Supplementary Materials

Summary

SUPPLEMENTARY NOTE

Description of replication cohorts

INGI-Val Borbera (INGI-VB): The INGI-Val Borbera population is a collection of 1,785 genotyped samples collected in the Val Borbera Valley, a geographically isolated valley located within the Appennine Mountains in Northwest Italy (Traglia et al.). The valley is inhabited by about 3,000 descendants from the original population, living in 7 villages along the valley and in the mountains. Participants were healthy people 18-102 years of age that had at least one grandfather living in the valley. A standard battery of tests were performed by the laboratory of ASL 22 - Novi Ligure (AL) on sera from fasting blood collected in the morning. The project was approved by the Ethical Committee of the San Raffaele Hospital and of the Piemonte Region. All participants signed an informed consent. hsCRP was measured using turbidimetric analysis in a Beckman Coulter SYNCHRON.

INGI-FVG: The INGI Friuli Venezia Giulia (FVG) cohort comprises about 1,700 samples from six isolated villages covering a total area of 7858 km^2 in a hilly part of Friuli-Venezia Giulia (FVG) county located in North-Eastern Italy (Esko et al.). Genotyping and phenotypic data for 1,590 samples are available. Participants were randomly selected people 3-92 years of age. People with age < 18 were excluded from analyses. Ethics approval was obtained from the Ethics Committee of the Burlo Garofolo children hospital in Trieste. Written informed consent was obtained from every participant to the study.

HELIC MANOLIS (HA): The HELIC (Hellenic Isolated Cohorts; www.helic.org) MANOLIS (Minoan Isolates) collection focuses on Anogia and surrounding Mylopotamos villages. Recruitment of this population-based sample was primarily carried out at the village medical centres. All individuals were older than 17 years and had to have at least one parent from the Mylopotamos area. The study includes biological sample collection for DNA extraction and lab-based blood measurements, and interview-based questionnaire filling. The phenotypes collected include anthropometric and biometric measurements, clinical evaluation data, biochemical and haematological profiles, self-reported medical history, demographic, socioeconomic and lifestyle information. The study was approved by the Harokopio University Bioethics Committee and informed consent was obtained from every participant. hsCRP was measured using an immunoturbidimetric assay on a COBAS 8000 analyser (Roche).

HELIC Pomak (HP). The HELIC (Hellenic Isolated Cohorts; www.helic.org) Pomak collection focuses on the Pomak villages, a set of isolated mountainous villages in the North of Greece. Recruitment of this population-based sample was primarily carried out at the village medical centres. The study includes biological sample collection for DNA extraction and lab-based blood measurements, and interview-based questionnaire filling. The phenotypes collected include anthropometric and biometric measurements, clinical evaluation data, biochemical and haematological profiles, self-reported medical history, demographic, socioeconomic and lifestyle information. The study was approved by the Harokopio University Bioethics Committee and informed consent was obtained from every participant. hsCRP measurements were performed as in HA

Nature Genetics: doi:10.1038/ng.3368

TwinsUK: TwinsUK is a cohort of identical and non-identical twins living in the United Kingdom. Samples used here are those sequenced within the UK10K consortium. Only one individual for each pair of identical twins was used in the analysis.

ALSPAC: ALSPAC is a geographically based UK cohort that recruited pregnant women residing in Avon (Southwest England) with an expected date of delivery between April 1, 1991, and December 31, 1992. A total of 15,247 pregnancies were enrolled, with 14,775 children born (see www.alspac.bris.ac.uk.). Samples used here are those sequenced within the UK10K consortium. Blood samples were collected from nonfasting participants and were immediately spun and frozen at −80°C. Inflammatory markers were assayed in 2008 after a median of 7.5 years in storage with no previous freeze-thaw cycles during this period. Interleukin 6 was measured by enzyme-linked immunosorbent assay (R&D Systems), and high-sensitivity CRP was measured by automated particle-enhanced immunoturbidimetric assay (Roche) (eMethods in the Supplement). All interassay coefficients of variation were less than 5%.

INCIPE: For the INCIPE study, 6,200 randomly chosen individuals, all Caucasians and at least 40 years of age as of 1 January 2006, received a letter inviting them to participate in the study. A total of 3,870 subjects (62%) accepted and were enrolled. Two studies were included in the analysis:

1. INCIPE1: Individuals genotyped on HumanOmniExpress-12 v1.1 Illumina

2. INCIPE2: Individuals genotyped on HumanCoreExome-12v1

The ethics committees of the involved institutions approved the study protocol. Lipid measurements: Enzymatic determination of cholesterol and triglycerides was performed on Dimension RxL apparatus (Siemens Diagnostics). High sensitive serum C-reactive protein (HS-CRP) levels were measured using high-sensitivity immunonephelometry (Dade Behring, Marbung, Germany) according to the manufacturer's protocol

SUPPLEMENTARY TABLES

SUPPLEMENTARY TABLE 1. Quality assessment, through comparison with deep sequenced samples

To evaluate the quality of low pass sequencing results, we deep sequenced a nuclear family to average depth of >65-fold per individual (father 65-fold, mother 85-fold, child 82-fold). In the table below, we compare deep sequencing results with low pass results, stratified by allele frequency.

		Low pass results		
Grouping	Number of Variant Sites	Monomorphic Sites % (N sites)	Overall Genotype Discordance % (N genotypes)	Overall Heterozygote Discordance % (N genotypes)
All variants discovered by deep sequencing				
All variants.	4,105,003	0.24% (N = 10,302)	0.38% (N = 12,315,009)	0.71% (N = 5,558,176)
All variants discovered by deep sequencing, stratified by frequency among low pass samples				
Frequency < 0.5%	54,781	3.70% (N = 2,029)	4.32% (N = 164,343)	8.11% (N = 83,113)
Frequency 0.5 – 5.0%	306,449	0.69% (N = 2,138)	0.97% (N = 919,347)	1.55% (N = 467,188)
Frequency > 5%	3,743,773	0.14% (N = 5,586)	0.27% (N = 11,231,319)	0.51% (N = 5,007,875)

SUPPLEMENTARY TABLE 2. Quality assessment, through comparison with genotyping array results

Grouping	Number of Variants in the Genotyping Arrays	Number of Variants Also Found by Sequencing	Low pass results: Missed In Sequence Analysis % (N variants)	Low pass results: Overall Discordance % (N genotypes)	Low pass results: Heterozygote Discordance % (N genotypes)
All variants genotyped using arrays.					
All variants.	851,655	831,771	2.33% (N = 19,884)	0.22% (N = 886,640,361)	0.23% (N = 261,628,802)
All variants genotyped using arrays, stratified by frequency among low pass samples.					
Frequency 0.05% Singletons	9,196	3,832	58.33% (N = 5,364)	0.04% (N = 4,085,910)	24.92% (N = 4,473)
Frequency 0.1% Doubletons	10,044	5,851	41.75% (N = 4,193)	0.03% (N = 6,239,915)	6.19% (N = 11,074)
Frequency < 0.5% (including singletons and doubletons)	40,725	28,194	30.77% (N=12,531)	0.05% (N = 31,562,620)	4.31% (N = 134,683)
Frequency 0.5 – 5.0%	101,462	100,321	1.12% (N=1,141)	0.10% (N = 119,266,141)	0.51% (N = 5,940,571)
Frequency > 5%	709,468	703,256	0.88% (N=19,884)	0.25% (N = 735,811,600)	0.22% (N = 255,553,548)

SUPPLEMENTARY TABLE 3. Shared haplotype length around variants with different frequency

Length of shared haplotypes surrounding variants with increasing allele count (from 2 to 5) shared between one of our 2,120 sequenced individuals and one of 100 individuals from the Lanusei Valley, or one of our 2,120 sequenced individuals and 100 from the CSCT cohort. The subset of 100 samples were randomly selected for each group.

Alternative allele count	Shared with 100 Lanusei Valley samples	Shared with 100 CSCT samples
2	1,185 Kb	717 Kb
3	370 Kb	325 Kb
4	255 Kb	213 Kb
5	188 Kb	155 Kb

SUPPLEMENTARY TABLE 4. Descriptive statistics for the SardiNIA cohort

The table shows summary statistics for the 6,602 samples that were fully genotyped.

Characteristics		**Value**
N males/N females		2,823/3,779
Age - *mean (min-max)*		43.5 (14 - 101.3)
BMI - *mean (min-max)*		25.3 (13.9 - 53.3)
Smokers *(N)*		1247
LDL-c (mg/dl)	N males/N females	2,502/3,379
	mean (min-max)	125.5 (27.9 –293.3)
TG (mg/dl)	N males/N females	2,505/3,380
	mean (min-max)	87.7 (12.66 - 1608.1)
TC (mg/dl)	N males/N females	2,505/3,380
	mean (min-max)	208 (79.6 – 445.0)
HDL (mg/dl)	N males/N females	2,505/3,380
	mean (min-max)	64.2 (21.3 – 147.6)
ADPN (mg/ml)	N males/N females	2,486 / 3,350
	mean (min-max)	2.8 (0.2 - 46.4)
hsCRP (ng/ml)	N males/N females	2,411 / 3,219
	mean (min-max)	2.7 (0.1 - 119)

ESR (mm/h)	N males/N females	2,531 / 3,410
	mean (min-max)	10.6 (1 - 110)
MCP-1 (pg/ml)	N males/N females	2,497 / 3,347
	mean (min-max)	282.9 (2 - 6080)
IL-6 (pg/L)	N males/N females	2,492 / 3,346
	mean (min-max)	3.2 (0.1 - 41)

SUPPLEMENTARY TABLE 5. Reported loci associated with lipid levels

The table shows the association levels for the SNPs reported in (Teslovich et al, 2010 Nature. 2010 Aug 5;466(7307):707-13.) with pvalue between $5x10^{-5}$ and $5x10^{-8}$ in our study. Additionally, it reports association for the Sardinian specific variant rs72658864. Columns are defined as in **Table 2**.

Candidate Gene	Chr:position	rs name	Effect Allele / Other	Freq	Effect (StdErr)	pvalue	Variance Explained (%)	RSQR	Variant Consequence
LDL									
APOB	2:21384358	rs544450	T/C	0.178	-0.171(0.032)	8.71×10^{-8}	0.5	0.995	Intergenic
ABCG5/8	2:44072576	rs4299376	G/T	0.295	0.106(0.023)	2.74×10^{-6}	0.4	Genotyped	Intronic
OSBPL7	*17:4542511*	rs7206971	G/A	0.492	-0.08(0.021)	2.76×10^{-5}	0.3	0.999	Intronic
LDLR	19:11190873	rs73015013	T/C	0.137	-0.164(0.030)	6.35×10^{-8}	0.5	Genotyped	Intergenic
LDLR	19:11227562	rs72658864*	C/T	0.005	-0.633(0.150)	2.54×10^{-5}	0.5	Genotyped	Missense, V578A
TC									
SORT1	1:109818306	rs629301	G/T	0.15	-0.139(0.029)	2.48×10^{-6}	0.4	Genotyped	Intergenic
ABCG5/8	2:44072576	rs544450	T/C	0.295	0.097(0.023)	1.66×10^{-5}	0.3	Genotyped	Intronic
APOA1/5	11 116648917	rs964184	G/C	0.096	0.149(0.036)	3.59×10^{-5}	0.3	Genotyped	3’ UTR
LDLR	19:11202306	rs6511720	T/G	0.129	-0.135(0.031)	1.25×10^{-5}	0.3	Genotyped	Intronic
HDL									
LPL	8:19844222	rs12678919	G/A	0.139	0.158(0.034)	3.62×10^{-6}	0.4	Genotyped	Intergenic
TTC39B	9:15305378	rs581080	G/C	0.276	-0.106(0.024)	7.75×10^{-6}	0.3	Genotyped	Intronic
ABCA1	9:107664301	rs1883025	T/C	0.273	-0.104(0.024)	1.52×10^{-5}	0.3	Genotyped	Intronic
TG									
GCKR	2:27730940	rs1260326	T/C	0.458	0.114(0.021)	9.56×10^{-8}	0.5	Genotyped	Missense, L446P
MAP3K1	5:55861786	rs9686661	T/C	0.306	0.08(0.021)	4.79×10^{-5}	0.3	Genotyped	Intergenic

*Association parameters reported for this marker refer to a model that includes rs73015013 as additional covariate

SUPPLEMENTARY TABLE 6. Estimated genome-wide significance thresholds for single marker tests

The tables show the estimated genome-wide significance threshold and the number of independent tests calculated for each phenotype scenario (see Methods) when using 300 (A) or 1000 (B) simulations.

A)

H2	300 simulations			
	MAF>0%		MAF>=0.5%	
	GW threshold	N independent tests	GW threshold	N independent tests
20%	8.88×10^{-09}	5,628,193	1.05×10^{-08}	4,757,072
40%	6.41×10^{-09}	7,800,977	1.26×10^{-08}	3,953,082
70%	4.65×10^{-09}	10,750,999	7.16×10^{-09}	6,980,950
0%	8.09×10^{-09}	6,177,286	1.01×10^{-08}	4,940,901

B)

H2	1000 simulations			
	MAF>0%		MAF>=0.5%	
	GW threshold	N independent tests	GW threshold	N independent tests
20%	1.13×10^{-08}	4,408,735	1.49×10^{-08}	3,351,741
40%	7.83×10^{-09}	6,382,793	1.37×10^{-08}	3,653,381
70%	na	na	na	na
0%	6.91×10^{-09}	7,239,653	1.40×10^{-08}	3,363,078

SUPPLEMENTARY TABLE 7. Gender specific effects at variants associated with lipid levels

The table shows the association parameters of the SNPs listed in **Table 2** when analysed in males and females separately. Columns are defined as in **Table 2**. SNPs showing significant heterogeneity (HetPval column) between genders are marked in bold.

