

## Exome Chip Meta-Analysis Fine Maps Causal Variants and Elucidates the Genetic Architecture of Rare Coding Variants in Smoking and Alcohol Use

### Supplement 1

#### Complete summary statistics

Complete sets of summary statistics ~~will be made~~ are available for download here:  
<https://genome.psych.umn.edu/index.php/GSCAN>.

#### Analysis plan

The analysis plan used by all studies to generate summary statistics is here:  
<https://genome.psych.umn.edu/index.php/GSCAN>.

#### ~~Supplementary note on~~ Exploratory analyses of individuals of African ancestry

Using the same techniques as for individuals of European ancestry, we conducted a GWAS meta-analysis of three cohorts of African and African admixed ancestry. These cohorts were the UK Biobank, The Collaborative Study on the Genetics of Alcoholism (COGA), and the Health and Retirement Study (HRS). Sample sizes and genomic controls are provided in **Table S3**. African ancestry in the UK Biobank were identified through inspection of genetic principal component 1 against component 2. Individuals with values  $PC2 > 0$  and  $PC1 > 150$ .

Ultimately, one genome-wide significant hit (rs3806243,  $p = 2.3 \times 10^{-8}$ ) was associated with cigarettes per day at the conventional African-ancestry  $p < 2.5 \times 10^{-8}$  threshold. This locus had not been discovered in a prior larger meta-analysis of cigarettes per day in African American individuals [1]. Given the lack of replication in the larger sample and marginal statistical evidence, no further analyses were conducted. We encourage investigators to continue to build cohorts of non-European ancestry. QQ plots and Manhattan plots are provided in **Figures S3 and S4**.

### **Supplementary note on ~~s~~Sources of funding of individual studies:**

**COGA:** The Collaborative Study on the Genetics of Alcoholism (COGA), Principal Investigators B. Porjesz, V. Hesselbrock, H. Edenberg, L. Bierut, includes eleven different centers: University of Connecticut (V. Hesselbrock); Indiana University (H.J. Edenberg, J. Nurnberger Jr., T. Foroud); University of Iowa (S. Kuperman, J. Kramer); SUNY Downstate (B. Porjesz); Washington University in St. Louis (L. Bierut, J. Rice, K. Bucholz, A. Agrawal); University of California at San Diego (M. Schuckit); Rutgers University (J. Tischfield, A. Brooks); Department of Biomedical and Health Informatics, The Children's Hospital of Philadelphia; Department of Genetics, Perelman School of Medicine, University of Pennsylvania, Philadelphia PA (L. Almasy), Virginia Commonwealth University (D. Dick), Icahn School of Medicine at Mount Sinai (A. Goate), and Howard University (R. Taylor). Other COGA collaborators include: L. Bauer (University of Connecticut); J. McClintick, L. Wetherill, X. Xuei, Y. Liu, D. Lai, S. O'Connor, M. Plawecki, S. Lourens (Indiana University); G. Chan (University of Iowa; University of Connecticut); J. Meyers, D. Chorlian, C. Kamarajan, A. Pandey, J. Zhang (SUNY Downstate); J.-C. Wang, M. Kapoor, S. Bertelsen (Icahn School of Medicine at Mount Sinai); A. Anokhin, V. McCutcheon, S. Saccone (Washington University); J. Salvatore, F. Aliev, B. Cho (Virginia Commonwealth University); and Mark Kos (University of Texas Rio Grande Valley). A. Parsian and M. Reilly are the NIAAA Staff Collaborators.

We continue to be inspired by our memories of Henri Begleiter and Theodore Reich, founding PI and Co-PI of COGA, and also owe a debt of gratitude to other past organizers of COGA, including Ting-Kai Li, P. Michael Conneally, Raymond Crowe, and Wendy Reich, for their critical contributions. This national collaborative study is supported by NIH Grant U10AA008401 from the National Institute on Alcohol Abuse and Alcoholism (NIAAA) and the National Institute on Drug Abuse (NIDA).

**FTC:** Phenotyping and genotyping of the Finnish Twin Cohort (FTC) has been supported by the Academy of Finland Center of Excellence in Complex Disease Genetics (grants 213506, 129680), the Academy of Finland (grants 100499, 205585, 118555, 141054, 265240, 263278 and 264146 to J. Kaprio), National Institute for Health (grant DA12854 to P.A.F. Madden), National Institute of Alcohol Abuse and Alcoholism (grants AA-12502, AA-00145, and AA-09203 to R. J. Rose and AA15416 and K02AA018755 to D. M. Dick), Sigrid Juselius Foundation (to J. Kaprio), Global Research Award for Nicotine Dependence, Pfizer Inc. (to J. Kaprio), and the Wellcome Trust Sanger Institute, UK. Antti-Pekka Sarin and Samuli Ripatti are acknowledged for genotype data quality controls and imputation. Association analyses were run at the ELIXIR Finland node hosted at CSC – IT Center for Science for ICT resources.

**GECCO:** Support for this study came from the National Cancer Institute, National Institutes of Health, U.S. Department of Health and Human Services (U01 CA137088; R01CA059045). The authors also thank all those at the GECCO Coordinating Center for helping bring together the data and people that made this project possible.

#### Substudies of GECCO:

**ASTERISK:** a Hospital Clinical Research Program (PHRC-BRD09/C) from the University Hospital Center of Nantes (CHU de Nantes) and supported by the Regional Council of Pays de la Loire, the Groupement des Entreprises Françaises dans la Lutte contre le Cancer (GEFLUC), the Association Anne de Bretagne Génétique and the Ligue Régionale Contre le Cancer (LRCC). We are very grateful to Dr. Bruno Buecher without whom this project would not have existed. We also thank all those who agreed to participate in

this study, including the patients and the healthy control persons, as well as all the physicians, technicians and students.

**CPS-II:** The authors thank the CPS-II participants and Study Management Group for their invaluable contributions to this research. The authors would also like to acknowledge the contribution to this study from central cancer registries supported through the Centers for Disease Control and Prevention National Program of Cancer Registries, and cancer registries supported by the National Cancer Institute Surveillance Epidemiology and End Results program.

**HPFS, NHS:** We would like to acknowledge Patrice Soule and Hardeep Ranu of the Dana Farber Harvard Cancer Center High-Throughput Polymorphism Core who assisted in the genotyping for NHS, HPFS under the supervision of Dr. Immaculata Devivo and Dr. David Hunter, Qin (Carolyn) Guo and Lixue Zhu who assisted in programming for NHS and HPFS. We would like to thank the participants and staff of the Nurses' Health Study and the Health Professionals Follow-Up Study, for their valuable contributions as well as the following state cancer registries for their help: AL, AZ, AR, CA, CO, CT, DE, FL, GA, ID, IL, IN, IA, KY, LA, ME, MD, MA, MI, NE, NH, NJ, NY, NC, ND, OH, OK, OR, PA, RI, SC, TN, TX, VA, WA, WY. The authors assume full responsibility for analyses and interpretation of these data.

