Supplementary Material

Association of *SPI1* Haplotypes with Altered *SPI1* Gene Expression and Alzheimer's Disease Risk

Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset

We obtained genetic and clinical data from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (http://adni.loni.usc.edu). The ADNI was launched in 2003 as a public–private partnership led by Principal Investigator Michael W. Weiner, MD. Its primary goal is to test whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessments can be combined to measure the progression of mild cognitive impairment (MCI) and early Alzheimer's disease (AD). For up-to-date details, please visit http://www.adni-info.org. The genetic data we obtained included single-nucleotide polymorphism (SNP) array data from ADNI 1 (n = 756; Illumina Human610-Quad BeadChip), ADNI GO/2 (n = 793; Illumina HumanOmniExpress BeadChip), and ADNI 3 (n = 327) as well as whole-genome sequencing (WGS) data from ADNI WGS (n = 812). Some participants from ADNI WGS overlap with those from ADNI 1 (n = 260) and ADNI GO/2 (n = 427). The latest diagnostic records (updated March 2019) were used in the present study. Finally, we obtained genotypic data for 2,001 participants including 685 normal controls (NCs), 632 patients with MCI, and 684 patients with AD.

Alzheimer's Disease Genetics Consortium (ADGC) Genome Wide Association Study–National Institute on Aging (NIA) Alzheimer's Disease Centers (ADC) Cohort dataset

We obtained genotypic and phenotypic data from the National Institute on Aging **(**NIA) Alzheimer's Disease Centers (ADC) Cohort (dbGaP accession number: phs000372.v2.p1). The SNP array data of the ADC cohort groups 6,065 participants in 3 batches (i.e., ADC1, ADC2, and ADC3). Genotypes are generated using the Illumina Human660W-Quad BeadChip (ADC1) or Illumina HumanOmniExpress Array (ADC2 and ADC3). All autopsied subjects were ≥60 years old at time of death. Diagnosis was determined using DSM-IV criteria or a Clinical Dementia Rating (CDR) ≥1. Please refer to the corresponding database from the Genotypes and Phenotypes (dbGaP) Project for details [1,2]. We included 5,692 participants with a definite diagnosis of AD (n = 3,946) or NCs (n = 1,746) in the present study.

National Institute on Aging (NIA)–Late Onset Alzheimer's Disease (LOAD) Family Study dataset

We obtained genotypic and phenotypic data from the NIA–Late Onset Alzheimer's Disease (LOAD) Family Study (dbGaP accession number: phs000168.v2.p2), specifically data from 3 LOAD datasets: General Research Use, Disease-Specific (Alzheimer's Disease), and Disease-Specific (Alzheimer's Disease, Non-Profit Use). The genotypic data of 5,192 participants was generated using the Illumina Human 610Quadv1_B BeadChip. Please refer to the corresponding dbGaP Project for details [3]. We included the genetic information of 2,695 participants—including those with a definite diagnosis of AD (n = 464) and NCs (n = 2,232)—in the replication study.

Genotype-Tissue Expression (GTEx) Project dataset

We obtained the data used in the analyses described herein from dbGaP accession number phs000424.v8.p2 [4]. We retrieved the genotypic and transcriptomic profiles of all available tissues from the Genotype-Tissue Expression (GTEx) dataset.

BRAINEAC database

The *SPI1* transcript level data (probe id: 3372176) of the frontal cortex and *SPI1* genotypes (available variants residing within ±1Mb) was downloaded from the BRAINEAC database [5].

Single-nucleus RNA sequencing data of the human frontal cortex

The single-nucleus RNA sequencing (snRNA-seq) data used in this study was described in previous publication (GSE157827) [6]. The dataset comprises 21 prefrontal cortex tissue samples from patients with AD (n = 12) and NCs (n = 9). Log normalized transcript level was used in this study.

Cell type deconvolution analysis

We estimated the proportion of the different cell types of the GTEx brain tissues by CIBERSORTx [7]. We first utilized the aforementioned snRNA-seq data to calculate marker genes of brain cell types using the *FindAllMarkers* function from the Seurat package in R [8]. We then

subjected the expression matrix of the top 200 marker genes per cell type (i.e., neuron, astrocyte, microglia, oligodendrocyte, oligodendrocyte progenitor cell, and endothelial cell) to CIBERSORTx as the reference. We calculated relative cell type scores using S-mode batch correction.

Cardiogenics Study dataset

We obtained genetic and transcriptomic data from the Cardiogenics Study dataset (EGAC00001000088, European Genome-phenome Archive) [9], which included 758 monocytes and 599 macrophages. We generated the genotype data using Illumina Human Custom 1,2M and Human 610 Quad Custom arrays, and tested transcriptomic profiling using the Illumina HumanRef-8 v3.0 BeadChip.

Supplementary notes

For the Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset, data collection and sharing for this project was funded by ADNI (National Institutes of Health Grant U01 AG024904) and the Department of Defense (DOD) ADNI (DOD award number: W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging (NIA), the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie; Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC; Johnson & Johnson Pharmaceutical Research & Development LLC; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research provides funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (http://www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory of Neuro Imaging at the University of Southern California. For the Alzheimer's Disease Genetics Consortium (ADGC) Genome Wide Association Study–NIA Alzheimer's Disease Centers (ADC) Cohort dataset, funding support for the ADGC was provided through the Division of Neuroscience, NIA (grant number: U01-AG032984). For the NIA–Late Onset Alzheimer's Disease (LOAD) Family Study dataset, funding support for the "Genetic Consortium for Late Onset Alzheimer's Disease" was provided through the Division of Neuroscience, NIA. The Genetic Consortium for Late Onset Alzheimer's Disease includes a genome-wide association study funded as part of the Division of Neuroscience, NIA. Finally, the Genetic Consortium for Late Onset Alzheimer's Disease provided assistance with phenotype harmonization and genotype cleaning as well as general study coordination. The Genotype-Tissue Expression (GTEx) Project was supported by the Common Fund of the Office of the Director of the National Institutes of Health and by the NCI,

NHGRI, NHLBI, NIDA, NIMH, and NINDS. For the Cardiogenics Study, we thank Dr. David-Alexandre Trégouët for his help in obtaining the data for this study.