Chr:position	rs name	Candidate Gene	Effect Allele / Other	Males			Females			HetPVal
				Freq	Effect (StdErr)	pvalue	Freq	Effect (StdErr)	pvalue	
	LDL (2502 males/3379 females)									
1:55505647	rs11591147	*PCSK9*	T/G	0.036	-0.430(0.078)	4.62×10^{-08}	0.039	-0.423(0.066)	1.35×10^{-10}	0.946
1:109821307	rs583104	*SORT1*	G/T	0.182	-0.196(0.039)	5.21×10^{-07}	0.179	-0.140(0.035)	7.30×10^{-05}	0.288
11:5248004	rs11549407	*HBB*	A/G	0.050	-0.548(0.072)	5.38×10^{-14}	0.046	-0.390(0.065)	1.97×10^{-09}	0.105
19:19456917	**rs58489806**	***CILP2***	**T/C**	**0.075**	**-0.358(0.057)**	**5.32×10^{-10}**	**0.073**	**-0.131(0.053)**	**0.013**	**3.57×10^{-03}**
19:45412079	rs7412	*APOE*	T/C	0.032	-0.552(0.083)	4.56×10^{-11}	0.041	-0.689(0.065)	5.36×10^{-26}	0.194
19:45411941	rs429358	*APOE*	C/T	0.077	0.231(0.056)	3.88×10^{-05}	0.072	0.362(0.050)	3.07×10^{-13}	0.080
	TC (2505 males/3380 females)									
1:55505647	rs11591147	*PCSK9*	T/G	0.036	-0.405(0.079)	3.35×10^{-07}	0.039	-0.406(0.065)	6.02×10^{-10}	0.992
4:41980435	-	*TMEM33, DCAF4L1, SLC30A9*	G/A	0.011	-0.523(0.140)	1.89×10^{-04}	0.014	-0.579(0.111)	1.98×10^{-07}	0.754
11:5248004	rs11549407	*HBB*	A/G	0.051	-0.593(0.073)	6.15×10^{-16}	0.046	-0.412(0.065)	2.02×10^{-10}	0.064
19:19456917	**rs58489806**	***CILP2***	**T/C**	**0.075**	**-0.409(0.058)**	**2.40×10^{-12}**	**0.073**	**-0.156(0.052)**	**2.83×10^{-03}**	**1.24×10^{-03}**
19:45412079	rs7412	*APOE*	T/C	0.032	-0.479(0.084)	1.27×10^{-08}	0.041	-0.573(0.065)	1.72×10^{-18}	0.375
19:45411941	rs429358**	*APOE*	C/T	0.076	0.183(0.057)	1.22×10^{-03}	0.072	0.280(0.050)	1.76×10^{-08}	0.200
	HDL (2505 males/3380 females)									
8:19815256	**rs286**	***LPL***	**T/A**	**0.125**	**0.257(0.046)**	**2.70×10^{-08}**	**0.129**	**0.093(0.041)**	**0.023**	**8.14×10^{-03}**
15:58687603	rs174418	*LIPC*	T/C	0.470	0.181(0.029)	9.16×10^{-10}	0.465	0.116(0.026)	7.81×10^{-06}	0.103
16: 56989590	rs247616	*CETP*	T/C	0.282	0.196(0.032)	1.65×10^{-09}	0.258	0.187(0.029)	2.02×10^{-10}	0.828
18:3412386	**rs8092903**	***TGIF1***	**T/C**	**0.027**	**-0.030(0.093)**	**0.746**	**0.026**	**-0.448(0.082)**	**4.49×10^{-08}**	**8.45×10^{-04}**
	TG (2505 males/3380 females)									
8:19845376	rs7841189	*LPL*	T/C	0.201	-0.226(0.038)	2.77×10^{-09}	0.215	-0.103(0.034)	2.16×10^{-03}	0.0156
11:116661101	-	*APOA5*	T/G	0.026	-0.422(0.091)	3.65×10^{-06}	0.024	-0.551(0.086)	1.44×10^{-10}	0.304
11:116664040	rs10750097	*APOA5*	G/A	0.174	0.188(0.040)	2.22×10^{-06}	0.170	0.170(0.036)	2.00×10^{-06}	0.734
19:19456917	rs58489806	*CILP2*	T/C	0.075	-0.341(0.057)	2.06×10^{-09}	0.073	-0.200(0.052)	1.22×10^{-04}	0.068

SUPPLEMENTARY TABLE 8. Gender specific effects at variants associated with inflammatory markers

The table shows the association parameters of the SNPs listed in **Table 3** when analysed in males and females separately. Columns are defined as in **Table 3.** SNPs showing significant heterogeneity (HetPval column) between genders are marked in bold.

Chr:position	rs name	Candidate Gene	Effect Allele / Other	Males			Females			HetPVal
				Freq	Effect (StdErr)	pvalue	Freq	Effect (StdErr)	pvalue	
	ADPN (2486 males/3350 females)									
3:186559460	rs17300539	*ADIPOQ*	A/G	0.165	0.277 (0.038)	$5.15x10^{-13}$	0.150	0.256 (0.034)	$1.28x10^{-13}$	0.687
13:108884835	-	*ABDH13*	A/G	0.001	-1.982 (0.521)	$1.44x10^{-04}$	0.001	-1.426 (0.350)	$4.68x10^{-05}$	0.377
	hsCRP (2411 males/3219 females)									
1:159684665	rs3091244	*CRP*	A/G	0.414	0.196 (0.028)	$2.46x10^{-12}$	0.439	0.229 (0.025)	$3.11x10^{-20}$	0.393
8:17450500	rs73198138	*PDGFRL*	A/G	0.005	-0.899 (0.203)	$9.55x10^{-06}$	0.003	-0.884 (0.213)	$3.38x10^{-05}$	0.960
12:125533106	**rs183233091**	***BRI3BP, AACS***	**A/G**	**0.010**	**1.308 (0.134)**	$\mathbf{3.57x10^{-22}}$	**0.010**	**0.807 (0.125)**	$\mathbf{1.25x10^{-10}}$	$\mathbf{6.36x10^{-03}}$
12:121415293	*rs7139079*	*HNF1A*	G/A	0.375	-0.127 (0.029)	$9.60x10^{-06}$	0.379	-0.134 (0.025)	$1.19x10^{-07}$	0.849
19:45411941	rs429358	*APOE*	C/T	0.076	-0.248 (0.051)	$1.53x10^{-06}$	0.071	-0.228 (0.047)	$1.02x10^{-06}$	0.774
	ESR (2531 males/3410 females)									
1:25724005	rs71721472	*RHCE*	T/C	0.305	-0.095 (0.032)	$2.65x10^{-03}$	0.291	-0.161 (0.028)	$1.45x10^{-08}$	0.123
1:207684359	**rs11117956**	***CR1***	**T/G**	**0.397**	**-0.102 (0.029)**	$\mathbf{3.65x10^{-04}}$	**0.403**	**-0.215 (0.025)**	$\mathbf{8.51x10^{-18}}$	$\mathbf{2.88x10^{-03}}$
11:5248004	**rs76728603**	***HBB***	**A/G**	**0.050**	**-0.348 (0.067)**	$\mathbf{2.52x10^{-07}}$	**0.046**	**-0.589 (0.060)**	$\mathbf{1.31x10^{-22}}$	$\mathbf{7.75x10^{-03}}$
12:125406340	-	*AACS, MIR5188*	G/A	0.008	1.260 (0.152)	$2.00x10^{-16}$	0.006	0.806 (0.164)	$9.45x10^{-07}$	0.044
	MCP-1 (2497 males/3347 females)									
1:159175354	rs12075	*DARC*	G/A	0.447	-0.415 (0.029)	$4.91x10^{-46}$	0.445	-0.408 (0.024)	$3.15x10^{-60}$	0.859
1:159164454	*rs2852718*	*CADM3*	C/T	0.021	-0.223 (0.103)	0.03	0.023	-0.450 (0.084)	$8.33x10^{-08}$	0.087
1:159175494	*rs34599082*	*DARC*	T/C	0.037	-0.068 (0.079)	0.39	0.037	-0.162 (0.066)	0.014	0.359
3:46383906	rs113403743	*CCR2*	T/G	0.100	0.270 (0.050)	$7.00x10^{-08}$	0.098	0.273 (0.043)	$3.65x10^{-10}$	0.959
3:46399764	*rs200491743*	*CCR2*	A/T	0.005	1.115 (0.199)	$2.50x10^{-08}$	0.006	0.516 (0.170)	$2.40x10^{-03}$	0.022
16:49072490	**rs76135610**	***N4BP1, CBLN1***	**T/C**	**0.005**	**0.969 (0.172)**	$\mathbf{1.76x10^{-08}}$	**0.006**	**0.286 (0.192)**	**0.1378**	$\mathbf{8.06x10^{-03}}$
	IL-6 (2492 males/3346 females)									
1:154428283	rs12133641	*IL6R*	G/A	0.258	0.123 (0.030)	$3.98x10^{-05}$	0.253	0.117 (0.026)	$9.58x10^{-06}$	0.885
9:136142355	rs643434	*ABO*	A/G	0.267	-0.223 (0.030)	$8.29x10^{-14}$	0.260	-0.218 (0.026)	$2.18x10^{-16}$	0.900

SUPPLEMENTARY TABLE 9. Replication results

The table describes replication results in the 8 European cohorts. For each cohort, we listed the number of individuals involved, the effect allele and the other allele, the frequency of the effect allele, the effect size and the standard error, the pvalue for association with hsCRP, and the imputation quality score INFO. We also provided the combined one-tail pvalues using a sample size weighted meta-analysis.

Cohort	N	Effect Allele / Other	Freq	Effect (StdErr)	Pvalue	INFO
chr12:125533106 **/ rs183233091**						
ING-FVG	411	A/G	0.022	0.0490 (0.270)	0.854	0.60
ING-VBI	1162	A/G	0.023	0.2260 (0.165)	0.173	0.78
HA	1093	A/G	0.019	0.2980 (0.204)	0.145	0.64
HP	839	A/G	0.033	0.0219 (0.163)	0.892	0.74
TwinsUK	1167	A/G	na	na	na	na
ALSPAC	879	A/G	na	na	na	na
INCIPE1	807	A/G	0.018	0.1890 (0.205)	0.357	0.78
INCIPE2	1332	A/G	0.018	-0.0356 (0.178)	0.841	0.71
Combined ING-FVG, ING-VBI, HA, HP (3505 individuals) one tail-pvalue 0.042						
Combined all cohorts (5644 individuals) one tail-pvalue 0.053						
chr12:121139532 **/ rs1803508**						
ING-FVG	411	G/C	0.003	-1.140 (0.067)	0.918	0.97
ING-VBI	1162	G/C	0.012	-0.4510 (0.211)	0.034	0.86
HA	1093	G/C	0.004	-0.0078 (0.036)	0.827	0.98
HP	839	G/C	0.005	0.103 (0.037)	0.783	0.97
TwinsUK	1167	G/C	0.009	-0.3126 (0.256)	0.223	na
ALSPAC	879	G/C	0.005	-0.4127 (0.304)	0.175	na
INCIPE1	807	G/C	0.005	0.2154 (0.348)	0.536	0.97
INCIPE2	1332	G/C	0.004	0.2733 (0.303)	0.368	0.96
Combined ING-FVG, ING-VBI, HA, HP (3505 individuals) one tail-pvalue 0.032 (but opposite direction compared to Sardinia)						
Combined all cohorts (7689 individuals) one tail-pvalue 0.058						
chr12:125650151 **/ rs148280335**						
ING-FVG	411	C/T	0.014	0.0161 (0.314)	0.957	0.84
ING-VBI	1162	C/T	0.011	0.108 (0.218)	0.621	0.85

HA	1093	C/T	0.015	0.199 (0.199)	0.318	0.84
HP	839	C/T	0.018	0.131 (0.208)	0.539	0.83
TwinsUK	1167	C/T	0.010	0.1702 (0.239)	0.477	na
ALSPAC	879	C/T	0.012	0.237 (0.192)	0.217	na
INCIPE1	807	C/T	0.011	0.3232 (0.251)	0.198	0.85
INCIPE2	1332	C/T	0.010	0.1972 (0.213)	0.354	0.81

Combined ING-FVG, ING-VBI, HA, HP (3505 individuals) one tail-pvalue 0.062
Combined all cohorts (7689 individuals) one tail-pvalue 0.125

chr12:125758826 / rs117135060						
ING-FVG	411	A/G	0.005	0.563(0.610)	0.357	0.60
ING-VBI	1162	A/G	0.011	-0.204 (0.227)	0.368	0.79
HA	1093	A/G	0.017	-0.259 (0.172)	0.132	0.97
HP	839	A/G	0.004	0.420 (0.414)	0.311	0.96
TwinsUK	1167	A/G	0.006	-0.0816 (0.303)	0.788	na
ALSPAC	879	A/G	0.010	0.237 (0.193)	0.219	na
INCIPE1	807	A/G	0.012	-0.095 (0.222)	0.666	1.00
INCIPE2	1332	A/G	0.012	0.1385 (0.220)	0.529	0.66