**PLCO:** Intramural Research Program of the Division of Cancer Epidemiology and Genetics and supported by contracts from the Division of Cancer Prevention, National Cancer Institute, NIH, DHHS. Additionally, a subset of control samples were genotyped as part of the Cancer Genetic Markers of Susceptibility (CGEMS) Prostate Cancer GWAS (Yeager, M et al. Genome-wide association study of prostate cancer identifies a second risk locus at 8q24. *Nat Genet* 2007 May;39(5):645-9), CGEMS pancreatic cancer scan (PanScan) (Amundadottir, L et al. Genome-wide association study identifies variants in the ABO locus associated with susceptibility to pancreatic cancer. *Nat Genet*. 2009 Sep;41(9):986-90, and Petersen, GM et al. A genome-wide association study identifies pancreatic cancer susceptibility loci on chromosomes 13q22.1, 1q32.1 and 5p15.33. *Nat Genet*. 2010 Mar;42(3):224-8), and the Lung Cancer and Smoking study (Landi MT, et al. A genome-wide association study of lung cancer identifies a region of chromosome 5p15 associated with risk for adenocarcinoma. *Am J Hum Genet*. 2009 Nov;85(5):679-91). The prostate and PanScan study datasets were accessed with appropriate approval through the dbGaP online resource (<http://cgems.cancer.gov/data/>) accession numbers phs000207.v1.p1 and phs000206.v3.p2, respectively, and the lung datasets were accessed from the dbGaP website (<http://www.ncbi.nlm.nih.gov/gap>) through accession number phs000093.v2.p2. Funding for the Lung Cancer and Smoking study was provided by National Institutes of Health (NIH), Genes, Environment and Health Initiative (GEI) Z01 CP 010200, NIH U01 HG004446, and NIH GEI U01 HG 004438. For the lung study, the GENEVA Coordinating Center provided assistance with genotype cleaning and general study coordination, and the Johns Hopkins University Center for Inherited Disease Research conducted genotyping. The authors thank Drs. Christine Berg and Philip Prorok, Division of Cancer Prevention, National Cancer Institute, the Screening Center investigators and staff or the Prostate, Lung, Colorectal, and Ovarian (PLCO) Cancer Screening Trial, Mr. Tom Riley and staff, Information Management Services, Inc., Ms. Barbara O'Brien and staff, Westat, Inc., and Drs. Bill Kopp and staff, SAIC-Frederick. Most importantly, we acknowledge the study participants for their contributions to making this study possible. The statements contained herein are solely those of the authors and do not represent or imply concurrence or endorsement by NCI.

**PMH:** National Institutes of Health (R01 CA076366 to P.A. Newcomb). The authors would like to thank the study participants and staff of the Hormones and Colon Cancer study.

**CCFR:** This work was supported by grant UM1 CA167551 from the National Cancer Institute and through cooperative agreements with the following CCFR centers: **Ontario Familial Colorectal Cancer Registry** (U01/U24 CA074783)

**HRS:** HRS is supported by the National Institute on Aging (NIA U01AG009740). The genotyping was funded separately by the National Institute on Aging (RC2 AG036495, RC4 AG039029). Our genotyping was conducted by the NIH Center for Inherited Disease Research (CIDR) at Johns Hopkins University. Genotyping quality control and final preparation of the data were performed by the University of Michigan School of Public Health.

**MEC:** Support for this study came from the National Institutes of Health (R37CA54281, P01CA033619, R01CA63464).

**MCTFR:** Data collection and analysis was supported by National Institutes of Health awards DA036216, DA05147, and DA024417.

**MHI:** We thank all participants and staff of the André and France Desmarais Montreal Heart Institute's (MHI) Biobank. The genotyping of the MHI Biobank was done at the MHI Pharmacogenomic Centre and funded by the MHI Foundation. Valerie Turcot is supported by a postdoctoral fellowship from the Canadian Institutes of Health Research (CIHR). Jean-Claude Tardif and Guillaume Lettre are supported by the Canada Research Chair Program.

**NESCOG:** This work is supported by the Netherlands Organization for Scientific Research (NWO Brain & Cognition 433-09-228, NWO Complexity Project 645-000-003, NWO VICI 453-14-005). Statistical analyses were carried out on the Genetic Cluster Computer hosted by SURFsara and financially supported by the Netherlands Organization for Scientific Research (NWO 480-05-003 PI: Posthuma) along with a supplement from the Dutch Brain Foundation and the VU University Amsterdam.

**WHI:** The WHI program is funded by the National Heart, Lung, and Blood Institute, National Institutes of Health, U.S. Department of Health and Human Services through contracts HHSN268201100046C, HHSN268201100001C, HHSN268201100002C, HHSN268201100003C, HHSN268201100004C, and HHSN271201100004C. Personal funding for Sean P. David from National Institute on Minority Health and Health Disparities grant U54-MD010724. The authors thank the WHI investigators and staff for their dedication, and the study participants for making the program possible. A full listing of WHI investigators can be found at: <http://www.whi.org/researchers/Documents%20%20Write%20a%20Paper/WHI%20Investigator%20Short%20List.pdf>.

Figure S1. QQ plots of GWAS meta-analysis in individuals of European ancestry

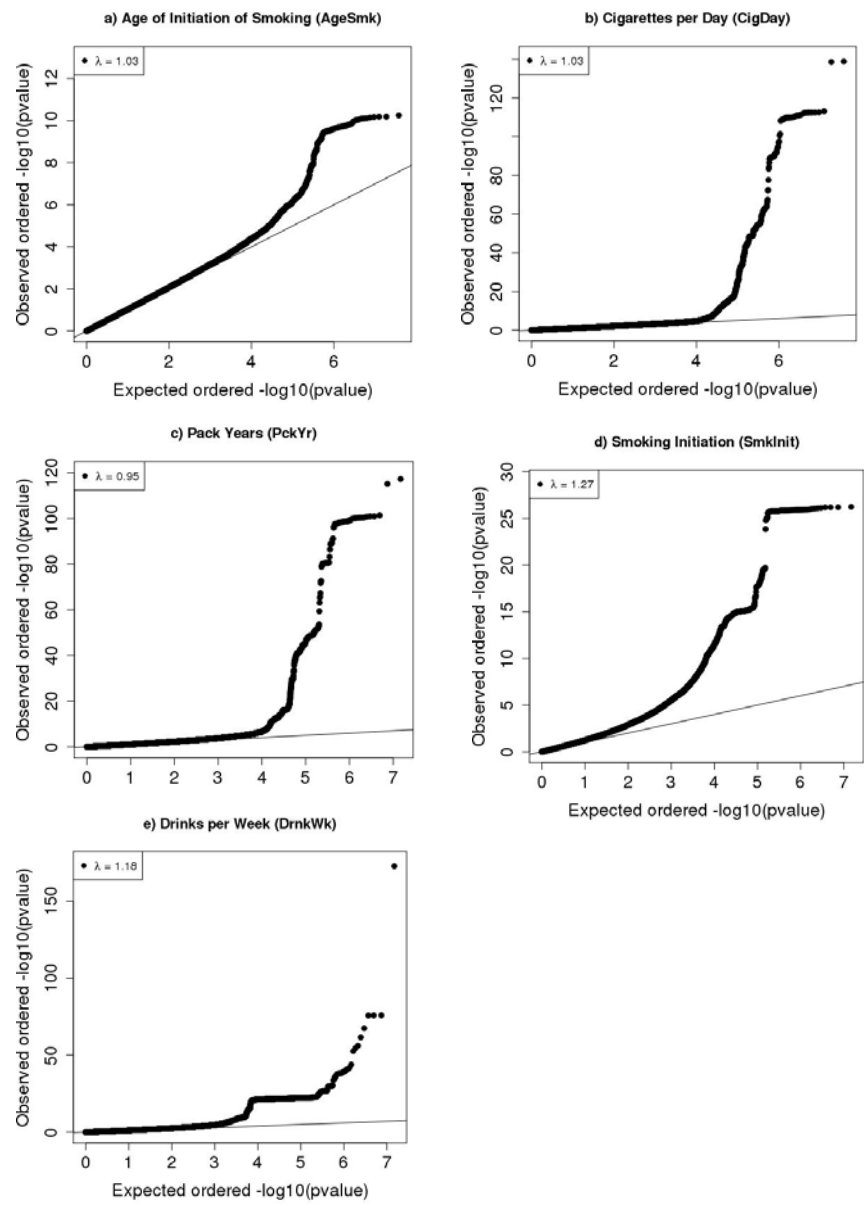


Figure S2. Manhattan plots of GWAS meta-analysis in individuals of European ancestry