Supplementary Table 1. Demographic characteristics of Alzheimer's disease cohorts

Cohort	Variable	NC	AD	*p*
Hong Kong	*n*	721	333	
	Male ratio	0.52	0.321	1.8E-09
	Age (±sd)	78.2 (±6.08)	80.5 (±6.02)	1.9E-08
	APOE ε2	0.092	0.054	8.4E-03
	APOE ε4	0.08	0.201	1.4E-14
ADC1	*N*	756	2,047	
	Male ratio	0.376	0.469	1.0E-05
	Age (±sd)	75.0 (±10.8)	79.4 (±7.76)	2.2E-16
	APOE ε2	0.082	0.03	1.1E-14
	APOE ε4	0.141	0.396	2.2E-16
ADC2	*N*	308	774	
	Male ratio	0.292	0.478	2.0E-08
	Age (±sd)	76.7 (±7.30)	80.4 (±6.69)	1.4E-14
	APOE ε2	0.0974	0.0368	3.8E-08
	APOE ε4	0.153	0.378	2.2E-16
ADC3	*N*	682	1,125	
	Male ratio	0.343	0.419	1.5E-03
	Age (±sd)	75.6 (±8.70)	79.9 (±8.11)	2.2E-16
	APOE ε2	0.08	0.04	3.3E-06
	APOE ε4	0.146	0.378	2.2E-16
ADNI	*N*	685	684	
	Male ratio	0.445	0.579	8.4E-07
	Age (±sd)	72.0 (±6.97)	74.6 (±7.37)	3.7E-11
	APOE ε2	0.067	0.034	2.1E-05
	APOE ε4	0.156	0.4	2.2E-16
LOAD	*N*	2,232	464	
	Male ratio	0.39	0.31	1.3E-03
	Age (±sd)	80.7 (±10.7)	83.8 (±6.62)	9.2E-16
	APOE ε2	0.069	0.028	2.9E-06
	APOE ε4	0.208	0.463	2.2E-16

AD, Alzheimer's disease; ADC1, Alzheimer's Disease Center Dataset 1; ADC2, Alzheimer's Disease Center Dataset 2; ADC3, Alzheimer's Disease Center Dataset 3; ADNI, Alzheimer's Disease Neuroimaging Initiative; LOAD, Late Onset Alzheimer's Disease Family Study; NC, normal control; SD, standard deviation.

Supplementary Table 2. Associations of the 3 *SPI1* single-nucleotide polymorphisms with Alzheimer's disease in the Chinese population

Group	SNP	Diagnosis	Genotypic test					Allelic test			
All	rs1057233		AA	Ag	gg	HWE	*p*	A	g	OR (95% CI)	*p*
		NC	387	270	62	0.13	1.1E-01	325	394	0.70 (0.54-0.93)	1.1E-02
		AD	201	107	23	0.12		178	153		
	rs3740688		TT	Tg	gg	HWE	*p*	T	g	OR (95% CI)	*p*
		NC	340	307	72	0.86	1.0E-02	268	451	0.60 (0.46-0.79)	2.1E-04
		AD	191	116	26	0.19		165	168		
	rs78245530		GG	Gt	tt	HWE	*p*	G	t	OR (95% CI)	*p*
		NC	667	53	1	1.00	8.0E-01	666	55	0.86 (0.49-1.46)	6.1E-01
		AD	311	22	0	1.00		311	22		
APOE33	rs1057233		AA	Ag	gg	HWE	*p*	A	g	OR (95% CI)	*p*
		NC	266	182	43	0.14	5.8E-01	223	268	0.79 (0.55-1.12)	1.9E-01
		AD	105	61	13	0.31		92	87		
	rs3740688		TT	Tg	gg	HWE	*p*	T	g	OR (95% CI)	*p*
		NC	229	214	49	1.00	6.6E-02	180	312	0.59 (0.41-0.85)	3.3E-03
		AD	102	65	13	0.56		89	91		
	rs78245530		GG	Gt	tt	HWE	*p*	G	t	OR (95% CI)	*p*
		NC	454	38	1	0.56	7.2E-01	453	40	0.74 (0.33-1.51)	5.1E-01
		AD	169	11	0	1.00		169	11		

AD, Alzheimer's disease; CI, confidence interval; HWE, Hardy–Weinberg equilibrium; NC, normal control; OR, odds ratio; SNP, single-nucleotide polymorphism. Letters in upper and lower case denote major and minor alleles, respectively.

Supplementary Table 3. Frequencies of identified *SPI1* haplotypes

Name	rs1057233-rs3740688-rs78245530	HK	ADC1	ADC2	ADC3	ADNI	LOAD
α	A-T-G	0.695	0.547	0.546	0.555	0.541	0.529
β	g-g-G	0.242	0.316	0.298	0.310	0.312	0.325
γ	A-g-t	0.027	0.106	0.130	0.112	0.117	0.112
δ	A-g-G	0.017	0.007	0.005	0.006	0.010	0.008
ε	A-T-t	0.242	0.019	0.019	0.014	0.015	0.023

ADC1, Alzheimer's Disease Center Dataset 1; ADC2, Alzheimer's Disease Center Dataset 2; ADC3, Alzheimer's Disease Center Dataset 3; ADNI, Alzheimer's Disease Neuroimaging Initiative; HK, Hong Kong Chinese AD cohort; LOAD, Late Onset Alzheimer's Disease Family Study

Supplementary Table 4. Meta-analysis of identified common *SPI1* haplotypes in Alzheimer's disease cohorts of European descent

Participants	Haplotype	Cohorts	*Beta*[a] (RE)	SE (RE)	*p* (RE2)	I^2 (%)[b]	Q[c]	*p* (Q)
All	β	5	−0.040	0.037	0.336	0.0	1.363	0.851
	γ	5	−0.060	0.054	0.325	0.0	3.132	0.536
	η	5	−0.269	0.223	0.134	58.7	9.681	0.046
APOE adjusted	β	5	−0.028	0.040	0.560	0.0	2.352	0.671
	γ	5	−0.109	0.067	0.084	21.7	5.110	0.276
	η	5	−0.165	0.271	0.195	66.7	12.021	0.017
APOE33	β	5	−0.028	0.057	0.687	0.0	0.437	0.979
	γ	5	−0.213	0.137	0.030	55.8	9.046	0.060
	η	5	0.008	0.233	0.959	13.5	4.625	0.328

[a] Estimated effect size; [b] percentage of variation across studies due to heterogeneity; [c] Cochran's Q-test for heterogeneity among studies. RE, random effects model; RE2, Han and Eskin's Random Effects model; SE, standard error.

