	*LPL*	rs328	rs328-?	CG
108	Type 2 diabetes	8q24.11	*SLC30A8*	rs13266634	rs13266634-?	CT
109	Waist circumference and related phenotypes	2p23.3	*GCKR*	rs1260326	rs1260326-?	CT

Table 6. Novel variants in longevity candidate genes (a number of variants are in common with Table 7).

Variant ID	Reference	Alternates	Genotype	Amino acid changes	Classification	Gene	Transcript	Prediction by Mutation Taster
1:237024416	C	T	C_T		Splicing	*MTR*	NM_000254	Polymorphism
2:31570510	T	C	C_T	p.Ser1052Gly	Nonsyn SNV	*XDH*	NM_000379	Disease causing
2:227659846	T	G	G_T	p.Pro1203Pro	Synonymous	*IRS1*	NM_005544	Polymorphism
4:141481144	C	T	C_T	p.Arg277Gln	Nonsyn SNV	*UCP1*	NM_021833	Disease causing
4:155488788	G	C	C_G	p.Lys178Asn	Nonsyn SNV	*FGB*	NM_001184741	Polymorphism
6:160499237	C	T	C_T	p.Thr1774Met	Nonsyn SNV	*IGF2R*	NM_000876	Polymorphism
11:102713159	T	A	A_T		Splicing	*MMP3*	NM_002422	Disease causing
13:99116004	C	A	A_C	p.Arg290Ser	Nonsyn SNV	*STK24*	NM_003576	Disease causing
15:58830608	T	G	G_T	p.Phe55Leu	Nonsyn SNV	*LIPC*	NM_000236	Polymorphism
17:28538344	C	A	A_C	p.Gly435Cys	Nonsyn SNV	*SLC6A4*	NM_001045	Disease causing
19:4179130	G	A	A_G	p.Asp116Asp	Synonymous	*SIRT6*	NM_001193285	Disease causing
19:45860787	G	A	A_G	p.Ser441Leu	Nonsyn SNV	*ERCC2*	NM_000400	Disease causing
22:42046844	C	A	A_C	p.His360Asn	Nonsyn SNV	*XRCC6*	NM_001469	Disease causing

Table 7. Novel variants in aging candidate genes (a number of variants are in common with Table 6).

Variant ID	Reference	Alternates	Genotype	Amino acid changes	Classification	Gene 1	Transcript 1	Prediction by Mutation Taster
2:170145602	A	G	A_G	p.Tyr326His	Nonsyn SNV	*LRP2*	NM_004525	Polymorphism
2:217498296	T	C	C_T	p.Leu17Pro	Nonsyn SNV	*IGFBP2*	NM_000597	Disease causing
2:227659846	T	G	G_T	p.Pro1203Pro	Synonymous	*IRS1*	NM_005544	Polymorphism
4:103517357	A	C	A_C	p.Thr455Pro	Nonsyn SNV	*NFKB1*	NM_001165412	Polymorphism
4:103531829	G	A	A_G	p.Thr774Thr	Synonymous	*NFKB1*	NM_001165412	Disease causing
4:141481144	C	T	C_T	p.Arg277Gln	Nonsyn SNV	*UCP1*	NM_021833	Disease causing
5:137803529	C	A	A_C	p.Thr464Asn	Nonsyn SNV	*EGR1*	NM_001964	Disease causing
6:10410466	T	G	G_T	p.Asn50His	Nonsyn SNV	*TFAP2A*	NM_001032280	Disease causing
6:170871043	G	A	A_G	p.Gln53Gln	Synonymous	*TBP*	NM_001172085	Disease causing
6:170871079	G	A	A_G	p.Gln65Gln	Synonymous	*TBP*	NM_001172085	Disease causing
6:170871085	G	A	A_G	p.Gln67Gln	Synonymous	*TBP*	NM_001172085	Polymorphism
9:32986031	A	-	-_-		Splicing	*APTX*	NM_001195249	Disease causing
17:40364119	G	T	G_T	p.Ser521Arg	Nonsyn SNV	*STAT5B*	NM_012448	Disease causing
19:4179130	G	A	A_G	p.Asp116Asp	Synonymous	*SIRT6*	NM_001193285	Disease causing
19:45860787	G	A	A_G	p.Ser441Leu	Nonsyn SNV	*ERCC2*	NM_000400	Disease causing
22:42046844	C	A	A_C	p.His360Asn	Nonsyn SNV	*XRCC6*	NM_001469	Disease causing
X:70586354	T	G	G_T		Splicing	*TAF1*	NM_004606	Polymorphism

Table 8. Comparison of novel longevity and aging variants among an Iranian population.