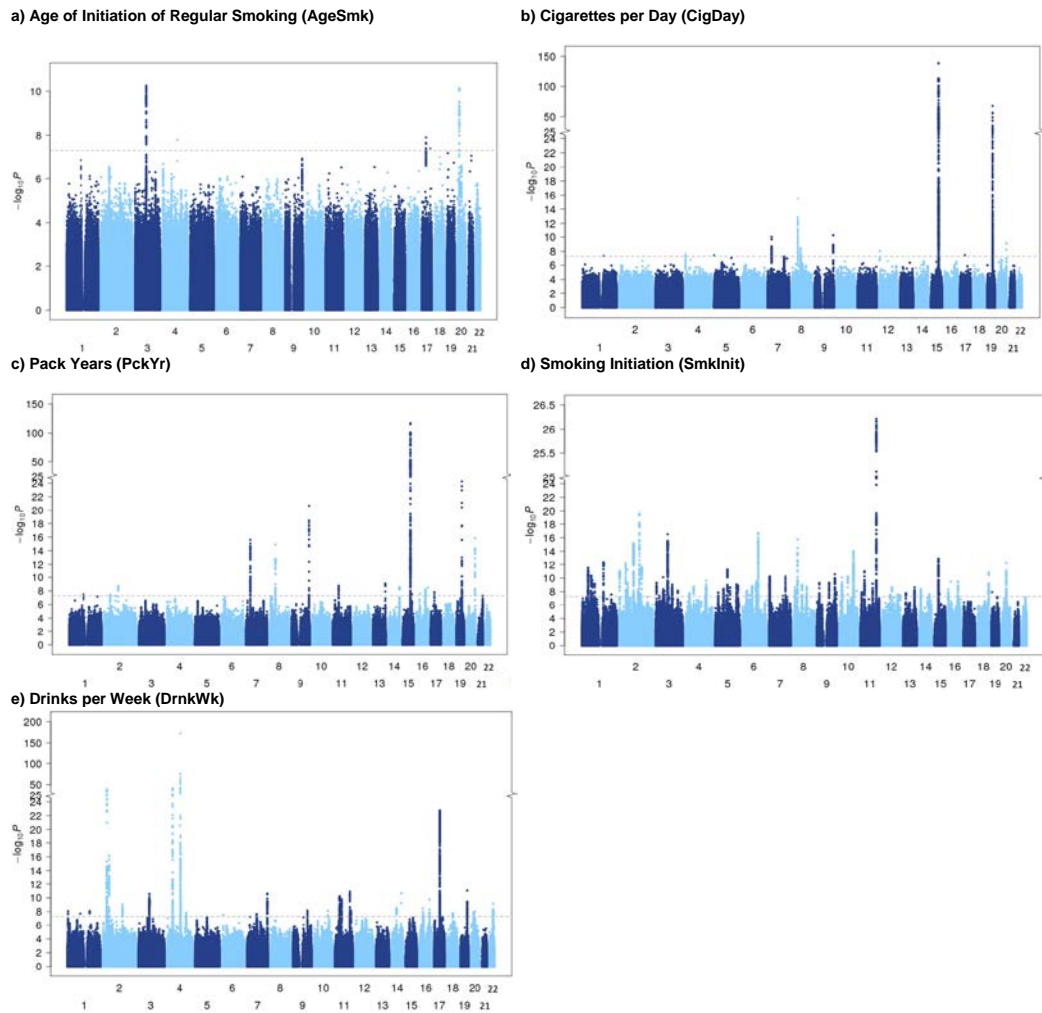


Figure S3. QQ plots of GWAS meta-analysis in individuals of African ancestry

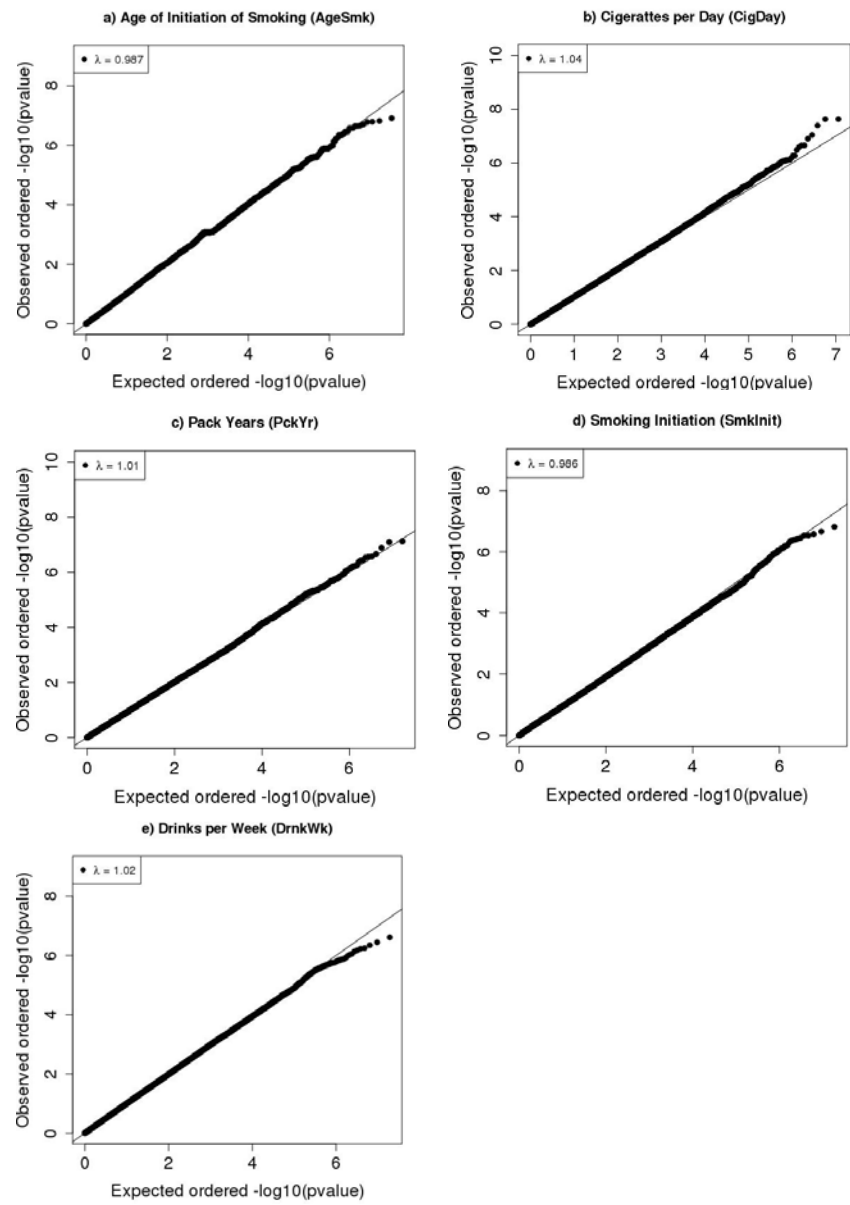
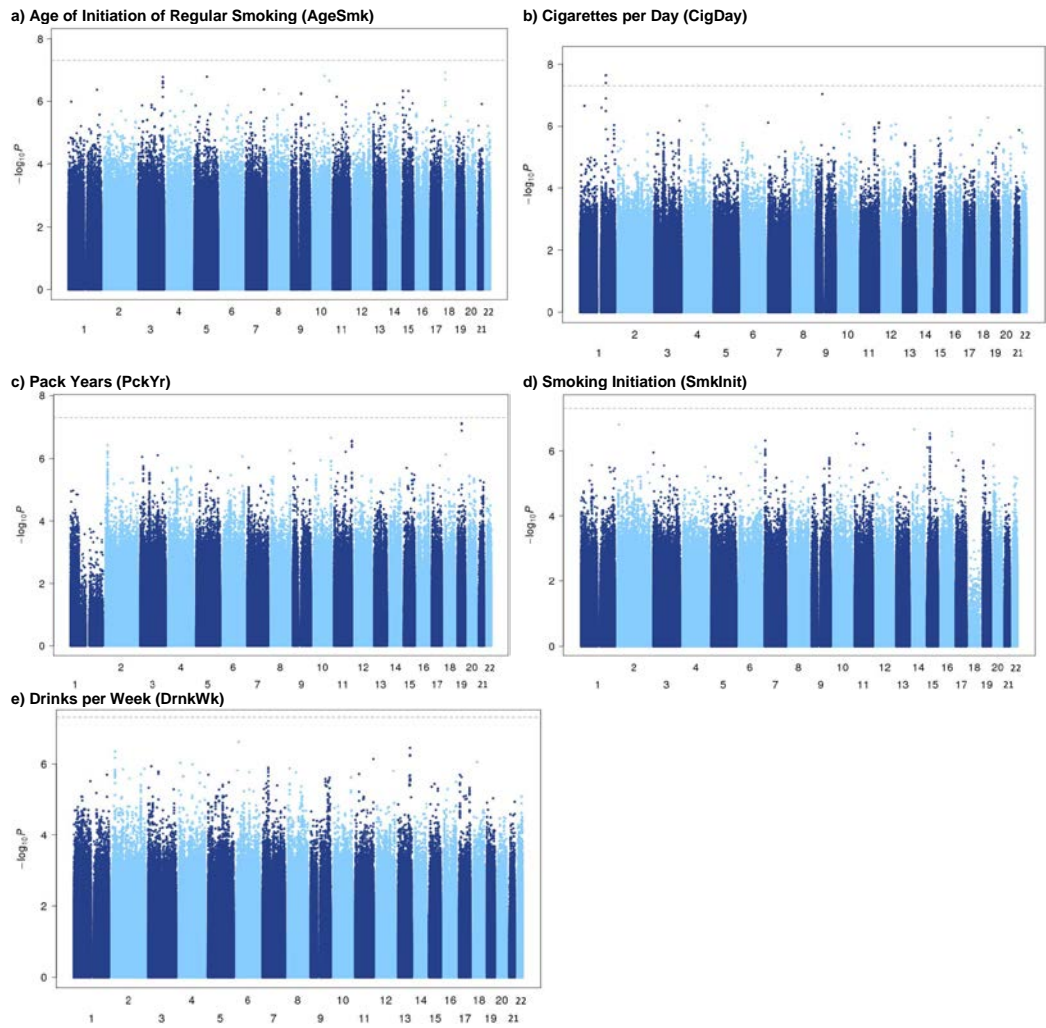