Combined ING-FVG, ING-VBI, HA, HP (3505individuals) one tail-pvalue 0.115
Combined all cohorts (7689 individuals) one tail-pvalue 0.391

chr12:125762779 / rs117818952						
ING-FVG	411	T/C	0.005	-0.567 (0.609)	0.353	0.60
ING-VBI	1162	T/C	0.011	-0.206 (0.227)	0.365	0.79
HA	1093	T/C	0.017	-0.259 (0.172)	0.132	0.97
HP	839	T/C	0.004	0.421 (0.414)	0.310	0.97
TwinsUK	1167	T/C	0.006	-0.0698 (0.303)	0.819	1.00
ALSPAC	879	T/C	0.010	0.2372 (0.193)	0.219	1.00
INCIPE1	807	T/C	0.012	-0.0984 (0.221)	0.657	0.66
INCIPE2	1332	T/C	0.012	0.1406 (0.220)	0.523	0.79

Combined ING-FVG, ING-VBI, HA, HP (3505 individuals) one tail-pvalue 0.115
Combined all cohorts (7689 individuals) one tail-pvalue 0.361

chr12:125765572 / rs149451744						
ING-FVG	411	G/C	0.005	-0.567 (0.610)	0.354	0.58
ING-VBI	1162	G/C	0.011	-0.206 (0.227	0.364	0.79
HA	1093	G/C	0.017	-0.259 (0.174)	0.136	0.96

HP	839	G/C	0.004	0.420 (0.414)	0.311	0.97
TwinsUK	1167	G/C	0.006	-0.0693 (0.304)	0.819	1.00
ALSPAC	879	G/C	0.010	0.2458 (0.194)	0.206	0.99
INCIPE1	807	G/C	0.012	-0.0993 (0.221)	0.654	0.66
INCIPE2	1332	G/C	0.012	0.1398 (0.220)	0.526	0.79

Combined ING-FVG, ING-VBI, HA, HP (3505 individuals) one tail-pvalue 0.128
Combined all cohorts (7689 individuals) one tail-pvalue 0.365

chr12:125766568 / rs142361132

ING-FVG	411	T/C	0.005	-0.546 (0.615)	0.375	0.56
ING-VBI	1162	T/C	0.011	-0.209 (0.228)	0.360	0.79
HA	1093	T/C	0.019	-0.214 (0.170)	0.210	0.91
HP	839	T/C	0.004	0.419 (0.414)	0.313	0.98
TwinsUK	1167	T/C	na	na	na	na
ALSPAC	879	T/C	na	na	na	na
INCIPE1	807	T/C	0.012	-0.105 (0.222)	0.634	0.99
INCIPE2	1332	T/C	0.012	0.176 (0.220)	0.422	0.65

Combined ING-FVG, ING-VBI, HA, HP (3505 individuals) one tail-pvalue 0.149
Combined all cohorts (5644 individuals) one tail-pvalue 0.272

chr8:17450500 / rs73198138

ING-FVG	411	A/G	0.005	0.2620 (0.588)	0.656	0.665
ING-VBI	1162	A/G	0.018	0.1920 (0.175)	0.272	0.826
HA	1093	A/G	0.009	-0.0829 (0.247)	0.737	0.694
HP	839	A/G	0.023	0.0124 (0.177)	0.484	0.933
TwinsUK	879	A/G	0.028	-0.1375 (0.149)	0.357	na
ALSPAC	1167	A/G	0.023	-0.0617 (0.141)	0.663	na
INCIPE1	807	A/G	0.015	-0.1720 (0.217)	0.430	0.870
INCIPE2	1332	A/G	0.016	0.1301 (0.178)	0.464	0.779

Combined ING-FVG, ING-VBI, HA, HP (3505 individuals) one tail-pvalue 0.173
Combined all cohorts (7689 individuals) one tail-pvalue 0.418

SUPPLEMENTARY TABLE 10. Variants contributing to significant signals in the rare variant tests.

The table show the chromosome and position (hg19) for variants that contribute to the significant rare variant association results at *STAB1* and *PTPRH* genes (see **Table 4**).

***STAB1* variants (chr:position)**	***PTPRH* variants (chr:position)**
3:52535766	19:55693469
3:52537010	19:55693470
3:52538507	19:55697255
3:52539378	19:55697712
3:52546403	19:55697718
3:52546872	19:55698955
3:52546910	19:55703094
3:52547252	19:55708649
3:52547915	19:55708761
3:52548194	19:55710068
3:52548465	19:55713439
3:52548787	19:55713661
3:52550173	19:55715342
3:52550722	19:55715359
3:52551566	19:55716713
3:52553377	
3:52553551	
3:52555958	
3:52556184	
3:52556385	
3:52557693	
3:52558237	

SUPPLEMENTARY TABLE 11. Validation of imputed variants newly associated with inflammatory markers.

For each SNP, we show the number of heterozygotes and homozygotes for the reference and alternative alleles that were imputed using the Sardinian panel, the number of these that were validated by Sanger sequencing along with the genotype mismatch rate, the pvalue observed in our primary analysis (as reported in **Table 3**) and the pvalue obtained replacing imputed genotypes with those derived by Sanger sequencing.

SNP	N hom ref/het/hom alt	N Sanger sequencing			Original pvalue	pvalue after validation
		Hom Ref (Mismatch %)	Het (Mismatch %)	Hom alt (Mismatch %)		
13:108884835	5824/12/0	12 (0%)	12 (0%)	0	3.35×10^{-08}	2.84×10^{-08}
8:17450500	5588/42/0	20 (0%)	42 (7%)	0	3.31×10^{-09}	2.65×10^{-09}
12:125533106	5524/105/1	20 (0%)	63 (0%)	1 (100%)	1.09×10^{-28}	1.80×10^{-28}
12:125406340	5864/77/0	20 (0%)	16 (0%)	0	4.40×10^{-23}	4.41×10^{-23}
16:49072490	3312/35/0	21 (0%)	33 (0%)	0	1.76×10^{-08}	1.76×10^{-08}

SUPPLEMENTARY TABLE 12. Association signals based on 1000G imputation for lipid levels

The table reports top association signals identified with 1000G imputation for lipid levels. Columns are the same as defined in **Table 2**. The signal for the *Q40X* mutation was missed and misplaced 122Kb away to SNP rs76053862, located between *OR52E* and *OR52A5* genes.

Candidate Gene	Chr:position	rs name	Effect Allele / Other	Freq	Effect (StdErr)	pvalue	R2	RSQR	Variant Consequence
LDL									
PCSK9	1:55505647	rs11591147	T/G	0.038	-0.406(0.053)	1.51×10^{-14}	1.0	Genotyped	Missense, R46L
SORT1	1:109821307	rs583104	G/T	0.180	-0.156(0.027)	2.06×10^{-08}	0.5	Genotyped	Downstream
OR52E, OR52A5	**11:5125982**	**rs76053862**	**T/C**	**0.049**	**-0.403(0.054)**	**1.44×10^{-13}**	**0.9**	**0.785**	**Intergenic**
CILP2	19:19456917	rs58489806	T/C	0.074	-0.233(0.042)	2.59×10^{-08}	0.5	Genotyped	Intronic
APOE	19:45412079	rs7412	T/C	0.037	-0.645(0.053)	2.47×10^{-33}	2.4	Genotyped	Missense, R176C
APOE	19:45411941	rs429358*	C/T	0.074	0.264(0.039)	1.21×10^{-11}	0.8	0.999	Missense, C130R
TC									
PCSK9	1:55505647	rs11591147	T/G	0.038	-0.38(0.053)	9.79×10^{-14}	1.0	Genotyped	Missense, R46L
OR52E, OR52A5	**11:5125982**	**rs76053862**	**T/C**	**0.048**	**-0.407(0.054)**	**5.95×10^{-14}**	**1.0**	**0.785**	**Intergenic**
CILP2	19:19456917	rs58489806	T/C	0.074	-0.264(0.041)	2.15×10^{-10}	0.6	Genotyped	Intronic
APOE	19:45412079	rs7412	T/C	0.036	-0.544(0.053)	2.06×10^{-24}	1.7	Genotyped	Missense, R176C
APOE	19:45411941	rs429358*	C/T	0.074	-0.215(0.038)	3.09×10^{-08}	0.5	0.999	Missense, C130R
HDL									
LPL	8:19815256	rs286**	T/A	0.125	0.257(0.046)	2.70×10^{-08}	1.2	Genotyped	Intronic
LIPC	15:58687603	rs174418	T/C	0.467	0.137(0.021)	1.19×10^{-10}	0.7	0.996	Intergenic
CETP	16:56987015	rs12446515	C/T	0.268	0.190(0.023)	1.96×10^{-16}	1.1	Genotyped	Intergenic
TG									
LPL	8:19938902	-	CAAAT/C	0.2078	-0.169(0.027)	3.14×10^{-10}	0.7	0.87	Intergenic
APOA5	**11:116661101**	**-**	**T/G**	**0.025**	**-0.450(0.064)**	**1.24×10^{-12}**	**0.9**	**Genotyped**	**Missense, R282S**

APOA5	11:116664040	rs10750097***	G/A	0.171	0.174(0.027)	1.13×10^{-10}	0.7	Genotyped	Upstream
CILP2	19:19456917	rs58489806	T/C	0.074	-0.260(0.039)	2.14×10^{-11}	0.8	Genotyped	Intronic

Notes:

* Association parameters reported for this marker refer to a model that includes rs7412 as additional covariate

** Results refer to the sex-specific analyses. See **Supplementary Table 7** for more details.

*** Association parameters reported for this marker refer to a model that includes 11:116661101 as additional covariate

SUPPLEMENTARY TABLE 13. Association signals based on 1000G imputation for the inflammatory markers

The table reports top association signals identified with 1000G imputation for inflammatory markers. Columns are the same as defined in **Table 3**. Signals at novel loci are in bold. Independent signals, indicated in italics, are reported along with the regression coefficients from the conditional analysis. More specifically, *PDGFRL*, *ABHD13* and *N4BP1*/*CBLN1* loci are not listed as none of the tested SNPs showed a pvalue < $5x10^{-8}$, whereas at the *AACS* locus association was observed with weaker evidence in the neighbouring region (near *SCARB1* gene). Variants in *PDGFRL*, *AACS* (rs183233091) and *N4BP1*/*CBLN1* are present in 1000G panels but are poorly imputed in our population. In further comparisons with the recent release of the UK10K project, marker chr13:108884835 at *ABHD13* is now reported, but with extremely low frequency (MAF=1/7562), whereas the variant chr12:125406340 near *AACS* remains absent outside Sardinia, and is thus either specific to Sardinians or extremely rare elsewhere in Europe. Finally, the association at the *HBB* locus was detected, but with a much weaker signal at a gene nearby.

SNP	rs name	*Nearest Gene*	Effect allele / Other	Freq	pvalue	Effect (StdErr)	RSQR	Type
ADPN								
3:186559460	rs17300539	*ADIPOQ*	A/G	0.156	$4.86x10^{-22}$	0.246 (0.025)	1.000	intergenic
hsCRP								
1:159684665	rs3091244	*CRP*	A/G	0.428	$5.28x10^{-27}$	0.207 (0.019)	Genotyped	intergenic
12:121428455	rs1169297	*HNF1A*	A/G	0.344	$3.25x10^{-08}$	-0.116 (0.021)	0.935	Intronic
12:125259484	**ss1372755559**	***SCARB1***	**AATTC/A**	**0.995**	**$9.66x10^{-21}$**	**-1.472 (0.157)**	**0.672**	**intergenic**
19:45411941	rs429358	*APOE*	C/T	0.074	$1.38x10^{-11}$	-0.240 (0.035)	0.988	nonsyn
ESR								
1:207688373	rs4433395	*CR1*	C/T	0.603	$1.15x10^{-17}$	0.156 (0.018)	0.984	intronic
11:5125982	rs76053862	*OR52A5*	C/T	0.049	$3.59x10^{-14}$	-0.343 (0.045)	0.785	Intergenic
12:125259484	**ss1372755559**	***SCARB1***	**AATTC/A**	**0. 995**	**$6.30x10^{-18}$**	**-1.269 (0.157)**	**0.672**	**intergenic**
MCP-1								
1:159175354	rs12075	*DARC*	A/G	0.554	$2.79x10^{-95}$	0.405 (0.019)	1.000	nonsyn
1:159162174	*rs2814767*	*CADM3*	*T/G*	*0.022*	*$1.049x10^{-16}$*	*-0.520 (0.062)*	*0.985*	*intronic*
1:159175494	***rs34599082***	***DARC***	***C/T***	***0.037***	*$1.02x10^{-11}$*	***-0.339 (0.050)***	***0.999***	***nonsyn***
3:46391788	rs17141006	*CCR2*	G/T	0.099	$3.64x10^{-15}$	0.271 (0.034)	1.000	intergenic
IL6								
1:154428283	rs12133641	*IL6R*	G/A	0.256	$7.23x10^{-09}$	0.119 (0.021)	0.999	intronic
9:136132908	N/A	*ABO*	T/TC	0.742	$1.66x10^{-26}$	-0.222 (0.021)	0.997	-

SUPPLEMENTARY TABLE 14. CADD score at loci associated with lipid levels

The table shows the CADD score for each lead SNP reported in **Table 2**, and indicates the SNP with the highest CADD among those in r^2>0.5 with the lead.