Variant ID	Chromosome	Position	Reference	Alternates	Genotype	Amino acid changes	Classification	Gene 1	Transcript 1	All samples in cohort | how many times is observed | in homozygote | in heterozygote
2:227659846	2	227659846	T	G	G_T	p.Pro1203Pro	Synonymous	*IRS1*	NM_005544	285|55|0|55
6:10410466	6	10410466	T	G	G_T	p.Asn46His	Nonsyn SNV	*TFAP2A*	NM_001032280	285|1|0|1
2:31570510	2	31570510	T	C	C_T	p.Ser1052Gly	Nonsyn SNV	*XDH*	NM_000379	285|1|0|1
4:155488788	4	155488788	G	C	C_G	p.Lys178Asn	Nonsyn SNV	*FGB*	NM_001184741	285|3|0|3
6:160499237	6	160499237	C	T	C_T	p.Thr1774Met	Nonsyn SNV	*IGF2R*	NM_000876	285|1|0|1
15:58830608	15	58830608	T	G	G_T	p.Phe55Leu	Nonsyn SNV	*LIPC*	NM_000236	285|1|0|1

Table 9. Seven longevity-associated variants and two variants that were reported as non-significant variants in previous GWAS.

rs number	Gene	Association	Population	Reference
rs1042713	*ADRB2*	Non-significant	American	[110]
		Non-significant	Chinese	[111]
rs1042714	*ADRB2*	Significant	American	[110]
		Non-significant	Chinese	[111]
rs4880	*SOD2*	Significant	Danish	[112]
rs1815739	*ACTN3*	Non-significant	Spanish	[113]
rs651922	*DCPS*	Significant	American	[1]
rs1280396	*CGNL1*	Significant	American	[112]
rs3024239	*WRN*	Significant	American	[113]
rs2229765	*IGF1R*	Significant	Italian	[14]
		Significant	Italian	[14]
		Non-significant	Italian (southern)	[14]
rs1042522	*TP53*	Significant	Italian	[114]
		Non-significant	Italian	[115]

ies,[8,14,22] in our study we decided to compare the sample's data with the NHGRI GWAS Catalog to provide a context for further investigations. In comparison with the NHGRI GWAS Catalog, it was revealed that our case had a homozygous genotype for the risk allele for serious medical conditions such as cardiovascular diseases, Type 2 diabetes, and a heterozygous genotype of the risk allele for conditions such as high blood pressure, cardiovascular diseases, and hypertriglyceridemia while he was healthy. Hence, these variants did not cause any significant clinical manifestations, which would be expected from variants identified through GWAS, and have a very small effect in general.

The rare variants hypothesis

Longevity is a rare and complex trait and the third hypothesis suggests that some novel or rare variants within a small number of individuals are important for shaping this phenotype. We decided to check aging and longevity candidate genes from the Genage[24] and Longevity map[25] databases and we found 26 novel variants within these genes. After investigation among the Iranian population, 25 rare variants remained. It is clear that more supercentenarian studies are needed in order to test this hypothesis and similar rare variants within these genes may shed light on new longevity candidate variants.

The enrichment of longevity-associated variants hypothesis

The fourth hypothesis suggests that long-lived individuals may

have a high number of longevity-associated variants. In comparison with previous GWAS studies, our case showed only seven of these variants. Even though, longevity-associated variants appear to play a role in the phenotype of longevity, it does not seem necessary to have a high number of these variants to develop this trait. The rs1042713 in *ADRB2* was described as a non-significant variant in two different studies among American and Chinese populations, but our case showed this variant. The rs2229765 in *IGF1R*, which was studied among the Italian population in three different studies, was found significant in two investigations but non-significant in another one, and our case showed rs2229765. So it seems that sequencing the genome of more supercentenarians can strengthen the accuracy of future GWAS.

This study supports the metabolic variant hypothesis and has a clue for rare variant hypothesis but it suggests that it is not necessary to have a high number of longevity associated variants for the phenotype of longevity.

Despite an increase in the number of old people in Iran, we did not find any genetic studies for longevity in our population; therefore, we recommend further investigation to provide more genetic information on the elderly society in Iran.

The advent of whole exome sequencing has opened a new insight to discover more about the secrets of healthy living and also longevity but we should acknowledge that this method cannot capture intronic variants and epigenetic alteration that play an important role in the process of longevity. Our study was performed on only one sample and the results cannot be interpreted as a generalized principle for other elderly societies. However, this is the first step in the investigation of the genetic basis of longevity in Iran and provides an insight for further studies in the field of longevity. These studies can increase our knowledge on the basic biological pathways interfering with life and extend our capability to gain a better understanding of life and age-related diseases.

