Molecular Psychiatry (2015) 20, 647–656

www.nature.com/mp

ORIGINAL ARTICLE

Genome-wide meta-analysis identifies six novel loci associated with habitual coffee consumption

The Coffee and Caffeine Genetics Consortium, MC Cornelis[1,2], EM Byrne[3,117], T Esko[4,5,6,7,117], MA Nalls[8,117], A Ganna[9], N Paynter[10], KL Monda[11], N Amin[12], K Fischer[4], F Renstrom[13], JS Ngwa[14], V Huikari[15], A Cavadino[16], IM Nolte[17], A Teumer[18], K Yu[19], P Marques-Vidal[20], R Rawal[21], A Manichaikul[22], MK Wojczynski[23], JM Vink[24], JH Zhao[25], G Burlutsky[26], J Lahti[27,28], V Mikkilä[29,30], RN Lemaitre[31], J Eriksson[32], SK Musani[33], T Tanaka[34], F Geller[35], J Luan[25], J Hui[36,37,38,39], R Mägi[4], M Dimitriou[40], ME Garcia[41], W-K Ho[42], MJ Wright[43], LM Rose[10], PKE Magnusson[9], NL Pedersen[9], D Couper[44], BA Oostra[45], A Hofman[12], MA Ikram[12,46,47], HW Tiemeier[12,48], AG Uitterlinden[12,49], FJA van Rooij[12], I Barroso[50,51], I Johansson[52], L Xue[14], M Kaakinen[15,53,54], L Milani[4], C Power[16], H Snieder[17], RP Stolk[17], SE Baumeister[55], R Biffar[56], F Gu[19], F Bastardot[57], Z Kutalik[58,59,60], DR Jacobs Jr[61], NG Forouhi[25], E Mihailov[4], L Lind[62], C Lindgren[63], K Michaëlsson[64], A Morris[63], M Jensen[2], K-T Khaw[42], RN Luben[42], JJ Wang[26], S Männistö[65], M-M Perälä[65], M Kähönen[66], T Lehtimäki[67], J Viikari[68], D Mozaffarian[1,2,69,70], K Mukamal[71], BM Psaty[31,72,73,74], A Döring[75], AC Heath[76], GW Montgomery[43], N Dahmen[77], T Carithers[78], KL Tucker[79], L Ferrucci[34], HA Boyd[35], M Melbye[35], JL Treur[24], D Mellström[32], JJ Hottenga[24], I Prokopenko[63,80], A Tönjes[81,82], P Deloukas[50,83,84], S Kanoni[83], M Lorentzon[32], DK Houston[85], Y Liu[86], J Danesh[42], A Rasheed[87], MA Mason[88], AB Zonderman[89], L Franke[90], BS Kristal[91,92], International Parkinson's Disease Genomics Consortium (IPDGC)[118], North American Brain Expression Consortium (NABEC)[118], UK Brain Expression Consortium (UKBEC)[118], J Karjalainen[90], DR Reed[93], H-J Westra[90], MK Evans[88], D Saleheen[42,87], TB Harris[41], G Dedoussis[40], G Curhan[1], M Stumvoll[81,82], J Beilby[36,37,38], LR Pasquale[1,94], B Feenstra[35], S Bandinelli[95], JM Ordovas[96], AT Chan[1,97], U Peters[98], C Ohlsson[32], C Gieger[21], NG Martin[43], M Waldenberger[99], DS Siscovick[31,72], O Raitakari[30,100], JG Eriksson[28,101,102], P Mitchell[26], DJ Hunter[1,103], P Kraft[103], EB Rimm[1,2,69], DI Boomsma[24], IB Borecki[23], RJF Loos[25,104,105], NJ Wareham[25], P Vollenweider[57], N Caporaso[19], HJ Grabe[106], ML Neuhouser[98], BHR Wolffenbuttel[107], FB Hu[1,2,69], E Hyppönen[16,108,109], M-R Järvelin[15,53,54,110,111], LA Cupples[14,112], PW Franks[2,13,113], PM Ridker[10], CM van Duijn[12,114], G Heiss[11], A Metspalu[4], KE North[11], E Ingelsson[62,63], JA Nettleton[115], RM van Dam[116] and DI Chasman[10]

[1]Channing Division of Network Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA; [2]Department of Nutrition, Harvard School of Public Health, Boston, MA, USA; [3]Queensland Brain Institute, The University of Queensland, Queensland, Australia; [4]Estonian Genome Center, University of Tartu, Tartu, Estonia; [5]Division of Endocrinology, Children's Hospital Boston, Boston, MA, USA; [6]Department of Genetics, Harvard Medical School, Boston, MA, USA; [7]Program in Medical and Population Genetics, Broad Institute, Cambridge, MA, USA; [8]Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health (NIH), Bethesda, MD, USA; [9]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Karolinska, Sweden; [10]Division of Preventive Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA; [11]Department of Epidemiology, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA; [12]Department of Epidemiology, Erasmus Medical Center, Rotterdam, The Netherlands; [13]Department of Clinical Sciences, Lund University, Malmö, Sweden; [14]Department of Biostatistics, Boston University School of Public Health, Boston, MA, USA; [15]Institute of Health Sciences, University of Oulu, Oulu, Finland; [16]Centre for Paediatric Epidemiology and Biostatistics, Medical Research Council (MRC) Centre of Epidemiology for Child Health, University College London Institute of Child Health, London, UK; [17]Department of Epidemiology, University of Groningen, University Medical Center Groningen, The Netherlands; [18]Interfaculty Institute for Genetics and Functional Genomics, University Medicine Greifswald, Germany; [19]Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD, USA; [20]Institute of Social and Preventive Medicine, Lausanne University Hospital, Lausanne, Switzerland; [21]Institute of Genetic Epidemiology, Helmholtz Zentrum-München, Munich-Neuherberg, Germany; [22]Center for Public Health Genomics, University of Virginia, Charlottesville, VA, USA; [23]Washington University School of Medicine, Department of Genetics, Division of Statistical Genomics, St Louis, MO, USA; [24]Department of Biological Psychology/Netherlands Twin Register, VU University, Amsterdam, The Netherlands; [25]Medical Research Council (MRC) Epidemiology Unit, University of Cambridge, Cambridge, UK; [26]Centre for Vision Research, Department of Ophthalmology and the Westmead Millennium Institute, University of Sydney, Sydney, New South Wales, Australia; [27]Institute of Behavioural Sciences, University of Helsinki, Helsinki, Finland; [28]Folkhälsan Research Centre, Helsinki, Finland; [29]Department of Food and Environmental Sciences, University of Helsinki, Helsinki, Finland; [30]Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku and Turku University Hospital, Turku, Finland; [31]Cardiovascular Health Research Unit, Department of Medicine, University of Washington, Seattle, WA, USA; [32]Centre for Bone and Arthritis Research, Institute of Medicine, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden; [33]University of Mississippi Medical Center, Jackson, MI, USA; [34]Translational Gerontology Branch, National Institute on Aging, NIH, Baltimore, MD, USA; [35]Statens Serum Institut, Department of Epidemiology Research, Copenhagen, Denmark; [36]Busselton Population Medical Research Foundation Inc., Busselton, Western Australia, Australia; [37]PathWest Laboratory Medicine WA, Nedlands, Western Australia, Australia; [38]School of Pathology and Laboratory Medicine, The University of Western Australia, Nedlands, Western Australia, Australia; [39]School of Population Health, The University of Western Australia, Nedlands, Western Australia, Australia; [40]Harokopio University, Athens, Greece; [41]Laboratory of Epidemiology and Population Sciences, National Institute on Aging, NIH, Bethesda, MD, USA; [42]Department of Public Health and Primary Care, Institute of Public Health, University of Cambridge, Cambridge, UK; [43]QIMR Berghofer Medical Research Institute, Queensland, Australia; [44]Department of Biostatistics, University of NC at Chapel Hill, Chapel Hill, NC, USA; [45]Department of Clinical Genetics, Erasmus MC, Rotterdam, The Netherlands; [46]Department of Radiology, Erasmus Medical Center, Rotterdam, The Netherlands; [47]Department of Neurology, Erasmus Medical Center, Rotterdam, The Netherlands; [48]Department of Psychiatry, Erasmus Medical Center, Rotterdam, The Netherlands; [49]Department of Internal Medicine, Erasmus Medical Center, Rotterdam, The Netherlands; [50]Wellcome Trust Sanger Institute, Hinxton, Cambridge, UK; [51]University of Cambridge Metabolic Research Laboratories and NIHR Cambridge Biomedical Research Centre, Cambridge, UK; [52]Department of Odontology, Umeå University, Umeå, Sweden; [53]Biocenter Oulu, University of Oulu, Oulu, Finland; [54]Department of Epidemiology and Biostatistics, MRC Health Protection Agency (HPE) Centre for Environment and Health, School of Public Health, Imperial College London, UK; [55]Institute for Community Medicine, University Medicine Greifswald, Germany; [56]Department of Prosthodontics, Gerodontology and Biomaterials, Center of Oral Health, University Medicine Greifswald, Germany; [57]Department of Internal Medicine, Centre Hospitalier Universitaire Vaudois and University of Lausanne, Lausanne, Switzerland; [58]Department of Medical Genetics, University of Lausanne, Lausanne, Switzerland; [59]Institute of Social and Preventive Medicine (IUMSP), Centre Hospitalier Universitaire Vaudois (CHUV), Lausanne, Switzerland; [60]Swiss Institute of Bioinformatics, Lausanne, Switzerland; [61]Division of Epidemiology and Community Health, School of Public Health, University of Minnesota, Minneapolis, MN, USA; [62]Department of Medical Sciences, Molecular Epidemiology and Science for Life Laboratory, Uppsala University, Uppsala, Sweden; [63]Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK; [64]Department of Surgical Sciences, Uppsala University, Uppsala, Sweden; [65]Department of Chronic Disease Prevention, National Institute for Health and Welfare, Helsinki, Finland; [66]Department of Clinical Physiology, Tampere University Hospital and School of Medicine University of Tampere, Tampere, Finland; [67]Department of Clinical Chemistry, Fimlab Laboratories, and School of Medicine, University of Tampere, Tampere, Finland; [68]Department of Medicine, University of Turku and Turku University Hospital, Turku, Finland; [69]Department of Epidemiology, Harvard School of Public Health, Boston, MA, USA; [70]Friedman School of Nutrition Science and Policy, Tufts University, Boston, MA, USA; [71]Beth Israel Deaconess Medical Center, Boston, MA, USA; [72]Cardiovascular Health Research Unit, Department of Epidemiology, University of

Coffee, a major dietary source of caffeine, is among the most widely consumed beverages in the world and has received considerable attention regarding health risks and benefits. We conducted a genome-wide (GW) meta-analysis of predominately regular-type coffee consumption (cups per day) among up to 91 462 coffee consumers of European ancestry with top single-nucleotide polymorphisms (SNPs) followed-up in ~30 062 and 7964 coffee consumers of European and African-American ancestry, respectively. Studies from both stages were combined in a trans-ethnic meta-analysis. Confirmed loci were examined for putative functional and biological relevance. Eight loci, including six novel loci, met GW significance ($\log_{10}$Bayes factor (BF) > 5.64) with per-allele effect sizes of 0.03–0.14 cups per day. Six are located in or near genes potentially involved in pharmacokinetics (*ABCG2*, *AHR*, *POR* and *CYP1A2*) and pharmacodynamics (*BDNF* and *SLC6A4*) of caffeine. Two map to *GCKR* and *MLXIPL* genes related to metabolic traits but lacking known roles in coffee consumption. Enhancer and promoter histone marks populate the regions of many confirmed loci and several potential regulatory SNPs are highly correlated with the lead SNP of each. SNP alleles near *GCKR*, *MLXIPL*, *BDNF* and *CYP1A2* that were associated with higher coffee consumption have previously been associated with smoking initiation, higher adiposity and fasting insulin and glucose but lower blood pressure and favorable lipid, inflammatory and liver enzyme profiles ($P < 5 \times 10^{-8}$).Our genetic findings among European and African-American adults reinforce the role of caffeine in mediating habitual coffee consumption and may point to molecular mechanisms underlying inter-individual variability in pharmacological and health effects of coffee.

Molecular Psychiatry (2015) **20,** 647–656; doi:10.1038/mp.2014.107; published online 7 October 2014

INTRODUCTION

Coffee is among the most widely consumed beverages in the world.[1] North American coffee drinkers typically consume ~2 cups per day while the norm is at least 4 cups in many European countries.[1] In prospective cohort studies, coffee consumption is consistently associated with lower risk of Parkinson's disease, liver disease and type 2 diabetes.[2] However, the effects of coffee on cancer development, cardiovascular and birth outcomes and other health conditions remain controversial.[2] For most populations, coffee is the primary source of caffeine, a stimulant also present in other beverages, foods and medications.[1,3] The fifth edition of the Diagnostic and Statistical Manual of Mental Disorders does not include a diagnosis of caffeine dependence or abuse due to a paucity of evidence but lists caffeine intoxication and withdrawal as disorders.[4] Knowledge of factors contributing to coffee's consumption and physiological effects may greatly advance the design and interpretation of population and clinical research on coffee and caffeine.[5] Genetic factors could be especially valuable as they offer ways to study the potential health effects of coffee *via* instrumental variables or gene–environment interactions.[5] Heritability estimates for coffee and caffeine use range between 36 and 58%.[6] Genome-wide association studies (GWAS) of habitual caffeine and coffee intake have identified variants near *CYP1A2* and aryl hydrocarbon receptor (*AHR*).[7–9] Cytochrome P450 (CYP)1A2 is responsible for ~95% of caffeine metabolism in humans and AHR has a regulatory role in basal and substrate-induced expression of target genes, including *CYP1A1* and *CYP1A2*.[10,11]

To identify additional loci, we conducted a staged genome-wide (GW) meta-analysis of coffee consumption including over 120 000 coffee consumers sourced from population-based studies of European and African-American ancestry.

Washington, Seattle, WA, USA; [73]Department of Health Services, University of Washington, Seattle, WA, USA; [74]Group Health Research Institute, Group Health Cooperative, Seattle, WA, USA; [75]Institute of Epidemiology, Helmholtz Zentrum-München, Munich-Neuherberg, Germany; [76]Department of Psychiatry, Washington University, St Louis, MO, USA; [77]Department for Psychiatry, Johannes-Gutenberg-University, Mainz, Germany; [78]School of Applied Sciences, University of Mississippi, Oxford, Mississippi, USA; [79]Clinical Laboratory and Nutritional Sciences, University of MA Lowell, Lowell, MA, USA; [80]Department of Genomics of Common Diseases, Imperial College London, London, UK; [81]Medical Department, University of Leipzig, Leipzig, Germany; [82]IFB Adiposity Diseases, University of Leipzig, Leipzig, Germany; [83]William Harvey Research Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, UK; [84]King Abdulaziz University, Jeddah, Saudi Arabia; [85]Wake Forest School of Medicine, Winston-Salem, NC, USA; [86]Wake Forest University Health Sciences, Winston-Salem, NC, USA; [87]Center for Non-Communicable Diseases, Pakistan; [88]Health Disparities Research Section, Clinical Research Branch, National Institute on Aging, NIH, Baltimore, MD, USA; [89]Laboratory of Personality and Cognition, National Institute on Aging, NIH, Baltimore, MD, USA; [90]Department of Genetics, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands; [91]Department of Neurosurgery, Brigham and Women's Hospital, Boston, MA, USA; [92]Department of Surgery, Harvard Medical School, Boston, MA, USA; [93]Monell Chemical Senses Center, Philadelphia, Pennsylvania, USA; [94]Mass Eye and Ear Infirmary, Boston, MA, USA; [95]Geriatric Unit, Azienda Sanitaria Firenze, Florence, Italy; [96]Jean Mayer USDA Human Nutrition Research Center on Aging, Tufts University, Boston, MA, USA; [97]Division of Gastroenterology, MA General Hospital, Boston, MA, USA; [98]Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, WA, USA; [99]Research Unit of Molecular Epidemiology, Helmholtz Zentrum-München, Munich-Neuherberg, Germany; [100]Department of Clinical Physiology and Nuclear Medicine, Turku University Hospital, Turku, Finland; [101]Department of General Practice and Primary Health Care, University of Helsinki, Helsinki, Finland; [102]Helsinki University Central Hospital, Unit of General Practice, Helsinki, Finland; [103]Program in Genetic Epidemiology and Statistical Genetics, Harvard School of Public Health, Boston, MA, USA; [104]The Genetics of Obesity and Related Metabolic Traits Program, The Icahn School of Medicine at Mount Sinai, New York, NY, USA; [105]The Charles Bronfman Institute for Personalized Medicine, The Icahn School of Medicine at Mount Sinai, New York, NY, USA; [106]Department of Psychiatry and Psychotherapy, University Medicine Greifswald, HELIOS Hospital Stralsund, Germany; [107]Department of Endocrinology, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands; [108]School of Population Health, University of South Australia, Adelaide, South Australia, Australia; [109]South Australian Health and Medical Research Institute, Adelaide, South Australia, Australia; [110]Department of Children and Young People and Families, National Institute for Health and Welfare, Oulu, Finland; [111]Unit of Primary Care, Oulu University Hospital, Oulu, Finland; [112]The Framingham Heart Study, Framingham, MA, USA; [113]Department of Public Health and Clinical Medicine, Section for Medicine, Umeå University, Umeå, Sweden; [114]Netherlands Consortium for Healthy Ageing and National Genomics Initiative, Leiden, The Netherlands; [115]Division of Epidemiology, Human Genetics and Environmental Sciences, University of Texas Health Science Center at Houston, Houston, TX, USA and [116]Saw Swee Hock School of Public Health and Department of Medicine, Yong Loo Lin School of Medicine, National University of Singapore and National University Health System, Singapore. Correspondence: Dr MC Cornelis, The Department of Nutrition, Harvard School of Public Health, 401, Park Drive, Boston, MA 02215 USA or Dr DI Chasman, Division of Preventive Medicine, Brigham and Women's Hospital, Harvard Medical School, 900 Commonwealth Avenue, Boston, MA 02215 USA.
E-mail: mcorneli@hsph.harvard.edu or dchasman@research.bwh.harvard.edu
[117]These authors contributed equally to this work.
[118]See Appendix.
Received 23 March 2014; revised 17 July 2014; accepted 22 July 2014; published online 7 October 2014

MATERIALS AND METHODS

Study design and populations

Supplementary Figure S1 depicts an overview of the current study. We performed a meta-analysis of GWAS summary statistics from 28 population-based studies of European ancestry to detect single-nucleotide polymorphisms (SNPs) that are associated with coffee consumption. Top loci were followed-up in studies of European (13 studies) and African-American (7 studies) ancestry and confirmed loci were explored in a single Pakistani population. Detailed information on study design, participant characteristics, genotyping and imputation for all contributing studies are provided in the Supplementary Information and Supplementary Tables S1–S6.

Phenotype

All phenotype data were previously collected *via* interviewer- or self-administered questionnaires (Supplementary Table S1). Our primary phenotype ('phenotype 1') was cups of predominately regular-type coffee consumed per day among coffee consumers. Coffee data collected categorically (for example, 2–3 cups per day) were converted to cups per day by taking the median value of each category (for example, 2.5 cups per day). A secondary analysis was performed comparing high with infrequent/non-coffee consumers ('phenotype 2'). A subset of stage 1 studies collected information on decaffeinated coffee consumption; which was examined in follow-up analysis of the confirmed loci.

Statistical analysis

Each stage 1 (discovery) study performed GWA testing for each phenotype across ~2.5 million genotyped or imputed autosomal SNPs (HapMap II, Centre d'Etude du Polymorphisme Humain (CEU) reference), based on linear (cups per day, phenotype 1) or logistic (high vs none/low, phenotype 2) regression under an additive genetic model. Analyses were adjusted for age, smoking status and, when applicable, sex, case–control status, study site, family structure and/or study-specific principal components of population substructure (Supplementary Table S7). SNPs with minor allele frequency < 0.02 or with low imputation quality scores were removed before meta-analysis (Supplementary Table S5). The GWAtoolbox (see Supplementary Information for URLs) was used for initial quality control. Minor allele frequencies and a plot comparing (1/median standard error of effect size) vs (square root of sample size) for each study were also reviewed for outliers and these were addressed before the final meta-analysis.

For both phenotypes, GW meta-analysis was conducted using a fixed-effects model and inverse-variance weighting with a single genomic control correction as implemented in METAL[12] and GWAMA[13] ($r > 0.99$ for correlation between METAL and GWAMA results). The phenotypic variance explained by additive SNP effects was estimated in the Women's Genome Health Study (WGHS, $n = 15\,987$ with identity-by-state < 0.025) using GCTA.[14] Stage 1 summary statistics were also subjected to pathway analysis using MAGENTA[15] (Supplementary Information).

For regions achieving association *P*-values $< 5 \times 10^{-8}$ (7p21, 7q23.11, 11p13 and 15q24), we performed conditional analysis using the summary statistics from the meta-analysis to test for the association of each SNP while conditioning on the top SNPs, with correlations between SNPs due to linkage disequilibrium (LD) estimated from the imputed genotype data from the Atherosclerosis Risk in Communities cohort,[16] a large and representative cohort of men and women of European ancestry.

Our approach to select SNPs for replication (stage 2) is described in Supplementary Information. Stage 2 meta-analyses were performed separately for European and African-American populations, using the same statistical models and methods as described for stage 1, but without genomic control (Supplementary Information).

Studies from all stages were included in an overall meta-analysis using MANTRA (Meta-ANalysis of TRans-ethnic Association) studies ;[17] which adopts a Bayesian framework to combine results from different ethnic groups by taking advantage of the expected similarity in allelic effects between the most closely related populations. MANTRA was limited to SNPs selected for replication thus no genomic control was applied. A random-effects analysis using GWAMA was performed in parallel to obtain effect estimates, which are not generated by MANTRA. The GW-significance threshold of $\log_{10}$ BF > 5.64 approximates a traditional GW *P*-value threshold of 5×10^{-8} under general assumptions.[18,19] Subgroup analysis and meta-regression were performed to investigate possible sources of between-study heterogeneity (Supplementary Information).

Fine-mapping. To assess the improvement in fine-mapping resolution due to trans-ethnic meta-analysis, we applied the methods of Franceschini *et al.*[17] to stage 1 and stage 2 (African Americans only) GW-summary level data (Supplementary Information).

Potential SNP function and biological and clinical inferences

Details pertaining to follow-up of confirmed loci are provided in the Supplementary Information. Briefly, all confirmed index SNPs and their correlated proxies were examined for putative function using publicly available resources. Bioinformatics and computational tools were used to systematically mine available knowledge and experimental databases to inform biological hypotheses underlying the link between loci and coffee consumption as well as connections between loci. For these analyses all genes mapping to the confirmed regions were considered as potential candidates. Finally, we searched the National Human Genome Research Institute GWAS catalog[20] and Metabolomics GWAS server[21] for all GW-significant associations with our confirmed coffee SNPs. Complete GWAS summary data for coffee-implicated diseases or traits were additionally queried.

RESULTS

SNPs associated with coffee consumption

Discovery stage. Results from the discovery stage are summarized in Supplementary Figures S2–S5. Little evidence for genomic inflation ($\lambda < 1.07$) was observed for either phenotype. The two analyses yielded similarly ranked loci and significant enrichment of 'xenobiotic' genes (MAGENTA's FDR < 0.006), suggesting no major difference in the genetic influence on coffee drinking initiation compared with the level of coffee consumption among coffee consumers at these loci. Overall, ~7.1% (standard error: 2%) of the variance in coffee cups consumed per day (phenotype 1) could be explained by additive and common SNP effects in the WGHS.

Conditioning on the index SNPs of each region achieving association *P*-values $< 5 \times 10^{-8}$ (7p21, 7q23.11, 11p13 and 15q24) in the discovery stage provided little evidence for multiple independent variants (Supplementary Figure S6). Only four of the SNPs on chromosome 7 were potentially independent and carried forward with other promising SNPs.

Replication and trans-ethnic meta-analysis. Forty-four SNPs spanning thirty-three genomic regions met significance criteria for candidate associations and were followed-up in stage 2 (Supplementary Tables S8–S13). Eight loci, including six novel, met our criteria for GW significance ($\log_{10}$ BF > 5.64) in a trans-ethnic meta-analysis of all discovery and replication studies (Table 1; Supplementary Tables S14–S16; Supplementary Figures S7 and S8). Confirmed loci have effect sizes of 0.03–0.14 cups per day per allele and together explain ~1.3% of the phenotypic variance of coffee intake. We were underpowered to replicate these associations in a Pakistani population (Supplementary Information).

Functional and biological inferences

Enhancer (H3K4me1) and promoter (H3K4me3) histone marks densely populate many of these regions and several non-synonymous and potential regulatory SNPs are highly correlated ($r^2 > 0.8$) with the lead SNP and thus strong candidates for being a causal variant (Table 2; Supplementary Information; Supplementary Tables S17–S19). Candidate genes form a highly connected network of interactions, featuring discernible clusters of genes around brain-derived neurotrophin factor (*BDNF*) and *AHR* (Figure 1; Supplementary Information; Supplementary Tables S20 and S21). At least one gene in each of the eight regions (i) is highly expressed in brain, liver and/or taste buds, (ii) results in phenotype abnormalities relevant to coffee consumption behavior when modified in mice and (iii) is differentially expressed in human hepatocytes when treated with high (7500 μM) but not low (1500 μM) doses of caffeine (Table 2; Supplementary Tables S22–S24).

Table 1. SNPs associated with cups of coffee consumed per day among coffee consumers

Locus	*Index SNP*[a]	*Closest gene*	*EA/NEA*	*EAF EUR/AA*	*Stage 1*[b] *EUR* n ⩽ *91 462*		*Stage 2*[b]				*Trans-ethnic meta-analysis*[c]			
							EUR n ⩽ *30 062*		*AA* n ⩽ *7964*					
					β (s.e.)	P	*β (s.e.)*	P	*β (s.e.)*	P	*β (s.e.)*	*Log*$_{10}$*BF*	*Post Prob*	n
2p24	**rs1260326**	*GCKR*	T/C	0.41/0.17	−0.04 (0.01)	1.06×10^{-07}	−0.03 (0.01)	0.02	−0.01 (0.03)	0.77	−0.04 (0.01)	6.48	0.07	129 417
4q22	**rs1481012**	*ABCG2*	A/G	0.89/0.95	0.06 (0.01)	1.13×10^{-06}	0.03 (0.02)	0.11	0.16 (0.05)	1.27×10^{-03}	0.06 (0.01)	6.08	0.23	126 019
7p21	rs4410790	*AHR*	T/C	0.37/0.52	−0.14 (0.01)	1.48×10^{-57}	−0.05 (0.01)	1.66×10^{-04}	−0.09 (0.02)	2.37×10^{-06}	−0.10 (0.01)	58.87	0.96	116 674
	rs6968554		A/G	0.39/0.33	−0.13 (0.01)	2.54×10^{-57}	−0.07 (0.01)	2.78×10^{-10}	−0.05 (0.02)	0.02	−0.10 (0.01)	69.69	1.00	124 849
7q11.23	**rs7800944**	*MLXIPL*	T/C	0.72/0.67	−0.05 (0.01)	7.82×10^{-09}	−0.06 (0.02)	4.20×10^{-04}	−0.02 (0.02)	0.37	−0.05 (0.01)	8.83	0.09	116 417
7q11.23	**rs17685**	*POR*	A/G	0.29/0.19	0.07 (0.01)	9.06×10^{-14}	0.05 (0.01)	1.01×10^{-03}	0.07 (0.03)	7.55×10^{-03}	0.07 (0.01)	15.12	0.08	115 465
11p13	**rs6265**	*BDNF*	T/C	0.19/0.07	−0.05 (0.01)	3.40×10^{-07}	−0.03 (0.01)	0.07	−0.05 (0.04)	0.25	−0.04 (0.01)	5.76	0.10	127 828
15q24	rs2470893	*CYP1A1*	T/C	0.31/0.06	0.12 (0.01)	6.89×10^{-44}	0.09 (0.01)	9.92×10^{-11}	0.20 (0.07)	4.23×10^{-03}	0.12 (0.01)	57.79	1.00	113 273
	rs2472297	*CYP1A2*	T/C	0.24/0.06	0.15 (0.01)	6.45×10^{-47}	0.11 (0.01)	3.26×10^{-16}	0.19 (0.05)	8.62×10^{-05}	0.14 (0.01)	62.77	0.97	116 272
17q11.2	**rs9902453**	*EFCAB5*	A/G	0.54/0.80	−0.04 (0.01)	2.26×10^{-06}	−0.03 (0.01)	9.13×10^{-03}	−0.04 (0.03)	0.17	−0.03 (0.01)	6.29	0.05	126 819

Abbreviations: AA, African-American ancestry; BF, Bayes-factor; EA, effect allele; EAF, effect allele frequency; EUR, European ancestry; NEA, non-effect allele; Post Prob, posterior probability; SNP, single-nucleotide polymorphism. [a]Genic SNPs are in boldface. [b]Effect coefficients (s.e.), representing cups per day per effect allele, and corresponding *P*-values from stage 1 fixed-effects meta-analysis (columns 6 and 7) and stage 2 fixed-effects meta-analyses (columns 8–11). [c]Effect coefficients (s.e.), representing cups per day per effect allele, from random-effects meta-analysis of all stage 1 and stage 2 studies (column 12). $Log_{10}BF$ (column 13) and the corresponding posterior probabilities (column 14) from trans-ethnic meta-analysis of all stage 1 and stage 2 studies. A posterior probability of >0.5 suggests heterogeneity in allelic effects.