Candidate gene	Lead SNP	CADD score	SNP with highest CADD score	CADD score	r^2 with lead
LDL					
PCSK9	1:55505647	11.46	-	-	-
SORT1	1:109821307	8.78	1:109821511	13.1	0.80
HBB	11:5248004	37	-	-	-
CILP2	19: 19456917	2.59	19:19379549	17.03	0.85
APOE	19:45412079	15.82	-	-	-
APOE	19:45411941	0.007	19:45410002	9.031	0.55
TC					
PCSK9	1:55505647	11.46	-	-	-
TMEM33, DCAF4L1, SLC30A9	4:41980435	4.075	-	-	-
HBB	11:5248004	37	-	-	-
CILP2	19:19456917	2.59	19:19379549	17.03	0.85
APOE	19:45412079	15.82	-	-	-
APOE	19:45411941	0.007	19:45410002	9.031	0.55
HDL					
LPL	8:19845376	1.991	8:19819724	42	0.63
LIPC	15:58687603	3.275	15:58678512	9.379	0.55
CETP	16: 56989590	0.229	16:56990716	7.471	0.99
TGIF1	18:3412386	1.244	-	-	-
TG					
LPL	8:19845376	8.273	8:19819724	42	0.63
APOA5	11:116661101	10.41	-	-	-
APOA5	11:116664040	3.350	11:116589652	10.27	0.50
CILP2	19:19456917	2.59	19:19379549	17.03	0.85

SUPPLEMENTARY TABLE 15. CADD score at loci associated with inflammatory markers

The table shows the CADD score for each lead SNP reported in **Table 3**, and indicates the SNP with the highest CADD among those in r^2>0.5 with the lead.

Candidate gene	Lead SNP	CADD score	SNP with highest CADD score	CADD score	r^2 with lead
ADPN					
ADIPOQ	3:186559460	2.871	3:186556037	9.058	0.95
ABDH13	13:108884835	6.903	-	-	-
hsCRP					
CRP	1:159684665	0.217	1:159713225	14.18	0.89
PDGFRL	8:17450500	0.306	-	-	-
BRI3BP, AACS	12:125533106	0.164	-	-	-
HNF1A	12:121415293	0.172	12:121438844	10.22	0.70
APOE	19:45411941	0.007	19:45410002	9.031	0.55
ESR					
RHCE, TMEM57	1:25724005	8.273	1:25561667	16.1	0.51
CR1	1:207684359	0.409	1:207789471	22.1	0.97
HBB	11:5248004	37	-	-	-
AACS, MIR5188	12:125406340	10.42	-	-	-
MCP-1					
DARC	1:159175354	0.168	-	-	-
CADM3	1:159164454	4.825	1:159090256	25.5	0.55
DARC	1:159175494	12.96	-	-	-
CCR2, CCR3	3:46383906	3.318	3:46412559	14.03	0.86
CCR2	3:46399764	15.88	-	-	-
N4BP1, CBLN1	16:49072490	0.903	16:49145757	6.255	0.56
IL-6					
IL6R	1:154428283	3.658	1:154414296	12.99	0.90
ABO	9:136142355	0.215	9:136149229	3.651	0.84

SUPPLEMENTARY TABLE 16. Variance explained by top hits for lipid levels

For each of the four lipid traits, the table shows the amount of phenotypic variance explained by the top signals identified in this work using the Sardinian (**Table 2**) and the 1000 Genomes reference panel. We calculated the log likelihood of each model and report the difference.

Trait	Heritability in our study	Variance explained by top hits identified with:		differences in Loglikelihood
		Sardinian	1000G	
LDL-c	42.5	7.77	6.95	32.91
TC	42.7	4.57	4.21	14.22
HDL	49.2	1.92	1.92	-0.04
TG	32.4	2.15	2.10	1.42

SUPPLEMENTARY TABLE 17. Variance explained by top hits for the inflammatory markers

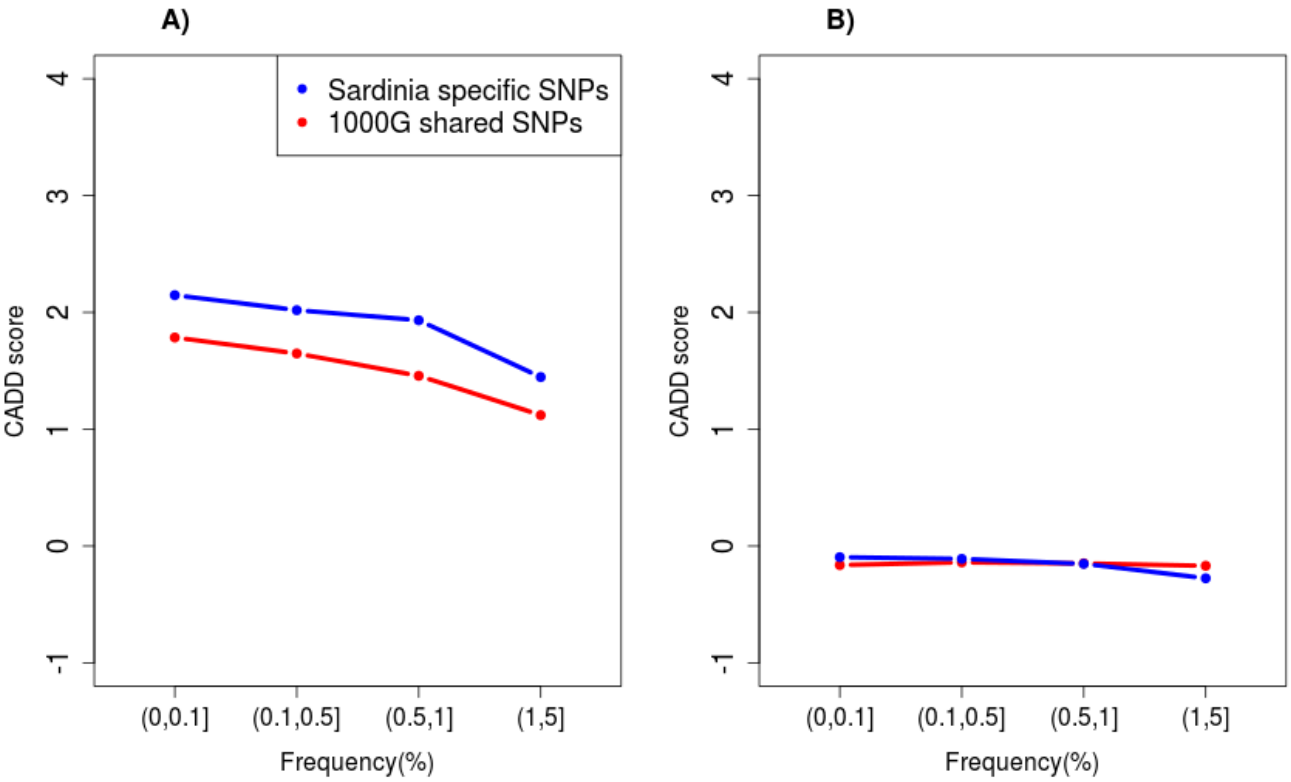
For each of the five inflammatory markers, the table shows the amount of phenotypic variance explained by the top signals identified in this work using the Sardinian and the 1000 Genomes reference panel, as well as the variance explained by loci detected in our previous report (Naitza S. et al, Plos Genetics 2012) based on HapMap2 variants (with calculations repeated using the same sample size and genotypes described in this manuscript). In addition, we calculated the log likelihood of each model and report the difference.

Trait	Heritability in our study	Variance explained by top hits identified with:			differences in Loglikelihood:	
		Sardinian	1000G	HapMap2	Sardinian - 1000G	Sardinian - HapMap2
ADPN	39.2	2.37	1.93	na	16.08	na
hsCRP	25.1	6.35	5.27	5.49	57.77	48.69
ESR	43.1	4.58	2.98	2.6	72.49	86.03
MCP-1	31	11.8	11.23	9.36	20.27	85.14
IL-6	15.3	2.49	2.54	2.49	-1.68	0.10

SUPPLEMENTARY FIGURES

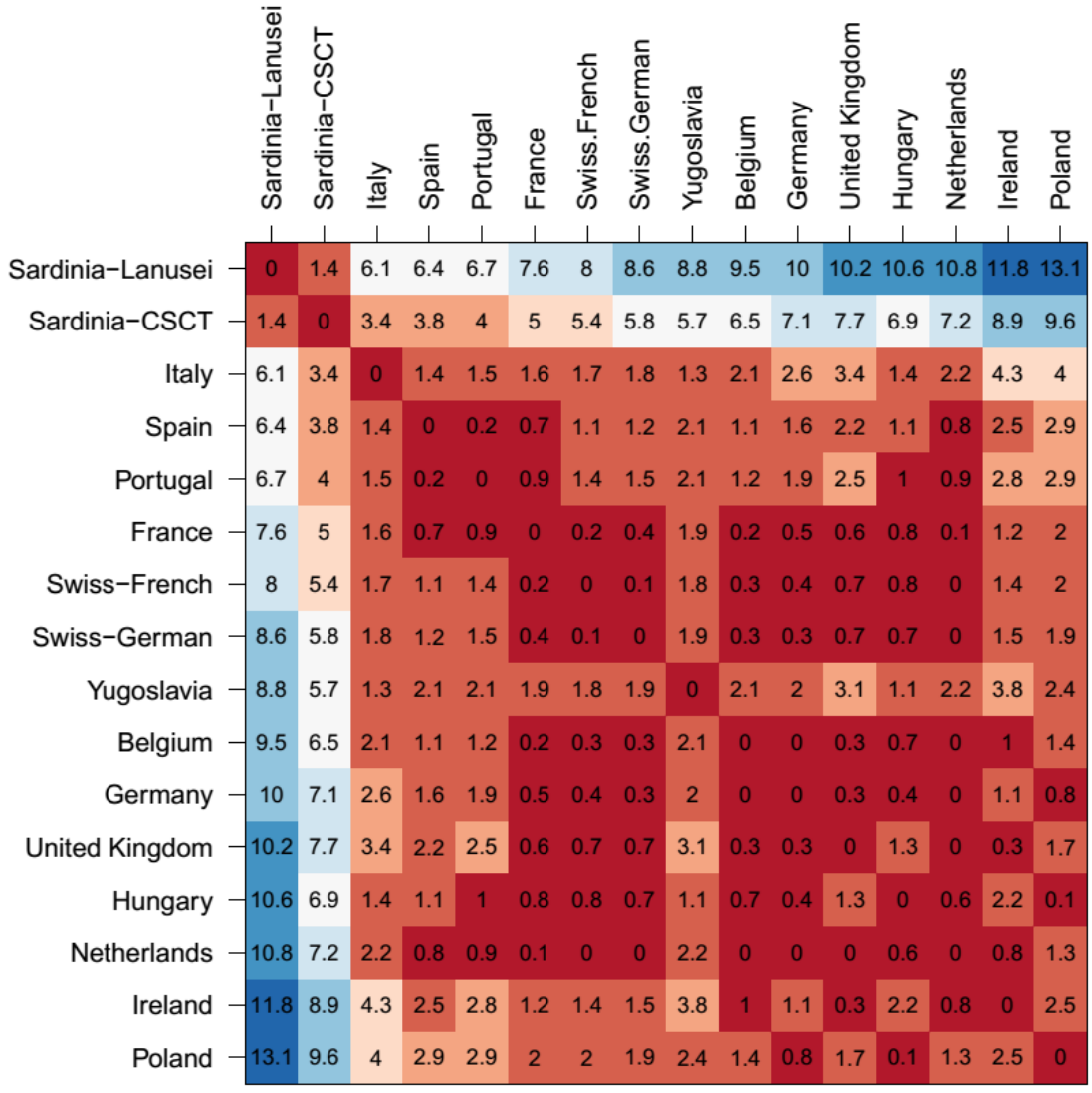
SUPPLEMENTARY FIGURE 1. Average deleteriousness of variants

Average deleteriousness of variants shared between 1000 Genomes and Sardinia (in red) or unique to Sardinia (in blue). Deleteriousness is summarized in a CADD score, where higher scores suggest larger deleterious effects. Panel A shows results for variants that alter protein coding sequences; Panel B shows results for remaining variants. The difference between the curves was significant only for coding variants (ANOVA test, pvalues =0.02 and 0.41 for coding and noncoding variants).