Supplementary Table 5. Meta-analysis of haplotypes β and γ in both populations of Chinese and European descent

Group	Haplotype	Cohorts	*Beta*[a] (RE)	SE (RE)	*p* (RE2)	I^2 (%)[b]	Q[c]	*p* (Q)
All	β	6	−0.073	0.043	0.079	29.9	7.135	0.211
	γ	6	−0.082	0.064	0.198	27.0	6.847	0.232
APOE adjusted	β	6	−0.070	0.053	0.135	46.6	9.360	0.096
	γ	6	−0.132	0.075	0.044	36.2	7.832	0.166
APOE33	β	6	−0.061	0.054	0.322	0.0	3.222	0.666
	γ	6	−0.248	0.133	0.015	52.1	10.442	0.064

[a] Estimated effect size; [b] percentage of variation across studies due to heterogeneity; [c] Cochran's Q-test for heterogeneity across studies. RE, random effects model; RE2, Han and Eskin's Random Effects model; SE, standard error.

Supplementary Table 6. Survival analysis of haplotypes β and γ with age of Alzheimer's disease onset in the Late Onset Alzheimer's Disease Family Study (LOAD) cohort

		Haplotype β				Haplotype γ			
Group	*n*	*Beta*[a]	SE	z	*p*	*Beta*	SE	z	*p*
All	2,696	−0.087	0.071	−1.224	0.221	−0.030	0.106	−0.287	0.774
APOE adjusted	2,696	−0.055	0.071	−0.780	0.436	−0.008	0.105	−0.076	0.939
APOE33	1,248	0.033	0.149	0.219	0.827	−0.611	0.290	−2.108	0.035

[a] Estimated effect size; SE, standard error.

Supplementary Table 7. Association analysis of haplotypes β and γ with cognitive performance in the Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset

		Haplotype β				Haplotype γ			
Group	*n*	*Beta*[a]	SE	t	*p*	*Beta*	SE	t	*p*
All	1,972	–0.008	0.030	–0.261	0.794	0.079	0.046	1.713	0.087
APOE adjusted	1,972	–0.018	0.029	–0.608	0.543	0.094	0.044	2.132	0.033
APOE33	935	–0.014	0.042	–0.336	0.737	0.173	0.066	2.621	0.009

[a] Estimated effect size; SE, standard error.

Supplementary Table 8. *SPI1* transcript level in human tissues obtained from the Genotype-Tissue Expression Project dataset

Tissue	Mean (TPM)	SD (TPM)	Median (TPM)
Whole Blood	1,018.2	627.6	934.7
Spleen	193.2	74.6	185.0
Lung	136.7	57.0	128.5
Cells-EBV-Transformed Lymphocytes	37.8	8.5	36.9
Adipose-Visceral (Omentum)	40.3	26.5	33.9
Adipose-Subcutaneous	36.3	23.3	31.8
Artery-Coronary	36.8	31.9	29.2
Small Intestine-Terminal Ileum	33.5	22.4	25.9
Artery-Aorta	28.8	22.7	23.6
Nerve-Tibial	24.4	10.4	22.8
Brain-Spinal Cord (Cervical C1)	23.6	18.7	19.1
Adrenal Gland	18.7	9.5	17.5
Breast-Mammary Tissue	21.0	19.9	17.2
Colon-Transverse	15.6	11.2	13.0
Minor Salivary Gland	16.4	13.6	11.9
Prostate	13.6	9.8	11.3
Skin-Not Sun Exposed (Suprapubic)	12.1	5.5	11.2
Skin-Sun Exposed (Lower Leg)	12.0	5.2	11.1
Colon-Sigmoid	12.9	8.1	11.0
Esophagus-Muscularis	12.9	8.0	11.0
Thyroid	13.1	8.9	10.6
Artery-Tibial	14.1	12.5	10.5
Esophagus-Gastroesophageal Junction	12.5	8.1	10.3
Uterus	10.8	6.8	9.5
Vagina	11.0	7.6	9.3
Brain-Substantia Nigra	13.1	11.2	9.1
Stomach	10.6	10.4	7.0
Brain-Hypothalamus	9.0	7.8	6.9
Heart-Atrial Appendage	8.8	7.1	6.9
Esophagus-Mucosa	8.2	7.9	6.7
Pituitary	10.5	12.4	6.4
Ovary	7.6	7.1	6.0
Liver	8.2	9.4	5.8
Brain-Caudate (Basal Ganglia)	7.3	6.3	5.6
Brain-Nucleus Accumbens (Basal Ganglia)	7.7	7.0	5.4
Kidney-Cortex	8.9	11.7	5.4
Brain-Amygdala	8.2	9.0	5.3
Brain-Hippocampus	7.9	12.5	4.8
Brain-Putamen (Basal Ganglia)	6.1	5.4	4.6
Heart-Left Ventricle	5.2	5.3	4.0
Brain-Anterior Cingulate Cortex (BA24)	5.7	7.5	3.9

Brain-Frontal Cortex (BA9)	6.4	10.0	3.6
Brain-Cortex	4.7	4.3	3.3
Brain-Cerebellar Hemisphere	3.7	3.3	2.9
Pancreas	4.0	4.0	2.6
Brain-Cerebellum	3.0	3.0	2.5
Muscle-Skeletal	2.5	1.9	2.0

SD, standard deviation; TPM, transcripts per kilobase million. Tissues are ranked according to median TPM.