Figure S4. Manhattan plots of GWAS meta-analysis in individuals of African ancestry





**Table S1.** Participating Cohort Descriptions

Study Abbreviation	Full Study Name	Design	Array Platform	Association Covariates
ARIC	Atherosclerosis Risk in Communities	Community sample of older adults	Illumina HumanExome	
COGA*	Collaborative Study on the Genetics of Alcoholism	Family study of alcoholism	Illumina HumanCoreExome*	Sex, age, sex*age, age <sup>2</sup> , birth cohort, DSM5 alcohol dependence
FTC	NAG-FIN, FinnTwin12, FinnTwin16, FITSA	Population-based twin samples from the Older and Younger Finnish Twin Cohorts	Illumina HumanCoreExome	Sex, age, age <sup>2</sup> , current or former smoker (cigarettes per day), year of birth, cohort status, BMI.
FUSION	Finland-United States Investigation of NIDDM Genetics	Type-2 diabetes case-control	Illumina HumanExome	Sex, age, age <sup>2</sup> , current v. former smoker (cigarettes per day), height and weight (drinks per week).
GECCO	Genetics and Epidemiology of Colorectal Cancer Consortium	Colorectal cancer case-control	Illumina HumanExome	
HRS	Health and Retirement Study	National representative sample of older adults	Illumina HumanExome	Age, age <sup>2</sup> , sex, age*sex, birth year, PCs 1-4 (European ancestry) or PCs 1-10 (African ancestry), current v. former smoker (for smoking outcomes), weight, bmi, bmi*gender, and current v. former drinker (for drinking outcomes)
ID1000	-	National representative sample of young adults	Illumina HumanExome	Age, age <sup>2</sup> , sex, age*sex, PCs 1-10, current v. former smoker (for cigarettes/day, pack years); bmi, weight, height, bmi*sex for (drinks per week)
MEC	Multi-Ethnic Cohort		Illumina HumanExome	

Study Abbreviation	Full Study Name	Design	Array Platform	Association Covariates
METSIM	Metabolic Syndrome in Men		Illumina HumanExome	Sex, age, age <sup>2</sup> , current v. former smoker (cigarettes per day), height and weight (drinks per week).
MHI	Montreal Heart Institute	Community sample of adults among visitors, patients and employees of the MHI.	Illumina HumanExome	Sex, age, age <sup>2</sup> , PCs 1-10, current or former smoker status (for cigarettes per day), height and weight (for drinks per week).
MCTFR	Minnesota Center for Twin and Family Research	Community-based family cohort	Illumina HumanExome	Sex, age, parent-child generation,
NAGOZALC			Illumina HumanExome	
NESCOG	Netherlands study of Cognition, Environment, and Genes	National representative sample of adults	Illumina HumanExome	Age, age <sup>2</sup> , sex, age*sex, PCs 1-10; current v. former smoker (for cigarettes/day, pack years); bmi, weight, height, bmi*sex (for drinks/week).
SardiNIA	-	Community-based Family Cohort	Illumina HumanExome	Sex, age, age <sup>2</sup> , current v. former smoker (cigarettes per day), height and weight (drinks per week).
TwinsUK	-	Twin cohort	Illumina HumanExome	
WHI	Womens Health Initiative	Complex design consisting of clinical trials and observational cohort.	Illumina HumanExome	Sex, age, age <sup>2</sup> , EV1, EV2, EV3 (all phenotypes); current v. former smoking (cigarettes per day), height and weight (drinks per week).
UK Biobank**	(Stratified by UK BiLEVE sample [N~50,000] and remainder).	Community sample of older adults, selected for heavy and non-smokers	UK BiLEVE and UK Axiom arrays	Sex, age, age <sup>2</sup> , current or former smoker (for cigarettes per day), PCs 1-15, height, and weight (for drinking

\*The exome array genotyping in COGA was performed in three broader groups comprised of 1059 founder subjects from 118 extended European American families and 2815 longitudinally ascertained subjects of mixed ethnicities. The 1059 subjects in 118 families were selected using the ExomePick program (<http://genome.sph.umich.edu/wiki/ExomePicks>) that uses the kinship information to suggests individuals to be sequenced in a large pedigree. Out of 2815

longitudinally ascertained subjects 538 subjects were also younger relatives of 1059 EA subjects from 118 extended families. There were around 726 subjects in these EA families that were not genotyped using the exome array. All of EA subjects from 118 families were previously genotyped using Illumina Human OmniExpress array 12.V1 (Illumina, San Diego, CA, USA). This gave us an opportunity to infer the dense SNPs in un-genotyped subjects using identity by descent information generated through the sparse array using publicly available long range phasing program ChromoPhase ([Daetwyler et al, 2014](#)). We phased genotyped subjects in each pedigree for each chromosome by combining the sparse and dense genotypes and used this IBD information to fill in the missing genotypes according to rules of Mendelian segregation. The phase of unambiguous SNPs were generated using the population frequency and were imputed according to population based imputation. Using this option we were able to guess > 98% missing haplotypes in all of the subjects. After infer process we removed the variants that didn't follow the rules of Mendelian segregation.

\*\*One member of all pairs of related individuals between first UKB release (150K) and second UKB release (350K) were removed.

**Table S2.** Per-study, per-phenotype sample size and genomic control (European ancestry only).

Study	Cigarettes per Day		Pack Years		Age of Initiation of Smoking		Smoking Initiation		Drinks per Week	
	N	GC	N	GC	N	GC	N	GC	N	GC
ARIC	5381	1.063	5304	1.045	5407	1.096	8970	1.064	5966	1.000
COGA	1465	.895	1435	1.050	1638	0.923	-	-	3398	0.953
FTC	819	1.048	767	1.012	769	1.059	1467	1.063	1242	0.995
FUSION	568	1.040	530	1.042	562	1.018	1153	1.016	830	0.997
GECCO	2916	1.018	2876	1.028	-	-	6459	0.993	2077	0.967
HRS	3303	0.988	3303	0.992	3303	0.998	6393	1.096	4507	0.988
ID1000	366	0.974	373	1.007	-	-	803	0.994	740	0.985
MEC	1087	0.979	1082	0.963	1086	0.999	1903	0.973	1271	1.064
METSIM	1374	1.028	1370	1.016	1370	1.026	8146	1.044	6303	1.099
MHI	4391	1.011	4400	1.016	4420	1.018	6820	1.025	4205	1.022
MCTFR	2043	0.991	-	-	-	-	-	-	4757	0.998
NAGOZALC	671	1.006	646	1.006	647	1.011	1038	1.004	663	0.994
NESCOG	217	1.004	220	1.000	-	-	486	1.038	437	0.980
SardinIA	1969	1.009	1967	1.064	1967	1.014	5069	1.082	2516	1.142
TwinsUK	358	1.039	358	1.010	358	1.006	878	0.971	603	0.989
WHI	6246	1.031	6236	1.006	-	-	-	-	7213	0.982
UK Biobank (MAF>1%)	120,744	1.10	120,126	1.08	124,590	1.03	383,631	1.15	311,126	1.06
UK Biobank (MAF≤1%)	<sup>a</sup>	1.03	<sup>a</sup>	1.01	<sup>a</sup>	.96	<sup>a</sup>	.98	<sup>a</sup>	1.02

Note: Study abbreviations are defined in Table S1.