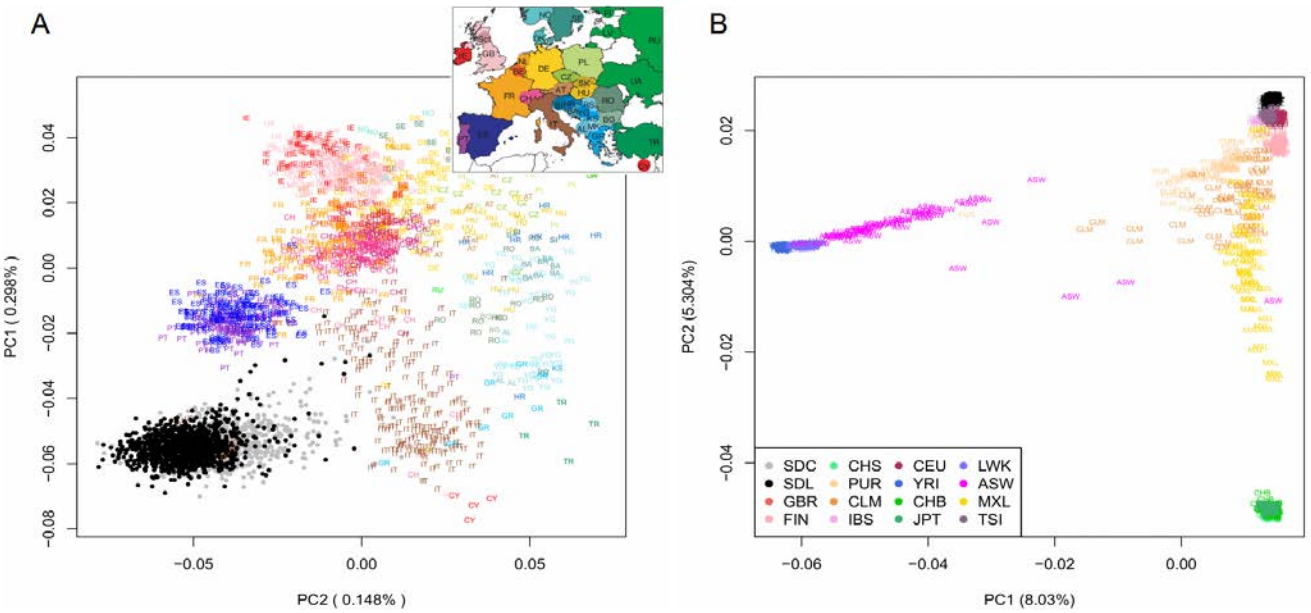
SUPPLEMENTARY FIGURE 2. FST differentiation between Sardinians and other Europeans

Differentiation between Sardinians and other European populations, summarized in a color-coded matrix of F_{ST} values. All tabulated F_{ST} values have been multiplied by 1,000. European populations are drawn from a previously curated set of 1,385 European individuals in the POPRES study (dbGAP accession number 2039). For all comparisons, the Fst values are significant with $P < 0.001$ (none of the 1000 random permutations that we generated had pvalues higher), with the exceptions of the following pairs: Belgium - Germany (p=0.005); Belgium – Netherlands (p=0.67); Germany – Netherlands (p=0.145); and Netherlands - United Kingdom (p=0.315).



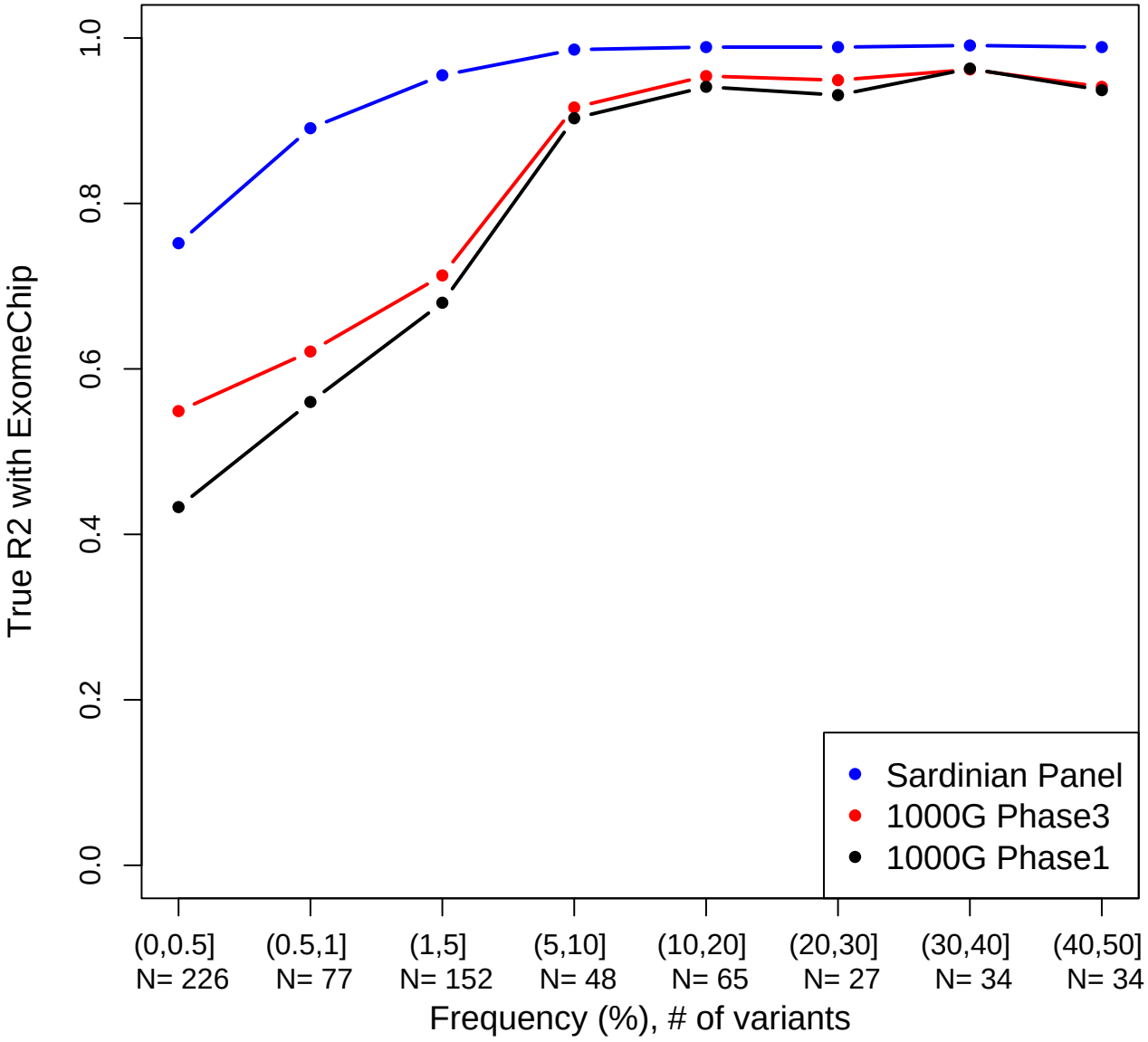
SUPPLEMENTARY FIGURE 3. Principal component ancestry map

Panel A shows a principal component ancestry map of Europe, including both our samples and a previously described set of diverse European individuals. In this analysis, the map was constructed using all available European POPRES samples, and all sequenced Sardinian samples were projected into an existing PCA coordinate space, one a time (see **Methods**). Sardinian samples from the Lanusei valley (SDL) and from case-control studies (SDC) are plotted in grey and black, respectively. Panel B graphically show a standard PCA analysis with Sardinians and individuals from the 1000 Genomes Phase 1.



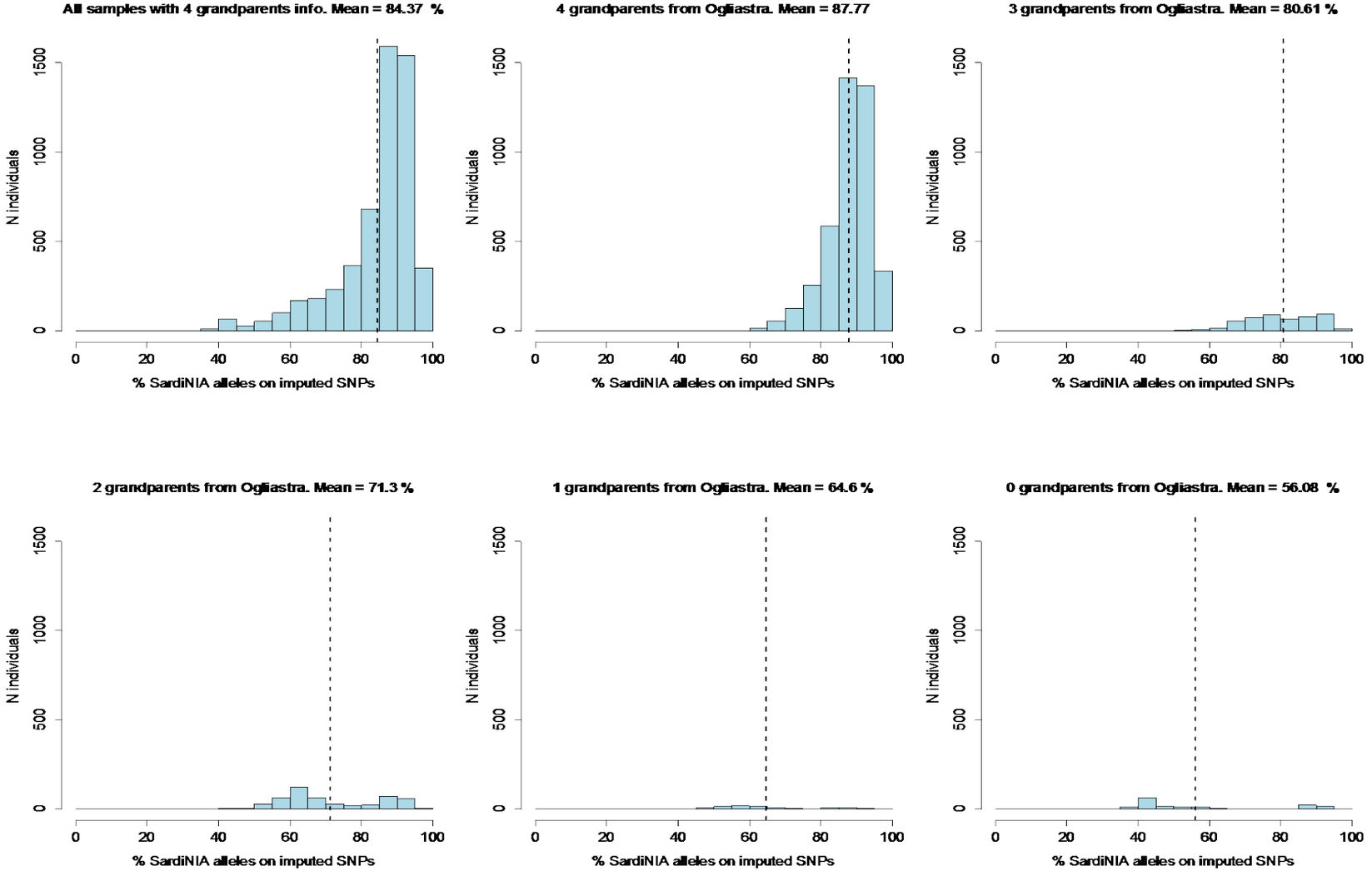
SUPPLEMENTARY FIGURE 4. Imputation accuracy comparisons

Performance of imputation based analyses using 1000 Genomes Phase 1 (black line), 1000 Genomes Phase 3 (red line) or our Sardinian sequences (blue line) as a reference panel. The quality is reported as a function of allele frequency; the number of variants used for estimating accuracy is reported below each frequency label.



SUPPLEMENTARY FIGURE 5. Proportion of haplotypes with nearest neighbour in the Lanusei valley, as a function of grandparental birthplace

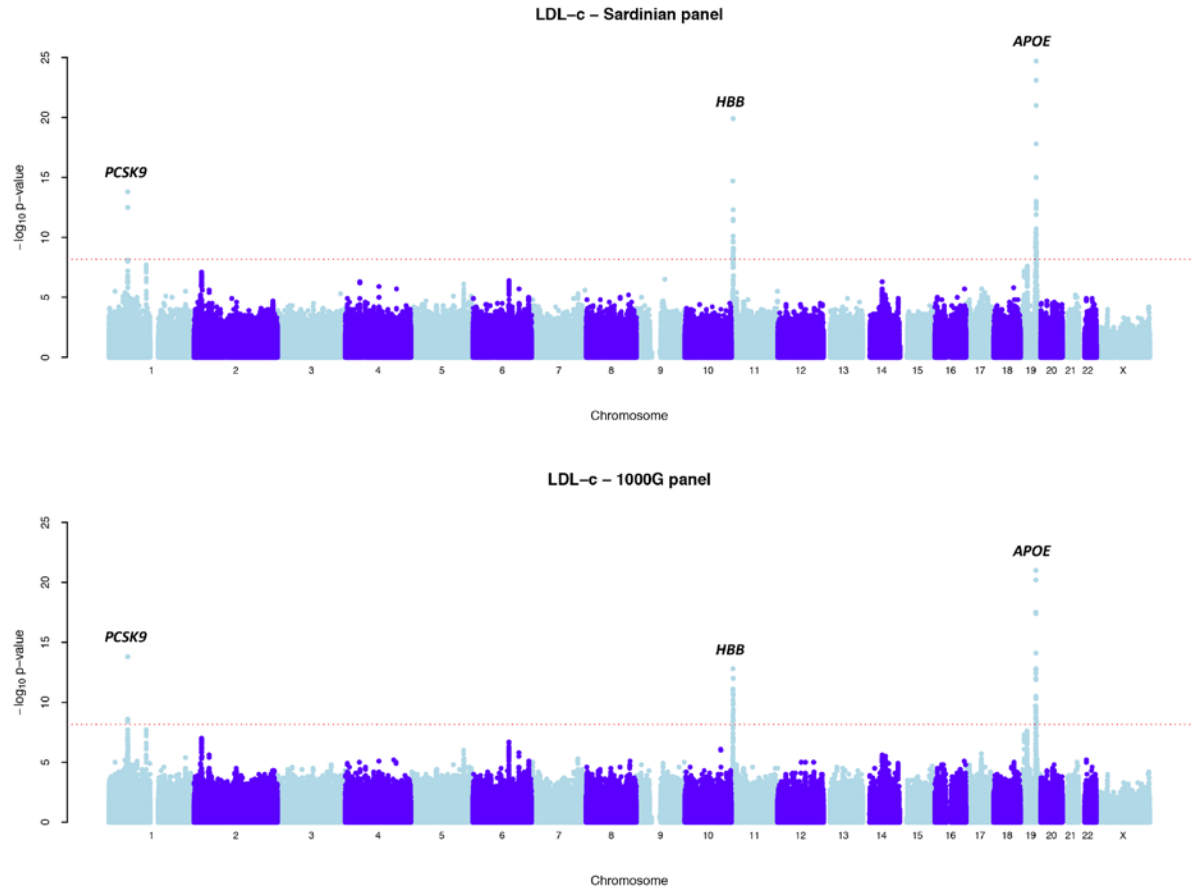
Panels show first all individuals in the SardiNIA study and, subsequently, stratify these individuals according to the number of haplotypes (0-4) known to originate in the valley. Note that the fraction of haplotypes where nearest neighbour originates in the valley closely tracks grandparental ancestry.