Acknowledgments

We gratefully thank the participant and his family for proper cooperation in this project. This study was carried out as a Master's degree thesis in Genetic Research Center, University of Social Welfare and Rehabilitation Sciences, Tehran, Iran.

References

1. Sebastiani P, Solovieff N, Dewan AT, Walsh KM, Puca A, Hartley SW, et al. Genetic signatures of exceptional longevity in humans. *PLoS One.* 2012; 7(1): e29848.
2. Young RD, Desjardins B, McLaughlin K, Poulain M, Perls TT. Typologies of extreme longevity myths. *Curr Gerontol Geriatr Res.* 2010; 2010: 423087.
3. Balistreri CR, Candore G, Accardi G, Bova M, Buffa S, Bulati M, et al. Genetics of longevity. data from the studies on Sicilian centenarians. *Immun Ageing.* 2012; 9(1): 8.
4. Herskind AM, McGue M, Holm NV, Sorensen TI, Harvald B, Vaupel JW. The heritability of human longevity: a population-based study of 2872 Danish twin pairs born 1870-1900. *Hum Genet.* 1996; 97(3): 319-23.
5. Perls TT, Wilmoth J, Levenson R, Drinkwater M, Cohen M, Bogan H, et al. Life-long sustained mortality advantage of siblings of centenarians. *Proc Natl Acad Sci U S A.* 2002; 99(12): 8442-7.
6. Atzmon G, Rincon M, Rabizadeh P, Barzilai N. Biological evidence for inheritance of exceptional longevity. *Mech Ageing Dev.* 2005; 126(2): 341-5.
7. Barzilai N, Atzmon G, Schechter C, Schaefer EJ, Cupples AL, Lipton R, et al. Unique lipoprotein phenotype and genotype associated with exceptional longevity. *JAMA.* 2003; 290(15): 2030-40.
8. Brooks-Wilson AR. Genetics of healthy aging and longevity. *Hum Genet.* 2013; 132(12): 1323-38.
9. Soerensen M. Genetic variation and human longevity. *Dan Med J.* 2012; 59(5): B4454.
10. Vaupel JW, Carey JR, Christensen K, Johnson TE, Yashin AI, Holm NV, et al. Biodemographic trajectories of longevity. *Science.* 1998; 280(5365): 855-60.
11. Ferrario A, Villa F, Malovini A, Araniti F, Puca AA. The application of genetics approaches to the study of exceptional longevity in humans: potential and limitations. *Immun Ageing.* 2012; 9(1): 7.
12. Perls TT. The different paths to 100. *Am J Clin Nutr* 2006; 83(2): 484s-7s.
13. Masnavi A, Kamrani AAA, Hosseini SA. Comparison of nursing home residents and family residing in Iran. *Iran J Ageing.* 2007; 2(2): 280-5.
14. Sebastiani P, Riva A, Montano M, Pham P, Torkamani A, Scherba E, et al. Whole genome sequences of a male and female supercentenarian, ages greater than 114 years. *Front Genet.* 2011; 2: 90.
15. Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* 2010; 26(5): 589-95.
16. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 2010; 20(9): 1297-303.
17. DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet.* 2011; 43(5): 491-8.
18. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* 2010; 38(16): e164.
19. McLaren W, Pritchard B, Rios D, Chen Y, Flicek P, Cunningham F. Deriving the consequences of genomic variants with the Ensembl API and SNP Effect Predictor. *Bioinformatics.* 2010; 26(16): 2069-70.
20. Thorvaldsdottir H, Robinson JT, Mesirov JP. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief Bioinform.* 2013; 14(2): 178-92.
21. Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G, et al. Integrative genomics viewer. *Nat Biotech.* 2011; 29(1): 24-6.
22. Beekman M, Nederstigt C, Suchiman HE, Kremer D, van der Breggen R, Lakenberg N, et al. Genome-wide association study (GWAS)-identified disease risk alleles do not compromise human longevity. *Proc Natl Acad Sci U S A.* 2010; 107(42): 18046-9.
23. Welter D, MacArthur J, Morales J, Burdett T, Hall P, Junkins H, et al. The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. *Nucleic Acids Res.* 2014; 42(Database issue): D1001-6.
24. de Magalhaes JP, Toussaint O. GenAge: a genomic and proteomic network map of human ageing. *FEBS Lett.* 2004; 571(1-3): 243-7.
25. Budovsky A, Craig T, Wang J, Tacutu R, Csordas A, Lourenco J, et al. LongevityMap: a database of human genetic variants associated with longevity. *Trends Genet.* 2013; 29(10): 559-60.
26. Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature.* 2007; 447(7145): 661-78.
27. Tang W, Schwienbacher C, Lopez LM, Ben-Shlomo Y, Oudot-Mellakh T, Johnson AD, et al. Genetic associations for activated partial thromboplastin time and prothrombin time, their gene expression profiles, and risk of coronary artery disease. *Am J Hum Genet.* 2012; 91(1): 152-62.
28. Middelberg RP, Ferreira MA, Henders AK, Heath AC, Madden PA, Montgomery GW, et al. Genetic variants in LPL, OASL and TOMM40/APOE-C1-C2-C4 genes are associated with multiple cardiovascular-related traits. *BMC Med Genet.* 2011; 12: 123.
29. Reiner AP, Beleza S, Franceschini N, Auer PL, Robinson JG, Kooperberg C, et al. Genome-wide association and population genetic analysis of C-reactive protein in African American and Hispanic American women. *Am J Hum Genet.* 2012; 91(3): 502-12.
30. Dehghan A, Dupuis J, Barbalic M, Bis JC, Eiriksdottir G, Lu C, et al. Meta-analysis of genome-wide association studies in >80 000 subjects identifies multiple loci for C-reactive protein levels. *Circulation.* 2011; 123(7): 731-8.
31. Franke A, McGovern DP, Barrett JC, Wang K, Radford-Smith GL, Ahmad T, et al. Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci. *Nat Genet.* 2010; 42(12): 1118-25.