<sup>a</sup>Sample sizes are the same for UK Biobank common and rare variants.

**Table S3.** Per-study, per-phenotype sample size and genomic control (African ancestry only).

Study	Cigarettes per Day		Pack Years		Age of Initiation of Smoking		Smoking Initiation		Drinks per Week	
	N	GC	N	GC	N	GC	N	GC	N	GC
COGA	476	0.93	457	0.99	494	0.91	-	-	1,182	0.94
HRS	961	1.03	961	1.02	961	1.01	1,746	1.03	980	0.99
UK Biobank (MAF>1%)	1,248	1.04	1,240	1.01	1,250	1.01	7,228	0.99	5432	1.04

Note: Study abbreviations are defined in Table S1.

Tables S4-S7 are available in ~~Xcel~~Excel spreadsheets for convenience. [See Supplement 2.](#)

**Table S8. Partition of  $h^2$  Heritability for Variants on Exome Array.** We estimate the “chip” heritability for variants on the exome array using LD Score Regression. We consider a model that consists of seven functional categories. We report estimates of heritability ( $\hat{h}^2$ ), their standard deviation  $se(\hat{h}^2)$  as well as the p-value and z-score.

Annotation	$(\hat{h}^2)$	$se(\hat{h}^2)$	p-value	z-Score
<b>Age of Initiation of Smoking</b>				
(Intercept)	1	0.022	0	47
3' UTR	0.0046	0.0013	0.26	1.1
5' UTR	0.0089	0.0019	0.14	1.5
Common Coding Variants	0.014	0.0016	0.0069	2.7
Intergenic	0.016	0.0036	0.15	1.4
Intron	0.0042	0.00072	0.066	1.8
Rare Coding	0.011	0.0015	0.028	2.2
Synonymous	0.0017	$7.00 \times 10^{-4}$	0.44	0.77
<b>Cigarettes per Day</b>				
(Intercept)	1	0.023	0	43
3' UTR	0.0044	0.00049	$1.90 \times 10^{-1}$	1.3
5' UTR	0.0061	0.00072	$2.10 \times 10^{-1}$	1.2
Common Coding Variants	0.025	$6.00 \times 10^{-4}$	$1.70 \times 10^{-9}$	6
Intergenic	0.027	0.0014	$3.10 \times 10^{-3}$	3
Intron	0.0022	0.00027	$2.30 \times 10^{-1}$	1.2
Rare Coding	0.0098	$6.00 \times 10^{-4}$	$1.70 \times 10^{-2}$	2.4
Synonymous	0.015	0.00058	$1.20 \times 10^{-4}$	3.8
<b>Drinks per Week</b>				
(Intercept)	1.1	0.027	0	42
3' UTR	0.015	0.0023	$3.00 \times 10^{-2}$	2.2
5' UTR	0.0095	0.0034	$3.60 \times 10^{-1}$	0.92
Common Coding Variants	0.035	0.0029	$5.10 \times 10^{-5}$	4.1
Intergenic	0.059	0.0065	$2.40 \times 10^{-3}$	3
Intron	0.0042	0.0013	$2.80 \times 10^{-1}$	1.1
Rare Coding	0.02	0.0013	$1.80 \times 10^{-7}$	5.2
Synonymous	0.017	0.0028	$4.30 \times 10^{-2}$	2

Annotation	$(\hat{h}^2)$	$se(\hat{h}^2)$	p-value	z-Score
<b>Pack Years</b>				
(Intercept)	1	0.024	0	43
3' UTR	0.0041	0.00056	$2.10 \times 10^{-1}$	1.3
5' UTR	0.0075	0.00082	$1.20 \times 10^{-1}$	1.6
Common Coding Variants	0.018	0.00069	$8.80 \times 10^{-6}$	4.4
Intergenic	0.038	0.0016	$3.70 \times 10^{-5}$	4.1
Intron	0.002	0.00031	$2.70 \times 10^{-1}$	1.1
Rare Coding	0.018	0.00068	$8.50 \times 10^{-6}$	4.5
Synonymous	0.012	0.00066	$1.30 \times 10^{-3}$	3.2
<b>Smoking Initiation</b>				
(Intercept)	1	$3.40 \times 10^{-2}$	$5.70 \times 10^{-206}$	31
3' UTR	0.019	$1.90 \times 10^{-4}$	$1.40 \times 10^{-17}$	8.5
5' UTR	0.04	$2.80 \times 10^{-4}$	$1.10 \times 10^{-33}$	12
Common Coding Variants	0.019	$2.30 \times 10^{-4}$	$5.40 \times 10^{-12}$	6.9
Intergenic	0.038	$5.30 \times 10^{-4}$	$6.20 \times 10^{-10}$	6.2
Intron	0.0024	$1.10 \times 10^{-4}$	$5.40 \times 10^{-2}$	1.9
Rare Coding	0.022	$2.30 \times 10^{-4}$	$3.90 \times 10^{-16}$	8.1
Synonymous	0.00025	$2.40 \times 10^{-4}$	$9.30 \times 10^{-1}$	0.09



**CHD Exome+ consortium members**

Praveen Surendran<sup>1</sup>, Robin Young<sup>1</sup>, Daniel R. Barnes<sup>1</sup>, Sune Fallgaard Nielsen<sup>2</sup>, Asif Rasheed<sup>3</sup>, Maria Samuel<sup>3</sup>, Wei Zhao<sup>4</sup>, Jukka Kontto<sup>5</sup>, Markus Perola<sup>5,6,7</sup>, Muriel Caslake<sup>8</sup>, Anton JM. de Craen<sup>9</sup>, Stella Trompet<sup>9,10</sup>, Maria Uria-Nickelsen<sup>11</sup>, Anders Malarstig<sup>12</sup>, Dermot F. Reilly<sup>13</sup>, Maarten Hoek<sup>14</sup>, Thomas Vogt<sup>14,15</sup>, J. Wouter Jukema<sup>11,16</sup>, Naveed Sattar<sup>17</sup>, Ian Ford<sup>8</sup>, Chris J. Packard<sup>8</sup>, Dewan S. Alam<sup>18</sup>, Abdulla al Shafi. Majumder<sup>19</sup>, Emanuele Di Angelantonio<sup>1,20</sup>, Rajiv Chowdhury<sup>1</sup>, Philippe Amouyel<sup>121,22,23,24</sup>, Dominique Arveiler<sup>25</sup>, Stefan Blankenberg<sup>26,27</sup>, Jean Ferrières<sup>28</sup>, Frank Kee<sup>29</sup>, Kari Kuulasmaa<sup>5</sup>, Martina Müller-Nurasyid<sup>30,31,32</sup>, Giovanni Veronesi<sup>33</sup>, Jarmo Virtamo<sup>5</sup>, EPIC-CVD Consortium, Philippe Frossard<sup>3</sup>, Børge Grønne Nordestgaard<sup>2</sup>, Danish Saleheen<sup>4,3,1</sup>, John Danesh<sup>1,35,20</sup>, Adam S. Butterworth<sup>1,20</sup>, Joanna MM. Howson<sup>1</sup>