Supplementary Table 9. Associations between haplotype γ and *SPI1* transcript level in human tissues obtained from the Genotype-Tissue Expression Project dataset

Tissue	*n*	Haplotype β			Haplotype γ		
		Beta	SE	*p*	*Beta*	SE	*p*
Adipose-Subcutaneous	582	−0.072	0.066	0.276	0.108	0.104	0.299
Adipose-Visceral (Omentum)	470	−0.052	0.072	0.469	0.016	0.106	0.883
Adrenal Gland	234	0.078	0.096	0.419	0.134	0.134	0.319
Artery-Aorta	388	0.016	0.080	0.839	0.002	0.120	0.986
Artery-Coronary	214	0.094	0.111	0.398	0.169	0.158	0.289
Artery-Tibial	585	0.052	0.064	0.419	−0.026	0.098	0.794
Brain-Amygdala	130	−0.090	0.142	0.526	−0.108	0.215	0.617
Brain-Anterior Cingulate Cortex (BA24)	148	0.120	0.147	0.416	−0.173	0.200	0.387
Brain-Caudate (Basal Ganglia)	195	0.204	0.116	0.082	−0.273	0.168	0.107
Brain-Cerebellar Hemisphere	176	0.033	0.126	0.791	0.047	0.173	0.787
Brain-Cerebellum	210	0.258	0.109	0.019	−0.266	0.159	0.096
Brain-Cortex	206	0.075	0.112	0.504	−0.447	0.166	0.008
Brain-Frontal Cortex (BA9)	176	0.058	0.125	0.646	−0.361	0.174	0.039
Brain-Hippocampus	166	−0.064	0.129	0.620	−0.045	0.182	0.804
Brain-Hypothalamus	171	0.072	0.119	0.545	−0.258	0.181	0.155
Brain-Nucleus Accumbens (Basal Ganglia)	203	0.018	0.111	0.869	−0.282	0.163	0.086
Brain-Putamen (Basal Ganglia)	171	0.053	0.120	0.662	−0.254	0.177	0.152
Brain-Spinal Cord (Cervical C1)	127	0.162	0.152	0.287	−0.086	0.206	0.676
Brain-Substantia Nigra	115	0.093	0.162	0.570	−0.251	0.230	0.278
Breast-Mammary Tissue	397	0.047	0.080	0.555	−0.193	0.122	0.115
Cells-EBV-Transformed Lymphocytes	148	0.097	0.126	0.443	0.063	0.194	0.745
Colon-Sigmoid	319	0.016	0.088	0.853	0.054	0.141	0.705
Colon-Transverse	369	0.063	0.081	0.435	−0.069	0.130	0.597
Esophagus-Gastroesophageal Junction	331	−0.171	0.086	0.046	0.228	0.127	0.073
Esophagus-Mucosa	498	−0.017	0.070	0.809	−0.023	0.106	0.827
Esophagus-Muscularis	466	−0.059	0.069	0.395	0.071	0.109	0.517
Heart-Atrial Appendage	373	−0.025	0.079	0.748	0.083	0.119	0.488
Heart-Left Ventricle	387	−0.002	0.080	0.982	−0.027	0.118	0.816
Kidney-Cortex	74	−0.020	0.162	0.902	−0.088	0.250	0.725
Liver	209	0.046	0.102	0.654	−0.127	0.147	0.391
Lung	516	0.013	0.071	0.853	−0.132	0.109	0.225
Minor Salivary Gland	145	−0.278	0.132	0.036	−0.095	0.215	0.661
Muscle-Skeletal	707	0.035	0.059	0.556	−0.046	0.095	0.629
Nerve-Tibial	533	0.031	0.066	0.646	0.066	0.105	0.528
Ovary	168	0.073	0.118	0.539	0.116	0.187	0.536
Pancreas	306	0.102	0.087	0.242	−0.105	0.134	0.431
Pituitary	238	−0.001	0.106	0.991	0.190	0.166	0.251
Prostate	222	0.111	0.110	0.315	−0.042	0.151	0.783
Skin-Not Sun Exposed (Suprapubic)	518	−0.030	0.070	0.671	−0.003	0.109	0.979
Skin-Sun Exposed (Lower Leg)	606	0.026	0.064	0.683	−0.048	0.100	0.636

Small Intestine-Terminal Ileum	175	0.157	0.126	0.214	−0.128	0.173	0.458
Spleen	228	−0.126	0.101	0.213	−0.127	0.158	0.425
Stomach	325	0.104	0.080	0.199	−0.043	0.116	0.709
Thyroid	575	0.088	0.066	0.185	−0.155	0.101	0.126
Uterus	130	−0.067	0.134	0.618	−0.216	0.200	0.282
Vagina	142	0.192	0.133	0.152	−0.164	0.192	0.394
Whole Blood	671	−0.119	0.060	0.047	0.083	0.093	0.372

[a] Estimated effect size. SE, standard error.

Supplementary Table 10. Associations between haplotype γ and *SPI1* transcript level controlling for the proportion of microglia

		Without microglia proportion adjustment				With microglia proportion adjustment			
Group	*n*	*Beta*[a]	SE	t	*p*	*Beta*	SE	t	*p*
All	206	−0.447	0.166	−2.694	0.008	−0.132	0.133	−0.994	0.321
APOE adjusted	206	−0.401	0.167	−2.398	0.017	−0.150	0.133	−1.133	0.259
APOE33	118	−0.693	0.223	−3.102	0.002	−0.468	0.180	−2.608	0.010

[a] Estimated effect size; SE, standard error.