32. Matteini AM, Li J, Lange EM, Tanaka T, Lange LA, Tracy RP, et al. Novel gene variants predict serum levels of the cytokines IL-18 and IL-1ra in older adults. *Cytokine.* 2014; 65(1): 10-6.
33. Grallert H, Dupuis J, Bis JC, Dehghan A, Barbalic M, Baumert J, et al. Eight genetic loci associated with variation in lipoprotein-associated phospholipase A2 mass and activity and coronary heart disease: meta-analysis of genome-wide association studies from five community-based studies. *Eur Heart J.* 2012; 33(2): 238-51.
34. Suchindran S, Rivedal D, Guyton JR, Milledge T, Gao X, Benjamin A, et al. Genome-wide association study of Lp-PLA(2) activity and mass in the Framingham Heart Study. *PLoS Genet.* 2010; 6(4): e1000928.
35. Wheeler E, Huang N, Bochukova EG, Keogh JM, Lindsay S, Garg S, et al. Genome-wide SNP and CNV analysis identifies common and low-frequency variants associated with severe early-onset obesity. *Nat Genet.* 2013; 45(5): 513-7.
36. Comuzzie AG, Cole SA, Laston SL, Voruganti VS, Haack K, Gibbs RA, et al. Novel genetic loci identified for the pathophysiology of childhood obesity in the Hispanic population. *PLoS One.* 2012; 7(12): e51954.
37. Repapi E, Sayers I, Wain LV, Burton PR, Johnson T, Obeidat M, et al. Genome-wide association study identifies five loci associated with lung function. *Nat Genet.* 2010; 42(1): 36-44.
38. Voight BF, Scott LJ, Steinthorsdottir V, Morris AP, Dina C, Welch RP, et al. Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. *Nat Genet.* 2010; 42(7): 579-89.
39. Evans DM, Spencer CC, Pointon JJ, Su Z, Harvey D, Kochan G, et al. Interaction between ERAP1 and HLA-B27 in ankylosing spondylitis implicates peptide handling in the mechanism for HLA-B27 in disease susceptibility. *Nat Genet.* 2011; 43(8): 761-7.
40. Reveille JD, Sims AM, Danoy P, Evans DM, Leo P, Pointon JJ, et al. Genome-wide association study of ankylosing spondylitis identifies non-MHC susceptibility loci. *Nat Genet.* 2010; 42(2): 123-7.
41. Scott LJ, Muglia P, Kong XQ, Guan W, Flickinger M, Upmanyu R, et al. Genome-wide association and meta-analysis of bipolar disorder in individuals of European ancestry. *Proc Natl Acad Sci U S A.* 2009; 106(18): 7501-6.
42. Wain LV, Verwoert GC, O'Reilly PF, Shi G, Johnson T, Johnson AD, et al. Genome-wide association study identifies six new loci influencing pulse pressure and mean arterial pressure. *Nat Genet.* 2011; 43(10): 1005-11.
43. Benyamin B, Middelberg RP, Lind PA, Valle AM, Gordon S, Nyholt DR, et al. GWAS of butyrylcholinesterase activity identifies four novel loci, independent effects within BCHE and secondary associations with metabolic risk factors. *Hum Mol Genet.* 2011; 20(22): 4504-14.
44. Willer CJ, Schmidt EM, Sengupta S, Peloso GM, Gustafsson S, Kanoni S, et al. Discovery and refinement of loci associated with lipid levels. *Nat Genet.* 2013; 45(11): 1274-83.
45. Kottgen A, Pattaro C, Boger CA, Fuchsberger C, Olden M, Glazer NL, et al. New loci associated with kidney function and chronic kidney disease. *Nat Genet.* 2010; 42(5): 376-84.