**Affiliations**

Cardiovascular Epidemiology Unit, Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK

1. Department of Clinical Biochemistry Herlev Hospital, Copenhagen University Hospital, Herlev, Denmark
1. Centre for Non Communicable Diseases, Karachi, Pakistan
1. Department of Biostatistics and Epidemiology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, USA
1. Department of Health, National Institute for Health and Welfare, Helsinki, Finland
1. Institute of Molecular Medicine FIMM, University of Helsinki, Finland
1. Estonian Genome Center, University of Tartu, Tartu, Estonia
1. University of Glasgow, Glasgow, UK
1. Department of Gerontology and Geriatrics, Leiden University Medical Center, Leiden, The Netherlands
1. Department of Cardiology, Leiden University Medical Center, Leiden, The Netherlands
1. Development Management and Planning, Pfizer Worldwide Research and Development
1. Pfizer Worldwide Research and Development, Stockholm, Sweden
1. Genetics and Pharmacogenomics, Merck Research Laboratories, Boston, Massachusetts, USA.
1. Merck Research Laboratories, Kenilworth, New Jersey, USA
1. CHDI Management/CHDI Foundation, Princeton, New Jersey, USA
1. The Interuniversity Cardiology Institute of the Netherlands, Utrecht, The Netherlands
1. Institute of Cardiovascular and Medical Sciences, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow, UK
1. ICDDR, B; Mohakhali, Dhaka, Bangladesh
1. National Institute of Cardiovascular Diseases, Sher-e-Bangla Nagar, Dhaka, Bangladesh
1. The National Institute for Health Research Blood and Transplant Research Unit in Donor Health and Genomics, University of Cambridge, Cambridge, UK
1. University of Lille, Risk Factors and Molecular Determinants of aging-related diseases, Lille, France
1. Inserm, Lille, France
1. Centre Hospitalier Universitaire Lille, Public Health, Lille, France
1. Institute Pasteur de Lille, Lille, France
1. Department of Epidemiology and Public Health, EA 3430, University of Strasbourg, Strasbourg, France
1. Department of General and Interventional Cardiology, University Heart Center Hamburg, Germany
- University Medical Center Hamburg-Eppendorf, Hamburg, Germany.

**Commented [R1]:** Since these two consortia are named as authors, this information must be present in the main text of the article. We have moved this on your behalf accordingly.

### Consortium for Genetics of Smoking Behaviour

A Mesut Erzurumluoglu<sup>1</sup>, Victoria E Jackson<sup>1</sup>, Carl A Melbourne<sup>1</sup>, Tibor V Varga<sup>2</sup>, Helen R Warren<sup>3,4</sup>, Vinicius Tragante<sup>5</sup>, Ioanna Tachmazidou<sup>6</sup>, Sarah E Harris<sup>7,8</sup>, Evangelos Evangelou<sup>9,10</sup>, Jonathan Marten<sup>11</sup>, Weihua Zhang<sup>12,13,14,15</sup>, Elisabeth Altmaier<sup>16</sup>, Jian'an Luan<sup>17</sup>, Claudia Langenberg<sup>17</sup>, Robert A Scott<sup>17</sup>, Hanieh Yaghootkar<sup>18</sup>, Kathleen Stirrups<sup>19,20</sup>, Stavroula Kanoni<sup>20,21</sup>, Eirini Marouli<sup>20,21</sup>, Fredrik Karpe<sup>22,23</sup>, Anna F Dominiczak<sup>24</sup>, Peter Sever<sup>25</sup>, Neil Poulter<sup>26</sup>, Olov Rolandsson<sup>27</sup>, Clemens Baumbach<sup>16</sup>, Saima Afaq<sup>12</sup>, John C Chambers<sup>12,13,28</sup>, Jaspal S Keener<sup>29,13,30,31</sup>, Nicholas J Wareham<sup>17</sup>, Frida Renström<sup>2,32</sup>, Göran Hallmans<sup>32</sup>, Riccardo E Marioni<sup>7,8</sup>, Janie Corley<sup>7,33</sup>, John M Starr<sup>7,34</sup>, Nick Verweij<sup>35,36</sup>, Rudolf A de Boer<sup>36</sup>, Peter van der Meer<sup>36</sup>, Ilonca Vaartjes<sup>37,38</sup>, Michiel L Bots<sup>37,38</sup>, Folkert W Asselbergs<sup>5,39</sup>, Hans J Grabe<sup>40</sup>, Henry Völzke<sup>41</sup>, Matthias Nauck<sup>42</sup>, Stefan Weiss<sup>43</sup>, Paul D P Pharoah<sup>44,45</sup>, Alison M Dunning<sup>45</sup>, Joe G Dennis<sup>44</sup>, Deborah J Thompson<sup>44</sup>, Kyriaki Michailidou<sup>46,44</sup>, Douglas F Easton<sup>44,45</sup>, Antonis C Antoniou<sup>44</sup>, Jessica Tyrrell<sup>48</sup>, Evelin Mihailov<sup>47</sup>, Nilesh J Samani<sup>48,49</sup>, Kaixin Zhou<sup>50</sup>, Matthew J Neville<sup>22,23</sup>, Andres Metspalu<sup>47</sup>, Colin N A Palmer<sup>51</sup>, Ian J Deary<sup>7,33</sup>, Tim M Frayling<sup>18</sup>, Caroline Hayward<sup>11</sup>, Pim van der Harst<sup>35,52</sup>, Eleftheria Zeggini<sup>6</sup>, Understanding Society Scientific Group\*, Patricia B Munroe<sup>3,4</sup>, Jan Håkan Janesson<sup>53</sup>, Paul W Franks<sup>2,54</sup>, Panos Deloukas<sup>55,56,57</sup>, Mark J Caulfield<sup>3,4</sup>, Martin D Tobin<sup>1</sup>

1. Department of Health Sciences, University of Leicester, Leicester, UK  
 2. Genetic and Molecular Epidemiology Unit, Lund University Diabetes Centre, Department of Clinical Sciences, Skåne University Hospital, Lund University, SE-214 28, Malmö, Sweden  
 3. Clinical Pharmacology, William Harvey Research Institute, Queen Mary University of London, London, EC1M 6BQ, UK  
 4. NIHR Barts Cardiovascular Biomedical Research Unit, Queen Mary University of London, London, EC1M 6BQ, UK  
 5. Department of Cardiology, Division Heart and Lungs, University Medical Center Utrecht, 3508GA Utrecht, the Netherlands  
 6. Wellcome Trust Sanger Institute, Hinxton, CB10 1SA, UK  
 7. Centre for Cognitive Ageing and Cognitive Epidemiology, University of Edinburgh, Edinburgh, UK, EH8 9JZ  
 8. Centre for Genomic and Experimental Medicine, University of Edinburgh, Edinburgh, UK, EH4 2XU  
 9. Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, London, UK  
 10. Department of Hygiene and Epidemiology, University of Ioannina Medical School, Ioannina, Greece  
 11. MRC Human Genetics Unit, MRC Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, UK  
 12. Department of Epidemiology and Biostatistics, Imperial College London, London W2 1PG, UK  
 13. Department of Cardiology, Ealing Hospital, London North West Healthcare NHS Trust, Middlesex UB4 3HW, UK  
 14. Biocenter Oulu, University of Oulu, Finland  
 15. Unit of Primary Care, Oulu University Hospital, Oulu, Finland  
 16. Research Unit of Molecular Epidemiology, Helmholtz Zentrum München German Research Center for Environmental Health, Neuherberg, Germany  
 17. MRC Epidemiology Unit, Institute of Metabolic Science, University of Cambridge School of Clinical Medicine, Cambridge, CB2 0QQ, UK  
 18. Genetics of Complex Traits, University of Exeter Medical School, Exeter, United Kingdom  
 19. Department of Haematology, University of Cambridge, Cambridge, UK, CB2 0PT  
 20. William Harvey Research Institute, Barts and The London School of Medicine and Dentistry, Queen Mary