Supplementary Table 11. Associations between *SPI* haplotypes and *SPI1* transcript level in the frontal cortex obtained from BRAINEAC database

		Haplotype β				Haplotype γ			
Group	*n*	*Beta*[a]	SE	t	*p*	*Beta*	SE	t	*p*
All	127	0.023	0.033	0.713	0.477	−0.059	0.069	−0.849	0.397
APOE adjusted	127	0.021	0.035	0.611	0.542	−0.073	0.070	−1.042	0.300
APOE33	82	0.037	0.046	0.810	0.420	−0.146	0.071	−2.069	0.042

[a] Estimated effect size; SE, standard error. *SPI1* transcript probe: 3372176

Supplementary Table 12. Associations between *SPI1* haplotypes and *SPI1* transcript level in monocytes and macrophages from the Cardiogenics study dataset

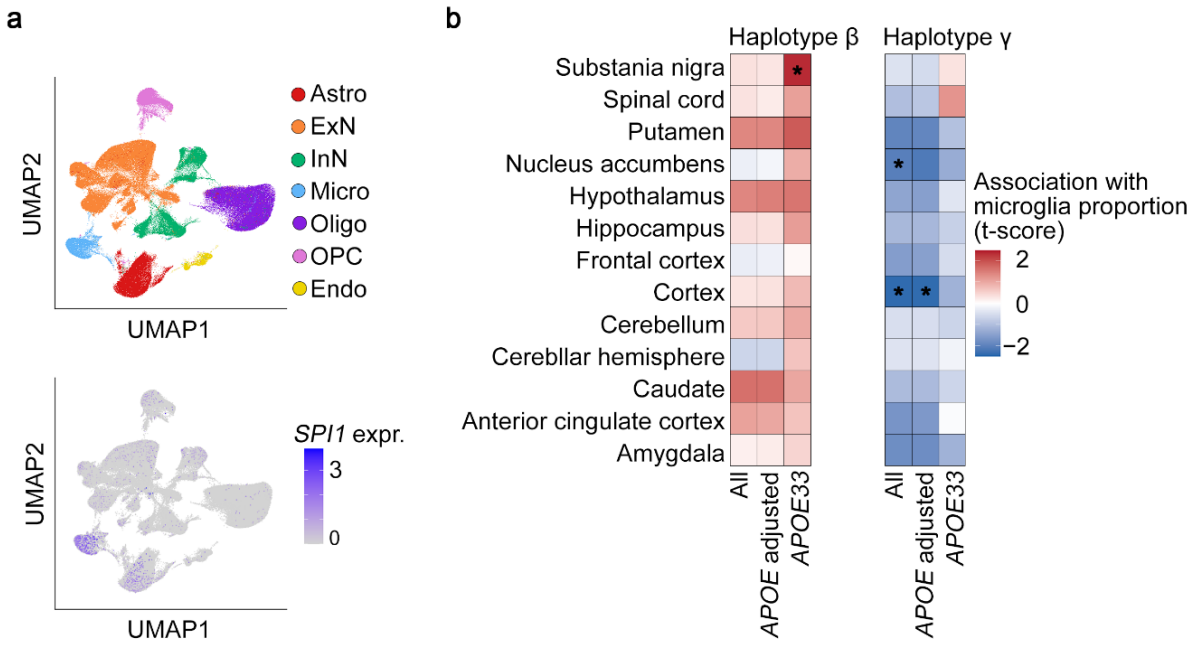
			Haplotype β				Haplotype γ			
Group	Cell type	*n*	*Beta*[a]	SE	t	*p*	*Beta*	SE	t	*p*
All	Monocyte	758	−0.257	0.032	−8.106	2.1E−15	0.059	0.045	1.315	1.9E−01
	Macrophage	599	−0.257	0.030	−8.550	2.0E−16	0.110	0.048	2.274	2.3E−02
APOE adjusted	Monocyte	758	−0.258	0.032	−8.175	1.3E−15	0.059	0.045	1.318	1.9E−01
	Macrophage	599	−0.254	0.029	−8.872	2.0E−16	0.111	0.049	2.288	2.3E−02
APOE33	Monocyte	459	−0.258	0.042	−6.191	1.3E−09	0.064	0.053	1.195	2.3E−01
	Macrophage	358	−0.254	0.041	−6.172	1.9E−09	0.081	0.060	1.348	1.8E−01

[a] Estimated effect size; SE, standard error. *SPI1* transcript probe: ILMN_1696463

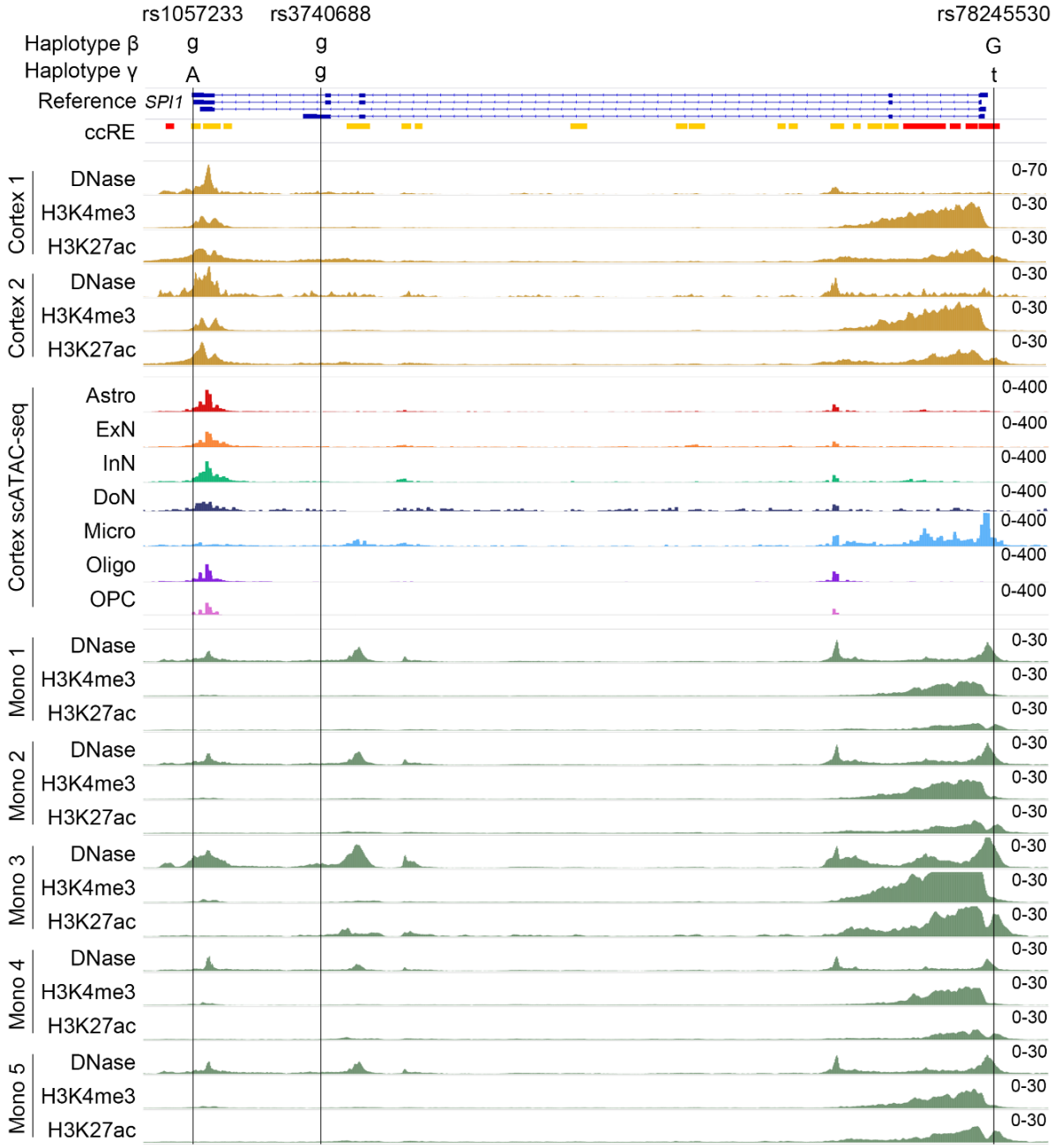
Supplementary Table 13. Micro RNA binding events potentially modulated by rs1057233