46. Ritchie MD, Denny JC, Zuvich RL, Crawford DC, Schildcrout JS, Bastarache L, et al. Genome- and phenome-wide analyses of cardiac conduction identifies markers of arrhythmia risk. *Circulation.* 2013; 127(13): 1377-85.
47. Chambers JC, Zhao J, Terracciano CM, Bezzina CR, Zhang W, Kaba R, et al. Genetic variation in SCN10A influences cardiac conduction. *Nat Genet.* 2010; 42(2): 149-52.
48. Holm H, Gudbjartsson DF, Arnar DO, Thorleifsson G, Thorgeirsson G, Stefansdottir H, et al. Several common variants modulate heart rate, PR interval and QRS duration. *Nat Genet.* 2010; 42(2): 117-22.
49. Nyholt DR, Low SK, Anderson CA, Painter JN, Uno S, Morris AP, et al. Genome-wide association meta-analysis identifies new endometriosis risk loci. *Nat Genet.* 2012; 44(12): 1355-9.
50. Hayes MG, Urbanek M, Hivert MF, Armstrong LL, Morrison J, Guo C, et al. Identification of HKDC1 and BACE2 as genes influencing glycemic traits during pregnancy through genome-wide association studies. *Diabetes.* 2013; 62(9): 3282-91.
51. Kathiresan S, Melander O, Guiducci C, Surti A, Burtt NP, Rieder MJ, et al. Six new loci associated with blood low-density lipoprotein cholesterol, high-density lipoprotein cholesterol or triglycerides in humans. *Nat Genet.* 2008; 40(2): 189-97.
52. Berndt SI, Gustafsson S, Magi R, Ganna A, Wheeler E, Feitosa MF, et al. Genome-wide meta-analysis identifies 11 new loci for anthropometric traits and provides insights into genetic architecture. *Nat Genet.* 2013; 45(5): 501-12.
53. Kamatani Y, Matsuda K, Okada Y, Kubo M, Hosono N, Daigo Y, et al. Genome-wide association study of hematological and biochemical traits in a Japanese population. *Nat Genet.* 2010; 42(3): 210-5.
54. Johansen CT, Wang J, Lanktree MB, Cao H, McIntyre AD, Ban MR, et al. Excess of rare variants in genes identified by genome-wide association study of hypertriglyceridemia. *Nat Genet.* 2010; 42(8): 684-7.
55. Fingerlin TE, Murphy E, Zhang W, Peljto AL, Brown KK, Steele MP, et al. Genome-wide association study identifies multiple susceptibility loci for pulmonary fibrosis. *Nat Genet.* 2013; 45(6): 613-20.
56. Chu AY, Guilianini F, Grallert H, Dupuis J, Ballantyne CM, Barratt BJ, et al. Genome-wide association study evaluating lipoprotein-associated phospholipase A2 mass and activity at baseline and after rosuvastatin therapy. *Circ Cardiovasc Genet.* 2012; 5(6): 676-85.
57. Chambers JC, Zhang W, Sehmi J, Li X, Wass MN, Van der Harst P, et al. Genome-wide association study identifies loci influencing concentrations of liver enzymes in plasma. *Nat Genet.* 2011; 43(11): 1131-8.
58. Meyer TE, Verwoert GC, Hwang SJ, Glazer NL, Smith AV, van Rooij FJ, et al. Genome-wide association studies of serum magnesium, potassium, and sodium concentrations identify six Loci influencing serum magnesium levels. *PLoS Genet.* 2010; 6(8).
59. Illig T, Gieger C, Zhai G, Romisch-Margl W, Wang-Sattler R, Prehn C, et al. A genome-wide perspective of genetic variation in human metabolism. *Nat Genet.* 2010; 42(2): 137-41.
60. Crosslin DR, McDavid A, Weston N, Zheng X, Hart E, de Andrade M, et al. Genetic variation associated with circulating monocyte count in the eMERGE Network. *Hum Mol Genet.* 2013; 22(10): 2119-27.