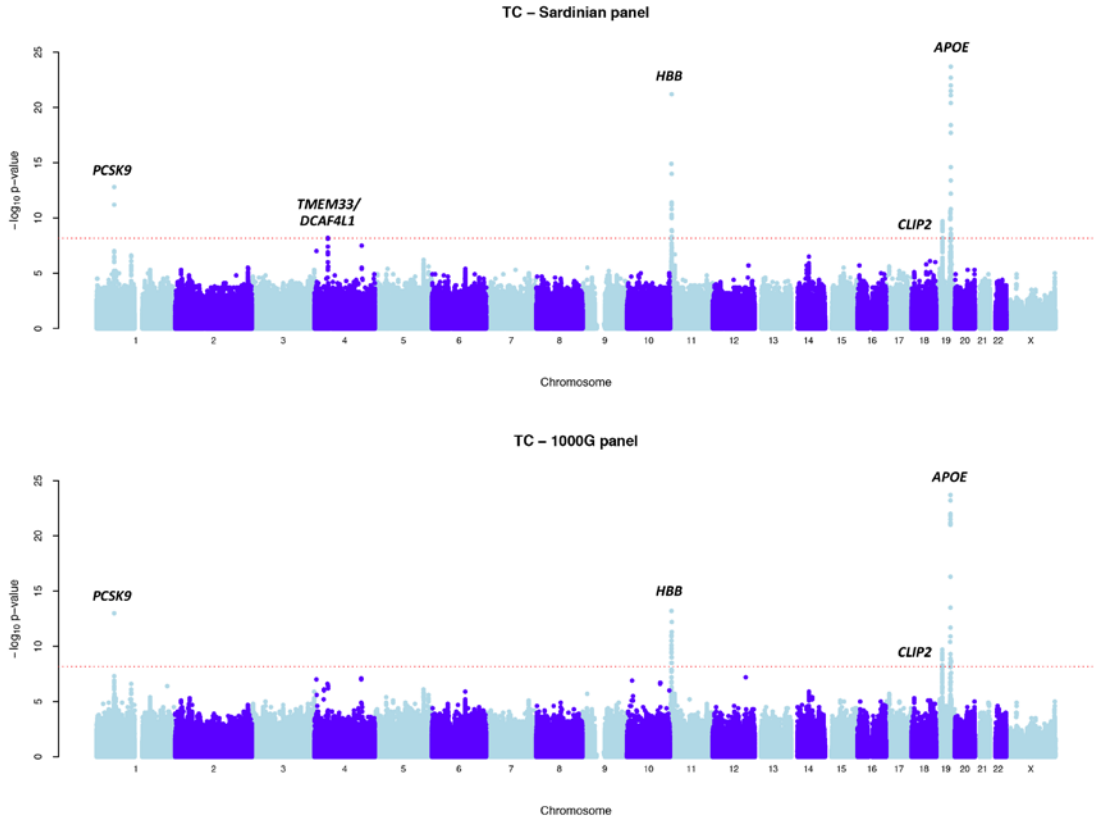
SUPPLEMENTARY FIGURE 6. Manhattan plots for studied traits

The figures show association results for the studied traits for all QCed genotyped and imputed markers at autosomes, and only genotyped markers on the X chromosome. For each trait, results refer to those obtained from imputation using the Sardinian (upper inset) and 1000G phase 3 (lower inset) sequencing data. For signals that reach the 6.9×10^{-9} threshold the corresponding candidate gene is annotated.

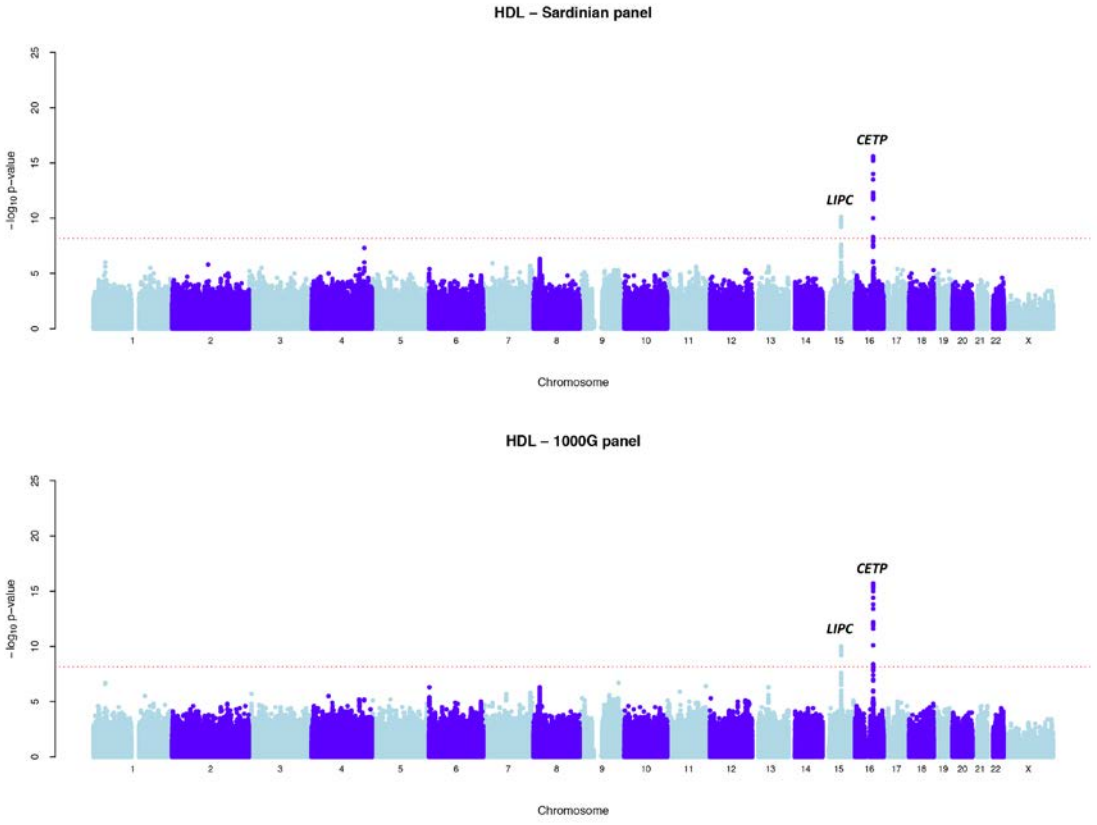
A)



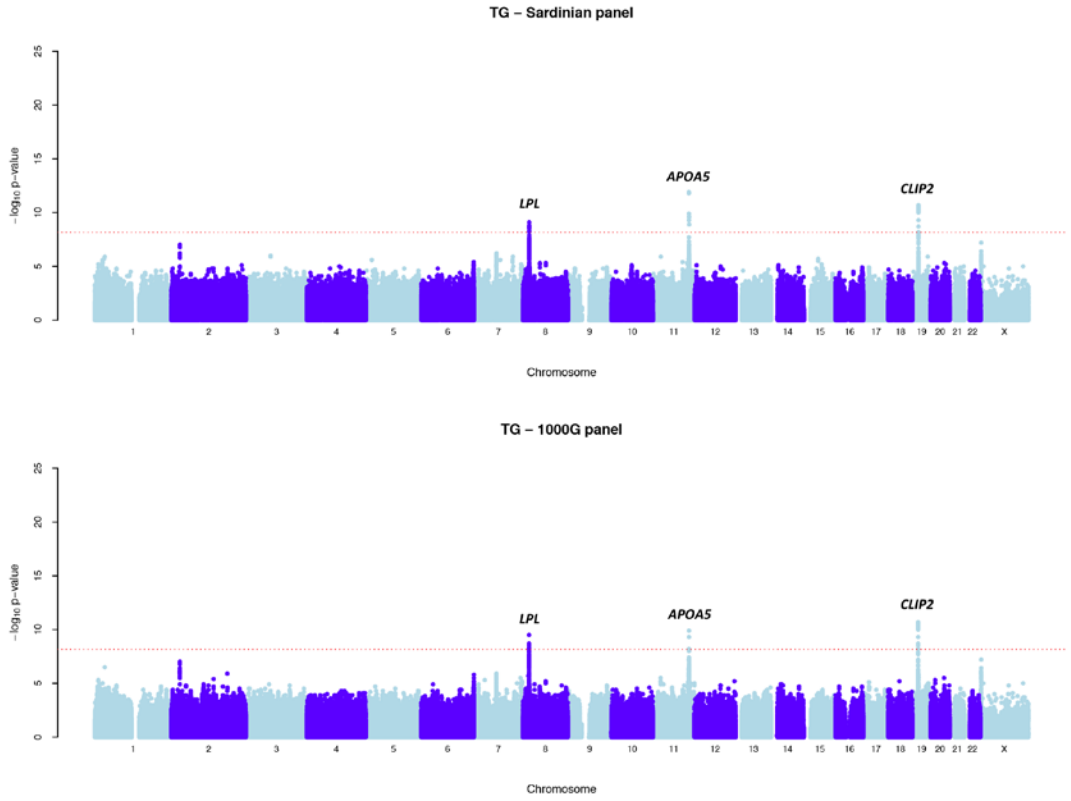
B)


TC – Sardinian panel
APOE
HBB
PCSK9
TMEM33/
DCAF4L1
CLIP2
−log10 p−value
Chromosome
TC – 1000G panel
APOE
PCSK9
HBB
CLIP2
−log10 p−value
Chromosome

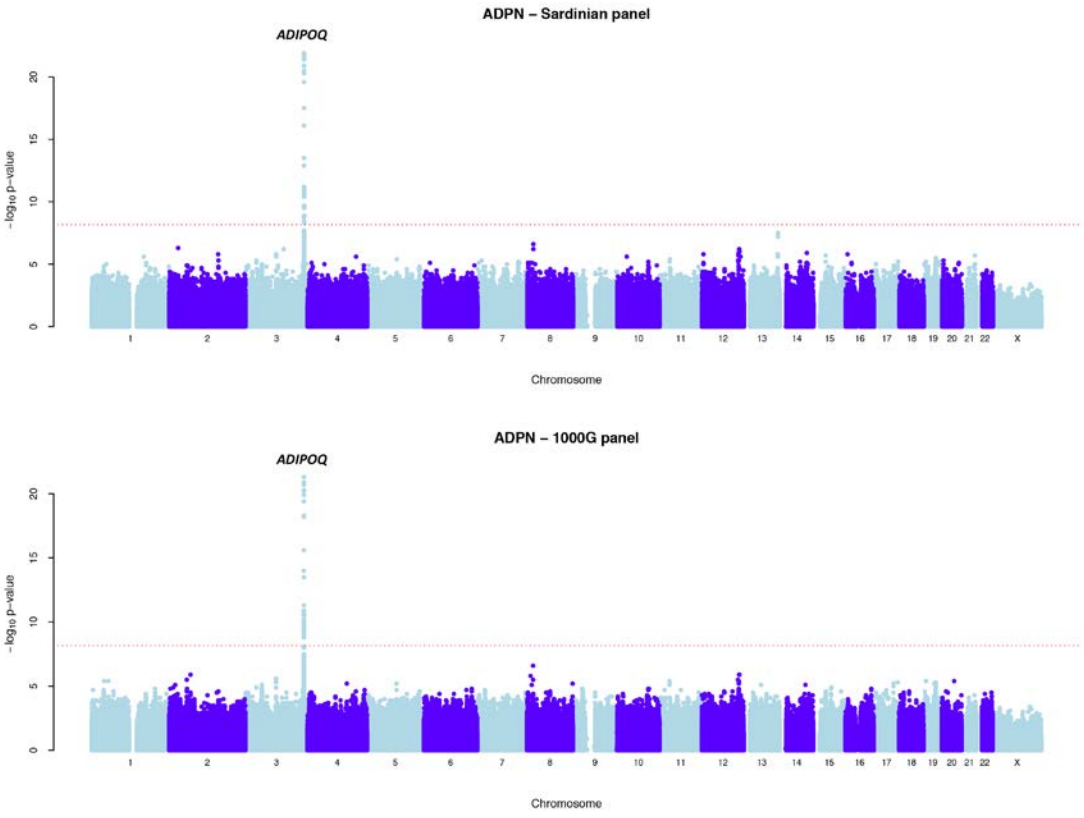
C)