SNP	EA	miRNA	Type	Binding region (hg38)	ΔG binding[a]	Score
rs1057233	g	hsa-miR-6086	Gain	chr11:47354893–47354898	−19.65	23.49
rs1057233	g	hsa-miR-3689d	Gain	chr11:47354892–47354897	−11.83	20.55
rs1057233	g	hsa-miR-154-5p	Gain	chr11:47354894–47354899	−8.6	23.4
rs1057233	g	hsa-miR-569	Gain	chr11:47354896–47354902	−8.74	21.83
rs1057233	g	hsa-miR-377-5p	Gain	chr11:47354893–47354898	−13.26	23.99
rs1057233	g	hsa-miR-655-5p	Gain	chr11:47354893–47354898	−12.06	23.03
rs1057233	g	hsa-miR-4779	Gain	chr11:47354891–47354896	−11.41	20.53
rs1057233	g	hsa-miR-6851-5p	Gain	chr11:47354892–47354897	−19.96	22.05
rs1057233	g	hsa-miR-4695-5p	Gain	chr11:47354891–47354896	−21.53	21.53
rs1057233	g	hsa-miR-4683	Loss	chr11:47354892–47354897	−9.8	21.05
rs1057233	g	hsa-miR-5186	Loss	chr11:47354893–47354898	−11.61	21.99
rs1057233	g	hsa-miR-6888-5p	Loss	chr11:47354891–47354897	−11.84	20.58

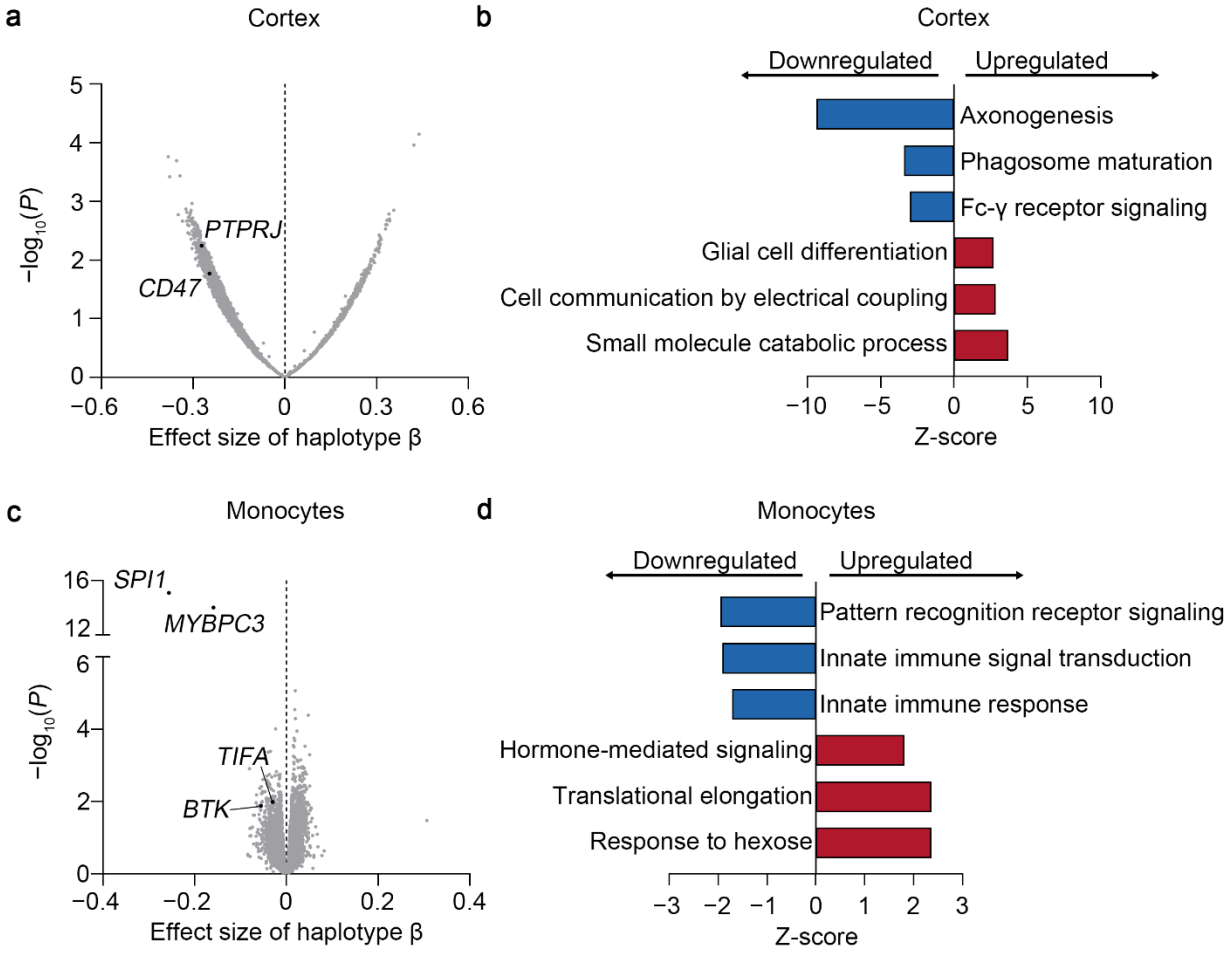
[a] Binding energy based on ensemble free energy. Data were obtained from the miRNASNP-v3 database [10]. EA, effect allele; miRNA, micro RNA; SNP, single-nucleotide polymorphism.