Additional genomic characterization of the top loci allows further biological inference as follows:

(i) Previously identified loci near AHR (7p21) and CYP1A2 (15q24). Consistent with previous reports in smaller samples,[7–9] the intergenic 7p21 and 15q24 loci near *AHR* and *CYP1A1/CYP1A2* respectively remained the most prominent and highly heterogeneous loci associated with coffee consumption. The same index SNPs were identified in European and African Americans, suggesting that they are robust HapMap proxies for causal variants in these two populations. Cohort-wide mean coffee consumption explained part of the heterogeneity in study results for both loci (Supplementary Table S25; Supplementary Information). The rs2472297 T and rs4410790 C alleles associated with increased coffee consumption have recently been associated with lower plasma caffeine levels[21] and shown to increase CYP1A2-mediated metabolism of olanzapine.[22] The C allele of rs4410790 is also positively correlated with cerebellum *AHR* methylation, suggesting a novel role of Ahr in motor or learning pathways that may trigger coffee consumption. The most significant variants at 15q24 reside in the *CYP1A1-CYP1A2* bidirectional promoter where *AHR* response elements have been identified and shown to be important for transcriptional activation of both *CYP1A1* and *CYP1A2*.[23] The rs2472297 T variant putatively weakens the binding of SP1, a co-activator in the Ahr–Arnt complex regulating CYP1 locus transcription[24] and is also implicated in the expression of several neighboring genes. The latter observation, together with this region's high LD and long range chromatin interactions (Supplementary Figure S9), suggests a regulatory network among these genes.

(ii) Novel loci at 7q11.23 (POR) and 4q22 (ABCG2) likely function in caffeine metabolism. Variants at 7q11.23 (rs17685) and 4q22 (rs1481012) map to novel yet biologically plausible candidate genes involved in xenobiotic metabolism. rs17685 maps to the 3′UTR of *POR*, encoding P450 oxidoreductase which transfers electrons to all microsomal CYP450 enzymes.[25] The rs17685 A variant associated with higher coffee consumption is linked to increased *POR* expression and potentially weakens the DNA binding of several transcriptional regulatory proteins including BHLHE40, which inhibits *POR* expression.[26] The same SNP is in LD (CEU: $r^2=0.93$) with *POR*28* (rs1057868 and Ala503Val), which is associated with differential CYP activity depending on the CYP isoform, substrate and experimental model used.[27] rs1481012 at 4q22 maps to *ABCG2*, encoding a xenobiotic efflux transporter. rs1481012 is in LD (CEU: $r^2=0.92$) with rs2231142 (Gln141Lys), a functional variant at an evolutionarily constrained residue.[28] However, fine-mapping of this region on the basis of reduced LD in the African-American sample limited an initial 189 102-kb region to a credible span of 6249 kb (Supplementary Table S16) that excluded rs2231142.

(iii) Novel loci at 11p13 (BDNF) and 17q11.2 ('SLC6A4') likely mediate the positive reinforcing properties of coffee constituents. The index SNP at 11p13 is the widely investigated missense mutation (rs6265 and Val66Met) in *BDNF* (Supplementary Table S26). BDNF modulates the activity of serotonin, dopamine and glutamate, and neurotransmitters involved in mood-related circuits and have a key role in memory and learning.[29] The Met66 allele impairs neuronal activity-dependent BDNF secretion[30] and thus may attenuate the rewarding effects of coffee and, in turn, motivation to consume coffee. The increasingly recognized roles of BDNF in the chemosensory system and conditioned taste preferences may also be relevant.[31] The index SNP (rs9902453) at 17q11.2 maps to the *EFCAB5* gene and is in LD (CEU: $r^2>0.8$) with SNPs that alter regulatory motifs for AhR[32] in the neighboring gene *NSRP1*, but neither gene is an obvious candidate for coffee consumption. Upstream of rs9902453 lies a possibly stronger candidate: *SLC6A4*

Table 2. Potential function of loci associated with coffee consumption[a]

Locus	Gene expression response to caffeine[b]	Lead-SNP, allele ↑coffee consumption[c]	Non-Syn SNPs in LD[d]	CR[e]	DNAse[f]	Proteins bound[g]	Histone marks[h]	Motifs changed[i]	eQTL[j]	mQTL[k]
2p24	*GCKR, CCDC121, FNDC4, ZNF513, SNX17, PPM1G, GPN1, SUPT7L, MPV17, SLC4A1AP, PREB, ATRAID, GTF3C2*	rs1260326, C	Leu446Pro	✓	✓	✓	Enhancer	NRSF	*EIF2B4, SNX17, NRBP1*	*KRTCAP3, PPM1G*
4q22	*ABCG2, SPP1*	rs1481012, A	✓	✓	✓	✓	Enhancer	AIRE, Zfp105		
7p21	*AHR*	rs4410790, C rs6968554, G		✓	✓			Cdx2, DMRT3, E4BP4, Foxa, GR, Hoxa10, Hoxa9, Hoxb13, Hoxb9, Hoxc9, Hoxd10, Myc, p300, TR4		*AHR*
7q11.23	*MLXIPL, BCL7B, DNAJC30, TBL2, WBSCR22*	rs7800944, C			✓	✓	Promoter enhancer	AP-4, BHLHE40, GATA,GR, Irf, Pax-5	*WBSCR22, MLXIPL*	*FZD9*
7q11.23	*RHBDD2, POR STYXL1, TMEM120A, MDH2, HSPB1*	rs17685, A	✓	✓	✓	✓		Arnt, BHLHE40, DEC,Ets, Mxi1,Myc, Pax-5, Sin3Ak-20, TFE	*RHBDD2, POR, TMEM120A, STYXL1, MDH2*	*STYXL1*
11p13	*CCDC34, LIN7C, METTL15,*	rs6265, C	Val66Met	✓	✓	✓	Promoter enhancer	BHLHE40, Myc, SREBP		
15q24	*PPCDC, ARID3B, ULK3, SEMA7A, EDC3, COX5A, CSK, RPP25, MPI*	rs2470893, T rs2472297, T						SP1	*MPI, SCAMP2, ULK3, ISLR, SNUPN, RPP25, CSK,*	*SCAMP2*
17q11.2	*TAOK1, SLC6A4 NSRP1, BLMH*	rs9902453, G	✓	✓	✓	✓	Promoter enhancer	STAT	*GIT1, ATAD5, SLC6A4*	*NSRP1, ANKRD13B, CRLF3, CORO6*

Abbreviations: CEU, Centre d'Etude du Polymorphisme Humain; CR, conserved region; eQTL, expression quantitative trait loci; LD, linkage disequilibrium; mQTL, methylation quantitative trait loci; SNP, single-nucleotide polymorphism. [a]See Supplementary Information for details and references to data resources. [b]*In vitro* human hepatic gene expression in response to caffeine. Red and green font corresponds to increased and decreased expression, respectively. [c]Lead SNP allele associated with *higher* coffee consumption. [d]Check marks (✓) denote the *presence* of non-synonymous SNPs in LD (CEU: $r^2 \geqslant 0.80$) with lead SNP (details provided for lead SNP only). [e]Check marks (✓) denote the *presence* of a conserved region (spanning lead SNP and its correlated proxies, CEU: $r^2 \geqslant 0.8$). [f]Check marks (✓) denote the *presence* of DNAse hypersensitivity sites at region spanning lead SNP and its correlated proxies, CEU: $r^2 \geqslant 0.8$. [g]Check marks (✓) denote the *presence* of proteins bound at region spanning lead SNP and its correlated proxies, CEU: $r^2 \geqslant 0.8$. [h]Enhancer (H3K4me1) or promoter (H3K4me3) histone marks (as defined by Ernst *et al.*[32]) spanning lead SNP and its correlated proxies, CEU: $r^2 \geqslant 0.8$. [i]Regulatory motifs altered by lead SNP. [j]Expression QTLs for lead SNP or perfect proxy (CEU: $r^2 = 1$) derived from lymphoblastoid cell lines, blood, or liver, adipose and brain tissues. Red and green font corresponds to increased and decreased expression, respectively, relative to allele associated with *higher* coffee consumption. Direction of *GITI* expression is not available. [k]Methylation QTLs for lead SNP derived from cerebellum and frontal cortex. Red and green font corresponds to increased and decreased expression, respectively, relative to allele associated with *higher* coffee consumption.

encoding the serotonin transporter. Serotonergic neurotransmission affects a wide range of behaviors including sensory processing and food intake.[33]

(iv) Novel loci at 2p24 (GCKR) and 17q11.2 (MLXIPL). Variants at 2p24 (rs1260326) and 7q11.23 (rs7800944) map to *GCKR* and *MLXIPL*, respectively. The former has been associated with plasma glucose and multiple metabolic traits and the latter with plasma triglycerides (Table 3; Supplementary Table S27). Adjustment of regression models for plasma lipids in the WGHS ($n \sim 17\,000$) and plasma glucose in TwinGene ($n \sim 8800$) did not significantly change the relationship between SNPs at these two loci and coffee consumption ($P > 0.48$, Supplementary Tables S28 and S29). The rs1260326 T allele encodes a non-synonymous change in the encoded, glucokinase regulatory protein leading to increased hepatic glucokinase activity.[34] Glucokinase regulatory protein and glucokinase may also cooperatively function in the glucose-sensing process of the brain[35] that may, in turn, influence central pathways responding to coffee constituents. A direct link between *MLXIPL* and coffee consumption remains unclear, except for the interactions with other candidate genes (Figure 1). Experimental evidence and results from formal prioritization analyses also warrants consideration of other candidates in these regions (Figure 1; Table 2; Supplementary Tables S23). For example, in the frontal cortex, the rs1260326 allele positively associated with coffee consumption correlates with lower methylation of *PPM1G*; a putative regulatory target for AhR and binding target for *PPP1R1B*, which mediates psychostimulant effects of caffeine.[36]

Pleiotropy and clinical inferences

None of the eight loci was significantly associated with caffeine taste intensity ($P > 0.02$) or caffeine-induced insomnia ($P > 0.08$), according to previously published GWAS of these traits.[37–39] SNPs near *AHR* associated with higher coffee consumption were also significantly associated with higher decaffeinated coffee consumption (~0.05 cups per day, $P < 0.0004$, $n = 24\,426$); perhaps a result of Pavlovian conditioning among individuals moderating their intake of regular coffee or the small amounts of caffeine in decaffeinated coffee.[1]

Across phenotypes in the GWAS catalog,[20] the alleles leading to higher coffee consumption at 2p24, 4q22, 7q11.23, 11p13 and 15q24 have been associated with one or more of the following: smoking initiation, higher adiposity and fasting insulin and glucose but lower blood pressure and favorable lipid, inflammatory and liver enzyme profiles ($P < 5 \times 10^{-8}$, Table 3; Supplementary Table S27). Focused on metabolic, neurologic and psychiatric traits for which coffee has been implicated (Table 3; Supplementary Table S32), there were additional sub-GW significant associations in published GWAS. Variants associated with higher coffee consumption increased adiposity (rs1481012, $P = 4.85 \times 10^{-3}$), birth weight (rs7800944, $P = 2.10 \times 10^{-3}$), plasma high-density lipoprotein (HDL, rs7800944, $P = 2.24 \times 10^{-3}$), risk of Parkinson's disease (rs1481012, $P = 7.11 \times 10^{-3}$), reduced blood pressure (rs6265, $P = 6.58 \times 10^{-4}$; rs2472297, $P < 6.80 \times 10^{-5}$ and rs9902453, $P = 6.05 \times 10^{-3}$), HDL (rs6968554, $P = 1.18 \times 10^{-3}$), risk of major depressive disorder (rs17685, $P = 6.98 \times 10^{-3}$) and bipolar disorder (rs1260326, $P = 2.31 \times 10^{-3}$). Associations with adiposity, birth weight, blood pressure, HDL and bipolar disorder remain significant after correcting for the number of SNPs tested.

DISCUSSION

Coffee's widespread popularity and availability has fostered public health concerns of the potential health consequences of regular coffee consumption. Findings from epidemiological studies of

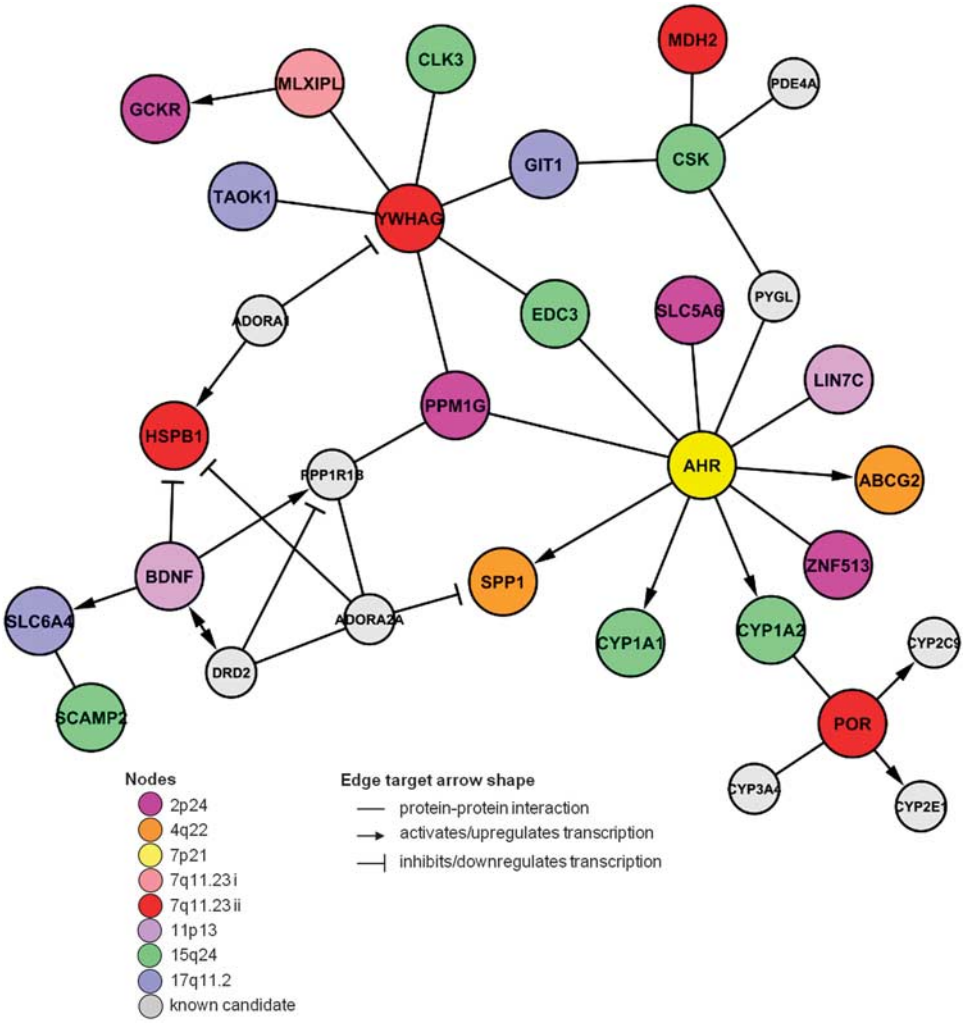


Figure 1. Network describing direct interactions between candidate genes of confirmed loci. Relationships were retrieved from databases of transcription regulation and protein–protein interaction experiments (Supplementary Table S21). Genes are represented as nodes that are colored according to locus. Candidate genes for loci identified in the current study were supplemented with known candidate genes related to caffeine pharmacology (gray nodes). Edges indicate known interactions.

coffee consumption and certain health conditions remain controversial.[2] Knowledge of genetic factors contributing to coffee's consumption and physiological effects may inform the design and interpretation of population and clinical research on coffee.[5] In the current report, we present results of the largest GWAS of coffee intake to-date and the first to include populations of African-American ancestry. In addition to confirming associations with *AHR* and *CYP1A2*, we have identified six new loci, not previously implicated in coffee drinking behavior.

Our findings highlight an important role of the pharmacokinetic and pharmacodynamic properties of the caffeine component of coffee underlying a genetic propensity to consume the beverage. Loci near *BDNF* and *SLC6A4* potentially impact consumption behavior by modulating the acute behavioral and reinforcing properties of caffeine. Others near *AHR*, *CYP1A2*, *POR* and *ABCG2* act indirectly by altering the metabolism of caffeine and thus the physiological levels of this stimulant. The strength of these four associations with coffee intake, along with results from pathway analysis showing significant enrichment for 'xenobiotic' genes, emphasize an especially pronounced role of caffeine metabolism in coffee drinking behavior. The current study is the first to link *GCKR* and *MLXIPL* variation to a behavioral trait. The non-synonymous rs1260326 SNP in *GCKR* has been a GW signal for various metabolic traits particularly those reflecting glucose homeostasis (Table 3). *GCKR* variation may impact the glucose-sensing process of the brain[35] that may, in turn, influence central pathways responding to coffee constituents. Methylation quantitative trait loci and binding motif analysis suggest that *PPM1G* may be another candidate underlying the association between rs1260326 and coffee consumption. Variants near *MLXIPL* have also topped the list of variants associated with plasma triglycerides (Table 3), but their link to coffee consumption remains unclear. Future studies on the potential pleiotropic effects of these two loci are clearly warranted. Interestingly, several candidate genes implicated in coffee consumption behavior, but not confirmed in our GWAS, interact with one or more of the eight confirmed loci (Figure 1). While these findings are encouraging for ongoing efforts they also emphasize the need to study sets or pathways of genes in the future.

Specific SNPs associated with higher coffee consumption have previously been associated with smoking initiation, higher adiposity and fasting insulin and glucose but lower blood pressure and favorable lipid, inflammatory and liver enzyme profiles. Whether these relationships reflect pleiotropy, confounding or

Table 3. Associations between coffee consumption loci and other traits

Lead SNP, allele ↑ coffee consumption[a]	*Other traits*[b]	
closest gene	*Higher levels/risk*[c]	*Lower levels/risk*[c]
rs1260326, C *GCKR*	Non-albumin protein Fasting glucose HOMA-IR Fasting insulin Mannose	Serum albumin 2-H glucose challenge Metabolic syndrome Glucose/mannose ratio Total cholesterol Triglycerides Hypertriglyceridemia Chronic kidney disease Uric acid SHBG Crohn's disease C-reactive protein Platelet counts GGT Docosapentaenoic acid Alanine/glutamine ratio Alanine
		LDL ($P = 2.33 \times 10^{-4}$) Waist-to-hip-ratio ($P = 3.40 \times 10^{-4}$) Bipolar disorder ($P = 2.31 \times 10^{-3}$)
rs1481012, A *ABCG2*		LDL response to statins ('responders') Uric acid
	Body mass index ($P = 4.85 \times 10^{-3}$)	
rs6968554, G *AHR*		Caffeine
		HDL ($P = 1.18 \times 10^{-3}$)
rs7800944, C *MLXIPL*		triglycerides
	HDL ($P = 2.24 \times 10^{-3}$) Birth weight ($P = 2.10 \times 10^{-3}$)	
rs6265, C *BDNF*	Smoking initiation Body mass index	
		DBP ($P = 6.58 \times 10^{-4}$)
rs2472297[d], T *CYP1A1_CYP1A2*		*Caffeine*[e]
		SBP ($P = 6.81 \times 10^{-5}$) DBP ($P = 6.75 \times 10^{-6}$)
rs9902453, G *EFCAB5*		SBP ($P = 6.05 \times 10^{-3}$)

Abbreviations: CEU, Centre d'Etude du Polymorphisme Humain; DBP, diastolic blood pressure; GGT, gamma-glutamyltransferase; HDL, high-density lipoprotein; LD, linkage disequilibrium; LDL, low-density lipoprotein; SBP, systolic blood pressure; SHBG, sex hormone binding globulin; SNP, single-nucleotide polymorphism. [a]Lead SNP allele associated with *higher* coffee consumption. [b]Traits associated with lead SNP (or close proxies: $r^2 > 0.80$) according to previous GWAS[20] (Shin *et al.*[21]). Gray cells denote all GW-significant significant associations ($P < 5.00 \times 10^{-8}$ [20] or $P < 1.03 \times 10^{-10}$ (Shin *et al.*[21]) and white cells denote coffee-relevant trait associations ($P < 6.25 \times 10^{-3}$). See Supplementary Information for details and references to original GWAS. [c]Relative to allele associated with *higher* coffee consumption. [d]rs1378942 A, also associated with higher coffee consumption ($P < 1.46 \times 10^{-17}$) in stage 1 of the current report but in low LD with rs2472297 (CEU: $r^2 = 0.10$), was previously associated with lower DBP in GWAS ($P < 5.00 \times 10^{-8}$). [e]Borderline significant ($P < 1.51 \times 10^{-10}$) according to Shin *et al.*[21]

offer insight to the potential causal role coffee plays in these traits merits further investigation. Future research, particularly Mendelian Randomization and gene–coffee interaction studies, will need to consider the direct and indirect roles that each SNP has in altering coffee drinking behavior as well as the potential for interactions between loci (Figure 1). The heterogeneous effects specific to *AHR*- and *CYP1A2*-coffee associations point to SNP-specific interactions with the environment or population characteristics that might also warrant consideration (Supplementary Information).

The strong cultural influences on norms of coffee drinking may have reduced our power for loci discovery. This might, in part, underlie our lack of replication in a Pakistani population, wherein coffee consumption is extremely rare. Methodological limitations specific to our approach may also have reduced our power for loci discovery or precision in estimating effect sizes (Supplementary Information). For example, some studies collected coffee data in categories of cups per day (for example, 2–3 cups per day) rendering a less precise record of intake as well as a non-Gaussian distributed trait for analysis. The precise chemical composition of different coffee preparations is also not captured by standard food frequency questionnaire and is likely to vary within and between populations. Nevertheless, the eight loci together explain ~1.3% of the phenotypic variance, a value *at least* as great as that reported for smoking behavior and alcohol consumption which are subjected to similar limitations in GWAS.[40,41]

The additive genetic variance (or narrow-sense heritability) of coffee intake as estimated by GCTA in WGHS (7%) is considerably lower than estimates based on pedigrees (36–57%).[6] The marked discrepancies between the GCTA and pedigree estimates of heritability may be due to one or more of the following: the potential contribution of rare variants to heritability (not captured by GCTA's 'chip-based heritability'), biases in pedigree analysis resulting in overestimates of heritability, differences in phenotype ascertainment or definition and cultural differences in the populations studied.[42]

In conclusion, our results support the hypothesis that metabolic and neurological mechanisms of caffeine contribute to coffee consumption habits. Individuals adapt their coffee consumption habits to balance perceived negative and reinforcing symptoms that are affected by genetic variation. Genetic control of this potential 'titrating' behavior would incidentally govern exposure to other potentially 'bioactive' constituents of coffee that may be related to the health effects of coffee or other sources of caffeine. Thus, our findings may point to molecular mechanisms underlying inter-individual variability in pharmacological and health effects of coffee and caffeine.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

Study-specific funding and acknowledgments are provided in the Supplementary Information. We collectively thank everyone who has contributed to the collection, genotyping and analysis of the individual cohorts, as well as all the study participants.

REFERENCES

1 Fredholm BB, Battig K, Holmen J, Nehlig A, Zvartau EE. Actions of caffeine in the brain with special reference to factors that contribute to its widespread use. *Pharmacol Rev* 1999; **51**: 83–133.
2 Cornelis MC. Gene-coffee interactions and health. *Curr Nutr Rep* 2014; **3**: 178–195.
3 Spiller MA. The chemical components of coffee. In: Spiller GA (eds) *Caffeine*. CRC: Boca Raton, 1998, pp 97–161.
4 American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*, 5th edn. American Psychiatric Publishing: Arlington, VA, 2013.
5 Cornelis MC. Coffee intake. *Prog Mol Biol Transl Sci* 2012; **108**: 293–322.