HDL – Sardinian panel
CETP
LIPC
−log10 p−value
Chromosome
HDL – 1000G panel
CETP
LIPC
−log10 p−value
Chromosome

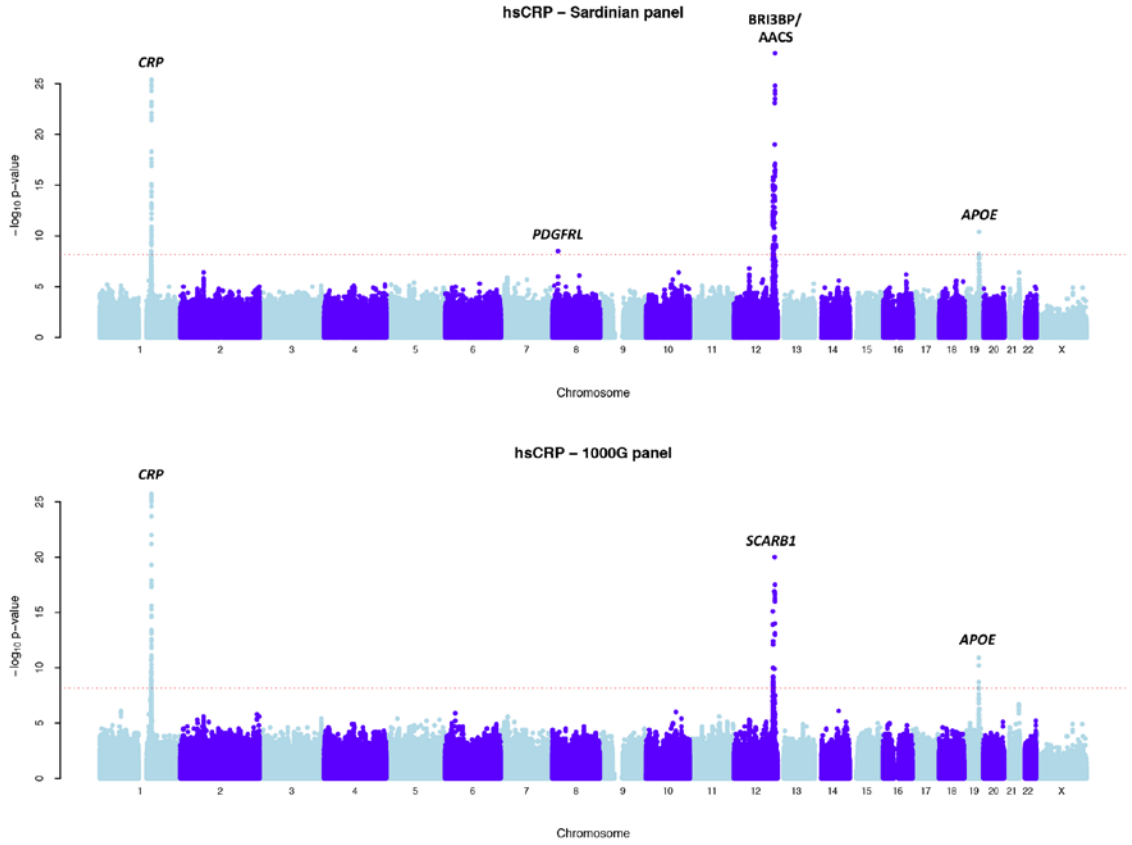
D)


TG – Sardinian panel
−log10 p−value
LPL
APOA5
CLIP2
Chromosome
TG – 1000G panel
−log10 p−value
LPL
APOA5
CLIP2
Chromosome

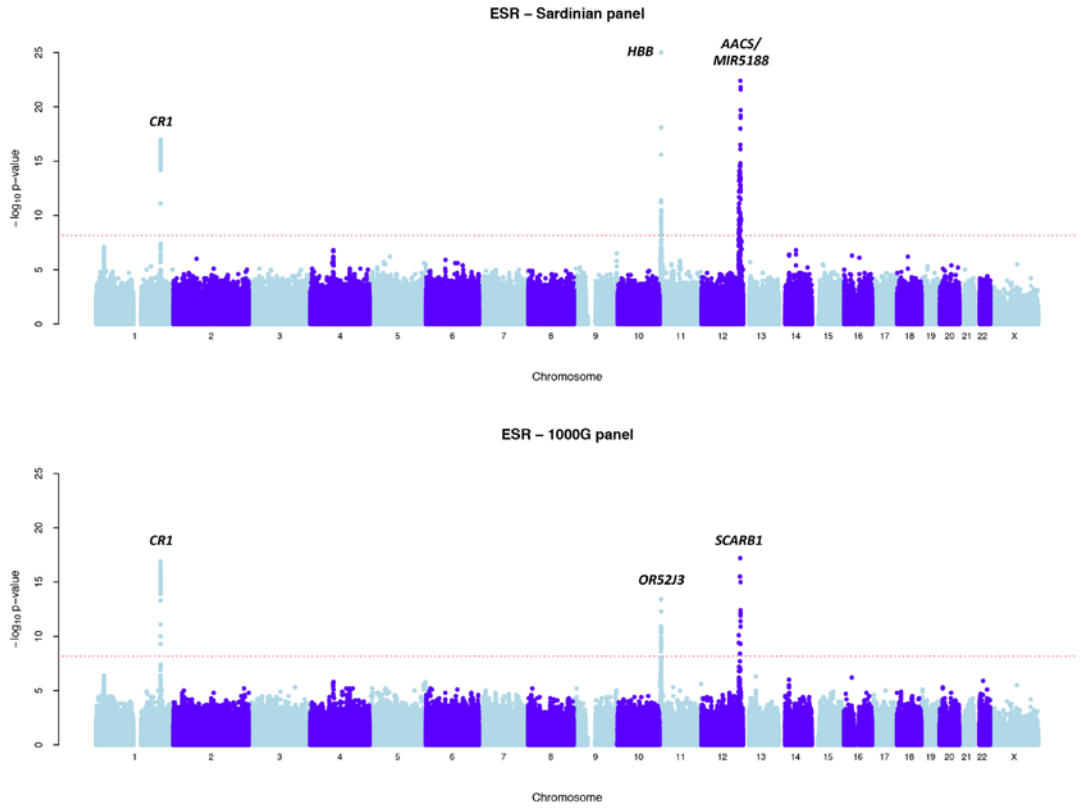
E)


ADPN – Sardinian panel
−log10 p−value
ADIPOQ
Chromosome
ADPN – 1000G panel
−log10 p−value
ADIPOQ
Chromosome

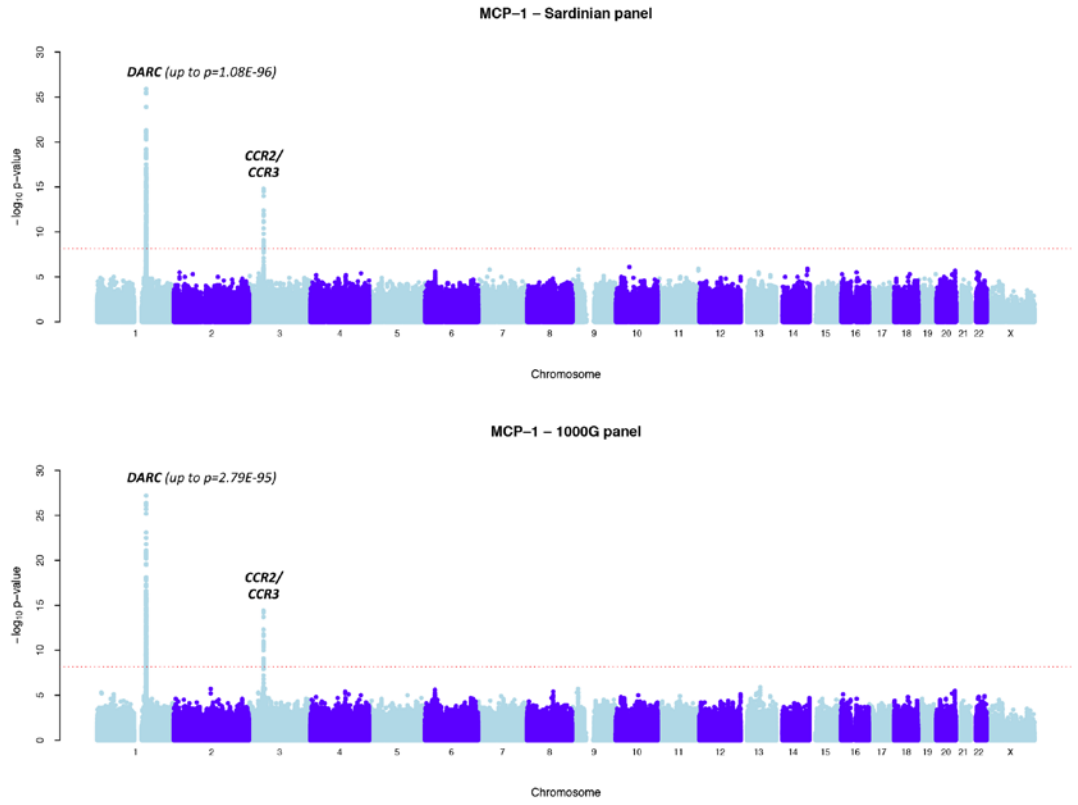
F)


hsCRP – Sardinian panel
CRP
BRI3BP/
AACS
PDGFRL
APOE
−log10 p-value
Chromosome
hsCRP – 1000G panel
CRP
SCARB1
APOE
−log10 p-value
Chromosome

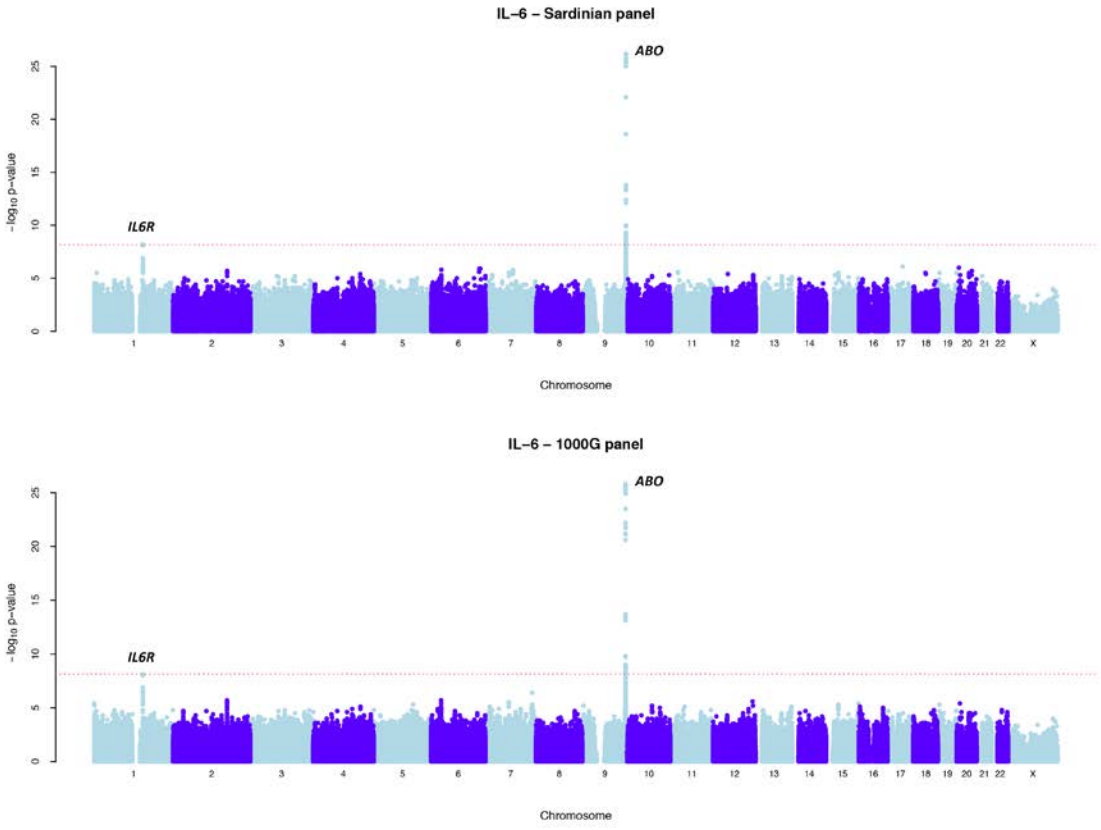
G)


ESR – Sardinian panel
HBB
AACS/
MIR5188
CR1
−log10 p-value
Chromosome
ESR – 1000G panel
CR1
SCARB1
OR52J3
−log10 p-value
Chromosome