Supplementary Figure 1. Cell type analysis of *SPI1* gene expression in the brain tissues. a) Dimension reduction plot showing cell type-specific *SPI1* transcript level revealed by single-nucleus RNA sequencing in the human frontal cortex. b) Heatmap showing associations between *SPI1* haplotypes and microglia proportion estimated by CIBERSORTx for Genotype-Tissue Expression Project brain tissues. *$p < 0.05$. UMAP, Uniform Manifold Approximation and Projection; Astro, astrocyte; ExN, excitatory neuron; InN, Inhibitory neuron; Micro, microglia; Oligo, oligodendrocyte; OPC, oligodendrocyte progenitor cell.



Supplementary Figure 2. Epigenetic modification of the single-nucleotide polymorphism (SNP)-harboring region for *SPI1* haplotypes' tag SNPs. Panels from top to bottom are the gene structure and coordinates of *SPI1*. Boxes and lines denote exons and introns, respectively. Red and yellow bars represent candidate cis-regulatory elements (ccREs) with high H3K4me3 and H3K27ac signals, respectively. Signals of DNase sequencing (DNase-seq), H3K4me3 chromatin immunoprecipitation sequencing (ChIP-seq), and H3K27ac ChIP-seq in the cerebral cortex. Signals of single-cell assay for transposase-accessible chromatin using sequencing (scATAC-seq) in the cerebral cortex. Signals of DNase-seq, H3K4me3 ChIP-seq, and H3K27ac ChIP-seq in blood mononuclear cells. Astro, astrocyte; ExN, excitatory neuron; InN, Inhibitory neuron; Micro, microglia; Oligo, oligodendrocyte; OPC, oligodendrocyte progenitor cell.



Supplementary Figure 3. Associations between haplotype β and transcriptomic changes in the cerebral cortex and monocytes. a) Volcano plot showing the associations between haplotype β and gene expression in the cerebral cortex. b) Gene Ontology analysis of genes modulated by haplotype β in the cerebral cortex. c) Volcano plot showing the associations between haplotype β and gene expression in monocytes. d) Gene Ontology analysis of genes modulated by haplotype β in monocytes.

REFERENCES

[1] Jun G, Naj AC, Beecham GW, Wang LS, Buros J, Gallins PJ, Buxbaum JD, Ertekin-Taner N, Fallin MD, Friedland R, Inzelberg R, Kramer P, Rogaeva E, St George-Hyslop P, Cantwell LB, Dombroski BA, Saykin AJ, Reiman EM, Bennett DA, Morris JC, Lunetta KL, Martin ER, Montine TJ, Goate AM, Blacker D, Tsuang DW, Beekly D, Cupples LA, Hakonarson H, Kukull W, Foroud TM, Haines J, Mayeux R, Farrer LA, Pericak-Vance MA, Schellenberg GD (2010) Meta-analysis confirms CR1, CLU, and PICALM as Alzheimer disease risk loci and reveals interactions with APOE genotypes. *Arch Neurol* **67**, 1473–1484.

[2] Naj AC, Jun G, Beecham GW, Wang LS, Vardarajan BN, Buros J, Gallins PJ, Buxbaum JD, Jarvik GP, Crane PK, Larson EB, Bird TD, Boeve BF, Graff-Radford NR, De Jager PL, Evans D, Schneider JA, Carrasquillo MM, Ertekin-Taner N, Younkin SG, Cruchaga C, Kauwe JSK, Nowotny P, Kramer P, Hardy J, Huentelman MJ, Myers AJ, Barmada MM, Demirci FY, Baldwin CT, Green RC, Rogaeva E, George-Hyslop PS, Arnold SE, Barber R, Beach T, Bigio EH, Bowen JD, Boxer A, Burke JR, Cairns NJ, Carlson CS, Carney RM, Carroll SL, Chui HC, Clark DG, Corneveaux J, Cotman CW, Cummings JL, Decarli C, Dekosky ST, Diaz-Arrastia R, Dick M, Dickson DW, Ellis WG, Faber KM, Fallon KB, Farlow MR, Ferris S, Frosch MP, Galasko DR, Ganguli M, Gearing M, Geschwind DH, Ghetti B, Gilbert JR, Gilman S, Giordani B, Glass JD, Growdon JH, Hamilton RL, Harrell LE, Head E, Honig LS, Hulette CM, Hyman BT, Jicha GA, Jin LW, Johnson N, Karlawish J, Karydas A, Kaye JA, Kim R, Koo EH, Kowall NW, Lah JJ, Levey AI, Lieberman AP, Lopez OL, MacK WJ, Marson DC, Martiniuk F, Mash DC, Masliah E, McCormick WC, McCurry SM, McDavid AN, McKee AC, Mesulam M, Miller BL, Miller CA, Miller JW, Parisi JE, Perl DP, Peskind E, Petersen RC, Poon WW, Quinn JF, Rajbhandary RA, Raskind M, Reisberg B, Ringman JM, Roberson ED, Rosenberg RN, Sano M, Schneider LS, Seeley W, Shelanski ML, Slifer MA, Smith CD, Sonnen JA, Spina S, Stern RA, Tanzi RE, Trojanowski JQ, Troncoso JC, Van Deerlin VM, Vinters H V., Vonsattel JP, Weintraub S, Welsh-Bohmer KA, Williamson J, Woltjer RL, Cantwell LB, Dombroski BA, Beekly D, Lunetta KL, Martin ER, Kamboh MI, Saykin AJ, Reiman EM, Bennett DA, Morris JC, Montine TJ, Goate AM, Blacker D, Tsuang DW, Hakonarson H, Kukull WA, Foroud TM, Haines JL, Mayeux R, Pericak-

Vance MA, Farrer LA, Schellenberg GD (2011) Common variants at MS4A4/MS4A6E, CD2AP, CD33 and EPHA1 are associated with late-onset Alzheimer's disease. *Nat Genet* **43**, 436–443.