University of London, London, UK, EC1M 6BQ  
21. Centre for Genomic Health, Queen Mary University of London, London EC1M 6BQ, UK  
22. Oxford Centre for Diabetes, Endocrinology and Metabolism, University of Oxford  
23. Oxford National Institute for Health Research, Biomedical Research Centre, Churchill Hospital, Oxford, UK  
24. Institute of Cardiovascular and Medical Sciences, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow, UK  
25. National Heart and Lung Institute, Imperial College London, W2 1PG, UK  
26. School of Public Health, Imperial College London, W2 1PG, UK  
27. Department of Public Health & Clinical Medicine, Section for Family Medicine, Umeå universitet, SE-90185 Umeå, Sweden  
28. School of Medicine and Pharmacology, The University of Western Australia, Crawley 6009, Australia  
29. National Heart and Lung Institute, Imperial College London, London W12 0NN, UK  
30. PathWest Laboratory Medicine of WA, Sir Charles Gairdner Hospital, Crawley WA 6009, Australia  
31. School of Pathology and Laboratory Medicine, The University of Western Australia, Crawley WA 6009, Australia  
32. Department of Biobank Research, Umeå University, SE-901 87, Umeå, Sweden  
33. Psychology, University of Edinburgh, Edinburgh, UK, EH8 9JZ  
34. Alzheimer Scotland Research Centre, University of Edinburgh, Edinburgh, UK, EH8 9JZ  
35. University Medical Center Groningen, University of Groningen, Department of Cardiology, the Netherlands  
36. Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, 301 Binney Street, Cambridge, MA 02142, USA  
37. Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, 3508GA Utrecht, the Netherlands  
38. Center for Circulatory Health, University Medical Center Utrecht, 3508GA Utrecht, the Netherlands  
39. Durrer Center for Cardiovascular Research, Netherlands Heart Institute, 3501DG Utrecht, Netherlands  
40. Department of Psychiatry and Psychotherapy, University Medicine Greifswald, 17475 Greifswald, Germany  
41. Institute for Community Medicine, University Medicine Greifswald, 17475 Greifswald  
42. Institute of Clinical Chemistry and Laboratory Medicine, University Medicine Greifswald, 17475 Greifswald, Germany  
43. Interfaculty Institute for Genetics and Functional Genomics; University Medicine and Ernst-Moritz-Arndt-University Greifswald, 17475 Greifswald, Germany  
44. Centre for Cancer Genetic Epidemiology, Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK, CB1 8RN  
45. Centre for Cancer Genetic Epidemiology, Department of Oncology, Cambridge Centre, University of Cambridge, Cambridge, UK, CB1 8RN  
46. Department of Electron Microscopy/Molecular Pathology, The Cyprus Institute of Neurology and Genetics, 1683 Nicosia, Cyprus  
47. Estonian Genome Center, University of Tartu, Tartu, Estonia  
48. Department of Cardiovascular Sciences, University of Leicester, Cardiovascular Research Centre, Glenfield Hospital, Leicester, LE3 9QP, UK  
49. NJS is supported by the British Heart Foundation and NJS is a NIHR Senior Investigator  
50. School of Medicine, University of Dundee, Dundee, UK  
51. Medical Research Institute, University of Dundee, Ninewells Hospital and Medical School, Dundee, UK.  
52. University of Groningen, University Medical Center Groningen, Department of Genetics, Groningen, The

**Netherlands**

53. Department of Medicine, Skellefteå Hospital, Skellefteå, Sweden

54. Department of Nutrition, Harvard T. H. Chan School of Public Health, Boston, MA 02115, USA

55. William Harvey Research Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, EC1M 6BQ UK

56. Wellcome Trust Sanger Institute, Hinxton, Cambridge CB10 1SA UK

57. Princess Al Jawhara Al Brahim Centre of Excellence in Research of Hereditary Disorders (PACER-HD), King Abdulaziz University, Jeddah 21589, Saudi Arabia

<sup>a</sup>Understanding Society: The UK Household Longitudinal Study: Michaela Benzeval<sup>a</sup>, Jonathan Burton<sup>a</sup>, Nicholas Buck<sup>a</sup>, Annette Jäckle<sup>a</sup>, Meena Kumari<sup>a</sup>, Heather Laurie<sup>a</sup>, Peter Lynn<sup>a</sup>, Stephen Pudney<sup>a</sup>, Birgitta Rabe<sup>a</sup>, Dieter Wolke<sup>b</sup>.

— Institute for Social and Economic Research, University of Essex, UK

— University of Warwick, UK

**Consortium for Genetics of Smoking Behaviour Funding statements**

Understanding Society Scientific Group is funded by the Economic and Social Research Council (ES/H020745/1) and the Wellcome Trust (WT098051). Paul D.P. Pharoah is funded by Cancer Research UK (C490/A16561). SHIP is funded by the German Federal Ministry of Education and Research (BMBF) and the German Research Foundation (DFG); see acknowledgements for details. F.W. Asselbergs is funded by the Netherlands Heart Foundation (2014T001) and UCL Hospitals NIHR Biomedical Research Centre. The LifeLines Cohort Study, and generation and management of GWAS genotype data for the LifeLines Cohort Study is supported by the Netherlands Organization of Scientific Research NWO (grant 175.010.2007.006), the Economic Structure Enhancing Fund (FES) of the Dutch government, the Ministry of Economic Affairs, the Ministry of Education, Culture and Science, the Ministry for Health, Welfare and Sports, the Northern Netherlands Collaboration of Provinces (SNN), the Province of Groningen, University Medical Center Groningen, the University of Groningen, Dutch Kidney Foundation and Dutch Diabetes Research Foundation. Nick Verweij is supported by Horizon 2020 (Marie Skłodowska-Curie, 661395) and ICIN-NHL LBC1921 and LBC1936 is supported by the MRC (MR/K026992/1). Paul W. Franks is supported by Novo Nordisk, the Swedish Research Council, Pålssons Foundation, Swedish Heart Lung Foundation (2020389), and Skåne Regional Health Authority. Nicholas J Wareham, Claudia Langenberg, Robert A Scott, and Jian'an Luan are supported by the MRC (MC\_U106179471 and MC\_UU\_12015/1). John C. Chambers and Jaspal S. Kooner are supported by the British Heart Foundation (SP/04/002), Medical Research Council (G0601966 and G0700931), Wellcome Trust (084723/Z/08/Z), NIHR (RP-PG-0407-10371), European Union FP7 (EpiMigrant, 279143), Action on Hearing Loss (G51), National Institute for Health Research (NIHR) Comprehensive Biomedical Research Centre Imperial College Healthcare NHS Trust, and iHealth-T2D (643774). The BRIGHT study was supported by the Medical Research Council of Great Britain (Grant Number G9521010D); and by the British Heart Foundation (Grant Number PG/02/128). The BRIGHT study is extremely grateful to all the patients who participated in the study and the BRIGHT nursing team. The Exome Chip genotyping was funded by Wellcome Trust Strategic Awards (083948 and 085475). We would also like to thank the Barts Genome Centre staff for their assistance with this project. The ASCOT study and the collection of the ASCOT DNA repository was supported by Pfizer, New York, NY, USA, Servier Research Group, Paris, France; and by Leo Laboratories, Copenhagen, Denmark. Genotyping of the Exome Chip in ASCOT-SC and ASCOT-UK was funded by the National Institutes of Health Research (NIHR). Anna F. Dominiczak was supported by the British

Heart Foundation (Grant Numbers RG/07/005/23633, SP/08/005/25115); and by the European Union Ingenious HyperCare Consortium: Integrated Genomics, Clinical Research, and Care in Hypertension (grant number LSHM-C7-2006-037093). Nilesh J. Samani is supported by the British Heart Foundation. Panos Deloukas is supported by the British Heart Foundation (RG/14/5/30893), and NIHR, where his work forms part of the research themes contributing to the translational research portfolio of Barts Cardiovascular Biomedical Research Centre which is funded by the National Institute for Health Research (NIHR).