61. Divaris K, Monda KL, North KE, Olshan AF, Lange EM, Moss K, et al. Genome-wide association study of periodontal pathogen colonization. *J Dent Res.* 2012; 91(7 Suppl): 21s-8s.
62. Gieger C, Radhakrishnan A, Cvejic A, Tang W, Porcu E, Pistis G, et al. New gene functions in megakaryopoiesis and platelet formation. *Nature.* 2011; 480(7376): 201-8.
63. Fatemifar G, Hoggart CJ, Paternoster L, Kemp JP, Prokopenko I, Horikoshi M, et al. Genome-wide association study of primary tooth eruption identifies pleiotropic loci associated with height and craniofacial distances. *Hum Mol Genet.* 2013; 22(18): 3807-17.
64. Newton-Cheh C, Eijgelsheim M, Rice KM, de Bakker PI, Yin X, Estrada K, et al. Common variants at ten loci influence QT interval duration in the QTGEN Study. *Nat Genet.* 2009; 41(4): 399-406.
65. Pfeufer A, Sanna S, Arking DE, Muller M, Gateva V, Fuchsberger C, et al. Common variants at ten loci modulate the QT interval duration in the QTSCD Study. *Nat Genet.* 2009; 41(4): 407-14.
66. Aberg KA, Liu Y, Bukszar J, McClay JL, Khachane AN, Andreassen OA, et al. A comprehensive family-based replication study of schizophrenia genes. *JAMA Psychiatry.* 2013; 70(6): 573-81.
67. Franceschini N, van Rooij FJ, Prins BP, Feitosa MF, Karakas M, Eckfeldt JH, et al. Discovery and fine mapping of serum protein loci through transethnic meta-analysis. *Am J Hum Genet.* 2012; 91(4): 744-53.
68. Yang W, Tang H, Zhang Y, Tang X, Zhang J, Sun L, et al. Meta-analysis followed by replication identifies loci in or near CDKN1B, TET3, CD80, DRAM1, and ARID5B as associated with systemic lupus erythematosus in Asians. *Am J Hum Genet.* 2013; 92(1): 41-51.
69. Teslovich TM, Musunuru K, Smith AV, Edmondson AC, Stylianou IM, Koseki M, et al. Biological, clinical and population relevance of 95 loci for blood lipids. *Nature.* 2010; 466(7307): 707-13.
70. Kathiresan S, Willer CJ, Peloso GM, Demissie S, Musunuru K, Schadt EE, et al. Common variants at 30 loci contribute to polygenic dyslipidemia. *Nat Genet.* 2009; 41(1): 56-65.
71. Saxena R, Hivert MF, Langenberg C, Tanaka T, Pankow JS, Vollenweider P, et al. Genetic variation in GIPR influences the glucose and insulin responses to an oral glucose challenge. *Nat Genet.* 2010; 42(2): 142-8.
72. Hara K, Fujita H, Johnson TA, Yamauchi T, Yasuda K, Horikoshi M, et al. Genome-wide association study identifies three novel loci for type 2 diabetes. *Hum Mol Genet.* 2014; 23(1): 239-46.
73. Sim X, Ong RT, Suo C, Tay WT, Liu J, Ng DP, et al. Transferability of type 2 diabetes implicated loci in multi-ethnic cohorts from Southeast Asia. *PLoS Genet.* 2011; 7(4): e1001363.
74. Takeuchi F, Serizawa M, Yamamoto K, Fujisawa T, Nakashima E, Ohnaka K, et al. Confirmation of multiple risk Loci and genetic impacts by a genome-wide association study of type 2 diabetes in the Japanese population. *Diabetes.* 2009; 58(7): 1690-9.
75. Saxena R, Voight BF, Lyssenko V, Burtt NP, de Bakker PI, Chen H, et al. Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science.* 2007; 316(5829): 1331-6.
76. Scott LJ, Mohlke KL, Bonnycastle LL, Willer CJ, Li Y, Duren WL,