[3] Lee JH, Cheng R, Graff-Radford N, Foroud T, Mayeux R (2008) Analyses of the national institute on aging late-onset Alzheimer's disease family study: Implication of additional loci. *Arch Neurol* **65**, 1518–1526.

[4] Aguet F, Barbeira AN, Bonazzola R, Brown A, Castel SE, Jo B, Kasela S, Kim-Hellmuth S, Liang Y, Oliva M, Flynn ED, Parsana P, Fresard L, Gamazon ER, Hamel AR, He Y, Hormozdiari F, Mohammadi P, Muñoz-Aguirre M, Park YS, Saha A, Segrè A V., Strober BJ, Wen X, Wucher V, Ardlie KG, Battle A, Brown CD, Cox N, Das S, Dermitzakis ET, Engelhardt BE, Garrido-Martín D, Gay NR, Getz GA, Guigó R, Handsaker RE, Hoffman PJ, Im HK, Kashin S, Kwong A, Lappalainen T, Li X, MacArthur DG, Montgomery SB, Rouhana JM, Stephens M, Stranger BE, Todres E, Viñuela A, Wang G, Zou Y, Anand S, Gabriel S, Graubert A, Hadley K, Huang KH, Meier SR, Nedzel JL, Nguyen DT, Balliu B, Conrad DF, Cotter DJ, deGoede OM, Einson J, Eskin E, Eulalio TY, Ferraro NM, Gloudemans MJ, Hou L, Kellis M, Li X, Mangul S, Nachun DC, Nobel AB, Park Y, Rao AS, Reverter F, Sabatti C, Skol AD, Teran NA, Wright F, Ferreira PG, Li G, Melé M, Yeger-Lotem E, Barcus ME, Bradbury D, Krubit T, McLean JA, Qi L, Robinson K, Roche N V., Smith AM, Sobin L, Tabor DE, Undale A, Bridge J, Brigham LE, Foster BA, Gillard BM, Hasz R, Hunter M, Johns C, Johnson M, Karasik E, Kopen G, Leinweber WF, McDonald A, Moser MT, Myer K, Ramsey KD, Roe B, Shad S, Thomas JA, Walters G, Washington M, Wheeler J, Jewell SD, Rohrer DC, Valley DR, Davis DA, Mash DC, Branton PA, Sobin L, Barker LK, Gardiner HM, Mosavel M, Siminoff LA, Flicek P, Haeussler M, Juettemann T, Kent WJ, Lee CM, Powell CC, Rosenbloom KR, Ruffier M, Sheppard D, Taylor K, Trevanion SJ, Zerbino DR, Abell NS, Akey J, Chen L, Demanelis K, Doherty JA, Feinberg AP, Hansen KD, Hickey PF, Hou L, Jasmine F, Jiang L, Kaul R, Kellis M, Kibriya MG, Li JB, Li Q, Lin S, Linder SE, Pierce BL, Rizzardi LF, Smith KS, Snyder M, Stamatoyannopoulos J, Tang H, Wang M, Branton PA, Carithers LJ, Guan P, Koester SE, Little AR, Moore HM, Nierras CR, Rao AK, Vaught JB, Volpi S (2020) The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science* **369**, 1318–1330.

[5] Ramasamy A, Trabzuni D, Guelfi S, Varghese V, Smith C, Walker R, De T, Consortium UBE, Consortium NABE, Coin L, Silva R de, Cookson MR, Singleton AB, Hardy J, Ryten M, Weale ME (2014) Genetic variability in the regulation of gene expression in ten regions of the human brain. *Nat Neurosci* **17**, 1418.

[6] Lau SF, Cao H, Fu AKY, Ip NY (2020) Single-nucleus transcriptome analysis reveals dysregulation of angiogenic endothelial cells and neuroprotective glia in Alzheimer's disease. *Proc Natl Acad Sci U S A* **117**, 25800–25809.

[7] Newman AM, Steen CB, Liu CL, Gentles AJ, Chaudhuri AA, Scherer F, Khodadoust MS, Esfahani MS, Luca BA, Steiner D, Diehn M, Alizadeh AA (2019) Determining cell type abundance and expression from bulk tissues with digital cytometry. *Nat Biotechnol* **37**, 773–782.

[8] Hao Y, Hao S, Andersen-Nissen E, Mauck WM, Zheng S, Butler A, Lee MJ, Wilk AJ, Darby C, Zager M, Hoffman P, Stoeckius M, Papalexi E, Mimitou EP, Jain J, Srivastava A, Stuart T, Fleming LM, Yeung B, Rogers AJ, McElrath JM, Blish CA, Gottardo R, Smibert P, Satija R (2021) Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573-3587.e29.

[9] Garnier S, Truong V, Brocheton J, Zeller T, Rovital M, Wild PS, Ziegler A, Munzel T, Tiret L, Blankenberg S, Deloukas P, Erdmann J, Hengstenberg C, Samani NJ, Schunkert H, Ouwehand WH, Goodall AH, Cambien F, Trégouët D-A (2013) Genome-wide haplotype analysis of Cis expression quantitative trait loci in monocytes. *PLoS Genet* **9**, e1003240.

[10] Liu CJ, Fu X, Xia M, Zhang Q, Gu Z, Guo AY (2021) MiRNASNP-v3: A comprehensive database for SNPs and disease-related variations in miRNAs and miRNA targets. *Nucleic Acids Res* **49**, D1276–D1281.