H)

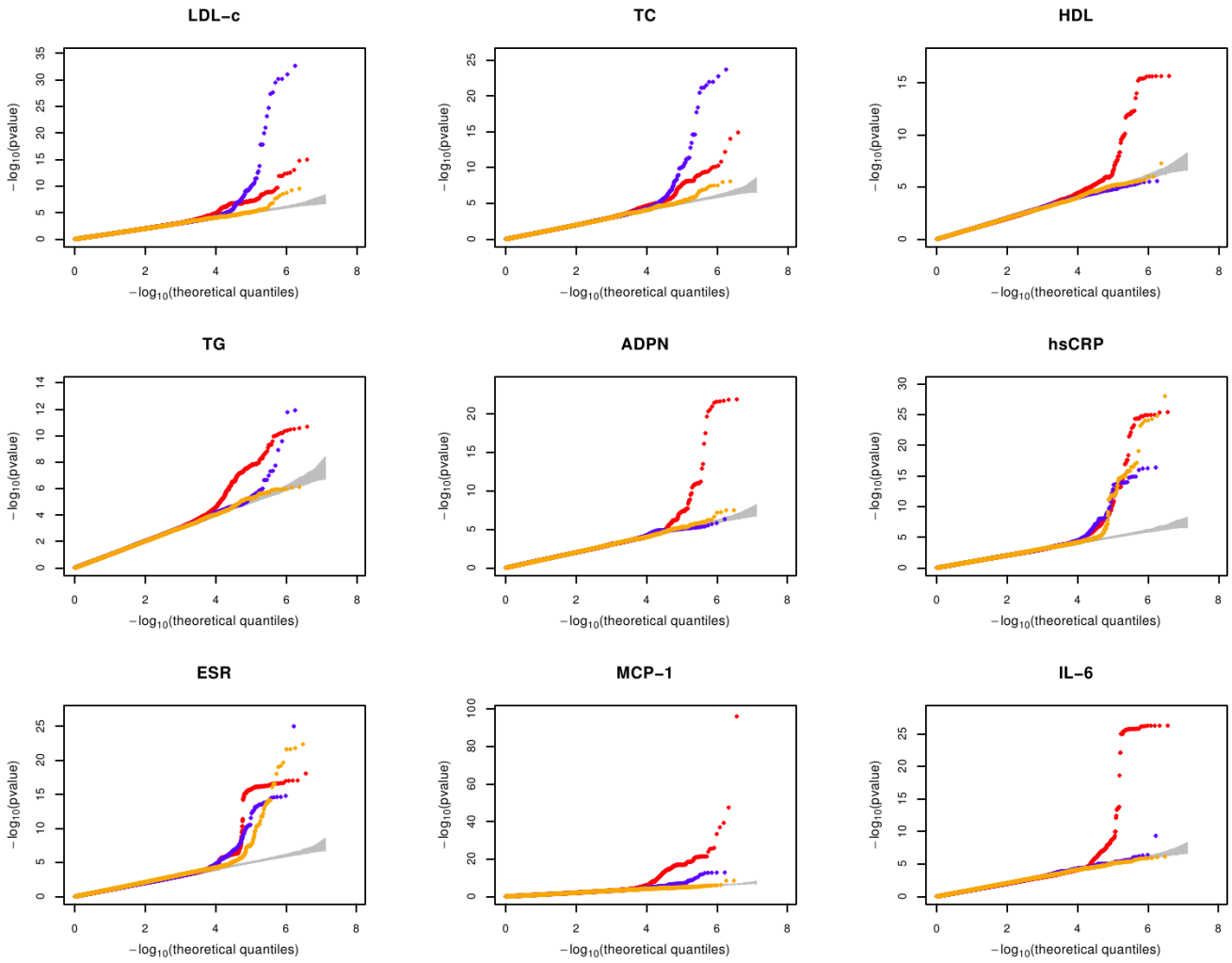

MCP–1 – Sardinian panel
DARC (up to p=1.08E-96)
CCR2/
CCR3
−log10 p−value
Chromosome
MCP–1 – 1000G panel
DARC (up to p=2.79E-95)
CCR2/
CCR3
−log10 p−value
Chromosome

I)


IL–6 – Sardinian panel
ABO
IL6R
−log10 p−value
Chromosome
IL–6 – 1000G panel
ABO
IL6R
−log10 p−value
Chromosome

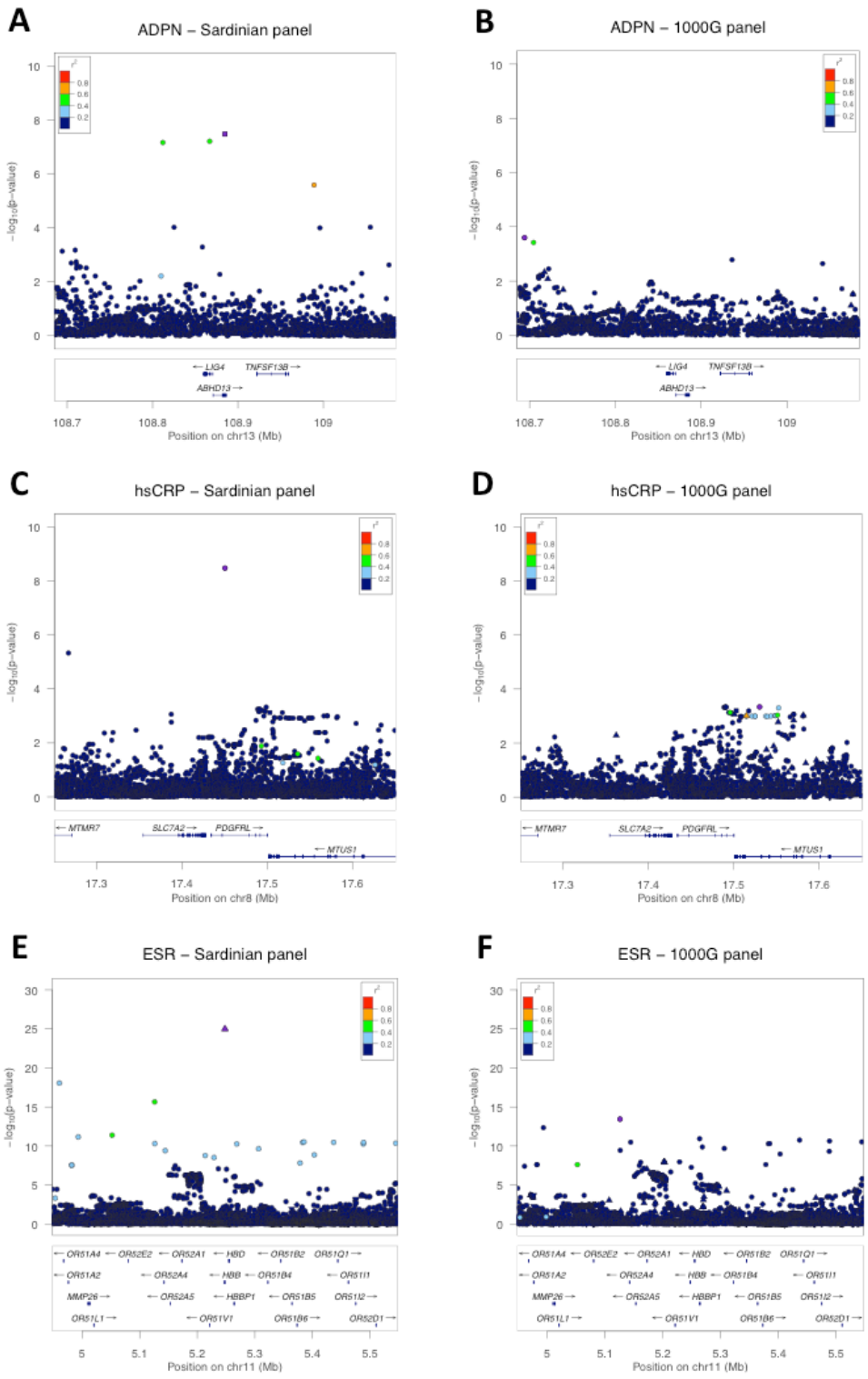
SUPPLEMENTARY FIGURE 7. QQplots for all studied traits

The figure illustrates the quantile-quantile plot of observed pvalues versus those expected under the null distribution, separately for common SNPs (MAF>5%, red), low frequency SNPs (MAF>1% and <5%, blue) and rare SNPs (MAF<1%, orange). Grey bars represent 90% confidence intervals.



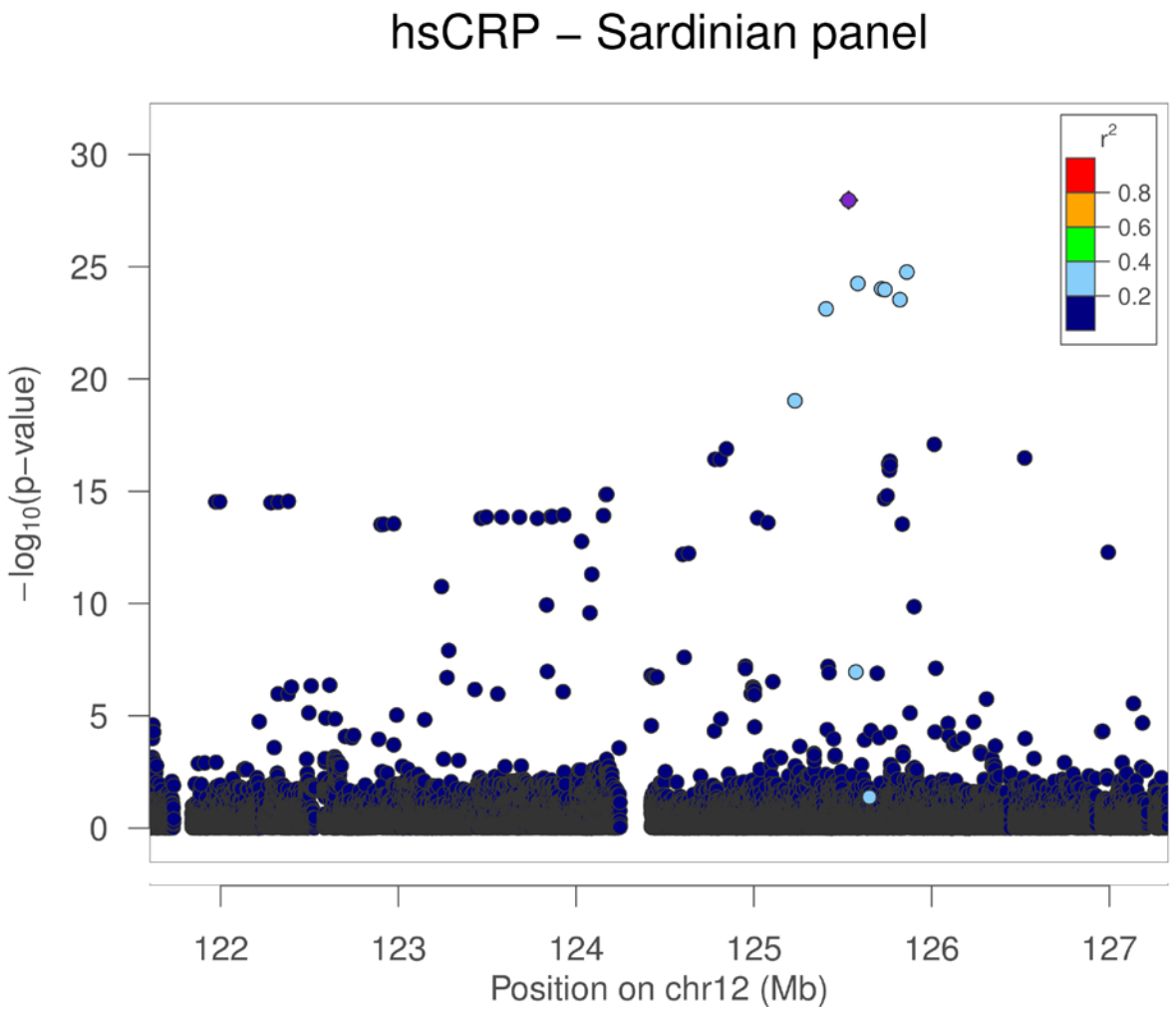
SUPPLEMENTARY FIGURE 8. Regional association plots

Regional association plots at the *ABDH13* locus for ADPN, at *PDGFRL* for hsCRP, and at *HBB* for ESR, for imputation performed using the Sardinian (panels A, C, E) and 1000 Genomes (panels B, D, F) reference panels, respectively. Figure style is identical to **Figure 3.**



SUPPLEMENTARY FIGURE 9. Extended associated region on chr12

Regional association plot of the 5.4Mb region on chr12. The results refer to association with hsCRP when the Sardinian panel is used for imputation. For a description of the figure style see Figure 3's legend.



SUPPLEMENTARY FIGURE 10. Variance explained by the accessible genome variants

In this plot, bars represent, for each assessed quantitative trait, the heritability and the variance explained by all variants in the accessible genome (see **Methods**), when the Sardinian or 1000 Genomes imputed data are considered (color coded blu and red, respectively). The white box indicates the estimated heritability in the SardiNIA sample.