et al. A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. *Science.* 2007; 316(5829): 1341-5.
77. Steinthorsdottir V, Thorleifsson G, Reynisdottir I, Benediktsson R, Jonsdottir T, Walters GB, et al. A variant in CDKAL1 influences insulin response and risk of type 2 diabetes. *Nat Genet.* 2007; 39(6): 770-5.
78. Zeggini E, Weedon MN, Lindgren CM, Frayling TM, Elliott KS, Lango H, et al. Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. *Science.* 2007; 316(5829): 1336-41.
79. Sladek R, Rocheleau G, Rung J, Dina C, Shen L, Serre D, et al. A genome-wide association study identifies novel risk loci for type 2 diabetes. *Nature.* 2007; 445(7130): 881-5.
80. Rung J, Cauchi S, Albrechtsen A, Shen L, Rocheleau G, Cavalcanti-Proenca C, et al. Genetic variant near IRS1 is associated with type 2 diabetes, insulin resistance and hyperinsulinemia. *Nat Genet.* 2009; 41(10): 1110-5.
81. McKay JD, Truong T, Gaborieau V, Chabrier A, Chuang SC, Byrnes G, et al. A genome-wide association study of upper aerodigestive tract cancers conducted within the INHANCE consortium. *PLoS Genet.* 2011; 7(3): e1001333.
82. Kottgen A, Albrecht E, Teumer A, Vitart V, Krumsiek J, Hundertmark C, et al. Genome-wide association analyses identify 18 new loci associated with serum urate concentrations. *Nat Genet.* 2013; 45(2): 145-54.
83. Sotoodehnia N, Isaacs A, de Bakker PI, Dorr M, Newton-Cheh C, Nolte IM, et al. Common variants in 22 loci are associated with QRS duration and cardiac ventricular conduction. *Nat Genet.* 2010; 42(12): 1068-76.
84. Kraja AT, Vaidya D, Pankow JS, Goodarzi MO, Assimes TL, Kullo IJ, et al. A bivariate genome-wide approach to metabolic syndrome: STAMPEED consortium. *Diabetes.* 2011; 60(4): 1329-39.
85. Kapoor M, Wang JC, Wetherill L, Le N, Bertelsen S, Hinrichs AL, et al. A meta-analysis of two genome-wide association studies to identify novel loci for maximum number of alcoholic drinks. *Hum Genet.* 2013; 132(10): 1141-51.
86. Park BL, Kim JW, Cheong HS, Kim LH, Lee BC, Seo CH, et al. Extended genetic effects of ADH cluster genes on the risk of alcohol dependence: from GWAS to replication. *Hum Genet.* 2013; 132(6): 657-68.
87. Dubois PC, Trynka G, Franke L, Hunt KA, Romanos J, Curtotti A, et al. Multiple common variants for celiac disease influencing immune gene expression. *Nat Genet.* 2010; 42(4): 295-302.
88. Bozaoglu K, Curran JE, Stocker CJ, Zaibi MS, Segal D, Konstantopoulos N, et al. Chemerin, a novel adipokine in the regulation of angiogenesis. *J Clin Endocrinol Metab.* 2010; 95(5): 2476-85.
89. Need AC, Attix DK, McEvoy JM, Cirulli ET, Linney KL, Hunt P, et al. A genome-wide study of common SNPs and CNVs in cognitive performance in the CANTAB. *Hum Mol Genet.* 2009; 18(23): 4650-61.
90. A genome-wide association study in Europeans and South Asians identifies five new loci for coronary artery disease. *Nat Genet.* 2011; 43(4): 339-44.
91. Lettre G, Palmer CD, Young T, Ejebe KG, Allayee H, Benjamin EJ, et al. Genome-wide association study of coronary heart disease and its risk factors in 8,090 African Americans: the NHLBI CARe Project. *PLoS Genet.* 2011; 7(2): e1001300.
92. Lango Allen H, Estrada K, Lettre G, Berndt SI, Weedon MN, Rivadeneira F, et al. Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature.* 2010; 467(7317): 832-8.
93. Weissglas-Volkov D, Aguilar-Salinas CA, Nikkola E, Deere KA, Cruz-Bautista I, Arellano-Campos O, et al. Genomic study in Mexicans identifies a new locus for triglycerides and refines European lipid loci. *J Med Genet.* 2013; 50(5): 298-308.
94. Benyamin B, McRae AF, Zhu G, Gordon S, Henders AK, Palotie A, et al. Variants in TF and HFE explain approximately 40% of genetic variation in serum-transferrin levels. *Am J Hum Genet.* 2009; 84(1): 60-5.
95. Chasman DI, Pare G, Mora S, Hopewell JC, Peloso G, Clarke R, et al. Forty-three loci associated with plasma lipoprotein size, concentration, and cholesterol content in genome-wide analysis. *PLoS Genet.* 2009; 5(11): e1000730.
96. McMahon FJ, Akula N, Schulze TG, Muglia P, Tozzi F, Detera-Wadleigh SD, et al. Meta-analysis of genome-wide association data identifies a risk locus for major mood disorders on 3p21.1. *Nat Genet.* 2010; 42(2): 128-31.
97. Inouye M, Ripatti S, Kettunen J, Lyytikainen LP, Oksala N, Laurila PP, et al. Novel Loci for metabolic networks and multi-tissue expression studies reveal genes for atherosclerosis. *PLoS Genet.* 2012; 8(8): e1002907.