6 Yang A, Palmer AA, de Wit H. Genetics of caffeine consumption and responses to caffeine. *Psychopharmacology* 2010; **211**: 245–257.
7 Cornelis MC, Monda KL, Yu K, Paynter N, Azzato EM, Bennett SN *et al.* Genome-wide meta-analysis identifies regions on 7p21 (AHR) and 15q24 (CYP1A2) as determinants of habitual caffeine consumption. *PLoS Genet* 2011; **7**: e1002033.
8 Sulem P, Gudbjartsson DF, Geller F, Prokopenko I, Feenstra B, Aben KK *et al.* Sequence variants at CYP1A1-CYP1A2 and AHR associate with coffee consumption. *Hum Mol Genet* 2011; **20**: 2071–2077.
9 Amin N, Byrne E, Johnson J, Chenevix-Trench G, Walter S, Nolte IM *et al.* Genome-wide association analysis of coffee drinking suggests association with CYP1A1/CYP1A2 and NRCAM. *Mol Psychiatry* 2011; **17**: 1116–1129.
10 Kot M, Daniel WA. The relative contribution of human cytochrome P450 isoforms to the four caffeine oxidation pathways: an in vitro comparative study with cDNA-expressed P450s including CYP2C isoforms. *Biochem Pharmacol* 2008; **76**: 543–551.
11 Le Vee M, Jouan E, Fardel O. Involvement of aryl hydrocarbon receptor in basal and 2,3,7,8-tetrachlorodibenzo-p-dioxin-induced expression of target genes in primary human hepatocytes. *Toxicol In Vitro* 2010; **24**: 1775–1781.
12 Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genome-wide association scans. *Bioinformatics* 2010; **26**: 2190–2191.
13 Mägi R, Morris AP. GWAMA: software for genome-wide association meta-analysis. *BMC Bioinformatics* 2010; **11**: 288.
14 Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet* 2011; **88**: 76–82.
15 Segre AV, Groop L, Mootha VK, Daly MJ, Altshuler D. Common inherited variation in mitochondrial genes is not enriched for associations with type 2 diabetes or related glycemic traits. *PLoS Genet* 2010; **6**: e1001058.
16 Yang J, Ferreira T, Morris AP, Medland SE, Madden PA, Heath AC *et al.* Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nat Genet* 2012; **44**: 369–375, S361–S363.
17 Franceschini N, van Rooij FJ, Prins BP, Feitosa MF, Karakas M, Eckfeldt JH *et al.* Discovery and fine mapping of serum protein loci through transethnic meta-analysis. *Am J Hum Genet* 2012; **91**: 744–753.
18 Stephens M, Balding DJ. Bayesian statistical methods for genetic association studies. *Nat Rev Genet* 2009; **10**: 681–690.
19 Sellke T, Bayarri M, Berger J. Calibration of *P* values for testing precise null hypotheses. *Am Stat* 2001; **55**: 62–71.
20 Hindorff LA, MacArthur J, Morales J, Junkins HA, Hall PN, Klemm AK *et al.* A catalog of published genome-wide association studies. Available at: www.genome.gov/gwastudies; accessed 1 January 2013.
21 Shin SY, Fauman EB, Petersen AK, Krumsiek J, Santos R, Huang J *et al.* An atlas of genetic influences on human blood metabolites. *Nat Genet* 2014; **46**: 543–550.
22 Soderberg MM, Haslemo T, Molden E, Dahl ML. Influence of CYP1A1/CYP1A2 and AHR polymorphisms on systemic olanzapine exposure. *Pharmacogenet Genomics* 2013; **23**: 279–285.
23 Jorge-Nebert LF, Jiang Z, Chakraborty R, Watson J, Jin L, McGarvey ST *et al.* Analysis of human CYP1A1 and CYP1A2 genes and their shared bidirectional promoter in eight world populations. *Hum Mutat* 2010; **31**: 27–40.
24 Swanson HI. DNA binding and protein interactions of the AHR/ARNT heterodimer that facilitate gene activation. *Chem Biol Interact* 2002; **141**: 63–76.
25 Hu L, Zhuo W, He YJ, Zhou HH, Fan L. Pharmacogenetics of P450 oxidoreductase: implications in drug metabolism and therapy. *Pharmacogenet Genomics* 2012; **22**: 812–819.
26 Rome S, Meugnier E, Lecomte V, Berbe V, Besson J, Cerutti C *et al.* Microarray analysis of genes with impaired insulin regulation in the skeletal muscle of type 2 diabetic patients indicates the involvement of basic helix-loop-helix domain-containing, class B, 2 protein (BHLHB2). *Diabetologia* 2009; **52**: 1899–1912.
27 Pandey AV, Fluck CE. NADPH P450 oxidoreductase: structure, function, and pathology of diseases. *Pharmacol Ther* 2013; **138**: 229–254.
28 Woodward OM, Tukaye DN, Cui J, Greenwell P, Constantoulakis LM, Parker BS *et al.* Gout-causing Q141K mutation in ABCG2 leads to instability of the nucleotide-binding domain and can be corrected with small molecules. *Proc Natl Acad Sci USA* 2013; **110**: 5223–5228.
29 Numakawa T, Suzuki S, Kumamaru E, Adachi N, Richards M, Kunugi H. BDNF function and intracellular signaling in neurons. *Histol Histopathol* 2010; **25**: 237–258.
30 Egan MF, Kojima M, Callicott JH, Goldberg TE, Kolachana BS, Bertolino A *et al.* The BDNF val66met polymorphism affects activity-dependent secretion of BDNF and human memory and hippocampal function. *Cell* 2003; **112**: 257–269.
31 Nosrat IV, Margolskee RF, Nosrat CA. Targeted taste cell-specific overexpression of brain-derived neurotrophic factor in adult taste buds elevates phosphorylated TrkB protein levels in taste cells, increases taste bud size, and promotes gustatory innervation. *J Biol Chem* 2012; **287**: 16791–16800.
32 Ernst J, Kheradpour P, Mikkelsen TS, Shoresh N, Ward LD, Epstein CB *et al.* Mapping and analysis of chromatin state dynamics in nine human cell types. *Nature* 2011; **473**: 43–49.
33 Canli T, Lesch KP. Long story short: the serotonin transporter in emotion regulation and social cognition. *Nat Neurosci* 2007; **10**: 1103–1109.
34 Beer NL, Tribble ND, McCulloch LJ, Roos C, Johnson PR, Orho-Melander M *et al.* The P446L variant in GCKR associated with fasting plasma glucose and triglyceride levels exerts its effect through increased glucokinase activity in liver. *Hum Mol Genet* 2009; **18**: 4081–4088.
35 Alvarez E, Roncero I, Chowen JA, Vazquez P, Blazquez E. Evidence that glucokinase regulatory protein is expressed and interacts with glucokinase in rat brain. *J Neurochem* 2002; **80**: 45–53.
36 Lindskog M, Svenningsson P, Pozzi L, Kim Y, Fienberg AA, Bibb JA *et al.* Involvement of DARPP-32 phosphorylation in the stimulant action of caffeine. *Nature* 2002; **418**: 774–778.
37 Reed DR, Zhu G, Breslin PA, Duke FF, Henders AK, Campbell MJ *et al.* The perception of quinine taste intensity is associated with common genetic variants in a bitter receptor cluster on chromosome 12. *Hum Mol Genet* 2010; **19**: 4278–4285.
38 Ledda M, Kutalik Z, Souza Destito MC, Souza MM, Cirillo CA, Zamboni A *et al.* GWAS of human bitter taste perception identifies new loci and reveals additional complexity of bitter taste genetics. *Hum Mol Genet* 2013; **23**: 259–267.
39 Byrne EM, Johnson J, McRae AF, Nyholt DR, Medland SE, Gehrman PR *et al.* A genome-wide association study of caffeine-related sleep disturbance: confirmation of a role for a common variant in the adenosine receptor. *Sleep* 2012; **35**: 967–975.
40 Tobacco and Genetics Consortium. Genome-wide meta-analyses identify multiple loci associated with smoking behavior. *Nat Genet* 2010; **42**: 441–447.
41 Schumann G, Coin LJ, Lourdusamy A, Charoen P, Berger KH, Stacey D *et al.* Genome-wide association and genetic functional studies identify autism susceptibility candidate 2 gene (AUTS2) in the regulation of alcohol consumption. *Proc Natl Acad Sci USA* 2011; **108**: 7119–7124.
42 Vinkhuyzen AA, Wray NR, Yang J, Goddard ME, Visscher PM. Estimation and partition of heritability in human populations using whole-genome analysis methods. *Annu Rev Genet* 2013; **47**: 75–95.

Supplementary Information accompanies the paper on the Molecular Psychiatry website (http://www.nature.com/mp)

APPENDIX

The members and affiliations of the International Parkinson Disease Genomics Consortium (IPDGC) are as follows: Michael A Nalls (Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, MD, USA), Vincent Plagnol (UCL Genetics Institute, London, UK), Dena G Hernandez (Laboratory of Neurogenetics, National Institute on Aging; and Department of Molecular Neuroscience, UCL Institute of Neurology, London, UK), Manu Sharma (Department for Neurodegenerative Diseases, Hertie Institute for Clinical Brain Research, University of Tübingen, and DZNE, German Center for Neurodegenerative Diseases, Tübingen, Germany), Una-Marie Sheerin (Department of Molecular Neuroscience, UCL Institute of Neurology), Mohamad Saad (INSERM U563, CPTP, Toulouse, France; and Paul Sabatier University, Toulouse, France), Javier Simón-Sánchez (Department of Clinical Genetics, Section of Medical Genomics, VU University Medical Centre, Amsterdam, Netherlands), Claudia Schulte (Department for Neurodegenerative Diseases, Hertie Institute for Clinical Brain Research), Suzanne Lesage (INSERM, UMR_S975 (formerly UMR_S679), Paris, France; Université Pierre et Marie Curie-Paris, Centre de Recherche de l'Institut du Cerveau et de la Moelle épinière, Paris, France; and CNRS, Paris, France), Sigurlaug Sveinbjörnsdóttir (Department of Neurology, Landspítali University Hospital, Reykjavík, Iceland; Department of Neurology, MEHT Broomfield Hospital, Chelmsford, Essex, UK; and Queen Mary College, University of London, London, UK), Sampath Arepalli (Laboratory of Neurogenetics, National

Institute on Aging), Roger Barker (Department of Neurology, Addenbrooke's Hospital, University of Cambridge, Cambridge, UK), Yoav Ben-Shlomo (School of Social and Community Medicine, University of Bristol), Henk W Berendse (Department of Neurology and Alzheimer Center, VU University Medical Center), Daniela Berg (Department for Neurodegenerative Diseases, Hertie Institute for Clinical Brain Research and DZNE, German Center for Neurodegenerative diseases), Kailash Bhatia (Department of Motor Neuroscience, UCL Institute of Neurology), Rob M A de Bie (Department of Neurology, Academic Medical Center, University of Amsterdam, Amsterdam, Netherlands), Alessandro Biffi (Center for Human Genetic Research and Department of Neurology, Massachusetts General Hospital, Boston, MA, USA; and Program in Medical and Population Genetics, Broad Institute, Cambridge, MA, USA), Bas Bloem (Department of Neurology, Radboud University Nijmegen Medical Centre, Nijmegen, Netherlands), Zoltan Bochdanovits (Department of Clinical Genetics, Section of Medical Genomics, VU University Medical Centre), Michael Bonin (Department of Medical Genetics, Institute of Human Genetics, University of Tübingen, Tübingen, Germany), Jose M Bras (Department of Molecular Neuroscience, UCL Institute of Neurology), Kathrin Brockmann (Department for Neurodegenerative Diseases, Hertie Institute for Clinical Brain Research and DZNE, German Center for Neurodegenerative diseases), Janet Brooks (Laboratory of Neurogenetics, National Institute on Aging), David J Burn (Newcastle University Clinical Ageing Research Unit, Campus for Ageing and Vitality, Newcastle upon Tyne, UK), Gavin Charlesworth (Department of Molecular Neuroscience, UCL Institute of Neurology), Honglei Chen (Epidemiology Branch, National Institute of Environmental Health Sciences, National Institutes of Health, NC, USA), Patrick F Chinnery (Neurology M4104, The Medical School, Framlington Place, Newcastle upon Tyne, UK), Sean Chong (Laboratory of Neurogenetics, National Institute on Aging), Carl E Clarke (School of Clinical and Experimental Medicine, University of Birmingham, Birmingham, UK; and Department of Neurology, City Hospital, Sandwell and West Birmingham Hospitals NHS Trust, Birmingham, UK), Mark R Cookson (Laboratory of Neurogenetics, National Institute on Aging), J Mark Cooper (Department of Clinical Neurosciences, UCL Institute of Neurology), Jean Christophe Corvol (INSERM, UMR_S975; Université Pierre et Marie Curie-Paris; CNRS; and INSERM CIC-9503, Hôpital Pitié-Salpêtrière, Paris, France), Carl Counsell (University of Aberdeen, Division of Applied Health Sciences, Population Health Section, Aberdeen, UK), Philippe Damier (CHU Nantes, CIC0004, Service de Neurologie, Nantes, France), Jean-François Dartigues (INSERM U897, Université Victor Segalen, Bordeaux, France), Panos Deloukas (Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Cambridge, UK), Günther Deuschl (Klinik für Neurologie, Universitätsklinikum Schleswig-Holstein, Campus Kiel, Christian-Albrechts-Universität Kiel, Kiel, Germany), David T Dexter (Parkinson's Disease Research Group, Faculty of Medicine, Imperial College London, London, UK), Karin D van Dijk (Department of Neurology and Alzheimer Center, VU University Medical Center), Allissa Dillman (Laboratory of Neurogenetics, National Institute on Aging), Frank Durif (Service de Neurologie, Hôpital Gabriel Montpied, Clermont-Ferrand, France), Alexandra Dürr (INSERM, UMR_S975; Université Pierre et Marie Curie-Paris; CNRS; and AP-HP, Pitié-Salpêtrière Hospital), Sarah Edkins (Wellcome Trust Sanger Institute), Jonathan R Evans (Cambridge Centre for Brain Repair, Cambridge, UK), Thomas Foltynie (UCL Institute of Neurology), Jing Dong (Epidemiology Branch, National Institute of Environmental Health Sciences), Michelle Gardner (Department of Molecular Neuroscience, UCL Institute of Neurology), J Raphael Gibbs (Laboratory of Neurogenetics, National Institute on Aging; and Department of Molecular Neuroscience, UCL Institute of Neurology), Alison Goate (Department of Psychiatry, Department of Neurology, Washington University School of Medicine, MI, USA), Emma Gray (Wellcome Trust Sanger Institute), Rita Guerreiro (Department of Molecular Neuroscience, UCL Institute of Neurology), Clare Harris (University of Aberdeen), Jacobus J van Hilten (Department of Neurology, Leiden University Medical Center, Leiden, Netherlands), Albert Hofman (Department of Epidemiology, Erasmus University Medical Center, Rotterdam, Netherlands), Albert Hollenbeck (AARP, Washington DC, USA), Janice Holton (Queen Square Brain Bank for Neurological Disorders, UCL Institute of Neurology), Michele Hu (Department of Clinical Neurology, John Radcliffe Hospital, Oxford, UK), Xuemei Huang (Departments of Neurology, Radiology, Neurosurgery, Pharmacology, Kinesiology, and Bioengineering, Pennsylvania State University– Milton S Hershey Medical Center, Hershey, PA, USA), Isabel Wurster (Department for Neurodegenerative Diseases, Hertie Institute for Clinical Brain Research and German Center for Neurodegenerative diseases), Walter Mätzler (Department for Neurodegenerative Diseases, Hertie Institute for Clinical Brain Research and German Center for Neurodegenerative diseases), Gavin Hudson (Neurology M4104, The Medical School, Newcastle upon Tyne, UK), Sarah E Hunt (Wellcome Trust Sanger Institute), Johanna Huttenlocher (deCODE genetics), Thomas Illig (Institute of Epidemiology, Helmholtz Zentrum München, German Research Centre for Environmental Health, Neuherberg, Germany), Pálmi V Jónsson (Department of Geriatrics, Landspítali University Hospital, Reykjavík, Iceland), Jean-Charles Lambert (INSERM U744, Lille, France; and Institut Pasteur de Lille, Université de Lille Nord, Lille, France), Cordelia Langford (Cambridge Centre for Brain Repair), Andrew Lees (Queen Square Brain Bank for Neurological Disorders), Peter Lichtner (Institute of Human Genetics, Helmholtz Zentrum München, German Research Centre for Environmental Health, Neuherberg, Germany), Patricia Limousin (Institute of Neurology, Sobell Department, Unit of Functional Neurosurgery, London, UK), Grisel Lopez (Section on Molecular Neurogenetics, Medical Genetics Branch, NHGRI, National Institutes of Health), Delia Lorenz (Klinik für Neurologie, Universitätsklinikum Schleswig-Holstein), Alisdair McNeill (Department of Clinical Neurosciences, UCL Institute of Neurology), Catriona Moorby (School of Clinical and Experimental Medicine, University of Birmingham), Matthew Moore (Laboratory of Neurogenetics, National Institute on Aging), Huw R Morris (MRC Centre for Neuropsychiatric Genetics and Genomics, Cardiff University School of Medicine, Cardiff, UK), Karen E Morrison (School of Clinical and Experimental Medicine, University of Birmingham; and Neurosciences Department, Queen Elizabeth Hospital, University Hospitals Birmingham NHS Foundation Trust, Birmingham, UK), Ese Mudanohwo (Neurogenetics Unit, UCL Institute of Neurology and National Hospital for Neurology and Neurosurgery), Sean S O'Sullivan (Queen Square Brain Bank for Neurological Disorders), Justin Pearson (MRC Centre for Neuropsychiatric Genetics and Genomics), Joel S Perlmutter (Department of Neurology, Radiology, and Neurobiology at Washington University, St Louis), Hjörvar Pétursson (deCODE genetics; and Department of Medical Genetics, Institute of Human Genetics, University of Tübingen), Pierre Pollak (Service de Neurologie, CHU de Grenoble, Grenoble, France), Bart Post (Department of Neurology, Radboud University Nijmegen Medical Centre), Simon Potter (Wellcome Trust Sanger Institute), Bernard Ravina (Translational Neurology, Biogen Idec, MA, USA), Tamas Revesz (Queen Square Brain Bank for Neurological Disorders), Olaf Riess (Department of Medical Genetics, Institute of Human Genetics, University of Tübingen), Fernando Rivadeneira (Departments of Epidemiology and Internal Medicine, Erasmus University Medical Center), Patrizia Rizzu (Department of Clinical Genetics, Section of Medical Genomics, VU University Medical Centre), Mina Ryten (Department of Molecular Neuroscience, UCL Institute of Neurology), Stephen Sawcer (University of Cambridge, Department of Clinical Neurosciences, Addenbrooke's hospital, Cambridge, UK), Anthony Schapira (Department of Clinical Neurosciences, UCL Institute of Neurology), Hans Scheffer (Department of Human Genetics, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands), Karen Shaw (Queen Square Brain Bank for Neurological Disorders), Ira Shoulson

(Department of Neurology, University of Rochester, Rochester, NY, USA), Ellen Sidransky (Section on Molecular Neurogenetics, Medical Genetics Branch, NHGRI), Colin Smith (Department of Pathology, University of Edinburgh, Edinburgh, UK), Chris CA Spencer (Wellcome Trust Centre for Human Genetics, Oxford, UK), Hreinn Stefánsson (deCODE genetics), Francesco Bettella (deCODE genetics), Joanna D Stockton (School of Clinical and Experimental Medicine), Amy Strange (Wellcome Trust Centre for Human Genetics), Kevin Talbot (University of Oxford, Department of Clinical Neurology, John Radcliffe Hospital, Oxford, UK), Carlie M Tanner (Clinical Research Department, The Parkinson's Institute and Clinical Center, Sunnyvale, CA, USA), Avazeh Tashakkori-Ghanbaria (Wellcome Trust Sanger Institute), François Tison (Service de Neurologie, Hôpital Haut-Lévêque, Pessac, France), Daniah Trabzuni (Department of Molecular Neuroscience, UCL Institute of Neurology), Bryan J Traynor (Laboratory of Neurogenetics, National Institute on Aging), André G Uitterlinden (Departments of Epidemiology and Internal Medicine, Erasmus University Medical Center), Daan Velseboer (Department of Neurology, Academic Medical Center), Marie Vidailhet (INSERM, UMR_S975, Université Pierre et Marie Curie-Paris, CNRS, UMR 7225), Robert Walker (Department of Pathology, University of Edinburgh), Bart van de Warrenburg (Department of Neurology, Radboud University Nijmegen Medical Centre), Mirdhu Wickremaratchi (Department of Neurology, Cardiff University, Cardiff, UK), Nigel Williams (MRC Centre for Neuropsychiatric Genetics and Genomics), Caroline H Williams-Gray (Department of Neurology, Addenbrooke's Hospital), Sophie Winder-Rhodes (Department of Psychiatry and Medical Research Council and Wellcome Trust Behavioural and Clinical Neurosciences Institute, University of Cambridge), Kári Stefánsson (deCODE genetics), Maria Martinez (INSERM UMR 1043; and Paul Sabatier University), Nicholas W Wood (UCL Genetics Institute; and Department of Molecular Neuroscience, UCL Institute of Neurology), John Hardy (Department of Molecular Neuroscience, UCL Institute of Neurology), Peter Heutink (Department of Clinical Genetics, Section of Medical Genomics, VU University Medical Centre), Alexis Brice (INSERM, UMR_S975, Université Pierre et Marie Curie-Paris, CNRS, UMR 7225, AP-HP, Pitié-Salpêtrière Hospital), Thomas Gasser (Department for Neurodegenerative Diseases, Hertie Institute for Clinical Brain Research, and DZNE, German Center for Neurodegenerative Diseases), Andrew B Singleton (Laboratory of Neurogenetics, National Institute on Aging).

The members and affiliations of the North American Brain Expression Consortium (NABEC) are as follows: Andrew Singleton (Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, MD, USA); Mark Cookson (Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, MD, USA); J. Raphael Gibbs (Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, MD, USA and Reta Lila Weston Institute and Department of Molecular Neuroscience, UCL Institute of Neurology, Queen Square, London WC1N 3BG, UK); Dena Hernandez (Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, MD, USA and Reta Lila Weston Institute and Department of Molecular Neuroscience, UCL Institute of Neurology, Queen Square, London WC1N 3BG, UK); Allissa Dillman (Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, MD, USA and Department of Neuroscience, Karolinska Institutet, 171 77 Stockholm, Sweden); Michael Nalls (Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, MD, USA) Alan Zonderman (Research Resources Branch, National Institute on Aging, National Institutes of Health, Bethesda, MD, USA); Sampath Arepalli (Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, MD, USA); Luigi Ferrucci (Clinical Research Branch, National Institute on Aging, Baltimore, MD, USA); Robert Johnson (NICHD Brain and Tissue Bank for Developmental Disorders, University of Maryland Medical School, Baltimore, MD 21201, USA); Dan Longo (Lymphocyte Cell Biology Unit, Laboratory of Immunology, National Institute on Aging, National Institutes of Health, Baltimore, MD, USA); Richard O'Brien (Brain Resource Center, Johns Hopkins University, Baltimore, MD, USA); Bryan Traynor (Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, MD, USA); Juan Troncoso (Brain Resource Center, Johns Hopkins University, Baltimore, MD, USA); Marcel van der Brug (Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, MD, USA and ITGR Biomarker Discovery Group, Genentech, South San Francisco, CA, USA); Ronald Zielke (NICHD Brain and Tissue Bank for Developmental Disorders, University of Maryland Medical School, Baltimore, MD 21201, USA).

The members and affiliations of the United Kingdom Brain Expression Consortium are as follows (UKBEC): John Hardy (Reta Lila Weston Institute and Department of Molecular Neuroscience, UCL Institute of Neurology, Queen Square, London WC1N 3BG, UK); Michael Weale (Department of Medical and Molecular Genetics, King's College London, 8th Floor, Tower Wing, Guy's Hospital, London SE1 9RT, UK); Mina Ryten (Reta Lila Weston Institute and Department of Molecular Neuroscience, UCL Institute of Neurology, Queen Square, London WC1N 3BG, UK); Adaikalavan Ramasamy (Department of Medical and Molecular Genetics, King's College London, 8th Floor, Tower Wing, Guy's Hospital, London SE1 9RT, UK and Reta Lila Weston Institute and Department of Molecular Neuroscience, UCL Institute of Neurology, Queen Square, London WC1N 3BG, UK); Daniah Trabzuni (Reta Lila Weston Institute and Department of Molecular Neuroscience, UCL Institute of Neurology, Queen Square, London WC1N 3BG, UK and Department of Genetics, King Faisal Specialist Hospital and Research Centre, PO Box 3354, Riyadh 11211, Saudi Arabia); Colin Smith (Department of Neuropathology, MRC Sudden Death Brain Bank Project, University of Edinburgh, Wilkie Building, Teviot Place, Edinburgh EH8 9AG); Robert Walker (Department of Neuropathology, MRC Sudden Death Brain Bank Project, University of Edinburgh, Wilkie Building, Teviot Place, Edinburgh EH8 9AG).

Supplementary Information for:

Genome-wide meta-analysis identifies six novel loci associated with habitual coffee consumption

The Coffee and Caffeine Genetics Consortium

Marilyn C Cornelis, Ph.D., Enda M Byrne[1], Ph.D., Tõnu Esko[1], Ph.D., Michael A Nalls[1], Ph.D., Andrea Ganna, M.Sc., Nina Paynter, Ph.D., Keri L Monda, Ph.D., Najaf Amin, Ph.D., Krista Fischer, Ph.D., Frida Renstrom, Ph.D., Julius S Ngwa, M.S., Ville Huikari, M.Sc., Alana Cavadino, M.Sc., Ilja M Nolte, Ph.D., Alexander Teumer, Ph.D., Kai Yu, Ph.D., Pedro Marques-Vidal, M.D., Ph.D.,FESC, Rajesh Rawal, Dr.Sc, Ani Manichaikul, Ph.D., Mary K Wojczynski, Ph.D., M.P.H., Jacqueline M Vink, Ph.D., Jing Hua Zhao, M.D., Ph.D., George Burlutsky, M.Appl.Stat., Jari Lahti, Ph.D., Vera Mikkilä, Ph.D., Rozenn N Lemaitre, Ph.D., M.P.H., Joel Eriksson, M.Sc., Solomon K Musani, Ph.D., Toshiko Tanaka, Ph.D., Frank Geller, M.Sc.., Jian'an Luan, Ph.D., Jennie Hui, Ph.D., Reedik Mägi, Ph.D., Maria Dimitriou, M.Med.Sci, Melissa E Garcia, M.P.H., Weang-Kee Ho, Ph.D., Margaret J Wright, Ph.D., Lynda M Rose, M.S., Patrik KE Magnusson, Ph.D., Nancy L Pedersen, Ph.D., David Couper, Ph.D., Ben A Oostra, Ph.D., Albert Hofman, Ph.D., Mohammad Arfan Ikram, Ph.D., Henning W Tiemeier, Ph.D., Andre G Uitterlinden, Ph.D., Frank JA van Rooij, M.Sc., Inês Barroso, Ph.D., Ingegerd Johansson, D.D.S., Ph.D., Luting Xue, M.S., Marika Kaakinen, Ph.D., Lili Milani, Ph.D., Christine Power, Ph.D., Harold Snieder, Ph.D., Ronald P Stolk, M.D., Ph.D., Sebastian E Baumeister, Ph.D., Reiner Biffar, M.D., Fangyi Gu, Sc.D., François Bastardot, M.D.., Zoltán Kutalik, Ph.D., David R Jacobs Jr , Ph.D., Nita G Forouhi, Ph.D., M.R.C.P., F.F.P.H.M., Evelin Mihailov, M.Sc., Lars Lind, M.D., Ph.D., Cecilia Lindgren, Ph.D., Karl Michaëlsson, M.D., Ph.D., Andrew Morris, Ph.D., Majken Jensen, Ph.D., Kay-Tee Khaw, M.D., Ph.D., F.R.C.P., Robert N Luben, B.Sc., Jie Jin Wang, M.Med., M.Appl.Stat, Ph.D., Satu Männistö, Ph.D., Mia-Maria Perälä, M.Sc., Mika Kähönen, M.D., Ph.D., Terho Lehtimäki, M.D., Ph.D., Jorma Viikari, M.D., Ph.D., Dariush Mozaffarian, M.D., Dr.P.H., Kenneth Mukamal, M.D., M.P.H., Bruce M Psaty, M.D., Ph.D., Angela Döring, M.D., Andrew C Heath, Dr.Phil., Grant W Montgomery, Ph.D., Norbert Dahmen, M.D., Teresa Carithers, Ph.D., R.D., L.D., Katherine L Tucker, Ph.D., Luigi Ferrucci, M.D., Ph.D., Heather A Boyd, Ph.D., Mads Melbye, M.D., Dr.Med.Sci, Jorien L Treur, M.Sc., Dan Mellström, M.D., Ph.D., Jouke Jan Hottenga, Ph.D., Inga Prokopenko, Ph.D., Anke Tönjes, M.D., Panos Deloukas, Ph.D., Stavroula Kanoni, Ph.D., Mattias Lorentzon, M.D., Ph.D., Denise K Houston, Ph.D., Yongmei Liu, M.D., Ph.D., John Danesh, M.B.Ch.B, D.Sc., Asif Rasheed, M.B.B.S., Marc A Mason, M.S., Alan B Zonderman, Ph.D., Lude Franke, Ph.D., Bruce S Kristal, Ph.D., International Parkinson's Disease Genomics Consortium (IPDGC)†, North American Brain Expression Consortium (NABEC)†, UK Brain Expression Consortium (UKBEC)†, Juha Karjalainen, Ph.D., Danielle R Reed, Ph.D., Harm-Jan Westra, Ph.D., Michele K Evans, M.D., Danish Saleheen, M.B.B.S., Ph.D., Tamara B Harris, M.D., George Dedoussis, Ph.D., Gary Curhan, M.D., Sc.D., Michael Stumvoll, M.D., John Beilby, Ph.D., Louis R Pasquale, M.D., Bjarke Feenstra, Ph.D., Stefania Bandinelli, M.D., Jose M Ordovas, Ph.D., Andrew T Chan, M.D., Ulrike Peters, Ph.D., Claes Ohlsson, M.D., Ph.D., Christian Gieger,

Ph.D., Nicholas G Martin, Ph.D., Melanie Waldenberger, Ph.D., M.P.H., David S Siscovick, M.D., M.P.H., Olli Raitakari, M.D., Ph.D., Johan G Eriksson, M.D., Ph.D., Paul Mitchell, M.D., Ph.D., David J Hunter, Sc.D., Peter Kraft, Ph.D., Eric B Rimm, Sc.D., Dorret I Boomsma, Ph.D., Ingrid B Borecki, Ph.D., Ruth JF Loos, Ph.D., Nicholas J Wareham, M.D., Ph.D., F.R.C.P., Peter Vollenweider, M.D., Neil Caporaso, M.D., Hans Jörgen Grabe, M.D., Marian L Neuhouser, Ph.D., Bruce HR Wolffenbuttel, M.D., Ph.D., Frank B Hu, M.D., Ph.D., Elina Hyppönen, Ph.D., Marjo-Riitta Järvelin, M.D., Ph.D., L Adrienne Cupples, Ph.D., Paul W Franks, Ph.D., Paul M Ridker, Ph.D., Cornelia M van Duijn, Ph.D., Gerardo Heiss, M.D., Ph.D., Andres Metspalu, M.D., Ph.D., Kari E North, Ph.D., Erik Ingelsson, M.D., Ph.D., Jennifer A Nettleton, Ph.D., Rob M van Dam, Ph.D., and Daniel I Chasman, Ph.D.