#### **Consortium for Genetics of Smoking Behaviour—Acknowledgements**

The authors would like to thank the many colleagues who contributed to collection and phenotypic characterisation of the clinical samples, as well as genotyping and analysis of the GWA data. Special mentions are as follows:

Some of the data utilised in this study were provided by the Understanding Society: The UK Household Longitudinal Study, which is led by the Institute for Social and Economic Research at the University of Essex and funded by the Economic and Social Research Council. The data were collected by NatCen and the genome-wide scan data were analysed by the Wellcome Trust Sanger Institute. The Understanding Society DAC have an application system for genetics data and all use of the data should be approved by them. The application form is at: <https://www.understandingsociety.ac.uk/about/health/data>.

SHIP (Study of Health in Pomerania) and SHIP-TREND both represent population-based studies. SHIP is supported by the German Federal Ministry of Education and Research (Bundesministerium für Bildung und Forschung (BMBF); grants 01ZZ9603, 01ZZ0403, and 01ZZ0403) and the German Research Foundation (Deutsche Forschungsgemeinschaft (DFG); grant GR 1912/5-1). SHIP and SHIP-TREND are part of the Community Medicine Research-net (CMR) of the Ernst-Moritz-Arndt-Universität Greifswald (EMAU) which is funded by the BMBF as well as the Ministry for Education, Science and Culture and the Ministry of Labor, Equal Opportunities, and Social Affairs of the Federal State of Mecklenburg-West Pomerania. The CMR encompasses several research projects that share data from SHIP. The EMAU is a member of the Center of Knowledge-Interchange (CKI) program of the Siemens AG. SNP typing of SHIP and SHIP-TREND using the Illumina Infinium HumanExome BeadChip (version v1.0) was supported by the BMBF (grant 03Z1CN22). LifeLines authors thank Behrooz Alizadeh, Annemieke Boesjes, Marcel Bruinenberg, Noortje Festen, Ilja Nolte, Lude Franke, Mitra Valimohammadi for their help in creating the GWAS database, and Rob Bieringa, Joost Keers, René Oostergo, Rosalie Visser, Judith Vonk for their work related to data collection and validation. The authors are grateful to the study participants, the staff from the LifeLines Cohort Study and Medical Biobank Northern Netherlands, and the participating general practitioners and pharmacists. LifeLines Scientific Protocol Preparation: Rudolf de Boer, Hans Hillegge, Melanie van der Klauw, Gerjan Navis, Hans Ormöl, Dirkje Postma, Judith Rosmalen, Joris Slaets, Ronald Stolk, Bruce Wolffenbuttel; LifeLines GWAS Working Group: Behrooz Alizadeh, Marika Boezen, Marcel Bruinenberg, Noortje Festen, Lude Franke, Pim van der Harst, Gerjan Navis, Dirkje Postma, Harold Snieder, Cisca Wijmenga, Bruce Wolffenbuttel. The authors wish to acknowledge the services of the LifeLines Cohort Study, the contributing research centres delivering data to LifeLines, and all the study participants.

Fenland authors thank Fenland Study volunteers for their time and help, Fenland Study general Practitioners and practice staff for assistance with recruitment, and Fenland Study Investigators, Co-ordination team and the Epidemiology Field, Data and Laboratory teams for study design, sample/data collection and genotyping.

We thank all ASCOT trial participants, physicians, nurses, and practices in the participating countries for their important contribution to the study. In particular we thank Clare Muckian and David Toomey for their help in DNA extraction, storage, and handling. We would also like to acknowledge the Barts and The London Genome Centre staff for genotyping the Exome Chip array.

The BRIGHT study is extremely grateful to all the patients who participated in the study and the BRIGHT nursing team. We would also like to thank the Barts Genome Centre staff for their assistance with this project. Patricia B. Munroe, Mark J. Caulfield, and Helen R. Warren wish to acknowledge the NIHR Cardiovascular Biomedical Research Unit at Barts and The London, Queen Mary University of London, UK for support. Nilesh J. Samani and Mark J. Caulfield are Senior National Institute for Health Research Investigators. EMBRACE Collaborating Centres are: Coordinating Centre, Cambridge: Daniel Barrowdale, Dobra Frost, Jo Perkins. North of Scotland Regional Genetics Service, Aberdeen: Zosia Miedzybrodzka, Helen Gregory. Northern Ireland Regional Genetics Service, Belfast: Patrick Morrison, Lisa Jeffers. West Midlands Regional Clinical Genetics Service, Birmingham: Kai ren Ong, Jonathan Hoffman. South West Regional Genetics Service, Bristol: Alan Donaldson, Margaret James. East Anglian Regional Genetics Service, Cambridge: Joan Paterson, Marc Tischkowitz, Sarah Downing, Amy Taylor. Medical Genetics Services for Wales, Cardiff: Alexandra Murray, Mark T. Rogers, Emma McCann. St James's Hospital, Dublin & National Centre for Medical Genetics, Dublin: M. John Kennedy, David Barton. South East of Scotland Regional Genetics Service, Edinburgh: Mary Porteous, Sarah Drummond. Peninsula Clinical Genetics Service, Exeter: Carole Brewer, Emma Kivuva, Anne Searle, Selina Goodman, Kathryn Hill. West of Scotland Regional Genetics Service, Glasgow: Rosemarie Davidson, Victoria Murday, Nicola Bradshaw, Lesley Snadden, Mark Longmuir, Catherine Watt, Sarah Gibson, Eshika Haque, Ed Tobias, Alexis Duncan. South East Thames Regional Genetics Service, Guy's Hospital London: Louise Izatt, Chris Jacobs, Caroline Langman. North West Thames Regional Genetics Service, Harrow: Huw Dorkins. Leicestershire Clinical Genetics Service, Leicester: Julian Barwell. Yorkshire Regional Genetics Service, Leeds: Julian Adlard, Gemma Serra-Feliu. Cheshire & Merseyside Clinical Genetics Service, Liverpool: Ian Ellis, Claire Foo. Manchester Regional Genetics Service, Manchester: D Gareth Evans, Fiona Laloo, Jane Taylor. North East Thames Regional Genetics Service, NE Thames, London: Lucy Side, Alison Male, Cheryl Berlin. Nottingham Centre for Medical Genetics, Nottingham: Jacqueline Eason, Rebecca Collier. Northern Clinical Genetics Service, Newcastle: Alex Henderson, Oonagh Claber, Irene Jobson. Oxford Regional Genetics Service, Oxford: Lisa Walker, Diane McLeod, Dorothy Halliday, Sarah Durell, Barbara Stayner. The Institute of Cancer Research and Royal Marsden NHS Foundation Trust: Ros Eccles, Nazneen Rahman, Elizabeth Bancroft, Elizabeth Page, Audrey Arden Jones, Kelly Kohut, Jennifer Wiggins, Jenny Pope, Sibel Saya, Natalie Taylor, Zoe Kemp and Angela George. North Trent Clinical Genetics Service, Sheffield: Jackie Cook, Oliver Quarrell, Cathryn Bardsley. South West Thames Regional Genetics Service, London: Shirley Hodgson, Sheila Goff, Glen Brice, Lizzie Winchester, Charlotte Eddy, Vishakha Tripathi, Virginia Attard. Wessex Clinical Genetics Service, Princess Anne Hospital, Southampton: Diana Eccles, Anneke Lucassen, Gillian Crawford, Donna McBride, Sarah Smalley.

#### **Consortium for Genetics of Smoking Behaviour Conflict of Interest statements**

Paul W. Franks has been a paid consultant for Eli Lilly and Sanofi-Aventis and has received research support from several pharmaceutical companies as part of European Union Innovative Medicines Initiative (IMI) projects. Neil Poulter has received financial support from several pharmaceutical companies that manufacture either blood pressure lowering or lipid lowering agents or both and consultancy fees. Peter Sever has received

~~research awards from Pfizer. Mark J. Caulfield is Chief Scientist for Genomics England, a UK government company.~~

**Supplemental References**

1. David, S.P., et al., *Genome-wide meta-analyses of smoking behaviors in African Americans*. Translational Psychiatry, 2012. **2**.
2. Daetwyler, H.D., et al., *Imputation of missing genotypes from sparse to high density using long-range phasing*. Genetics, 2011, **1**.

**Commented [R2]:** Please add the complete reference for Daetwyler et al.

**Commented [s3R2]:** Done