98. Kettunen J, Tukiainen T, Sarin AP, Ortega-Alonso A, Tikkanen E, Lyytikainen LP, et al. Genome-wide association study identifies multiple loci influencing human serum metabolite levels. *Nat Genet.* 2012; 44(3): 269-76.
99. Meng W, Butterworth J, Bradley DT, Hughes AE, Soler V, Calvas P, et al. A genome-wide association study provides evidence for association of chromosome 8p23 (MYP10) and 10q21.1 (MYP15) with high myopia in the French Population. *Invest Ophthalmol Vis Sci.* 2012; 53(13): 7983-8.
100. Otowa T, Yoshida E, Sugaya N, Yasuda S, Nishimura Y, Inoue K, et al. Genome-wide association study of panic disorder in the Japanese population. *J Hum Genet.* 2009; 54(2): 122-6.
101. Melzer D, Perry JR, Hernandez D, Corsi AM, Stevens K, Rafferty I, et al. A genome-wide association study identifies protein quantitative trait loci (pQTLs). *PLoS Genet.* 2008; 4(5): e1000072.
102. Hancock DB, Artigas MS, Gharib SA, Henry A, Manichaikul A, Ramasamy A, et al. Genome-wide joint meta-analysis of SNP and SNP-by-smoking interaction identifies novel loci for pulmonary function. *PLoS Genet.* 2012; 8(12): e1003098.
103. McClay JL, Adkins DE, Aberg K, Bukszar J, Khachane AN, Keefe RS, et al. Genome-wide pharmacogenomic study of neurocognition as an indicator of antipsychotic treatment response in schizophrenia. *Neuropsychopharmacology.* 2011; 36(3): 616-26.
104. Terao C, Yamada R, Ohmura K, Takahashi M, Kawaguchi T, Kochi Y, et al. The human AIRE gene at chromosome 21q22 is a genetic determinant for the predisposition to rheumatoid arthritis in Japanese population. *Hum Mol Genet.* 2011; 20(13): 2680-5.
105. Freudenberg J, Lee HS, Han BG, Shin HD, Kang YM, Sung YK, et al. Genome-wide association study of rheumatoid arthritis in Koreans: population-specific loci as well as overlap with European susceptibility loci. *Arthritis Rheum.* 2011; 63(4): 884-93.
106. Matarin M, Brown WM, Scholz S, Simon-Sanchez J, Fung HC, Hernandez D, et al. A genome-wide genotyping study in patients with ischaemic stroke: initial analysis and data release. *Lancet Neurol.* 2007; 6(5): 414-20.
107. Tan A, Sun J, Xia N, Qin X, Hu Y, Zhang S, et al. A genome-wide association and gene-environment interaction study for serum triglycerides levels in a healthy Chinese male population. *Hum Mol Genet.* 2012; 21(7): 1658-64.
108. Timpson NJ, Lindgren CM, Weedon MN, Randall J, Ouwehand WH, Strachan DP, et al. Adiposity-related heterogeneity in patterns of type 2 diabetes susceptibility observed in genome-wide association data. *Diabetes.* 2009; 58(2): 505-10.
109. Chambers JC, Elliott P, Zabaneh D, Zhang W, Li Y, Froguel P, et al. Common genetic variation near MC4R is associated with waist circumference and insulin resistance. *Nat Genet.* 2008; 40(6): 716-8.
110. Kulminski AM, Culminskaya I, Ukraintseva SV, Arbeev KG, Land KC, Yashin AI. Beta2-adrenergic receptor gene polymorphisms as systemic determinants of healthy aging in an evolutionary context. *Mech Ageing Dev.* 2010; 131(5): 338-45.
111. Zhao L, Yang F, Xu K, Cao H, Zheng GY, Zhang Y, et al. Common genetic variants of the beta2-adrenergic receptor affect its translational efficiency and are associated with human longevity. *Aging Cell.* 2012; 11(6): 1094-101.
112. Soerensen M, Christensen K, Stevnsner T, Christiansen L. The Mn-superoxide dismutase single nucleotide polymorphism rs4880 and the glutathione peroxidase 1 single nucleotide polymorphism rs1050450 are associated with aging and longevity in the oldest old. *Mech Ageing Dev.* 2009; 130(5): 308-14.
113. Fiuza-Luces C, Ruiz JR, Rodriguez-Romo G, Santiago C, Gomez-Gallego F, Yvert T, et al. Are 'endurance' alleles 'survival' alleles? Insights from the ACTN3 R577X polymorphism. *PLoS One.* 2011; 6(3): e17558.
114. Barbieri M, Boccardi V, Esposito A, Papa M, Vestini F, Rizzo MR, et al. A/ASP/VAL allele combination of IGF1R, IRS2, and UCP2 genes is associated with better metabolic profile, preserved energy expenditure parameters, and low mortality rate in longevity. *Age (Dordr).* 2012; 34(1): 235-45.
115. Albani D, Mazzuco S, Polito L, Batelli S, Biella G, Ongaro F, et al. Insulin-like growth factor 1 receptor polymorphism rs2229765 and circulating interleukin-6 level affect male longevity in a population-based prospective study (Treviso Longeva--TRELONG). *Aging Male.* 2011; 14(4): 257-64.