[1]These authors contributed equally to this work.

Correspondance:

Marilyn C Cornelis, PhD
Department of Nutrition
Harvard School of Public Health
401 Park Drive, Boston, MA 02215 USA
Email: mcorneli@hsph.harvard.edu

Daniel I Chasman, PhD
Division of Preventive Medicine
Brigham and Women's Hospital, Harvard Medical School
900 Commonwealth Ave, Boston, MA 02446 USA
Email: dchasman@research.bwh.harvard.edu.

Table of Contents

Supplementary Methods

1. Description of Participating Studies

Following are brief descriptions of participating studies. Methods for addressing population structure are also provided when appropriate. Additional details regarding phenotype and genotype measures pertinent to the current study are provided in Supplementary Tables S1-S7. All genetic analysis utilized existing GW data. None of the studies was initially designed to identify genetic predictors of coffee intake. All studies were conducted under local institutional review board-approved protocols and written or verbal informed consent was obtained from each study participant.

1a. Stage 1 (discovery)

Atherosclerosis Risk in Communities (ARIC) Study: The ARIC study is a multi-center prospective investigation of atherosclerotic disease[1]. Men and women aged 45-64 years at baseline (1987-1989) were recruited from 4 communities: Forsyth County, North Carolina; Jackson, Mississippi; suburban areas of Minneapolis, Minnesota; and Washington County, Maryland. A total of 15,792 individuals participated in the baseline examination in 1987-1989, with three follow-up examinations: 1990-1992, 1993-1995, and 1996-1998. EIGENSTRAT[2] was used to carryout ancestry analyses and any study samples with substantial similarity to non-European reference samples (either the HapMap YRI or CHB+JPT samples) were excluded for subsequent analyses. EIGENSTRAT was also used to generate informative principal components for population sub-structure.

The Busselton Health Study (Busselton): Residents of the town of Busselton in the southwest of Western Australia have been involved in a series of health surveys since 1966[3]. In 1994/1995 there was a follow-up study involving a subset of those who had attended any of the previous surveys. The participants included in the current GWAS of coffee consumption were initially selected for a nested-case control GWAS of asthma and have been described in detail previously[4]. EIGENSTRAT[2] was used to carryout ancestry analyses and any study samples with substantial similarity to non-European reference samples (either the HapMap YRI or CHB+JPT samples) were excluded for subsequent stage 1 analyses. EIGENSTRAT was also used to generate informative principal components for population sub-structure.

Cardiovascular Health Study (CHS): CHS is an ongoing cohort study from four centers across the U.S. Adults aged 65 years and older were recruited from Medicare eligibility lists in Sacramento, CA; Pittsburgh, PA; Hagerstown, MD; and Forsyth County, NC. An initial cohort of 5,201 older adults enrolled in 1989-1990 and was supplemented with an additional 687 African-Americans in 1992-1993. Participants underwent yearly examinations at the field centers through 1998-1999. Only men and women of European ancestry with quality GW scans were considered for stage 1 of the current study. EIGENSTRAT [2] was used to carryout ancestry analyses and any study samples with substantial similarity to non-European reference samples (either the HapMap YRI or CHB+JPT samples) were excluded for subsequent stage 1 analyses. EIGENSTRAT was also used to generate informative principal components for population sub-structure.

Cohorte Lausannoise (COLAUS): The CoLaus Study is a cross-sectional study aimed at assessing the prevalence of cardiovascular disease (CVD) risk factors as well as the molecular determinants of CVD in the Caucasian population aged between 35 and 75 years in Lausanne, Switzerland [5](www.colaus.ch). Recruitment began in June 2003 and ended in May 2006, and included 6733 participants. The current study includes a subset of individuals participating in the first follow-up of the cohort which took place between 2009 and 2012. EIGENSTRAT [2] was used to carryout ancestry analyses and any study samples with substantial similarity to non-European reference samples (either the HapMap YRI or CHB+JPT samples) were excluded for subsequent analyses. EIGENSTRAT was also used to generate informative principal components for population sub-structure.

Danish National Birth Cohort: Preterm Birth Study (DNBC): DNBC is a population-based cohort of 101,042 pregnancies, recruited in the years 1996–2002 [6]. All participating women underwent thorough phenotype characterization based on information from four computer-assisted telephone interviews conducted during pregnancy (two interviews) and after delivery (two interviews). Genome-wide data was available for 3,840 mothers and children (Illumina Human660w-Quadv1_A array) through the preterm birth study conducted within the Gene Environment Association Studies (GENEVA) Consortium. To ensure a high degree of genetic homogeneity in the genotyped sample, we obtained birthplace information from the Danish Civil Registry, and only included individuals who themselves as well as their parents were born in Scandinavia. Outliers in the multidimensional scaling analysis (as implemented in PLINK[7]) were

excluded, the remaining samples clustered tightly together near the Caucasian HapMap populations from Utah and Tuscany as expected, and our subsequent GWAS repeatedly showed no evidence of population stratification.

Estonian Genome Center of the University of Tartu (EGCUT1): EGCUT is a population-based biobank of the Estonian Genome Project of University of Tartu (www.biobank.ee)[8]. The cohort currently includes over 51,515 men and women 18 years of age or older, which reflects closely the age distribution in the adult Estonian population. The samples included in this study form a random subset of the cohort, with the exception of 500 female individuals aged 83+ which were specifically selected according to age and gender. Subjects are recruited by the general practitioners and physicians in the hospitals. Each participant completed a Computer Assisted Personal interview, including personal data (place of birth, place(s) of living, nationality etc.), genealogical data (family history, three generations), educational and occupational history and lifestyle data (physical activity, dietary habits, smoking, alcohol consumption, women's health, quality of life). EIGENSTRAT [2] was used to carryout ancestry analyses and any study samples with substantial similarity to non-European reference samples (either the HapMap YRI or CHB+JPT samples) were excluded for subsequent analyses. EIGENSTRAT was also used to generate informative principal components for population sub-structure.

European Prospective Investigation in Cancer and Nutrition Norfolk Study (EPIC-Norfolk): The EPIC-Norfolk study is a population-based cohort study of 25,639 men and women aged 40-79 years, residing in Norfolk, U.K and recruited from general practice registers between 1993 and 1997. The study is part of a 10-country European collaboration that was originally initiated to examine the prospective association between diet and cancer, and the scope of the study was subsequently broadened to include other end points such as diabetes and coronary heart disease. Participants attended a study clinic visit after an overnight fast where they underwent clinical examination, provided blood samples, and completed health and lifestyle questionnaires, including a food frequency questionnaire for dietary assessment. Probable ethnic outliers were identified using the method described by McGinnis *et al* [9]($p<0.001$) or PLINK[7] (z-score<2.5) and excluded from analysis. The individuals after QC and used in GWAS have shown no evidence of population stratification as through EIGENSTRAT[2] and analysis on a variety of other traits, therefore no adjustment for population stratification was carried out in this study.

Erasmus Rucphen Family study (ERF): The ERF study is a family based study embedded in the Genetic Research in Isolated Populations (GRIP) program in the South West of the Netherlands. ERF includes over 3000 participants descending from 22 couples living in the Rucphen region between 1850-1900. All living descendants of these couples and their spouses were invited to take part in the study, which began in 2002. Participants with both phenotype and GW genotype data were available for the current study and have been described in detail previously [10]. Probable ethnic outliers were identified by principal component analysis with HapMap II population samples and excluded from analysis. Subsequent GWAS repeatedly showed no evidence of population stratification, therefore no adjustment for population stratification was carried out in this study.

Family Heart Study (FamHS): The Family Heart Study (https://dsgweb.wustl.edu/PROJECTS/MP1.html) began in 1992 with the ascertainment of 1,200 families, half randomly sampled, and half selected because of an excess of coronary heart disease (CHD) or risk factor abnormalities as compared with age- and sex-specific population rates [11]. The families, with approximately 6,000 individuals, were sampled on the basis of information on probands from four population-based parent studies: the Framingham Heart Study (field center at Boston University), the Utah Family Tree Study (at University of Utah), and two ARIC study centers (Minneapolis, University of Minnesota, and Forsyth County, NC, University of North Carolina). In general, the proband, and the proband's spouse, children, brothers, sisters, and parents were recruited thus producing three-generation pedigrees. Information regarding medical and lifestyle history was collected between 1994 and 1996 when subjects attended a clinic visit. Study participants belonging to the largest pedigrees were invited for a second clinical exam between 2002 and 2004. Approximately 82% of the recruited subjects returned for the second visit and 275 newly eligible family members were also recruited. In addition, a sample of African-American families was recruited at an additional ARIC field center at the University of Alabama in Birmingham. EIGENSTRAT [2] was used to carryout ancestry analyses and any study samples with substantial similarity to non-European reference samples (either the HapMap YRI or CHB+JPT samples) were excluded for subsequent stage 1 analyses. The individuals who were used for GWAS repeatedly showed no evidence of population stratification. Consequently, we have not adjusted for population stratification.

Fenland Study (Fenland): The Fenland study is a population-based study that was initiated in 2005 and is currently ongoing, having recruited 10,411 men and women aged 30 to 62 years as of March 2013. The aim of the Fenland study is to examine the interactions between lifestyle and genetic factors on the risk of obesity, glycemia and related metabolic traits. The sampling frame included all individuals born within the birth cohort 1950 – 1975 and resident in Cambridgeshire, U.K. at the time of study recruitment, and was constructed from the lists of local general practices in Ely, Wisbech, Cambridge and their surrounding villages. The first 1,500 volunteers (European descent) with complete anthropometric data were genotyped and included in the current analyses. The individuals after QC and used in GWAS have shown no evidence of population stratification as through EIGENSTRAT[2] and analysis on a variety of other traits, therefore no adjustment for population stratification was carried out in this study.

Framingham Heart Study (FHS): The Framingham Offspring Study is a community-based longitudinal study designed to examine CVD risk in the offspring of the original participants and their spouses of the Framingham Heart Study cohort. In 1971, 5,124 individuals were enrolled in the study; since then, the cohort has been examined every 3–4 years. Between 1991 and 1995, during the 5th examination cycle, 3,799 adults underwent a standardized medical history and physical examination. Beginning in 2002, 4,095 Gen 3 participants, who had at least one parent in the offspring cohort, were enrolled in the Framingham Heart Study. At the first cycle of the Gen 3 study, 4,095 individuals with a mean age of 40 y, underwent the standard clinic examination. For the present study both cohorts were combined for the analysis. A total of 5,835 adults with available DNA, valid dietary information, and consent to share genetic data were eligible for the current study. EIGENSTRAT[2] was used to obtain principal components for possible population substructure. Principal components associated with the outcome were included as covariates in statistical models. We used linear mixed effects models with the family kinship matrix (lmekin in R 2.6.1) for phenotype 1 and logistic regression for phenotype 2 with generalized estimating equations clustering on pedigrees and robust standard errors.

Gothenburg Osteoporosis and Obesity Determinants Study (GOOD): The GOOD study was initiated to determine both environmental and genetic factors involved in the regulation of bone and fat mass. Young, Caucasian men were randomly identified in the greater Gothenburg area in Sweden using national population registers, contacted by telephone, and invited to participate. Enrolled subjects were between 18 and 20 years of age. There were no other exclusion criteria,

and 49% of the study candidates agreed to participate (n =1,068). Principal components analysis was performed using IBS/IBD distance analysis in PLINK[7]. The individuals who were used for GWAS repeatedly showed no evidence of population stratification. Consequently, we have not adjusted for population stratification.

Helsinki Birth Cohort Study (HBCS): The HBCS is composed of 8,760 individuals born between the years 1934 and 1944 in one of the two main maternity hospitals in Helsinki, Finland. Between 2001 and 2003, a randomly selected sample of 928 males and 1075 females participated in a clinical follow-up study with a focus on cardiovascular, metabolic and reproductive health, cognitive function and depressive symptoms. Individuals of this latter subset of the cohort were included in the current GWAS of coffee consumption. Principal components analysis was performed using IBS/IBD distance analysis in PLINK[7] (along with HapMap II population samples) and any probable ethnic outliers were excluded from analysis. Subsequent GWAS repeatedly showed no evidence of population stratification, therefore no adjustment for population stratification was carried out in this study.

Health, Aging and Body Composition Study (HealthABC): HealthABC is a prospective cohort study investigating the associations between body composition, weight-related health conditions, and incident functional limitation in older adults. HealthABC enrolled well-functioning, community-dwelling black (n=1281) and white (n=1794) men and women aged 70-79 years between April 1997 and June 1998. Participants were recruited from a random sample of white and all black Medicare eligible residents in the Pittsburgh, PA, and Memphis, TN, metropolitan areas. Participants have undergone annual exams and semi-annual phone interviews. Only men and women of European ancestry with quality GW scans were considered for stage 1 of the current study. EIGENSTRAT [2] was used to carryout ancestry analyses and any study samples with substantial similarity to non-European HapMap 3 reference samples were excluded. EIGENSTRAT was also used to generate informative principal components for population sub-structure.

Health Professionals Follow-up Study (HPFS T2D, HPFS CHD, HPFS GA, HPFS KS): The HPFS was initiated in 1986 when 51,529 male health professionals between 40 and 75 years of age years and residing in the U.S. completed a food frequency questionnaire (FFQ) and a questionnaire on lifestyle and medical history. The participants have been followed with repeated

questionnaires on lifestyle and health every 2 years and FFQs every 4 years. Genome-wide scans that contribute to this meta-analysis were obtained from 4 independent GWAS of the HPFS cohort, initially designed for outcomes of type 2 diabetes (T2D, dbGaP:phs000091.v2.p1), CHD, kidney stone (KS, dbGaP:phs000460.v1.p1) disease and open-angle glaucoma (GA, dbGaP:phs000308.v1.p1). Both cases and controls were included for analysis. Each subset was processed and analyzed separately. EIGENSTRAT [2] was used to carryout ancestry analyses and any study samples with substantial similarity to non-European reference samples (either the HapMap YRI or CHB+JPT samples) were excluded for subsequent stage 1 analyses. EIGENSTRAT was also used to generate informative principal components for population sub-structure.

Invecchiare in Chianti (inCHIANTI): The InCHIANTI study is a population-based epidemiological study aimed at evaluating the factors that influence mobility in the older population living in the Chianti region in Tuscany, Italy. The details of the study have been previously reported [12]. Briefly, 1616 residents were selected from the population registry of Greve in Chianti (a rural area: 11,709 residents with 19.3% of the population greater than 65 years of age), and Bagno a Ripoli (Antella village near Florence; 4,704 inhabitants, with 20.3% greater than 65 years of age). The participation rate was 90% (n=1453), and the subjects ranged between 21-102 years of age. A 236 item, interviewer administered FFQ was used to investigate how frequently (weekly, monthly, yearly) a specific food is generally consumed. Nutrient data for specific foods were obtained from the Food Composition Database for Epidemiological Studies in Italy [13]. No ethnic outliers were identified based on principal component analysis using EIGENSTRAT[2]. Further adjustment for population stratification was achieved through genomic control.

Cooperative Health Research in the Augsburg Region (KORA_F3, KORA_S4): The focus of KORA ("Kooperative Gesundheitsforschung in der Region Augsburg") is to survey the development and course of chronic diseases such as CHD and T2D. All survey participants are residents of German nationality identified through the registration office. KORA S4 is the KORA-Survey which was completed in 1999-2000 (S4, 1999/2000) and included a total of 1,814 subjects. KORA F3 is a follow-up study to the KORA-Survey 1994-1995 (S3, n=4856)**.** A total of 3,006 subjects participated in a 10-year follow-up examination of S3 and for genetic analysis we randomly selected 1,644 subjects of these participants. The individuals who were used for

GWAS repeatedly showed no evidence of population stratification. Consequently, we have not adjusted for population stratification.

The Multi-Ethnic Study of Atherosclerosis (MESA): MESA is a multi-center longitudinal study of subclinical atherosclerosis and risk factors that predict progression to clinically overt CVD or progression of the subclinical disease [14]. In 2000-2002 (baseline clinical exam) a total of 6,814 participants, all free of CVD, were recruited from six field centers across the United States. Approximately 38% of the recruited participants were Caucasian, 28% African-American, 22% Hispanic, and 12% Asian, predominantly of Chinese descent. Only men and women of European ancestry with quality GW scans were considered for stage 1 of the current study: EIGENSTRAT [2] was used to carryout ancestry analyses and any study samples with substantial similarity to non-European reference samples (either the HapMap YRI or CHB+JPT samples) were excluded for subsequent stage 1 analyses. EIGENSTRAT was also used to generate informative principal components for population sub-structure.

Northern Finland Birth Cohort 1966 (NFBC 1966): NFBC 1966 was initiated in 1965 by enrolling mothers living in the two Northernmost provinces of Finland (Oulu and Lapland) and with expected dates of delivery in 1966 [15]. A total of 12231 children were born into the cohort, 12058 of them live-born. Baseline data have been supplemented by data collected with postal questionnaires at the ages of 1, 14 and 31 years and various hospital records and national register data. Coffee consumption and smoking habits were enquired via a postal questionnaire as part of the 31-year follow-up study. At the same time, those still living in the original target area (Northern Finland) or in the capital (Helsinki) area were invited to a clinical examination, in which 71% (n=6033) participated. Blood samples were drawn and DNA was extracted successfully for 5753 of these subjects. Following GW genotyping, self-reported European ancestry was verified on the basis of multidimensional scaling analysis of identity by state using PLINK[7]. The eigen function in R[16] was used to generate informative principal components for population sub-structure. Those subjects with both phenotype and GW genotype data were included for the current study.

Nurses' Health Study (NHS T2D, NHS CHD, NHS GA, NHS KS, NHS BrCa): The NHS was established in 1976 when 121,700 female registered nurses aged 30-55 years and residing in 11 large U.S. states completed a mailed questionnaire on medical history and lifestyle

characteristics [17]. Every two years, follow-up questionnaires have been sent to update information on exposures and newly diagnosed diseases and every 2 to 4 years diet was assessed using a validated semi-quantitative FFQ [18]. Genome-wide scans that contribute to this meta-analysis were obtained from 5 independent GWAS of the NHS cohort, initially designed for outcomes of T2D (dbGaP:phs000091.v2.p1), CHD, breast cancer (BrCa, dbGaP:phs000147.v1.p1), KS disease (dbGaP:phs000460.v1.p1) and open-angle glaucoma (GA, dbGaP:phs000308.v1.p1). Both cases and controls were included for analysis. Each subset was processed and analyzed separately. EIGENSTRAT [2] was used to carryout ancestry analyses and any study samples with substantial similarity to non-European reference samples (either the HapMap YRI or CHB+JPT samples) were excluded for subsequent stage 1 analyses. EIGENSTRAT was also used to generate informative principal components for population sub-structure.

Netherlands Twin Register (NTR s1): The NTR study is an ongoing twin-family study on health-related behaviour that assesses families with adolescent and (young) adult twins since 1991 [19, 20]. Participants are invited every two or three years to complete a survey that contains questions about health, lifestyle, personality and psychopathology. Information on coffee consumption was obtained from the fifth survey which occurred in 2000. Because it is a family-based sample, we randomly selected 1 person per family to create a sample of unrelated individuals. Blood for DNA extraction and genotyping was collected via our biobank project, which is described elsewhere [21]. A detailed population structure analysis of NTR has been published previously[22]. Individuals with non-European ancestry were identified by projecting PCs from the 1000 Genomes samples on the study samples. EIGENSTRAT[2] was used to generate informative principal components for population sub-structure.

Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial (PLCO): PLCO is a large, randomized controlled trial investigating the efficacy of cancer screening to prevent early death from prostate, lung, colorectal and ovarian cancer[46]. Between 1992 and 2001, approximately 155,000 men and women in ten U.S. cities were enrolled in PLCO and randomized to either a screening or control arm. Eligibility criteria included an age at enrollment between 55 and 74 years and no history of prostate, lung, colon and ovarian cancer, although prior diagnoses of other cancers were acceptable. Individuals were followed up for all cancer diagnoses by annual mailed questionnaire and, additionally for trial disease outcomes, by screening examinations

during the first six years of follow-up. Blood specimens were collected annually from screening-arm participants as part of the screening examinations. Buccal cell specimens were collected from control-arm participants. In total, approximately 112,500 participants provided blood or buccal cell specimens. Caucasian subjects included for the current study were selected from both arms of the trial based on availability of a valid baseline questionnaire, FFQ, consent, and GWAS data. STRUCTURE[23] was used to carryout ancestry analyses and any subjects with <80% European ancestry (relative to HapMap II populations) were excluded for subsequent analyses.

Rotterdam Study (RS-I, RS-II): RS-I is a prospective population-based cohort study of 7,983 residents aged 55 years and older living in Ommoord, a suburb of Rotterdam, the Netherlands. RS-II is a second prospective population-based cohort study of respectively 3,011 residents also aged 55 years and older. A detailed description of RS-I and RS-II has been published previously [24]. Probable ethnic outliers were identified by principal component analysis with HapMap II population samples and excluded from analysis. Subsequent GWAS repeatedly showed no evidence of population stratification, therefore no adjustment for population stratification was carried out in this study.

Study of Health in Pomerania (SHIP): SHIP is a population-based cross-sectional study, sampled from the northeastern German region West Pomerania. Initially, 6,267 subjects were selected from population registries and 4,308 participated (response of 68.8%) in the baseline examination between 1997 and 2001. Since the sample consists of a homogenous European population (validated by MDS including the HapMap III samples), no adjustments for population stratification was performed.

Sorbs (SORBS): All subjects are part of a sample from an extensively phenotyped self-contained population from Germany, the Sorbs [25]. The Sorbs, who reside in the Lusatia region of eastern Germany, are a population isolate defined by their use of a west Slavic language (Sorbian) in an area with a majority of Germanic speakers. At present, about 1000 individuals speaking Upper Sorbian are enrolled in the study. Extensive phenotyping included amongst others a standardised interview for past medical history and family history, eating behaviour, consumption of natural stimulants, physical activities as well as collection of anthropometric data and a glucose-tolerance-test. EIGENSTRAT's *smartpca*[2] was used to carryout ancestry

analyses for outlier detection. The sample consists of a homogenous European population and in GWAS repeatedly showed no evidence of population stratification. Consequently, no adjustment for population stratification was applied for the current analysis.

The Hellenic study of Interactions between Snps and Eating in Atherosclerosis Susceptibility (THISEAS): THISEAS is a case- control study designed to investigate the association between genetic and lifestyle environmental factors and the risk of coronary artery disease (CAD) in adults [26]. Case (presenting with either acute coronary syndrome or stable CAD) and control participants (n= 1838) were recruited from 8 hospitals and from Open Care Centers for the elderly in the region of Athens in Greece. Hematological, biochemical and anthropometric measurements were obtained from all participants. Diet and physical activity data were collected through face-to-face interview by well trained scientists. The individuals who were used for GWAS repeatedly showed no evidence of population stratification. Consequently, no adjustment for population stratification was applied for the current analysis.

TwinGene: The TwinGene project, conducted between 2004 and 2008, is a population-based Swedish study of twins born between 1911 and 1958. The study participants have previously participated in a telephone interview called *Screening Across the Lifespan Twin Study*, conducted between 1998 and 2002. To be included in TwinGene, both twins within a pair had to be alive. The zygosity of the twins was based on self-reported childhood resemblance, or by using DNA markers (for 18% of the total sample). In total, 12591 individuals participated by donating blood to the study, and by answering questionnaires about life style and health. Blood from 9896 subjects (all available dizygous twins + one twin from each available monozygous twin pair) was sent to SNP&SEQ Technology Platform Uppsala, Sweden for GW genotyping. Both dizygous twins are included in the current analysis. A single record for each monozygous twin is included and his/her phenotype is the phenotype average among the twin pair. EIGENSTRAT's *smartpca*[2] was used to carryout ancestry analyses and any study samples with substantial similarity to non-European reference samples (either the HapMap YRI or CHB+JPT samples) were excluded for subsequent stage 1 analyses. Informative principal components were also generated for population sub-structure.

Women's Genome Health Study (WGHS): WGHS is a prospective cohort of female healthcare professionals, aged 45 or older at baseline, who provided baseline blood sample and consent for

blood based analysis in the Women's Health Study (WHS), a randomized, placebo controlled trial of aspirin and vitamin E in the primary prevention of CVD and cancer [27]. Genotyping in the WGHS sample was performed using the HumanHap300 Duo ''+ '' chips or the combination of the HumanHap300 Duo and iSelect chips with the Infinium II protocol. In either case, the custom SNP content was the same; these custom SNPs were chosen without regard to minor allele frequency (MAF) to saturate candidate genes for cardiovascular disease as well as to increase coverage of SNPs with known or suspected biological function, e.g. disease association, non-synonymous changes, substitutions at splice sites, etc. A subset of individuals were identified with self-reported European ancestry that could be verified on the basis of multidimensional scaling analysis of identity by state using 1443 ancestry informative markers in PLINK[7].

1b. Stage 2 (replication)

Blue Mountain Eye Study (BMES): The BMES is a population-based cohort study of eye diseases and other health outcomes in an urban population aged 49 years or older. Between 1992 and 1994, 3654 residents (82.4% of those eligible) aged 49+ years, living in two postcode areas near Sydney, Australia, participated; 2335 (75.1% of survivors) were re-examined after 5 years between 1997 and 1999, and 1952 (76% of survivors) were re-examined after 10 years between 2002 and 2004. The BMES extension study was conducted shortly after the BMES 5-year follow-up examinations, when additional 1174 (85.2% of the newly eligible) participants were examined once during 1999-2000. BMES participants who had DNA available were genotyped at the Wellcome Trust Centre for Human Genetics, Sanger Institute, Cambridge as part of the Wellcome Trust Case Control Consortium 2. EIGENSTRAT[2] was used to carryout ancestry analyses. No significant population stratification was evident in this predominantly European-ancestry sample, and therefore no adjustment for Eigenstrat principal components was made in the analyses.

British Birth Cohort 1958 (1958BC T1DGC, 1958BC WTCCC2): The 1958BC is a population based cohort study initially including all children born in England, Scotland or Wales during one week in March 1958. Biomedical assessment including DNA collection was done at age 45 years. From a target sample of 11,971 cohort members still living in Britain, 9,377 participants provided data, including self-reported information on coffee consumption as part of

a lifestyle questionnaire. Genome-wide data for the 1958BC was obtained through two sub-studies, both using the 1958BC members as a control population. First, 3,000 DNA samples were randomly selected as part of the Wellcome Trust Case Control Consortium (WTCCC)[28] and genotyped on the Affymetrix SNP 6.0 platform. Secondly, 2,592 DNA samples from the 1958BC were used as controls for a type 1 diabetes case-control study (T1DGC)[29], with samples genotyped through the JDRF/WT Diabetes and Inflammation Laboratory using the Illumina Infinium 550K chip. Participants of white European ancestry were included in the sub-studies, and principal components were generated using multidimensional scaling in PLINK in order to control for population sub-structure.

Estonian Genome Center of the University of Tartu (EGCUT2, EGCUT3): Additional genetic data from EGCUT became available after stage 1. Samples contributing to EGCUT2 and EGCUT3 were selected from the same source population as EGCUT1 (described above). No samples overlapped across the three study sets. EGCUT2 and EGCUT3 were genotyped on different platforms and therefore processed and analyzed separately. All participants were of European ancestry.

Gene-Lifestyle interactions And Complex traits Involved in Elevated disease Risk Study (GLACIER): GLACIER is a prospective, population-based cohort study comprised of 19,547 adults from the Northern Swedish county of Västerbotten, nested within the Northern Sweden Health and Disease Study. All GLACIER participants underwent detailed health and lifestyle examinations as part of the Västerbotten Health Survey, an ongoing population-based prospective cohort study focused on type 2 diabetes and cardiovascular disease. Since 1985, all residents of the county of Västerbotten have been invited to visit their primary care centre for a clinical examination within the year of their 40th, 50th, and/or 60th birthday. The protocol is standardized across study centers and conducted by trained nurses**.** Within GLACIER, 6,064 participants have been genotyped with the MetaboChip array, of which 5,742 participants were eligible for the current analyses.

Health Professionals Follow-up Study (HPFS CC): A fifth GWAS nested in the HPFS (described above) became available after stage 1 of the current study and was therefore included in stage 2. Initially designed for colon cancer (CC), HPFS CC (dbGaP: in progress) was processed and analyzed independently of other HPFS GWAS subsets but following similar

protocols. EIGENSTRAT [2] was used to carryout ancestry analyses and any study samples with substantial similarity to non-European reference samples (either the HapMap YRI or CHB+JPT samples) were excluded for subsequent stage 2 analyses. EIGENSTRAT was also used to generate informative principal components for population sub-structure. No HPFS CC samples included in stage 2 overlapped with other HPFS GWAS subsets.

Life Lines (LifeLines): The LifeLines Cohort Study (www.lifelines.net) is a multi-disciplinary prospective population-based cohort study examining in a unique three-generation design the health and health-related behaviors of 165,000 persons living in the North East region of The Netherlands [30]. It employs a broad range of investigative procedures in assessing the biomedical, socio-demographic, behavioral, physical and psychological factors which contribute to the health and disease of the general population, with a special focus on multimorbidity. Participants were recruited through their general practitioner. Recruitment has been going on since the end of 2006, and as of February 2013 over 120,000 participants have been included and a subset of these has GW data. EIGENSTRAT [2] was used to carryout ancestry analyses and any study samples with substantial similarity to non-European reference samples (either the HapMap YRI or CHB+JPT samples) were excluded for subsequent stage 2 analyses. EIGENSTRAT was also used to generate informative principal components for population sub-structure.

Osteoporotic Fractures in Men (MrOS) Sweden: The Osteoporotic Fractures in Men (MrOS) study is a prospective multicenter study including older men in Sweden, Hong Kong and the United States. The MrOS Sweden cohort consist of three subcohorts from three different Swedish cities (n=1,005 in Malmö, n=1,010 in Göteborg, and n=999 in Uppsala). In the present study, only participants from Göteborg, Sweden, were included. Study subjects were randomly identified using national population registers, contacted and asked to participate. To be eligible for the study, the subjects had to be able to walk without assistance, provide self-reported data, and sign an informed consent; there were no other exclusion criteria. EIGENSTRAT [2] was used to verify European ancestry and to generate informative principal components for population sub-structure.

Netherlands Twin Register (NTR s2): Within the NTR study (described above), extra genotype and phenotype data became available after stage 1. In November 2012, participants with genotype data but without data on coffee consumption were approached for a brief online survey

asking participants a variety of questions such as age, coffee consumption and smoking behavior. Within a week, 47% of the participants who were approached completed the survey. For the stage 2 analyses, we selected participants with genotype and phenotype data who did not take part in stage 1, and who were not biologically related to participants from the stage 1 sample. We randomly selected 1 person per family to create a sample with unrelated individuals. Data on coffee consumption were available from the fifth NTR survey collected in 2000 (n=494) and from the additional online data collection in November 2012 (n=427). A detailed population structure analysis of NTR has been published previously[22].

Nurses' Health Study (NHS CC): A sixth GWAS nested in the NHS (described above) became available after the discovery stage of the current GWAS and was therefore included in stage 2 . Initially designed for colon cancer (CC), NHS CC (dbGaP: in progress) was processed and analyzed independently of other NHS GWAS subsets but following similar protocols. EIGENSTRAT [2] was used to carryout ancestry analyses and any study samples with substantial similarity to non-European reference samples (either the HapMap YRI or CHB+JPT samples) were excluded for subsequent stage 2 analyses. EIGENSTRAT was also used to generate informative principal components for population sub-structure. No NHS CC samples included in stage 2 overlapped with other NHS GWAS subsets.

Prospective Investigation of the Vasculature in Uppsala Seniors (PIVUS): All 70-year-old individuals living in Uppsala, Sweden, in 2001–2004 were eligible for PIVUS, and 2025 randomly selected individuals were invited within 2 months of their 70th birthday from April 2001 to June 2004. Of these, 1016 (50%) participated in the study. At the examination the participants underwent a blood pressure measurement and anthropometry, blood sampling after an overnight fast, routine medical history, questionnaire and a 7-day diet registration. Caucasian ethnicity was confirmed using multidimensional scaling analysis in PLINK[7].

Queensland Institute of Medical Research (QIMR): Samples for QIMR were twins selected from a population-based twin registry-the Australian Twin Registry (est. 1981). The registry enrolls Australian twins of any type, age, or state of health. A Health and Lifestyle Questionnaire was mailed to registered twins between 1980 and 1982. Participants were genotyped in one of 4 separate genotyping events on either the Illumina 317K, 370K, or 610K platform. Several samples were included in multiple projects to allow for quality control across genotype events.

STRUCTURE[23] was used with the HapMap populations as a reference, to identify those of non-European ancestry. Principal components were included as covariates in the analysis.

Uppsala Longitudinal Study of Adult Men (ULSAM): Subjects born between 1920 and 1924 in Uppsala, Sweden were invited to participate at age 50 in this longitudinal cohort study, which began in 1970. Subjects were followed up at the ages of 60, 70, 77, 82 and 88 years. Information collected includes a medical and lifestyle questionnaire, blood pressure and anthropometric measurements, glucose tolerance test and 24-hour ambulatory blood pressure. Blood samples for DNA extraction and determination of established cardiovascular risk factors were available from the investigation at 70 years of age (1991-1994). EIGENSTRAT [2] was used to carryout ancestry analyses and any study samples with substantial similarity to non-European reference samples (either the HapMap YRI or CHB+JPT samples) were excluded for subsequent stage 2 analyses.

Cardiovascular Risk in Young Finns (YFS): In 1980, over 3,500 children and adolescents all around Finland participated in the baseline study. Thereafter these subjects have been followed with several examinations including comprehensive risk factor assessments. The follow-up studies have been conducted mainly with 3-year intervals. Caucasian ethnicity was confirmed using multidimensional scaling analysis in PLINK[7].

Atherosclerosis Risk in Communities-African Americans (AA_ARIC): African Americans from the ARIC cohort (described above) with quality GW scans were considered for stage 2 of the current study. To control for sub-structure, principal components were generated using the EIGENSTRAT software [2].

Cardiovascular Health Study-African Americans (AA_CHS): In 1992-1993 the CHS (described above) was supplemented with an additional 687 African Americans. Those with genetic and phenotype data were included in stage 2 of the current study of coffee consumption. We computed PCs based on 97,404 genotyped SNPs from the Illumina Omni 1M chip. We included the first ten PCs to control for population substructure.

Healthy Aging in Neighborhoods of Diversity across the Life Span-African Americans (AA_HANDLS): HANDLS is an interdisciplinary, community-based, prospective longitudinal epidemiologic study examining the influences of race and socioeconomic status on the development of age-related health disparities among socioeconomically diverse African Americans and whites (total n=3,722) in Baltimore MD. Only African Americans with genetic

and phenotype data were included in stage 2 of the current study of coffee consumption. EIGENSTRAT [2] was used to carryout ancestry analyses and only study samples with substantial similarity to the YRI and ASW HapMap 3 reference samples were included. EIGENSTRAT was also used to generate informative principal components for population sub-structure.

Health, Aging and Body Composition Study-African Americans (AA_HealthABC): African Americans from the Health ABC study (described above) with quality GW scans were considered for stage 2 of the current study. EIGENSTRAT [2] was used to carryout ancestry analyses and only study samples with substantial similarity to the YRI and ASW HapMap 3 reference samples were included. EIGENSTRAT was also used to generate informative principal components for population sub-structure.

The Jackson Heart Study-African Americans (AA_JHS): JHS is a prospective population-based study to seek the causes of the high prevalence of common complex diseases among African Americans in the Jackson, Mississippi metropolitan area [31]. During the baseline examination period (2000-2004) 5,301 self-identified African Americans were recruited from four sources, including (1) randomly sampled households from a commercial listing; (2) ARIC participants; (3) a structured volunteer sample that was designed to mirror the eligible population; and (4) a nested family cohort. Unrelated participants were between 35 and 84 years old, and members of the family cohort were ≥ 21 years old when consent for genetic testing was obtained and blood was drawn for DNA extraction. In the current study, JHS participants who were also enrolled in the ARIC study were analyzed with the AA_ARIC dataset. To control for sub-structure, principal components were generated using the EIGENSTRAT software [2].

The Multi-Ethnic Study of Atherosclerosis-African American (AA_MESA): African Americans from the MESA cohort (described above) with quality GW scans were considered for stage 2 of the current study. To control for sub-structure, principal components were generated using the EIGENSTRAT software [2].

The Women's Health Initiative-African Americans (AA_WHI): The SNP Health Association Resource (SHARe) minority cohort included self-identified African American and Hispanic women from the Women's Health Initiative Study [32] who provided consent for genetic research. A total of 8,515 African American women were genotyped; a subset of which had available coffee consumption information and were included for stage 2 of the current GWAS. We

computed eigenvectors using EIGENSTRAT[2] at 178,101 markers that were in common between our samples and the reference panels (475 publically available samples from ancestral populations -YRI, CEP, HGDP East Asian and Native Americans). We attempted to control for population stratification by including the first 10 principal components.

2. Statistical Analysis

2a. Pathway analysis of stage 1 meta-analysis

We used Meta-Analysis Gene-set Enrichment of variaNT Associations (MAGENTA, version 2.4)[33] to test whether the stage 1 GW meta-analysis results for phenotype 1 and phenotype 2 were enriched for members of specific biological pathways. MAGENTA calculates a gene association P-value based on the most significant SNP association P-value of all SNPs in a gene region (defined as 110 kb upstream to 40 kb downstream from transcript start/stop). The software corrects each P-value for gene size, number of SNPs/gene, and recombination. Results were tested against the following databases: Gene Ontology, KEGG, PANTHER Biological Processes, PANTHER Molecular Function, Reactome, BioCarta and Ingenuity Pathway. For each pathway, enrichment of highly ranked gene scores above the 95th percentile of all gene scores in the meta-analyses was evaluated compared to 10,000 randomly sampled gene sets of identical size from the genome.

2b. SNP-selection for replication

Stage 1 phenotype 1 (coffee cups/d among coffee consumers) results for SNPs based on less than 50% of maximum sample size or SNPs imputed by all studies were excluded, leaving 1,452,203 SNPs for further consideration. We followed-up SNPs based on two selection approaches:

i) *SNP-level approach:* Among SNPs showing evidence for association ($P<1\times10^{-5}$) and in strong LD, we elected to follow up only the most significant SNP. SNPs with $r^2<0.3$ (HapMap CEU) and at a distance 500kb or greater were treated as independent association signals. Of the 62 SNPs that remained, only those showing directional consistency across at least two thirds of the contributing studies were selected which yielded a list of 48 SNPs.

ii) *Gene-level approach:* Meta-analysis summary level data for all 1,452,203 SNPs was subject to gene-based analysis using Versatile Gene-based Association Study

(VEGAS)[34]. The software applies a test that incorporates information from a set of markers within a gene and accounts for LD between markers by using simulations from the multivariate normal distribution. A Bonferroni-corrected threshold of $P < 2.8 \times 10^{-6}$ [~0.05/17,787 (number of autosomal genes)] is used to indicate significant GW gene-based association. For the current study, the software's 'BestSNP's from each gene showing nominal gene-level significance ($P < 1 \times 10^{-4}$) and directional consistency across at least two thirds of the contributing studies were selected which yielded a list of 43 SNPs.

Combining these two approaches yielded 57 'independent' SNPs. Of these SNPs, 21 and 13 mapped to common regions on chromosome 7 and 15, respectively. Based on a second round meta-GWAS conditioning on the GW significant loci in these regions (see Statistical analysis section of main paper)[35], four of the SNPs on chromosome 7 were potentially independent. Thus, a total of 29 SNPs were considered for follow-up in stage 2. The identical SNP-selection approach was applied to our Stage 1 phenotype 2 (high vs none/low) results yielding 19 SNPs; four overlapping with the prior list of 29 SNPs for phenotype 1. Many SNPs selected for phenotype 2 and phenotype 1 were in strong LD but no additional pruning was applied to the combined lists. Thus, a total of 44 SNPs were selected for follow-up in stage 2; although proxies were provided to the follow-up groups in case the lead SNP failed quality control.

2c. Stage 2 (replication)

All 44 SNPs underwent *in silico* replication in population-based studies of European and African American Ancestry, herein referred to as 'European' and 'African American' studies, respectively, for simplicity. Four stage 2 studies included new samples drawn from studies in stage 1. Two studies were limited to metabochip content; while GW-data were available for the remainder. Only high quality imputed (quality matrix > 0.5) or genotyped SNPs were analyzed. Perfect proxies were selected if the 'best' index SNP from stage 1 was not available or failed quality control (details provided in Supplementary Table S9). Ethnicity-specific meta-analyses were performed using the same statistical models and methods as described for stage 1 in section 4a above but without GC correction. Since all 44 SNPs were at least nominally associated with each phenotype, all 44 SNPs were tested for association with each phenotype in stage 2 regardless of which phenotype the signal initially came from.

Due to potential differences in LD pattern between European and African American populations, we considered both SNP-level (described above) and gene-level associations for the

latter. SNP-level analysis and meta-analysis proceeded as described for stage 2 European studies. For gene-level analysis, candidate genes spanning the 44 loci were selected in three *exclusive* stages (additional details are provided with results in Supplementary Tables S12 and S13). First, for any SNP selected for replication based on stage 1 gene-level results, only the matching gene was tested for that locus. Second, we submitted our SNP-list to Gene Relationships Across Implicated Loci (GRAIL)[36] to quantify the functional relationships between genes mapping near each SNP using a text-based similarity metric and for any SNP yielding a significant (corrected P-value <0.05) candidate gene, only the candidate gene was tested for that locus. The GRAIL analysis parameters were set as follows: genome assembly: HapMap Release 22/Hg18; HapMap population: CEU; functional data-source: PubMed text (August 2012); gene-size correction: on; gene list: all human genes within database. Finally, for the remaining loci, all query genes for each, as identified by GRAIL (see section 6a. 'defining candidate regions and genes' below), were selected for gene-level testing.

2d. Between-study heterogeneity analysis

Post-hoc subgroup analysis and meta-regressions were performed to investigate possible sources of between-study heterogeneity in our primary (phenotype 1) meta-analysis of loci mapping to 7p21 (rs6968554) and 15q25 (rs2472297). Covariates considered included cohort mean age, proportion of females, proportion of current smokers, mean coffee consumption, geographic regions (North America, Europe/Australia) and SNP quality. We restricted our investigation to stage 1 and stage 2 studies of European ancestry to improve efforts to identify sources of heterogeneity unrelated to population stratification. We performed random-effects meta-analyses and meta-regression as implemented by the *metan* and *metareg* commands, respectively, in Stata (version 10.0, Stata Corp, College Station, TX, USA). No GC was applied. In all meta-analyses, between-study heterogeneity was tested by the Q statistic and quantified by the I^2 value[37]. Characteristics yielding differential heterogeneity in subgroup analysis or moderate associations (P<0.1) with effect sizes according to univariate meta-regressions were considered for multivariate meta-regressions.

2e. Fine-mapping

To assess the improvement in fine-mapping resolution due to trans-ethnic meta-analysis we applied the methods of Franceschini *et al* [38] to the GW-summary results of stage 1 European and stage 2 African American studies (GW-summary results were not available for stage 2 European

studies). We defined 95 and 99% ‘‘credible sets’’ of SNPs with the strongest signals of association on the basis of the European-only trans-ethnic meta-analysis and then after inclusion of the African American studies. At each locus, defined by the genomic region 500 kb up and downstream of the lead SNP, we calculated the posterior probability that the j^{th} SNP is ‘‘causal’’ by

$$\phi_j = \frac{BF_j}{\sum_k BF_k}$$

where BF_j denotes the BF in favor of association of the j^{th} SNP from the trans-ethnic analysis, and the summation in the denominator is over all SNPs across the locus. A 100ω% credible set at the locus was then constructed through (1) ranking all SNPs according to their BF and (2) combining ranked SNPs until their cumulative posterior probability exceeded ω (i.e. 0.95 or 0.99).

3. Potential SNP-Function

For all eight confirmed loci (Table 1), index SNPs and their correlated proxies (defined as $r^2 \geq 0.8$ with index SNP) were examined for putative function using publicly available resources.

3a. Regulatory elements

Using the customized track features of the UCSC browser, the stage 1 results were aligned with the genome according to Hg19 position, enabling an integrative view of regulation data from ENCODE[39]. HaploReg[40] and RegulomeDB[41] were also used to retrieve specific locus functional annotations derived from ENCODE[39], TRANSFAC[42], JASPAR[43], chromatin state segmentation experiments[44], protein-binding microarray experiments[45-47] and 1000 Genome Project (March 2012)[48]. Conservation predictions from GERP[49] and SyPhi[50, 51]; were also extracted via HaploReg.

3b. Expression and methylation quantitative trait loci (eQTL, mQTL)

We examined associations between coffee consumption SNPs and expression of nearby genes in lymphoblastoid cell lines[52-58], and liver[59, 60], adipose[54] and brain cortex[61] tissues, via the eQTL Chicago and GENe Expression VARiation integrative analysis and visualization tools[62]. eQTLs for whole blood were obtained from a recent meta-analysis of 5,300 samples (GEO:GSE36382, GSE20142, GSE20332, GSE33828, GSE33321, GSE47729; ArrayExpress:E-TABM-1036, E-

MTAB-945, E-MTAB-1708)[63]. eQTL and mQTL results from cerebellum and frontal cortex were obtained via collaboration with the North American Brain Expression Consortium (NABEC) and UK Brain Expression Consortium (UKBEC), although data are also publically available (GEO:GSE15745; dbGaP:phs000249.v1.p1)[64, 65]. Study specific significance thresholds for eQTLs are described in the original papers. For mQTLs, a Bonferroni corrected P value of 6.25×10^{-3} (0.05/8 loci) was our threshold criteria for significance.

4. Biological Inferences

4a. Defining candidate regions and genes

GRAIL[36] and Disease Association Protein-Protein Link Evaluator (DAPPLE)[66] were used to identify a broad set of candidate genes mapping to the eight confirmed loci associated with coffee consumption. For each query SNP both bioinformatic tools find the furthest neighboring SNPs in the 3′ and 5′ direction in LD (r^2>0.5, HapMap CEU) and then proceed outward in each direction to the nearest recombination hotspot. All genes that overlap that interval are considered implicated by the SNP. If there are no genes in that region, GRAIL will extend the interval an additional 250 kb in either direction. We considered all genes identified by GRAIL *or* DAPPLE. This gene list was also supplemented with *RHBDD2* (7q11.23)*, CCDC34* (11p13), *ISLR* (15q24), *SNUPN* (15q24) and *ATAD5* (17q11.2); genes whose expression correlates with loci of interest (Supplementary Table S17) but were not identified as a regional candidate by GRAIL or DAPPLE. A total of 91 genes (listed in Supplementary Table S20) were considered candidates for the eight confirmed loci.

4b. Tissue expression

Relative expression levels of candidate genes in human liver, human brain and primate taste buds (tissues relevant to our phenotype of interest and for which data was available) were obtained from the following publically available and/or published resources:

i) Illumina Human Body Map 2.0 Project: The Broad Institute's Integrative Genomics Viewer [67, 68] was used to examine RNA-seq data from the Illumina Human Body Map 2.0 Project, accessible from ArrayExpress (E-MTAB-513). Reads were 75 base pairs long and came from individual and mixtures of 16 human tissue RNA. The samples were prepared using the Illumina mRNA-seq kit. They were made with a random priming

process and are not stranded. For the current study, 'low' (<log 500) and 'high' (>log 500) expression was defined relative to normalized and absolute log values of expression.

ii) Allen Human Brain Atlas- the Allen Human Brain Atlas[69] is a multimodal atlas of gene expression and anatomy comprising a comprehensive "all genes, all structures" array-based dataset of gene expression and complementary in situ hybridization studies targeting selected genes in specific brain regions. All data are publicly available online (http://human.brain-map.org/) along with a suite of integrated data visualization and mining tools. To compliment the brain expression data provided by Illumina Human Body Map 2.0 Project, we retrieved and downloaded normalized expression data from six human donors (H0351.2001, H0351.2002, H0351.1009, H0351.1012, H0351.1015, H0351.1016, Ages 24 to 57, protocol details: http://help.brain-map.org/display/humanbrain/Documentation). We ignored expression data for sulici and spaces, since data was available for only three of the six subjects. For each brain region, we calculated the mean expression z-score across all six subjects. A mean z-score >1.5 was our threshold for defining regions highly expressing the probe/gene of interest.

iii) Taste Buds. Comprehensive gene-expression data for human taste buds is not currently available. We therefore obtained taste bud gene expression data from a database generated using laser capture microdissection procured fungiform (FG) and circumvallate (CV) taste buds from primates (GEO:GSE16485, Rhesus Macaque Genome Array)[70]. We examined results as provided in the supplementary materials of the original paper[70]. Accordingly, genes were defined as taste bud-associated if they met the following inclusion criteria: minimum mean taste bud expression ≥25 fold and fold expression difference [taste bud versus lingual epithelium (LE)] ≥5 with a p value ≤0.05. Genes were selected from each set of pairwise comparisons to give 4 lists (CV versus LE, FG versus LE, CV bottom versus LE and CV top versus LE) that were subsequently combined to give a single non-redundant master list consisting of 2382 taste bud-genes.

4c. Literature mining

SNIPPER (http://csg.sph.umich.edu/boehnke/snipper) was used to systematically scan the literature for all gene-term combinations. The search terms were selected on the basis of their relationship to coffee consumption behavior and included 'caffeine', 'coffee', 'taste', 'smell', 'gustation', 'olfaction', 'psychostimulant', and 'addiction'. As a product of the blinded-nature of

the software, papers with no relevance may be extracted. Examples: *CAD* for 'coronary artery disease'; *AHR* for 'Amer. His. Rev'; *POR* for 'prevalence odds ratio'; and search terms included as covariates in multivariate regressions. Therefore all top scientific articles were manually reviewed for relevance and specific texts extracted as evidence supporting the gene-term connection (Supplementary Table S26). To highlight results that point towards functional candidacy of genes in the region, we further excluded any references to GWAS or other genetic association studies; with the exception of functional studies.

4d. Mouse phenotypes

Phenotype information on candidate genes having mouse homologues was retrieved from the Mouse Genome Database, Mouse Genome Informatics, The Jackson Laboratory, Bar Harbor, Maine (http://www.informatics.jax.org, March 2013).

4e. Network construction

To inform hypotheses underlying the link between loci and coffee consumption as well as connections between loci, MetaCore™ (GeneGO, Thomson Reuters, New York, NY), Ingenuity Pathway Analysis™ (Ingenuity Systems), STRING 9.0[71] and STITCH 3.1[72] were used to mine available *experimental* databases for *direct* relationships *between* candidate regions. Given strong priors implicating six of the eight loci with caffeine metabolism/response pathways, we chose to leverage, rather than ignore, the vast amount of credible research on these pathways by supplementing our candidate list of 91 genes with known candidates in these pathways[73-77]. All relationships identified were manually curated and references cited in Supplementary Table S21. Connections among gene members of the same locus were ignored. Cytoscape[78] was used to build and visualize the final network. Two networks consisted of two nodes and were not displayed.

4f. Gene prioritization analysis

GRAIL[36] and DAPPLE[66] were also used to formally identify the gene most likely to underlie each of the eight confirmed associations while leveraging connectivity between all genes mapping to these regions. We carried out a PubMed literature analysis using GRAIL including all eight sentinel SNPs simultaneously and with gene size correction enabled. We used the 2006 PubMed data set to avoid confounding from GWAS results arising after that date. In a second iteration, we leveraged connectivity between probable candidates ('seed' genes) and novel

regions ('query'- all genes near SNP) to inform gene prioritization for the latter. More specifically, we examined only 2 'query' SNPs (rs1260326 and rs7800944) and included the following genes as 'seeds': *AHR* for rs6968554, *CYP1A2* for rs2472297, *POR* for rs17685, *ABCG2* for rs1481012, *BDNF* for rs6265 and *SLC6A4* for rs9902453. We employed DAPPLE[66], which uses a refined database of high-confidence protein-protein interactions (InWeb)[79, 80], to assess the amount of physical interactions connecting the genes near the eight loci. Following a similar approach to that applied with GRAIL, we performed a second analysis prioritizing six of the eight loci to further inform candidate gene identification for the remaining 2 loci. Both direct and indirect (through first-order common interaction partners) were measured and compared to a random expectation over 10,000 permutations.

4g. In vitro response to caffeine

Gene expression in response to caffeine was obtained from the 'Toxicogenomics Project-Genomics Assisted Toxicity Evaluation system' *(Open TG-GATEs, http://toxico.nibio.go.jp)*, a large-scale database of transcriptomics and pathology data potentially useful for predicting the toxicity of new chemical entities. The overall experiment (E-MTAB-798) protocol is described in detail http://www.ebi.ac.uk/arrayexpress/experiments/E-MTAB-798/protocols/?sortby=accession&sortorder=ascending . Briefly, cryopreserved human hepatocytes were thawed, washed twice with medium and seeded. Following an attachment period of 3 hours, the medium was replaced and kept overnight before drug exposure at 37°C in an atmosphere of 5% CO_2. Caffeine was administered at two dose levels: control (vehicle only), 1500 μM, and 7500 μM and cells were exposed to the compound for 8 and 24 hours. Experiments were conducted in duplicate. Following compound exposure, hepatocytes were then lysed with RLT buffer and collected for expression profiling on the Affymetrix GeneChip Human Genome U133 Plus 2.0 array. Total RNA was isolated from the hepatocyte lysate using an RNeasy kit (Qiagen). Labeling was carried out according to manufacturer's instructions. 10 μg of fragmented cRNA was hybridized to the probe array for 18 hours at 45°C at 60 rpm, after which the array was washed and stained by streptavidin-phycoerythrin using Fluidics Station 400 (Affymetrix) and scanned by Gene Array Scanner (Affymetrix). Affymetrix raw expression data were processed in R using the Bioconductor *limma* package[81]. Raw data were normalized by robust multiarray averaging. Linear models were used to average data between replicate arrays and to also look for variability between them. Using the complete expression data set (54,675

probes), linear models were fit to contrast caffeine vs vehicle treatment using Benjamini-Hochberg adjustment for false discovery rate. Log ratios were converted to fold-change for ease of interpretation. Of the 91 candidate genes spanning the 8 confirmed loci, 77 are expressed in the liver (Supplementary Table S22). Probes for 75 of the latter were present on the array; probes for *SNORA14A* and *FTHIP3*, were absent.

We additionally tested whether the complete expression data set was enriched for differential expression of our set of 75 gene candidates using Gene Set Enrichment Analysis (GSEA, Broad Institute)[82, 83]. Ratio of classes was used for ranking genes and we permuted by gene-set rather than phenotype since there were not enough random permutations of sample labels to generate a sufficient null distribution.

5. Pleiotropy and Clinical Inferences

We searched the National Human Genome Research Institute (NHGRI) GWAS catalogue[84] and the Metabolomics GWAS server (results from '*KORA + Twins UK (Meta)*')[85] for GW-significant associations with variants in strong LD (CEU: $r^2>0.8$) with the eight index SNPs associated with coffee consumption. GW-significance was defined as $P<5\times10^{-8}$ for studies in the NHGRI GWAS catalogue[84] and $P<1.03\times10^{-10}$ (as estimated by Shin et al)[85]) for metabolite GWAS[85]. Look-ups in published GWAS of caffeine taste-intensity[86, 87] and caffeine-induced insomnia[88] were provided by their corresponding authors. Complete GWAS summary data for coffee-implicated diseases or traits were additionally queried (Supplementary Table S32). For these specific look-ups, a Bonferroni corrected *P* value of 6.25×10^{-3} (0.05/8 loci) was our threshold criteria for significance.

Supplementary Notes

1. Follow-up of Confirmed Loci in Pakistani Population

To explore whether the candidate loci maintained robust associations in cultures where coffee drinking is less prominent than in Europe or the US, the eight confirmed loci were additionally examined in the Pakistan Risk of Myocardial Infarction Study (PROMIS), an ongoing retrospective case-control study designed to identity genetic, lifestyle and other determinants of CHD in South Asia[89]. Each participant completed a questionnaire on lifestyle and dietary habits (tailored to local dietary patterns). Detailed information on participant characteristics, genotyping and imputation for the 12,236 individuals with phenotype information and quality GW-scan data are provided in Supplementary Tables S1-S3, S5 and S6. STATA was used to conduct each association analysis and statistical models were adjusted for age, sex, myocardial infarction status, smoking (never, former, current <15 cig/d, current ≥15 cig/d) and ten PCs. The latter were calculated using IBS matrix analyses in PLINK[7]. The frequency of alleles associated with increased coffee consumption in European and African Americans ranged from 0.03 (rs2472297) to 0.90 (rs1481012) in PROMIS. Coffee consumption was less common among this Pakistani population compared to the participating European and African American studies. Only 365 individuals reported drinking coffee and among these none of the eight loci were associated with level of coffee consumed (phenotype 1, lowest P=0.01 for rs1481012). Moreover, none of the eight loci were associated coffee drinking initiation based on a comparison of these 365 coffee consumers to the rest of the population sample (n=11,871)(phenotype 2, lowest P=0.12 for rs2472297). Further studies will be needed to determine whether SNP-coffee associations confirmed in Europeans and African Americans also apply to other populations.

2. Between-Study Heterogeneity at 7p21 and 15q24

The current study as well as our previous GWAS of caffeine and coffee intake[90-92] identified significant between-study heterogeneity in effects for SNPs mapping to 7p21 and 15q24 (I^2≤70% regardless of stage or ancestry, Supplementary Tables S8, S10 and S11). The low heterogeneity observed for the remaining six confirmed loci suggested interactions with individual cohort characteristics at these two SNPs might be the cause of their heterogeneity, rather than a systematic feature of the meta-analysis. In post-hoc subgroup meta-analyses, heterogeneity in

effects sizes for rs6968554 and rs2472297 was driven by studies with a larger proportion of females, with a higher mean coffee consumption and of populations from Europe and Australia (Supplementary Table S25A). Thus, univariate and multivariate meta-regressions revealed that cohort-wide mean level of coffee consumption was the most significant source of heterogeneity for both loci, such that study mean coffee consumption was positively correlated with variant effect size (Supplementary Table S25B). Whether level of coffee consumption *per se* is the source of heterogeneity (i.e. dose-dependent effect of SNP) or instead reflects a separate confounding factor distinguishing low from high coffee drinkers requires further study. Study geographical region and proportion of current smokers were highly correlated with study mean coffee consumption (Spearman's rank correlation coefficient >0.4) and thus not included in multivariate regressions. Nevertheless, smoking behavior remains a strong candidate source of heterogeneity and merits independent investigation.

3. Methodological Considerations

The current GWAS focused on consumption of regular-type coffee since this was the predominate type of coffee consumed by participants of the contributing studies or was the only type of coffee captured by study-specific FFQs. Total dietary caffeine intake was not examined since information on non-coffee sources of caffeine was not available for all contributing studies. Moreover, previous studies suggest that some of the heritability underlying specific caffeine sources (i.e. coffee, tea and soda) may be distinct in relation to total caffeine intake [93]. Thus our phenotype was chosen to both maximize sample size and reduce phenotype heterogeneity. Nevertheless, imprecision in phenotypic assessment and differences across studies could have limited our power for discovery. Although dietary intake obtained by FFQ is subject to misclassification, validation studies in subsamples of participating studies indicate that assessment of coffee consumption is remarkably accurate[94-98]. For example, correlations (r) between intake of coffee from FFQs and multiple dietary records ranged between 0.78 and 0.90[96-98]. However, the precise chemical composition of different coffee preparations is not captured by standard FFQs and is likely to vary within and between populations.

Many studies collected coffee data in categories of cups/day (e.g. 2-3 cups/day) and the median value of each category (e.g. 2.5 cups/day) used for our GW analysis was not strictly non-Gaussian distribution. Our previous GWAS of caffeine and coffee consumption employed

different trait modeling strategies directed at normalizing the categorical data for linear regression analysis and yet revealed the same top loci as found here (i.e. *AHR* and *CYP1A2*)[90-92]. In the WGHS, an alternative modeling strategy applied to the lead SNPs (Table 1) using ordinal regression of coffee consumption categories and including non-coffee drinkers yielded association statistics that were highly comparable in significance to the linear regression approach (Supplementary Table S33)**.** Thus, the lead SNPs appear to be robust predictors of coffee intake. With the exception of possibly rare alleles, the linear modeling approach taken in the current study is likely not especially susceptible to spurious association.

In addition to nearly doubling the discovery sample size included in our previous reports[90-92], we have extended our study to include African Americans for the first time. To cope with heterogeneous effect sizes between populations, studies from all stages were included in an overall meta-analysis using MANTRA[38]. Although random-effects methods are designed to address heterogeneity explicitly, they rely on a conservative assumption that the effect sizes are different across studies[99]. Applied in the context of GWAS, MANTRA is robust to heterogeneous effect sizes at a genuinely associated genetic locus that will attenuate the overall signal under conventional models, particularly in the presence of diverse population genetic architecture or when some studies do not carry the association due to different study designs or environmental modification[100]. At the recommended $\log_{10}$ BF >5, the false-positive rate in simulations is zero[100]. The BF criterion is not meant to be calibrated against the definition of statistical significance[101], but given conventional reliance on p-values, we chose a significance threshold of $\log_{10}$ BF >5.64 which approximates a traditional GW P-value threshold of 5×10^{-8} under general assumptions[101, 102]. This threshold is more conservative than the recommended $\log_{10}$ BF >5 threshold but more lenient than the $\log_{10}$ BF >6.1 threshold suggested in simulation[100]. Had the recommended $\log_{10}$ BF >5 threshold been applied, a second SNP near *BDNF*, rs12288512 ($\log_{10}$ BF=5.24), but in low LD with rs6265 (CEU: r^2=0.05) would have reached significance. Conversely, had the more conservative threshold of 6.1 been used, rs6265 ($\log_{10}$ BF=5.76) and rs1481012 ($\log_{10}$ BF=6.08) would not have met GW-significance. For these latter loci, we cite a wealth of independent functional and biological evidence supporting their link to coffee drinking behavior and are confident these loci are not false positives.

4. Acknowledgements

Atherosclerosis Risk in Communities (ARIC, AA_ARIC): The ARIC Study is carried out as a collaborative study supported by National Heart, Lung, and Blood Institute contracts HHSN268201100005C, HHSN268201100006C, HHSN268201100007C, HHSN268201100008C, HHSN268201100009C, HHSN268201100010C, HHSN268201100011C, and HHSN268201100012C, R01HL087641, R01HL59367 and R01HL086694; National Human Genome Research Institute contract U01HG004402; and National Institutes of Health contract HHSN268200625226C. The authors thank the staff and participants of the ARIC study for their important contributions. Infrastructure was partly supported by Grant Number UL1RR025005, a component of the National Institutes of Health and NIH Roadmap for Medical Research**.**

The Busselton Health Study (Busselton): The Busselton Health Study acknowledges the generous support for the 1994/5 follow-up study from Healthway, Western Australia and the numerous Busselton community volunteers who assisted with data collection and the study participants from the Shire of Busselton. The Busselton Health Study is supported by The Great Wine Estates of the Margaret River region of Western Australia.

Cardiovascular Health Study (CHS, AA_CHS): CHS was supported by contracts HHSN268201200036C, HHSN268200800007C, N01 HC55222, N01HC85079, N01HC85080, N01HC85081, N01HC85082, N01HC85083, N01HC85086, and grants HL080295 and HL085251 from the National Heart, Lung, and Blood Institute (NHLBI), with additional contribution from the National Institute of Neurological Disorders and Stroke (NINDS). Additional support was provided by AG023629 from the National Institute on Aging (NIA). A full list of principal CHS investigators and institutions can be found at CHS-NHLBI.org. The provision of genotyping data was supported in part by the National Center for Advancing Translational Sciences, CTSI grant UL1TR000124, and the National Institute of Diabetes and Digestive and Kidney Disease Diabetes Research Center (DRC) grant DK063491 to the Southern California Diabetes Endocrinology Research Center.

Cohorte Lausannoise (COLAUS): The CoLaus study was supported by research grants from GlaxoSmithKline and from the Faculty of Biology and Medicine of Lausanne, Switzerland, and is currently supported by Swiss National Science Foundation (grant no: 33CSCO-122661 and 33CSC0-139468). The authors thank the other PIs of the study (Vincent Mooser, Gérard Waeber, Martin Preisig and Dawn Waterworth), the staff and all the participants of the CoLaus study for their important contributions.

Danish National Birth Cohort: Preterm Birth Study (DNBC): The Danish National Birth Cohort received major funding support from the Danish National Research Foundation, the Danish Pharmacists' Fund, the Egmont Foundation, the March of Dimes Birth Defects Foundation, the Augustinus Foundation, and the Health Fund of the Danish Health Insurance Societies. The generation of GWAS genotype data for the Danish National Birth Cohort samples was carried out within the Gene Environment Association Studies (GENEVA) consortium with

funding provided through the National Institutes of Health's Genes, Environment, and Health Initiative (U01HG004423; U01HG004446; U01HG004438). The study protocol was approved by the Danish Scientific Ethical Committee and the Danish Data Protection Agency.

Estonian Genome Center of the University of Tartu (EGCUT1, EGCUT2, EGCUT3): EGCUT received targeted financing from Estonian Government SF0180142s08, Center of Excellence in Genomics (EXCEGEN) and University of Tartu (SP1GVARENG). We acknowledge EGCUT technical personnel, especially Mr. V. Soo and S. Smit. Data analyses were carried out in part at the High Performance Computing Center of University of Tartu.

European Prospective Investigation in Cancer and Nutrition Norfolk Study (EPIC-Norfolk): The EPIC-Norfolk Study is supported by program grants from the Medical Research Council, and Cancer Research UK and with additional support from the European Union, Stroke Association, British Heart Foundation, Research into Ageing, Department of Health, The Wellcome Trust and the Food Standards Agency. We acknowledge the contribution of the staff and participants of the EPIC-Norfolk Study.

Erasmus Rucphen Family study (ERF): The genotyping for the ERF study was supported by EUROSPAN (European Special Populations Research Network) and the European Commission FP6 STRP grant (018947; LSHG-CT-2006-01947). The ERF study was further supported by grants from the Netherlands Organisation for Scientific Research, Erasmus MC, the Centre for Medical Systems Biology (CMSB) and the Netherlands Brain Foundation (HersenStichting Nederland). We are grateful to all patients and their relatives, general practitioners and neurologists for their contributions and to P. Veraart for her help in genealogy, Jeannette Vergeer for the supervision of the laboratory work and P. Snijders for his help in data collection.

Family Heart Study (FamHS): FamHS was supported by NIH grants R01-HL-087700 and R01-HL-088215 (Michael A. Province, PI) from NHLBI; and R01-DK-8925601 and R01-DK-075681 (Ingrid B. Borecki, PI) from NIDDK.

Fenland Study (Fenland): The Fenland Study is funded by the Wellcome Trust and the Medical Research Council. We are grateful to all the volunteers for their time and help, and to the General Practitioners and practice staff for help with recruitment. We thank the Fenland Study Investigators, Fenland Study Co-ordination team and the Epidemiology Field, Data and Laboratory teams. Biochemical assays were performed by the National Institute for Health Research, Cambridge Biomedical Research Centre, Core Biochemistry Assay Laboratory, and the Cambridge University Hospitals NHS Foundation Trust, Department of Clinical Biochemistry.

Framingham Heart Study (FHS): This research was conducted in part using data and resources from the Framingham Heart Study of the National Heart Lung and Blood Institute of the National Institutes of Health and Boston University School of Medicine. This work was partially supported by the National Heart, Lung and Blood Institute's Framingham Heart Study in collaboration with Boston University (Contract No. N01-HC-25195) and its contract with Affymetrix, Inc for genotyping services (Contract No. N02-HL-6-4278). A portion of this research utilized the Linux Cluster for Genetic Analysis (LinGA-II) funded by the Robert

Dawson Evans Endowment of the Department of Medicine at Boston University School of Medicine and Boston Medical Center.

Gothenburg Osteoporosis and Obesity Determinants Study (GOOD): Financial support was received from the Swedish Research Council (K2010-54X-09894-19-3, 2006-3832), the Swedish Foundation for Strategic Research, the ALF/LUA research grant in Gothenburg, the Lundberg Foundation, the Torsten and Ragnar Söderberg's Foundation, Petrus and Augusta Hedlunds Foundation, the Västra Götaland Foundation, the Göteborg Medical Society, the Novo Nordisk foundation and the European Commission grant HEALTH-F2-2008-201865-GEFOS. We would like to thank Dr. Tobias A. Knoch, Luc V. de Zeeuw, Anis Abuseiris, and Rob de Graaf as well as their institutions the Erasmus Computing Grid, Rotterdam, The Netherlands, and especially the national German MediGRID and Services@MediGRID part of the German DGrid, both funded by the German Bundesministerium fuer Forschung und Technology under grants #01 AK 803 A-H and # 01 IG 07015 G for access to their grid resources. We would also like to thank Karol Estrada, Department of Internal Medicine, Erasmus MC, Rotterdam, The Netherlands for advice regarding the grid resources.

Helsinki Birth Cohort Study (HBCS): We thank all study participants as well as everybody involved in the HBCS especially Katri Räikkönen, Aarno Palotie, and Elisabeth Widen. Helsinki Birth Cohort Study has been supported by grants from the Academy of Finland, the Finnish Diabetes Research Society, Folkhälsan Research Foundation, Novo Nordisk Foundation, Finska Läkaresällskapet, Signe and Ane Gyllenberg Foundation, University of Helsinki, Ministry of Education, Ahokas Foundation, Emil Aaltonen Foundation.

Health Aging and Body Composition Study (HealthABC, AA_HealthABC): This research was supported by NIA contracts N01AG62101, N01AG62103, and N01AG62106 and was supported in part by the Intramural Research Program of the NIH, National Institute on Aging (Z01 AG000949-02 and Z01 AG007390-07, Human subjects protocol UCSF IRB is H5254-12688-11). The genome-wide association study was funded by NIA grant 1R01AG032098-01A1 to Wake Forest University Health Sciences and genotyping services were provided by the Center for Inherited Disease Research (CIDR). CIDR is fully funded through a federal contract from the National Institutes of Health to The Johns Hopkins University, contract number HHSN268200782096C. This study utilized the high-performance computational capabilities of the Biowulf Linux cluster at the National Institutes of Health, Bethesda, MD. (http://biowulf.nih.gov).

Health Professionals Follow-up Study (HPFS T2D, HPFS CHD, HPFS GA, HPFS KS, HPFS CC) and Nurses' Health Study (NHS T2D, NHS CHD, NHS GA, NHS KS, NHS BrCa, NHS CC): The HPFS and NHS GWAS were supported by grants from the National Institutes of Health [NCI (CA40356, CA087969, CA055075, CA98233, U01 CA137088, R01 CA059045, R01 CA137178), NIDDK (DK058845, DK070756), NHGRI (HG004399, HG004728), NHLBI (HL35464)] with additional support from Merck/Rosetta Research Laboratories, North Wales, PA. MCC was funded, in part, by a NARSAD Young Investigator Award and the American Diabetes Association Grant #7-13-JF-15.

Invecchiare in Chianti (inCHIANTI): The InCHIANTI study baseline (1998-2000) was supported as a "targeted project" (ICS110.1/RF97.71) by the Italian Ministry of Health and in part by the U.S. National Institute on Aging (Contracts: 263 MD 9164 and 263 MD 821336).

Cooperative Health Research in the Region of Augsburg Study (KORA_F3, KORA_S4) The KORA Augsburg studies were financed by the Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany, and supported by grants from the German Federal Ministry of Education and Research (BMBF). Part of this work was financed by the German National Genome Research Network (NGFN). Our research was supported within the Munich Center of Health Sciences.

The Multi-Ethnic Study of Atherosclerosis (MESA, AA_MESA): The Multi-Ethnic Study of Atherosclerosis (MESA) and MESA SHARe project are conducted and supported by contracts N01- HC-95159 through N01-HC-95169 and RR-024156 from the National Heart, Lung, and Blood Institute (NHLBI). Funding for MESA SHARe genotyping was provided by NHLBI Contract N02-HL-6-4278. The authors thank the participants of the MESA study, the Coordinating Center, MESA investigators, and study staff for their valuable contributions. A full list of participating MESA investigators and institutions can be found at http://www.mesa-nhlbi.org.

Northern Finland Birth Cohort 1966 (NFBC 1966): NFBC1966 is supported by the Academy of Finland [project grants 104781, 120315, 129418, Center of Excellence in Complex Disease Genetics and Public Health Challenges Research Program (SALVE)], University Hospital Oulu, Biocenter, University of Oulu, Finland (75617), the European Commission [EUROBLCS, Framework 5 award QLG1-CT-2000-01643], The National Heart, Lung and Blood Institute [5R01HL087679-02] through the SNP Typing for Association with Multiple Phenotypes from Existing Epidemiologic Data (STAMPEED) program [1RL1MH083268-01], The National Institute of Health/The National Institute of Mental Health [5R01MH63706:02], European Network of Genomic and Genetic Epidemiology (ENGAGE) project and grant agreement [HEALTH-F4-2007-201413], and the Medical Research Council, UK [G0500539, G0600705, PrevMetSyn/Public Health Challenges Research Program (SALVE)].

Netherlands Twin Register (Stage 1 samples- NTR s1, Stage 2 samples- NTR s2): Funding was obtained from the Netherlands Organization for Scientific Research (NWO: MagW/ZonMW grants 904-61-090, 985-10-002, 904-61-193,480-04-004, 400-05-717, Addiction-31160008, Middelgroot-911-09-032, Spinozapremie 56-464-14192), Center for Medical Systems Biology (CSMB, NWO Genomics), NBIC/BioAssist/RK(2008.024), Biobanking and Biomolecular Resources Research Infrastructure (BBMRI –NL, 184.021.007), VU University's Institute for Health and Care Research (EMGO+) and Neuroscience Campus Amsterdam (NCA), European Community's Seventh Framework Program (FP7/2007-2013), ENGAGE (HEALTH-F4-2007-201413); European Research Council (ERC 230374, 284167), Rutgers University Cell and DNA Repository (NIMH U24 MH068457-06), the Avera Institute, Sioux Falls, South Dakota (USA), and the National Institutes of Health (NIH, R01D0042157-01A, Grand Opportunity grants 1RC2MH089951-01 and 1RC2 MH089995-01). Part of the genotyping and analyses were funded by the Genetic Association Information Network (GAIN) of the Foundation for the National Institutes of Health.

Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial (PLCO): The authors thank the National Cancer Institute for funding and providing the human material collected by the PLCO Screening Trial, particularly Maria Teresa Landi, Nilanjan Chatterjee, Stephen Chanock, Meredith Yeager, Kevin Jacobs, and Bill Wheeler.

Rotterdam Study (RS1, RS2): The Rotterdam Study is supported by the Erasmus Medical Center and Erasmus University Rotterdam, The Netherlands Organization for Scientific Research (NWO), The Netherlands Organization for Health Research and Development (ZonMw), The Research Institute for Diseases in the Elderly (RIDE), The Netherlands Genomics Initiative, the Ministry of Education, Culture and Science, the Ministry of Health, Welfare and Sports, the European Commission (DG XII), and the Municipality of Rotterdam. The contribution of inhabitants, general practitioners and pharmacists of the Ommoord district to the Rotterdam Study is gratefully acknowledged.

Study of Health in Pomerania (SHIP): SHIP is part of the Community Medicine Research net (www.community-medicine.de) and the Greifswald Approach to Individualized Medicine (GANI-MED) consortium (www.gani-med.de) of the University Medicine Greifswald, Germany, which are funded by the Federal Ministry of Education and Research (BMBF 01ZZ9603, 01ZZ0103 and 03IS2061A), the Ministry of Cultural Affairs as well as the Social Ministry of the Federal State of Mecklenburg-West Pomerania. Genome-wide data have been supported by the Federal Ministry of Education and Research (grant no. 03ZIK012) and a joint grant from Siemens Healthcare, Erlangen, Germany and the Federal State of Mecklenburg-West Pomerania. The University of Greifswald is a member of the 'Center of Knowledge Interchange' program of the Siemens AG and the Cache´ Campus program of the InterSystems GmbH.

SORBS: SORBS was supported by grants from the German Research Council (SFB- 1052 "Obesity mechanisms"), from the German Diabetes Association and from the DHFD (Diabetes Hilfs- und Forschungsfonds Deutschland). IFB Adiposity Diseases is supported by the Federal Ministry of Education and Research (BMBF), Germany, FKZ: 01EO1001. We would like to thank Knut Krohn (Microarray Core Facility of the Interdisciplinary Centre for Clinical Research, University of Leipzig) for the genotyping/analytical support and Joachim Thiery (Institute of Laboratory Medicine, Clinical Chemistry and Molecular Diagnostics, University of Leipzig) for clinical chemistry services. We thank Nigel W. Rayner (WTCHG, University of Oxford, UK) for the excellent bioinformatics support. Reedik Mägi is funded by European Commission under the Marie Curie Intra-European Fellowship and by Estonian Government (grant #SF0180142s08).

The Hellenic study of Interactions between SNPs & Eating in Atherosclerosis Susceptibility (THISEAS): This work was funded by the Wellcome Trust. We like to thank the members of the WTSI Genotyping Facility in particular Sarah Edkins and Cordelia Langford. Recruitment for THISEAS was partially funded by a research grant (PENED 2003) from the Greek General Secretary of Research and Technology; we thank all the dieticians and clinicians for their contribution to the project.

TWINGENE: This work was supported by grants from the Ministry for Higher Education, the Swedish Research Council (M-2005-1112 and 2009-2298), GenomEUtwin (EU/QLRT-2001-01254; QLG2-CT-2002-01254), NIH grant DK U01-066134, The Swedish Foundation for Strategic Research (SSF; ICA08-0047).

Women's Genome Health Study (WGHS): The WGHS is supported by HL 043851 and HL69757 from the NHLBI and CA 047988 from the NCI, the Donald W. Reynolds Foundation, and the Fondation Leducq, with collaborative scientific support and funding for genotyping provided by Amgen.

Blue Mountain Eye Study (BMES): The Blue Mountains Eye Study (BMES) was supported by the Australian National Health & Medical Research Council (NHMRC), Canberra Australia (NHMRC project grant IDs 974159, 211069, 302068, and Centre for Clinical Research Excellence in Translational Clinical Research in Eye Diseases, CCRE in TCR-Eye, grant ID 529923). The BMES GWAS and genotyping costs was supported by Australian NHMRC, Canberra Australia (NHMRC project grant IDs 512423, 475604 and 529912), and the Wellcome Trust, UK as part of Wellcome Trust Case Control Consortium 2 (A Viswanathan, P McGuffin, P Mitchell, F Topouzis, P Foster, grant IDs 085475/B/08/Z and 085475/08/Z).

British Birth Cohort 1958 (1958BC T1DGC, 1958BC WTCCC2): These studies were supported by the Medical Research Council (MRC G0601653 and SALVE/PrevMedsyn) and the Academy of Finland. Collection of DNA in the 1958 Birth Cohort was funded by the MRC grant G0000934 and Wellcome Trust grant 068545/Z/02. Dr Sue Ring and Dr Wendy McArdle (University of Bristol), and Mr Jon Johnson (Centre for Longitudinal Studies, Institute of Education, London) are thanked for help with data linkage. This research used resources provided by the Type 1 Diabetes Genetics Consortium, a collaborative clinical study sponsored by the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), National Institute of Allergy and Infectious Diseases, National Human Genome Research Institute, National Institute of Child Health and Human Development, and Juvenile Diabetes Research Foundation International (JDRF) and supported by U01 DK062418. This study makes use of data generated by the Wellcome Trust Case-Control Consortium II. A full list of investigators who contributed to generation of the data is available from the Wellcome Trust Case-Control Consortium website (www.wtccc.org.uk). Funding for the project was provided by the Wellcome Trust under award 083948. Work was undertaken at Great Ormond Street Hospital /University College London, Institute of Child Health which received a proportion of funding from the Department of Health's National Institute of Health Research ('Biomedical Research Centres' funding). The Medical Research Council provides funds for the MRC Centre of Epidemiology for Child Health.

Gene-Lifestyle interactions And Complex traits Involved in Elevated disease Risk Study (GLACIER): We are indebted to the study participants for dedicating their time and samples. We thank the Västerbotten's Health Survey and Umeå Medical Biobank staff for biomedical data collection and preparation. We specifically thank John Hutiainen, Åsa Ågren and Sara Nilsson (Umeå Medical Biobank) for data organization, Kerstin Enquist and Thore Johansson (Västerbottens County Council) for expert technical assistance with DNA preparation. The current study was funded by Novo Nordisk, the Swedish Research Council, Påhlssons

Foundation, the Swedish Heart Lung Foundation, and the Skåne Health Authority (all to P.W.F). I.B is funded by the Wellcome Trust grant WT098051 and United Kingdom NIHR Cambridge Biomedical Research Centre.

Life Lines (LifeLines): The LifeLines Cohort Study was supported by the Netherlands Organization for 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; the Dutch Kidney Foundation; and the Dutch Diabetes Research Foundation. This work was supported by the National Consortium for Healthy Ageing, and the BioSHaRE-EU consortium (KP7, project reference 261433). The authors are grateful to the study participants, the staff of the LifeLines Cohort Study and Biobank, and the participating general practitioners and pharmacists. Statistical analyses were carried out on the Genetic Cluster Computer (http://www.geneticcluster.org) which is financially supported by the Netherlands Scientific Organization (NWO 480-05-003) along with a supplement from the Dutch Brain Foundation and the VU University Amsterdam.

Osteoporotic Fractures in Men (MrOS, Sweden): Financial support was received from the Swedish Research Council (K2010-54X-09894-19-3, 2006-3832),the Swedish Foundation for Strategic Research, the ALF/LUA research grant in Gothenburg, the Lundberg Foundation, the Torsten and Ragnar Söderberg's Foundation, the Västra Götaland Foundation, the Göteborg Medical Society, the Novo Nordisk foundation and the European Commission grant HEALTH-F2-2008-201865-GEFOS.

Prospective Investigation of the Vasculature in Uppsala Seniors (PIVUS): This project was supported by grants from the Swedish Research Council, the Swedish Heart-Lung Foundation, the Swedish Foundation for Strategic Research, the Royal Swedish Academy of Sciences, Swedish Diabetes Foundation, Swedish Society of Medicine, and Novo Nordisk Fonden. Genotyping was performed by the SNP&SEQ Technology Platform in Uppsala (www.genotyping.se). We thank Tomas Axelsson, Ann-Christine Wiman and Caisa Pöntinen for their excellent assistance with genotyping. The SNP Technology Platform is supported by Uppsala University, Uppsala University Hospital and the Swedish Research Council for Infrastructures.

Queensland Institute of Medical Research (QIMR): Funding for phenotype and blood collection was from NHMRC grants to NGM and NIH grants to ACH. EMB is supported by NHMRC grant 613608. We also thank Dixie Statham, Bronwyn Morris and Megan Ferguson for coordinating the data collection for the twins; David Smyth, Olivia Zheng and Harry Beeby for data management of the ATR; Anjali Henders, Lisa Bowdler, Steven Crooks (DNA processing); Sarah Medland, Dale Nyholt and Scott Gordon (imputation and genotyping QC).

Uppsala Longitudinal Study of Adult Men (ULSAM): This project was supported by grants from the Swedish Research Council, the Swedish Heart-Lung Foundation, the Swedish Foundation for Strategic Research, the Royal Swedish Academy of Sciences, Swedish Diabetes Foundation, Swedish Society of Medicine, and Novo Nordisk Fonden. Genotyping was

performed by the SNP&SEQ Technology Platform in Uppsala (www.genotyping.se). We thank Tomas Axelsson, Ann-Christine Wiman and Caisa Pöntinen for their excellent assistance with genotyping. The SNP Technology Platform is supported by Uppsala University, Uppsala University Hospital and the Swedish Research Council for Infrastructures.

Cardiovascular Risk in Young Finns (YFS): YFS has been financially supported by the Academy of Finland: grants 134309 (Eye), 126925, 121584, 124282, 129378 (Salve), 117787 (Gendi), and 41071 (Skidi), the Social Insurance Institution of Finland, Kuopio, Tampere and Turku University Hospital Medical Funds (grant 9M048 for 9N035 for TeLeht), Juho Vainio Foundation, Paavo Nurmi Foundation, Finnish Foundation of Cardiovascular Research and Finnish Cultural Foundation, Tampere Tuberculosis Foundation and Emil Aaltonen Foundation. We thank Ville Aalto for her assistance with the genetic data analysis.

Healthy Aging in Neighborhoods of Diversity across the Life Span-African Americans (AA_HANDLS): This research was supported by the Intramural Research Program of the NIH, National Institute on Aging and the National Center on Minority Health and Health Disparities (project # Z01-AG000513 and human subjects protocol # 2009-149). Data analyses for the HANDLS study utilized the high-performance computational capabilities of the Biowulf Linux cluster at the National Institutes of Health, Bethesda, Md. (http://biowulf.nih.gov).

The Jackson Heart Study-African Americans (AA_JHS): The Jackson Heart Study is supported by contracts N01-HC-95170, N01-HC-95171, N01-HC-95172 from the National Heart, Lung, and Blood Institute and the National Institute on Minority Health and Health Disparities, with additional support from the National Institute on Biomedical Imaging and Bioengineering.

The Women's Health Initiative-African Americans (AA_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. This manuscript was prepared in collaboration with investigators of the WHI, and has been reviewed and/or approved by the Women's Health Initiative (WHI). WHI investigators are listed at https://cleo.whi.org/researchers/SitePages/WHI%20Investigators.aspx . Funding for WHI SHARe genotyping was provided by NHLBI Contract N02-HL-64278. Funding for the genetic analyses were provided by NIH contract A10-0347-001 and R01 DK075681 awarded to Dr. North. The datasets used for the analyses described in this manuscript were obtained from dbGaP (accession phs000386.v3.p2)

Pakistan Risk of Myocardial Infarction Study (PROMIS): PROMIS has been supported by grants from the Wellcome Trust (084711/Z/08/Z), the US National Institutes of Health (1R21NS064908), the British Heart Foundation (RG/08/014), and Pfizer. We thank the members of the Wellcome Trust Sanger Institute's Genotyping Facility for genotyping PROMIS samples. Funding: Wellcome Trust grants 083948/B/07/Z and 077016/Z/05/Z

Blood eQTLs (Groningen)[63]**:** The blood expression QTL component of the current report was supported by grants from the Celiac Disease Consortium (an innovative cluster approved by the Netherlands Genomics Initiative and partly funded by the Dutch Government (grant BSIK03009), the Netherlands Organization for Scientific Research (NWO-VICI grant 918.66.620, NWO-VENI grant 916.10.135 to L.F.), the Dutch Digestive Disease Foundation (MLDS WO11-30), and a Horizon Breakthrough grant from the Netherlands Genomics Initiative (grant 92519031 to L.F.). This project was supported by the Prinses Beatrix Fonds, VSB fonds, H. Kersten and M. Kersten (Kersten Foundation), The Netherlands ALS Foundation, and J.R. van Dijk and the Adessium Foundation. The research leading to these results has received funding from the European Community's Health Seventh Framework Programme (FP7/2007-2013) under grant agreement 259867.

International Parkinson Disease Genomics Consortium (IPDGC)
North American Brain Expression Consortium (NABEC)
UK Brain Expression Consortium (UKBEC):
Summary-level results for associations between confirmed coffee consumption loci and Parkinson's Disease[64] as well as brain eQTL/mQTL[64, 65] were provided by the IPDGC with additional contributions from NABEC and UKBEC. IPDGC was supported in part by the Intramural Research Programs of the National Institute of Neurological Disorders and Stroke (NINDS), the National Institute on Aging (NIA), and the National Institute of Environmental Health Sciences both part of the National Institutes of Health, Department of Health and Human Services; project numbers Z01-AG000949-02 and Z01-ES101986. Additional support was received from the Department of Defense (award W81XWH-09-2-0128), the Michael J Fox Foundation for Parkinson's Disease Research, National Institutes of Health grants NS057105 and RR024992, American Parkinson Disease Association (APDA); Barnes Jewish Hospital Foundation; Greater St Louis Chapter of the APDA; Hersenstichting Nederland; Neuroscience Campus Amsterdam; and the section of medical genomics, the Prinses Beatrix Fonds. The KORA (Cooperative Research in the Region of Augsburg) research platform was started and financed by the Forschungszentrum für Umwelt und Gesundheit, which is funded by the German Federal Ministry of Education, Science, Research, and Technology and by the State of Bavaria. This study was also funded by the German National Genome Network (NGFNplus number 01GS08134, German Ministry for Education and Research); by the German Federal Ministry of Education and Research (NGFN 01GR0468, PopGen); and 01EW0908 in the frame of ERA-NET NEURON and Helmholtz Alliance Mental Health in an Ageing Society (HA-215), which was funded by the Initiative and Networking Fund of the Helmholtz Association. The French GWAS work was supported by the French National Agency of Research (ANR-08-MNP-012). This study was also sponsored by the Landspitali University Hospital Research Fund (grant to SSv); Icelandic Research Council (grant to SSv); and European Community Framework Programme 7, People Programme, and IAPP on novel genetic and phenotypic markers of Parkinson's disease and Essential Tremor (MarkMD), contract number PIAP-GA-2008-230596 MarkMD (to HP and JHu). We used the Biowulf Linux cluster at the National Institutes of Health, Bethesda, MD, USA, and DNA panels, samples, and clinical data from the National Institute of Neurological Disorders and Stroke Human Genetics Resource Center DNA and Cell Line Repository. People who contributed samples are acknowledged in descriptions of every panel on the repository website. We thank the French Parkinson's Disease Genetics Study Group: Y Agid, M Anheim, A-M Bonnet, M Borg, A Brice, E Broussolle, J-C Corvol, P Damier,

A Destée, A Dürr, F Durif, S Klebe, E Lohmann, M Martinez, P Pollak, O Rascol, F Tison, C Tranchant, M Vérin, F Viallet, and M Vidailhet. We also thank the members of the French 3C Consortium: A Alpérovitch, C Berr, C Tzourio, and P Amouyel for allowing us to use part of the 3C cohort, and D Zelenika for support in generating the genome-wide molecular data. We thank P Tienari (Molecular Neurology Programme, Biomedicum, University of Helsinki), T Peuralinna (Department of Neurology, Helsinki University Central Hospital), L Myllykangas (Folkhalsan Institute of Genetics and Department of Pathology, University of Helsinki), and R Sulkava (Department of Public Health and General Practice Division of Geriatrics, University of Eastern Finland) for the Finnish controls (Vantaa85+ GWAS data). We used genome-wide association data generated by the Wellcome Trust Case-Control Consortium 2 (WTCCC2) from UK patients with Parkinson's disease and UK control individuals from the 1958 Birth Cohort and National Blood Service. Genotyping of UK replication cases on ImmunoChip was part of the WTCCC2 project, which was funded by the Wellcome Trust (083948/Z/07/Z). UK population control data was made available through WTCCC1. This study was supported by the Medical Research Council and Wellcome Trust disease centre (grant WT089698/Z/09/Z to NW, JHa, and ASc). As with previous IPDGC efforts, this study makes use of data generated by the Wellcome Trust Case-Control Consortium. A full list of the investigators who contributed to the generation of the data is available from www.wtccc.org.uk. Funding for the project was provided by the Wellcome Trust under award 076113, 085475 and 090355. This study was also supported by Parkinson's UK (grants 8047 and J-0804) and the Medical Research Council (G0700943). We thank Jeffrey Barrett for assistance with the design of the ImmunoChip. DNA extraction work that was done in the UK was undertaken at University College London Hospitals, University College London, who received a proportion of funding from the Department of Health's National Institute for Health Research Biomedical Research Centres funding. This study was supported in part by the Wellcome Trust/Medical Research Council Joint Call in Neurodegeneration award (WT089698) to the Parkinson's Disease Consortium (UKPDC), whose members are from the UCL Institute of Neurology, University of Sheffield, and the Medical Research Council Protein Phosphorylation Unit at the University of Dundee. We would like to thank the NINDS sponsored Neurogenetics Repository hosted by Coriell Cell Repositories for the use of both case and control samples. NABEC was supported in part by the Intramural Research Program of the National Institute on Aging, National Institutes of Health, part of the US Department of Health and Human Services; project number ZIA AG000932-04. In addition this work was supported by a Research Grant from the Department of Defense, W81XWH-09-2-0128. This study utilized the high-performance computational capabilities of the Biowulf Linux cluster at the National Institutes of Health, Bethesda, Md. (http://biowulf.nih.gov). UKBEC was supported by the MRC through the MRC Sudden Death Brain Bank (C.S.), by a Project Grant (G0901254 to J.H. and M.W.) and by a Fellowship award (G0802462 to M.R.). D.T. was supported by the King Faisal Specialist Hospital and Research Centre, Saudi Arabia. Computing facilities used at King's College London were supported by the National Institute for Health Research (NIHR) Biomedical Research Centre based at Guy's and St Thomas' NHS Foundation Trust and King's College London.

All authors of the current study would like to thank the following consortia/authors for allowing public access to their full GWAS results:

Meta-Analyses of Glucose and Insulin-related traits Consortium (MAGIC): Summary-level results for glycaemic traits[103-106] including HbAlc, fasting glucose, 2 hr glucose challenge, fasting insulin, HOMA-B, HOMA-IR and proinsulin were contributed by MAGIC investigators and were downloaded from www.magicinvestigators.org.

DIAbetes Genetics Replication And Meta-analysis (DIAGRAM) Consortium: Summary-level results for type 2 diabetes[107] were contributed by DIAGRAM+ investigators and were downloaded from http://www.well.ox.ac.uk/DIAGRAM.

Genetic Investigation of ANthropometric Traits (GIANT) consortium: Summary-level results for waist-to-hip ratio[108] and body-mass index[109] were contributed by GIANT investigators and were downloaded from http://www.broadinstitute.org/collaboration/giant/index.php/GIANT_consortium_data_files

Early Growth Genetics (EGG) Consortium: Summary-level results for birth weight[110] were contributed by the EGG Consortium and were downloaded from www.egg-consortium.org

Psychiatry Genomics Consortium Distribution Files (PGC): Summary-level results for Schizophrenia[111], Major Depression Disorder[112], Attention Deficit Hyperactivity Disorder[113] and Bipolar Disorder[114] were contributed by PGC investigators and were downloaded from https://pgc.unc.edu

Tobacco and Genetics Consortium (TAG): Summary-level results for smoking behavior traits[115] were contributed by TAG investigators and were downloaded from https://pgc.unc.edu

International Consortium for Blood Pressure (ICBP): Summary-level results for blood pressure were contributed by ICBP investigators and were downloaded from http://www.georgehretlab.org/icbp_088023401234-9812599.html Beta-coefficients were not made public and were thus obtained via personal communication with Dr. Georg Ehret on behalf of ICBP.

Blood Lipid Traits: Summary-level results for all blood lipid traits were contributed by Teslovich et al, *Nature* (2010)[116] and were downloaded from http://www.sph.umich.edu/csg/abecasis/public/lipids2010 .

5. Author Contributions

Study-specific data analysis were performed by: M.C.C., E.M.B, T.E, M.A.N., A.G., N.P., K.L.M., N.A., K.F., F.R., J.S.N., V.H., A.C., I.M.N, A.T., K.Y., P.M.V., R.R., A.M., M.K.W., J.M.V., J.Z., G.B., J.L., V.M., R.N.L., J.E., S.K.M., T.T., F.G., J. Luan, J.H., R.M., M.D., M.E.G., W.-K.H. and L.X. Study-specific design and management was performed by: M.C.C., E.M.B, T.E, M.A.N., L.M.R., P.K.E.M., N.L.P, D.C., B.A.O., A.H., M.A.I., H.W.T., A.G.U., F.J.A.v.R, I.B., I.J., M.K., L.M., C.P., H.S., R.P.S., S.E.B., R.B., F. Gu, F.B., Z.K., D.R.J., N.G.F., E.M., L.L., C.L., K.M., A. Morris, M.J., K.-T.K., R.N. Leuben, J.J.W., S.M., M.-M.P., M. Kähönen, T.L., J.V., D.M., K. Mukamal, B.M.P., A.D., A.C.H., G.W.M., N.D., T.C., K.L.T., L.F., H.A.B., M.M., J.L.T., D. Mellström, J.J.H., I.P., A. Tonjes, P.D., S.K., M.L., D.K.H., Y.L., J.D., A.R., M.A.M, A.B.Z., M.K.E., D.S., T.B.H., G.D., G.C., M.S., J.B., L.R.P., B.F., S.B., J.M.O., A.T.C, U.P., C.O., C.G., N.G.M., M.W., D.S.S, O.R., J.G.E., P. Mitchell, D.J.H., P.K., E.B.R, D.I.B., I.B.B., R.J.F.L., N.J.W., P.V., N.C., H.J.G., M.L.N., B.H.R.W., F.B.H., E.H., M.-R.J., L.A.C., P.F., P.M.R., C.M.v.D, G.H., A. Metspalu, K.E.N, E.I., J.A.N, R.M.v.D, and D.I.C. Overall study coordination was performed by: M.C.C., N.A., J.A.N., and D.I.C. Meta-analyses were carried out by M.C.C. Conditional meta-analyses were carried out by E.M.B. Blood and brain tissue eQTLs and mQTLs analyses were performed by M.A.N, IPDGC, NABEC, UKBEC, T.E., J.K., H.-J.W. and L. Franke. Results from GWAS of coffee-implicated traits were provided by M.A.N, IPDGC, Z.K., E.M.B., M.J.W., D.R.R., and N.G.M. Bioinformatic analyses were performed by M.C.C. Follow-up analysis in WGHS and TwinGene were performed by D.I.C. and A.G., respectively. M.C.C., N.P., K.L.M., N.C., K.Y., R.M.v.D, and D.I.C. conceived and designed the study. The manuscript draft was prepared by M.C.C. and D.I.C. All authors contributed to and have approved the final manuscript.

6. URLs

R statistical environment, http://cran.r-project.org/ ;
GWAtoolbox, http://cran.r-project.org/web/packages/GWAtoolbox/ ;
Bioconductor, http://www.bioconductor.org/ ;
PLINK, http://pngu.mgh.harvard.edu/~purcell/plink/ ;
MACH 1.0, http://www.sph.umich.edu/csg/abecasis/mach/ ;
International HapMap Project, http://www.hapmap.org/index.html;
METAL, http://www.sph.umich.edu/csg/abecasis/Metal ;
GWAMA, http://www.well.ox.ac.uk/gwama/index.shtml ;
GCTA, http://www.complextraitgenomics.com/software/gcta/ ;
SNAP, http://www.broadinstitute.org/mpg/snap/;
VEGAS, http://gump.qimr.edu.au/VEGAS/;
LocusZoom, http://csg.sph.umich.edu/locuszoom/ ;
eQTL Chicago, http://eqtl.uchicago.edu/cgi-bin/gbrowse/eqtl ;
Genvar, http://www.sanger.ac.uk/resources/software/genevar ;
GRAIL, http://www.broadinstitute.org/mpg/grail/grail.php ;
DAPPLE, http://www.broadinstitute.org/mpg/dapple/dapple.php;
MAGENTA, http://www.broadinstitute.org/mpg/magenta ;
MAGIC, http://www.magicinvestigators.org ;
DIAGRAM, http://www.well.ox.ac.uk/DIAGRAM ;
GIANT,
http://www.broadinstitute.org/collaboration/giant/index.php/GIANT_consortium_data_files;
PGC, https://pgc.unc.edu ;
IBPGC, http://www.georgehretlab.org/icbp_088023401234-9812599.html ;
EGG, http://www.egg-consortium.org ;
NHGRI GWAS catalogue, http://www.genome.gov/gwastudies/ ;
Metabolomics GWAS server, http://mips.helmholtz-muenchen.de/proj/GWAS/gwas/index.php ;
SNIPPER, http://csg.sph.umich.edu/boehnke/snipper;
Mouse Genome Database, http://www.informatics.jax.org;
STITCH (Search Tool for Interactions of Chemicals), http://stitch.embl.de/cgi ;
STRING (Search Tool for the Retrieval of Interacting Genes/Proteins), http://string-db.org ;
GSEA, http://www.broadinstitute.org/gsea ;
UCSC Genome Browser, http://genome.ucsc.edu ;
Cytoscape, http://www.cytoscape.org ;
HaploReg, http://www.broadinstitute.org/mammals/haploreg/haploreg.php ;
RegulomeDB, http://regulome.stanford.edu/ ;
ENCODE, http://genome.ucsc.edu/ENCODE/ ;
Roadmap Epigenomics Project, http://www.roadmapepigenomics.org/ ;
1000 Genome Project, http://www.1000genomes.org/ ;
Blood eQTL browser, http://www.genenetwork.nl/bloodeqtlbrowser/ ;

Supplementary Figures S1-S9

Figure S1. Study Design Overview

Overview of the genome-wide association study of coffee consumption and putative functional and biological follow-up of confirmed loci.

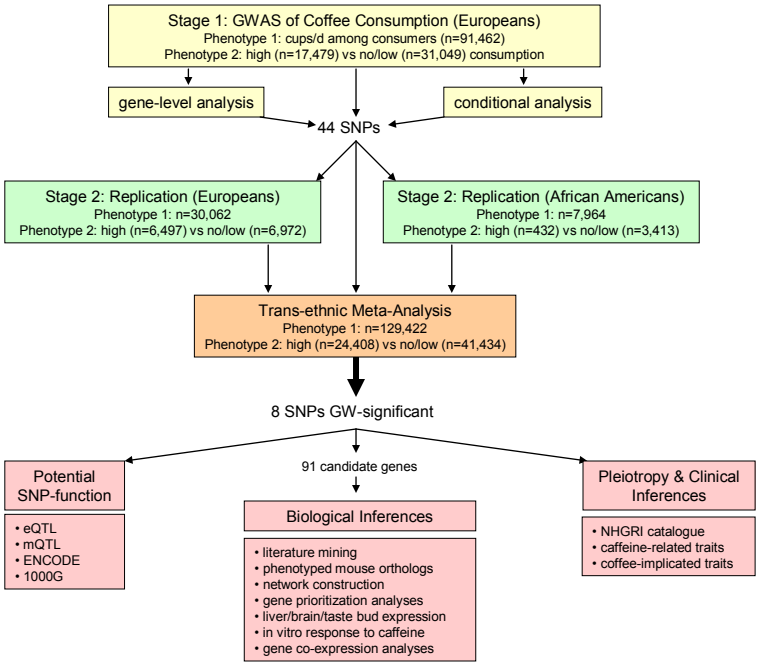


Figure S2. Genome-Wide Meta-Analysis of Coffee Consumption (Phenotype 1)

Manhattan plot for genome-wide meta-analyses of cups of coffee consumed per day among coffee consumers (phenotype 1, n ≤91,462). Data points correspond to $-\log_{10}$ p-values for 2,373,958 SNP-phenotype associations based on results from at least 50% of the maximum sample size and are ordered by chromosomal position. Data points in green span loci achieving genome-wide significance in the trans-ethnic meta-analysis of stage 1 and stage 2 studies (i.e. Table 1).

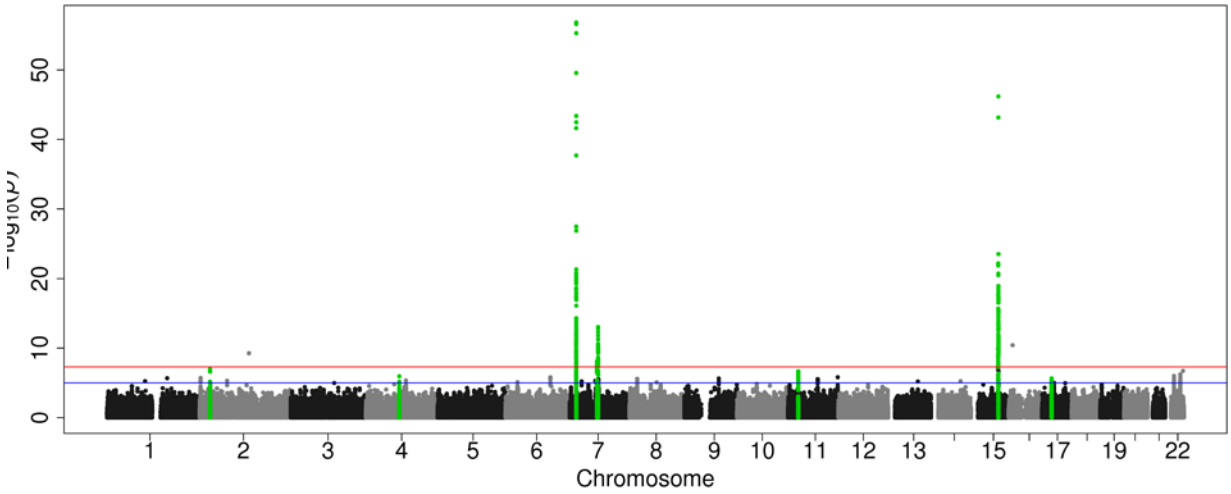
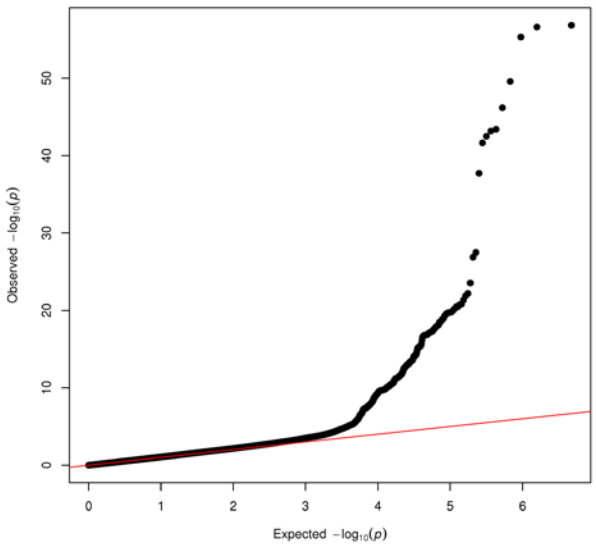


Figure S3. Quantile-Quantile Plot for Genome-Wide Meta-Analysis of Coffee Consumption (Phenotype 1)

Quantile-quantile plot of stage 1 association results from the genome-wide meta analysis of cups of coffee consumed per day among coffee consumers (phenotype 1, n≤91,462). Data points correspond to $-\log_{10}$ p-values for 2,373,958 SNP-phenotype associations based on results from at least 50% of the maximum sample size. Red line (y = x line) indicates instances where the observed (y) p-value is equal to the expected (x) p-value. Little evidence for genomic inflation was observed (λ<1.065).



.

Figure S4. Genome-Wide Meta-Analysis of High Coffee Consumption (Phenotype 2)

Manhattan plot for genome-wide meta-analyses of high vs. no/low coffee consumption (phenotype 2, n=48,528). Data points correspond to $-\log_{10}$ p-values for 2,376,205 SNP-phenotype associations based on results from at least 50% of the maximum sample size and are ordered by chromosomal position. Data points in green span loci achieving genome-wide significance in the trans-ethnic meta-analysis of stage 1 and stage 2 studies for *phenotype 1* (i.e. Table 1).

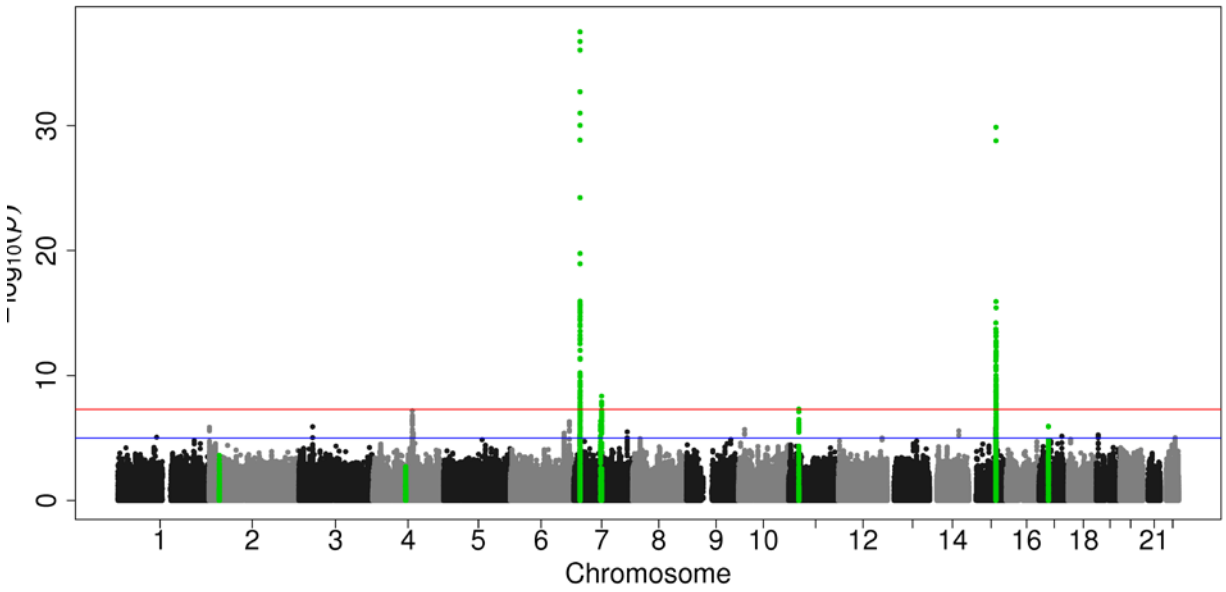


Figure S5. Quantile-Quantile Plot for Genome-Wide Meta-Analysis of High Coffee Consumption (Phenotype 2)

Quantile-quantile plot of stage 1 association results from the genome-wide meta analysis of high vs. no/low coffee consumption (phenotype 2, n=48,528). Data points correspond to $-\log_{10}$ p-values for 2,376,205 SNP-phenotype associations based on results from at least 50% of the maximum sample size. Red line (y = x line) indicates instances where the observed (y) p-value is equal to the expected (x) p-value. Little evidence for genomic inflation was observed ($\lambda<1.056$).

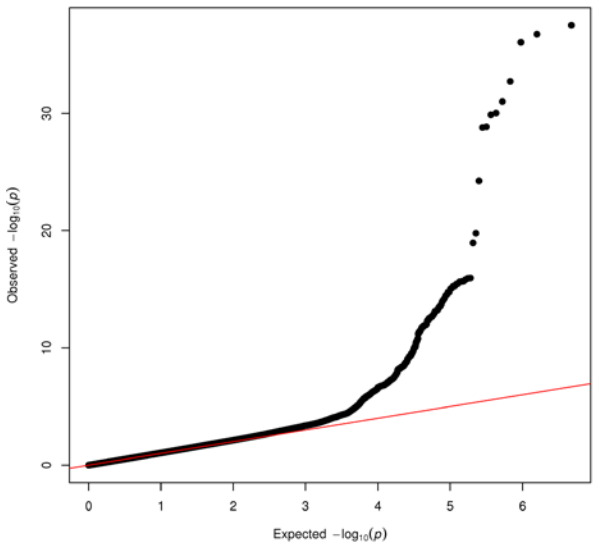
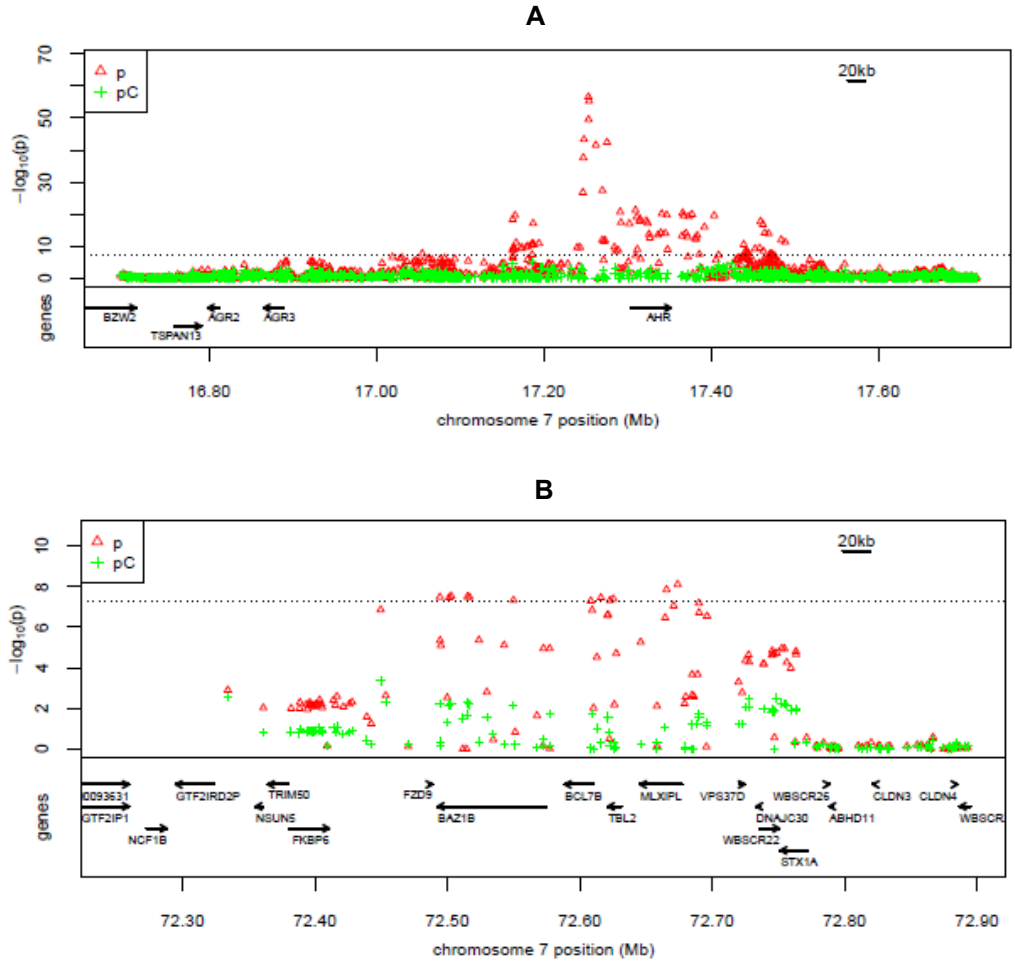
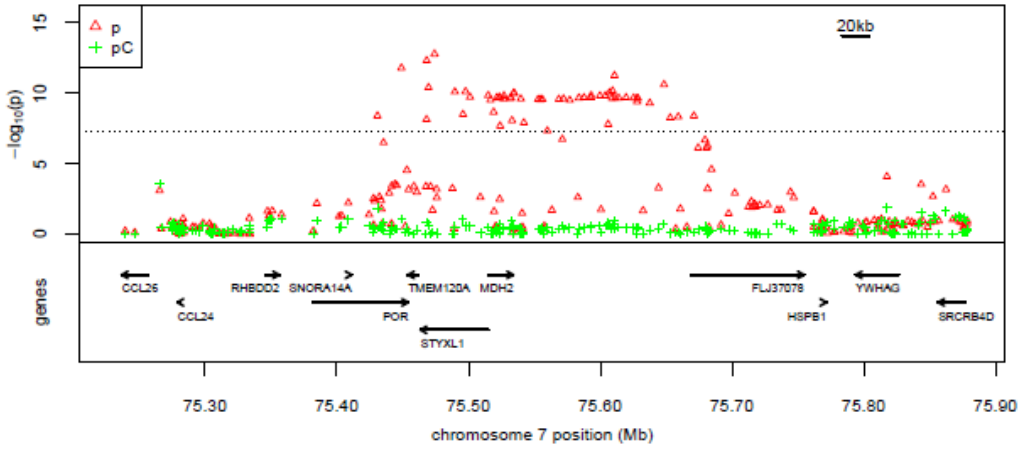


Figure S6. Summary-level Genome-Wide Meta-Analysis Conditioning on Stage 1 Genome-Wide Significant SNPs

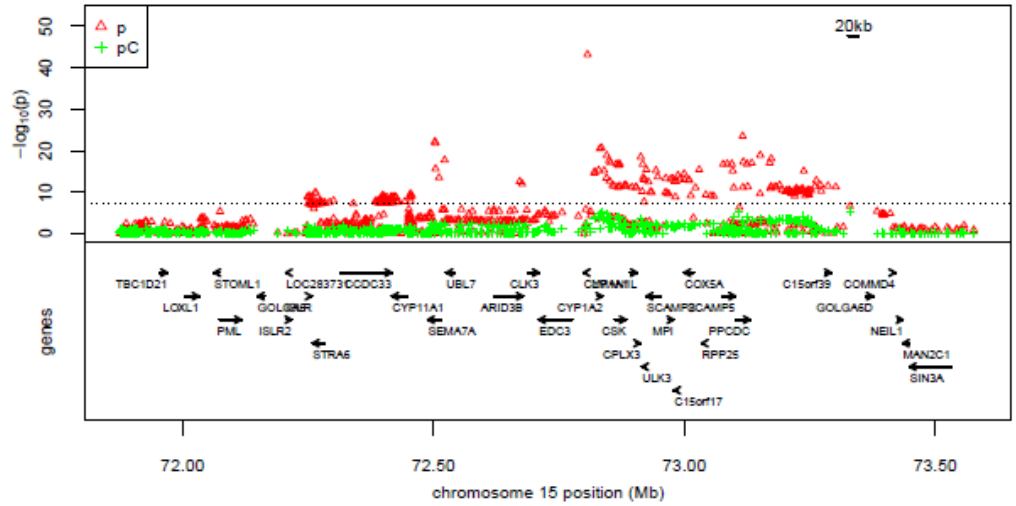
Conditional analysis of coffee consumption (phenotype 1) for regions A) 7p21 conditioning on rs4410790 B) 7q11.23 conditioning on rs7800944 C) 7q11.23 conditioning on rs17685 and D) 15q24 conditioning on rs2472297. Conditional analysis of high coffee consumption (phenotype 2) for regions E) 7p21 conditioning on rs6968554 F) 7q11.23 conditioning on rs17685 and G) 11p13 conditioning on rs6265 and H) 15q24 conditioning on rs2470893. Data points correspond to –log10 pvalues of SNP-phenotype associations before (red triangles) and after (green crosses) conditioning on index SNP.



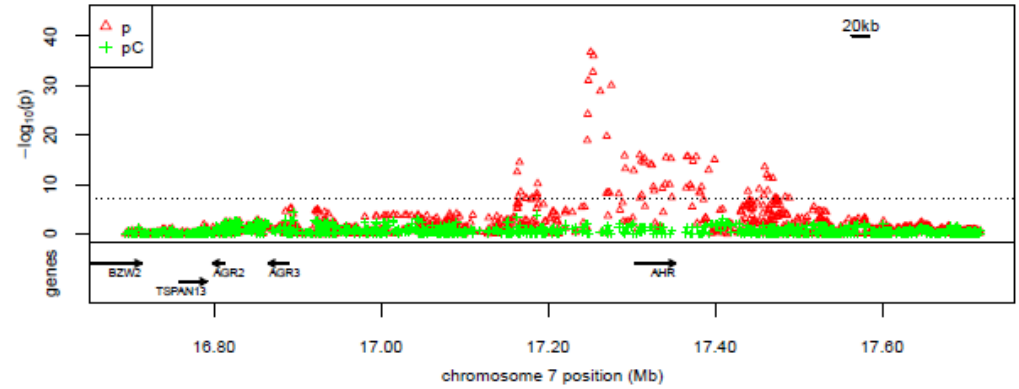
C


p
pC
20kb
−log10(p)
genes
CCL26
CCL24
RHBDD2
SNORA14A
POR
TMEM120A
STYXL1
MDH2
FLJ37078
HSPB1
YWHAG
SRCRB4D
75.30
75.40
75.50
75.60
75.70
75.80
75.90
chromosome 7 position (Mb)

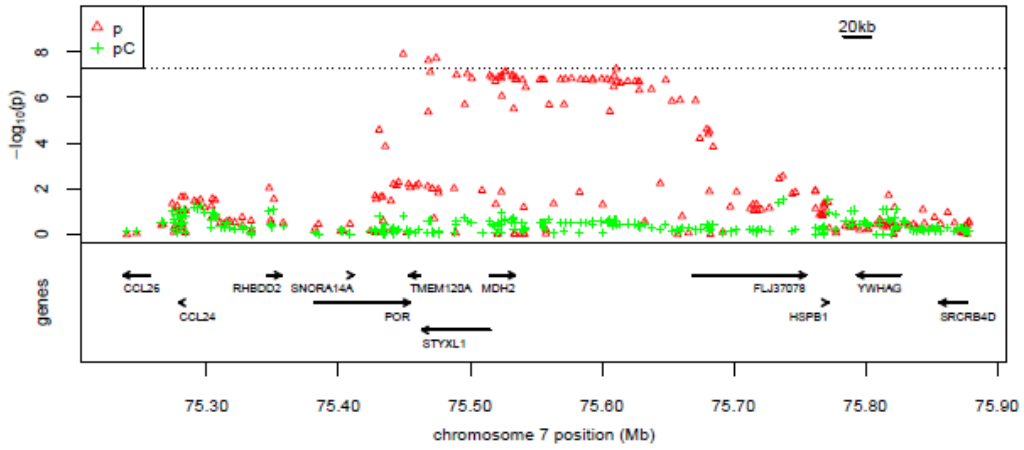
D


p
pC
20kb
−log10(p)
genes
TBC1D21
LOXL1
STOML1
PML
ISLR2
STRA6
LOC283731
CDC33
CYP11A1
SEMA7A
UBL7
ARID3B
CLK3
EDC3
CYP1A2
CSK
CPLX3
ULK3
C15orf17
MPI
COX5A
SCAMP5
PPCDC
RPP25
C15orf39
GOLGA6D
COMMD4
NEIL1
MAN2C1
SIN3A
72.00
72.50
73.00
73.50
chromosome 15 position (Mb)

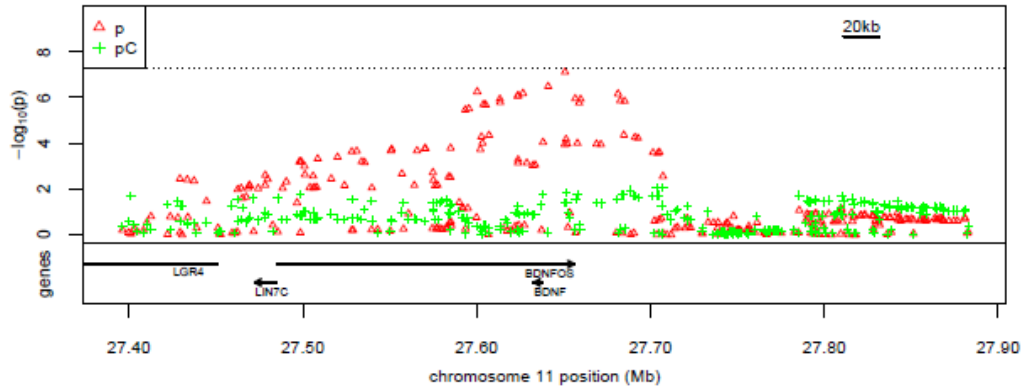
E


p
pC
20kb
−log10(p)
genes
BZW2
TSPAN13
AGR2
AGR3
AHR
16.80
17.00
17.20
17.40
17.60
chromosome 7 position (Mb)

F


p
pC
20kb
−log10(p)
0
2
4
6
8
genes
CCL26
CCL24
RHBDD2
SNORA14A
POR
STYXL1
TMEM120A
MDH2
FLJ37078
HSPB1
YWHAG
SRCRB4D
75.30
75.40
75.50
75.60
75.70
75.80
75.90
chromosome 7 position (Mb)

G


p
pC
20kb
−log10(p)
0
2
4
6
8
genes
LGR4
LIN7C
BDNFOS
BDNF
27.40
27.50
27.60
27.70
27.80
27.90
chromosome 11 position (Mb)

H

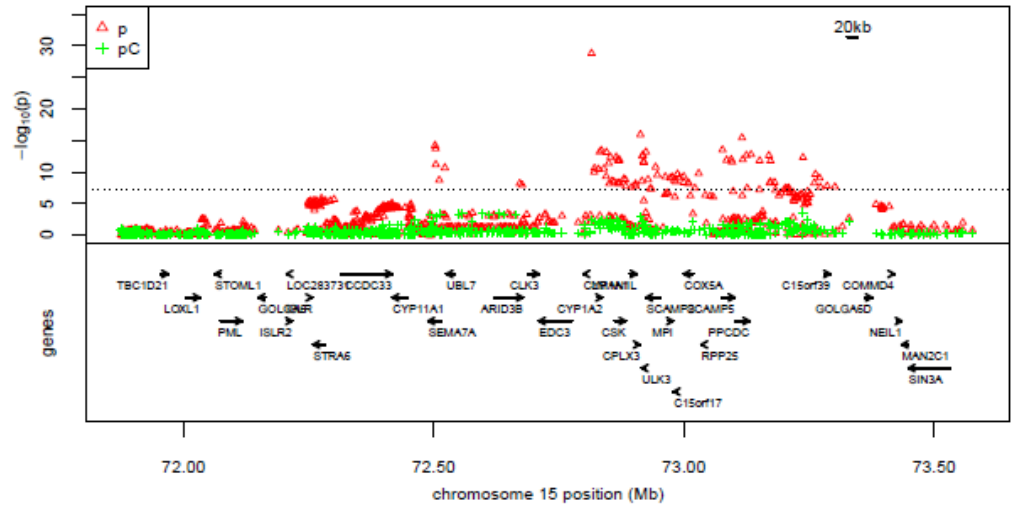
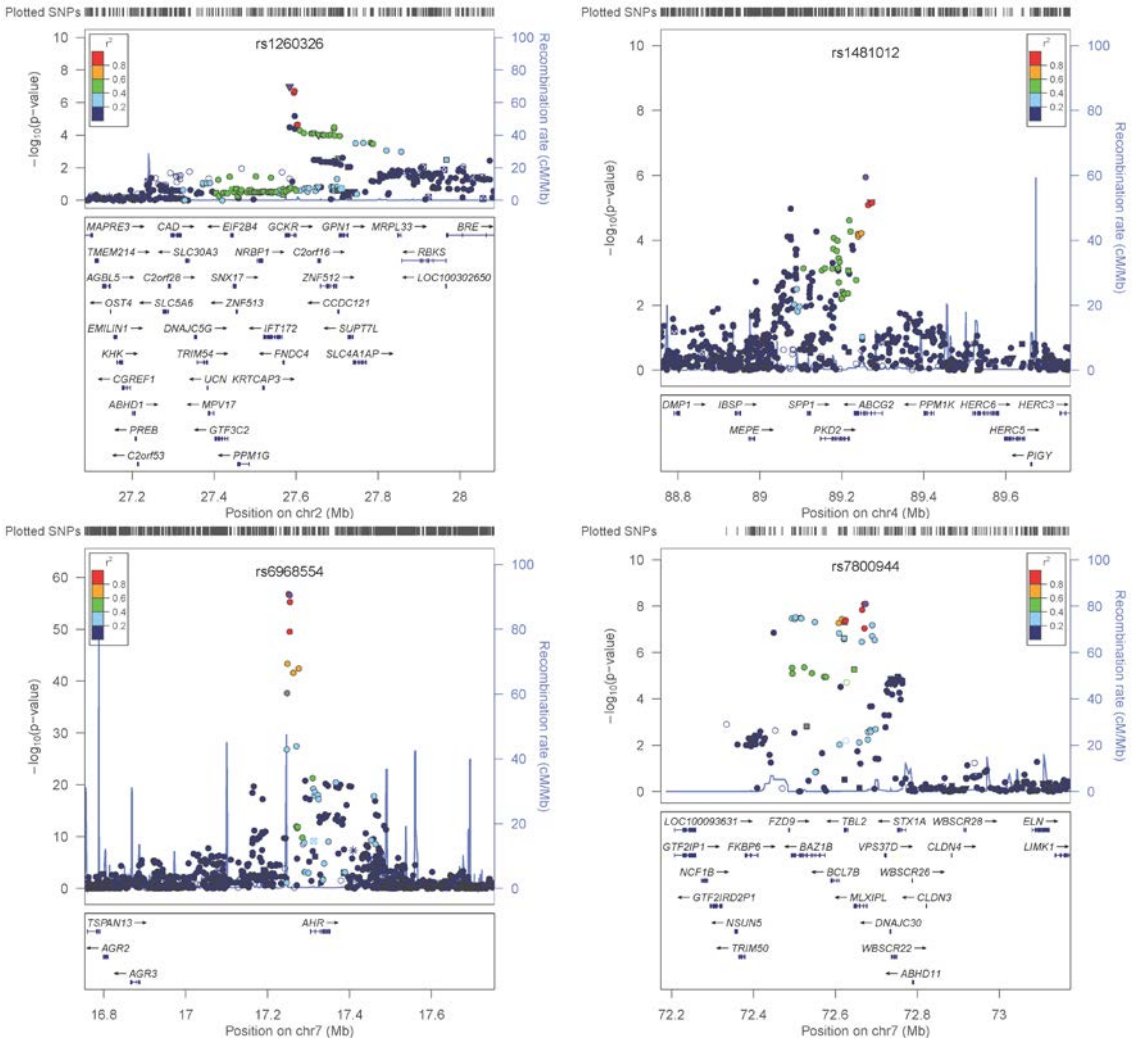
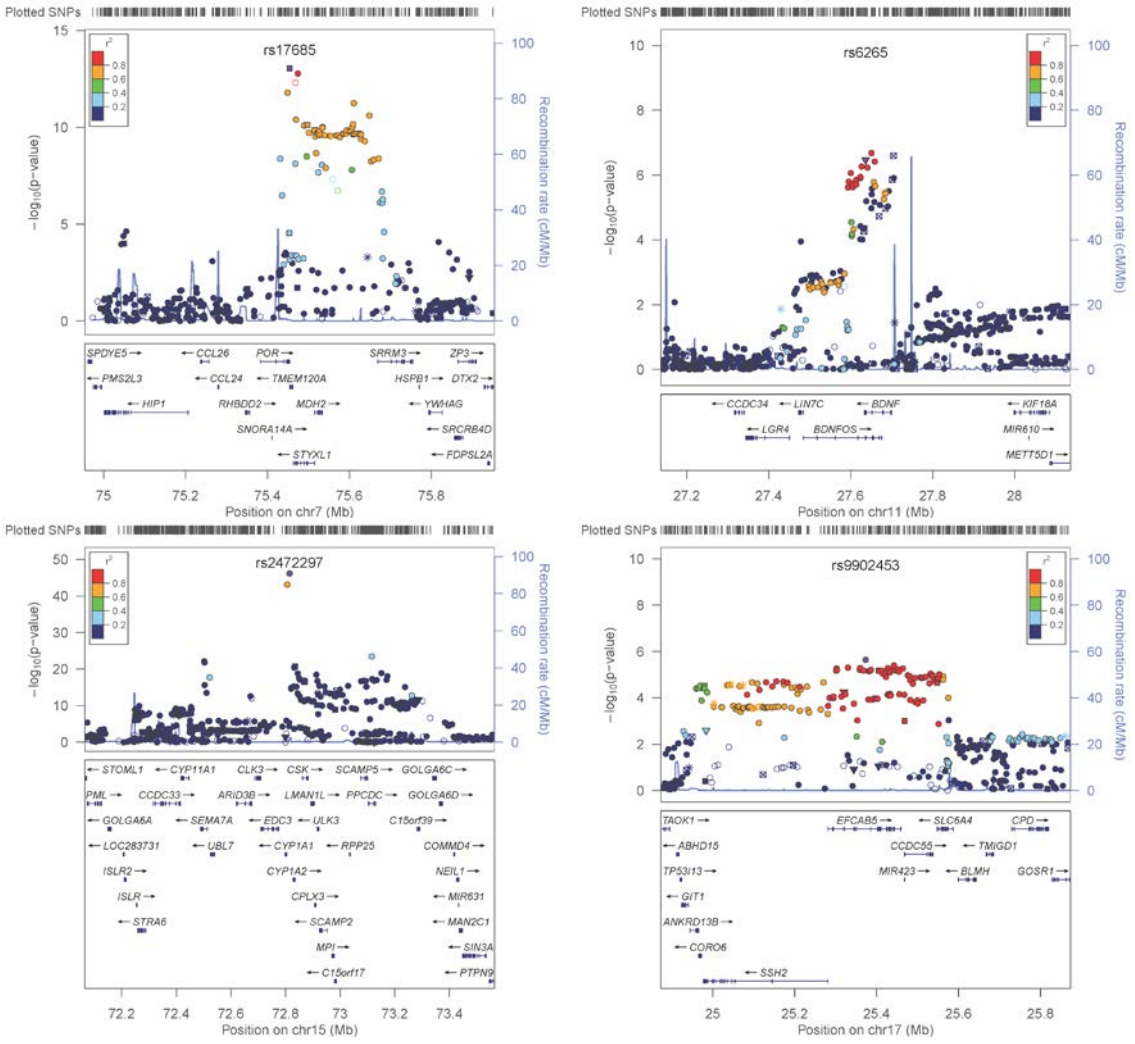

p
pC
20kb
−log10(p)
0
5
10
20
30
genes
TBC1D21
LOXL1
PML
STOML1
ISLR2
STRA6
LOC283731
GOLGA6L6
CDC33
CYP11A1
SEMA7A
UBL7
ARID3B
CLK3
EDC3
CYP1A2
CSK
CPLX3
ULK3
C15orf17
MPI
COX5A
SCAMP5
PPCDC
RPP25
C15orf39
GOLGA6D
COMMD4
NEIL1
MAN2C1
SIN3A
72.00
72.50
73.00
73.50
chromosome 15 position (Mb)

Figure S7. Regional Association Plots of Genome-Wide Significant Coffee Consumption Loci.

In each panel, SNPs are plotted with their stage 1 meta-analysis phenotype 1 $-\log_{10}$ p-values as a function of genomic position (NCBI Build 36). Estimated recombination rates (taken from HapMap CEU) are plotted to reflect the local LD structure. SNP color indicates LD with the index SNP (labeled in purple) according to a scale from r^2=0 to r^2=1 based on pairwise r^2 values from HapMap CEU. Plots were created using LocusZoom (see URLs).




Plotted SNPs
rs17685
rs6265
rs2472297
rs9902453
−log10(p-value)
Recombination rate (cM/Mb)
r²
Position on chr7 (Mb)
Position on chr11 (Mb)
Position on chr15 (Mb)
Position on chr17 (Mb)

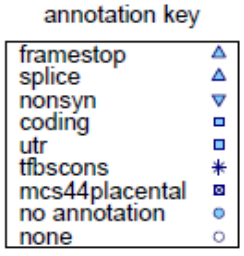
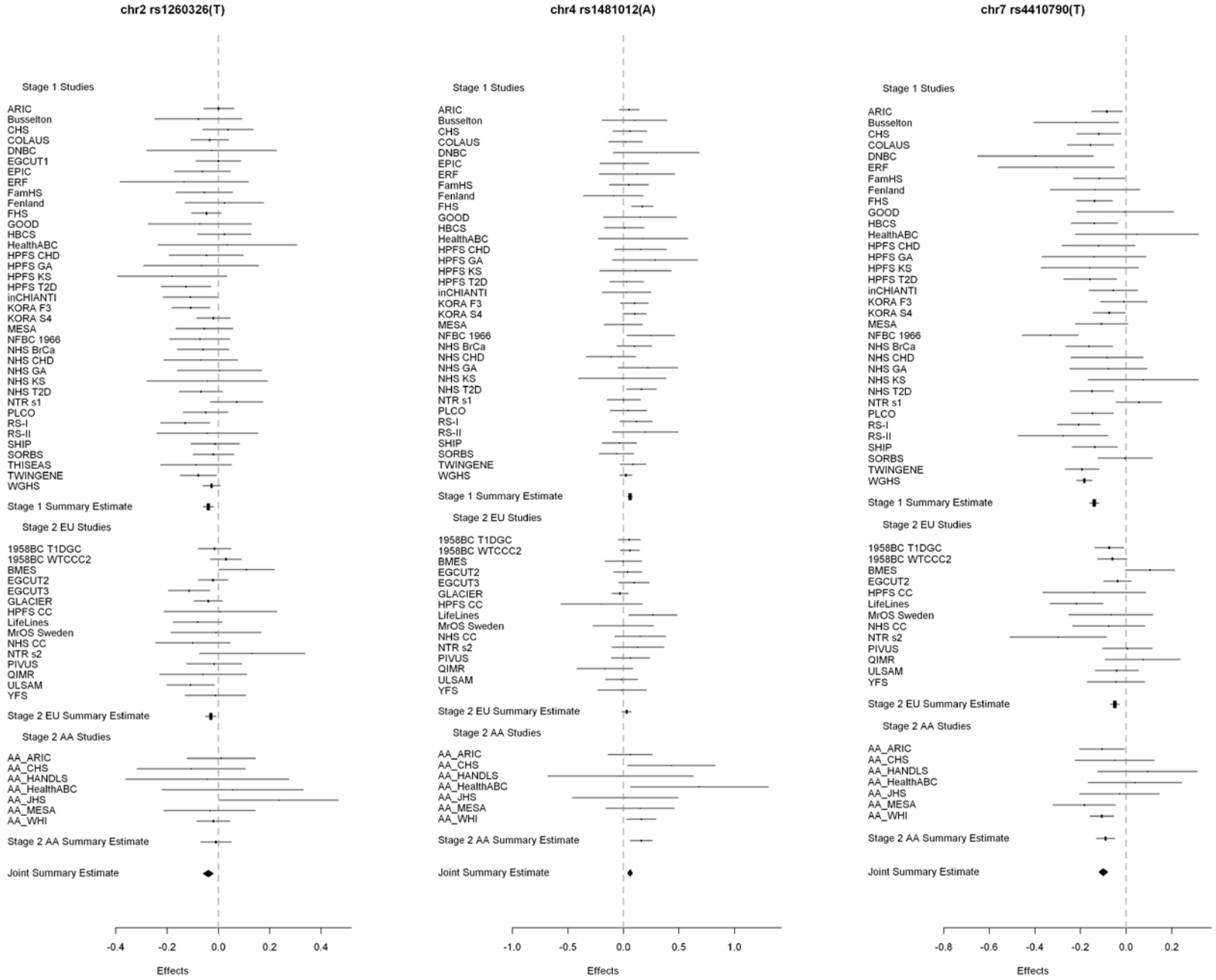
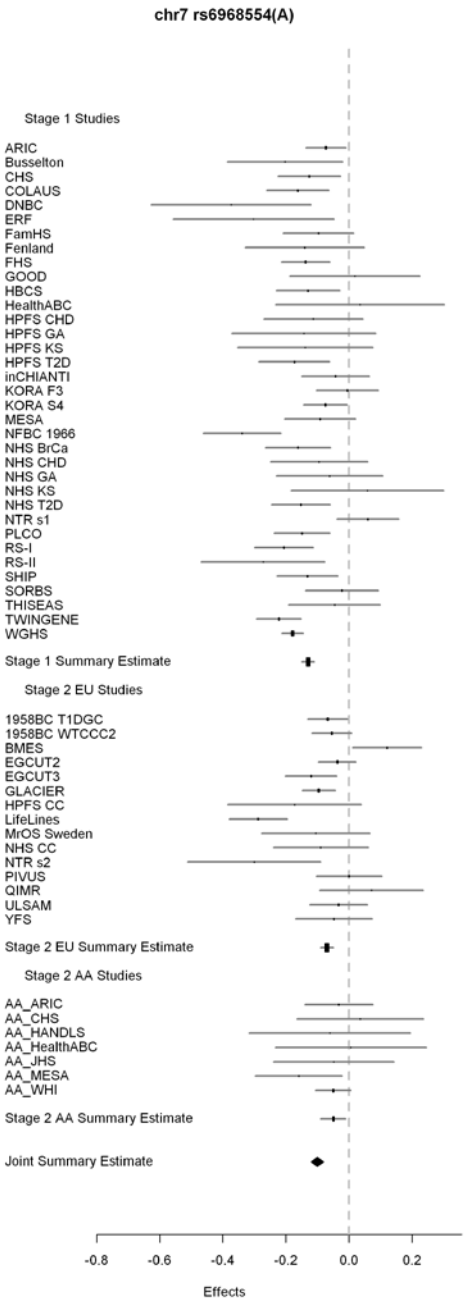
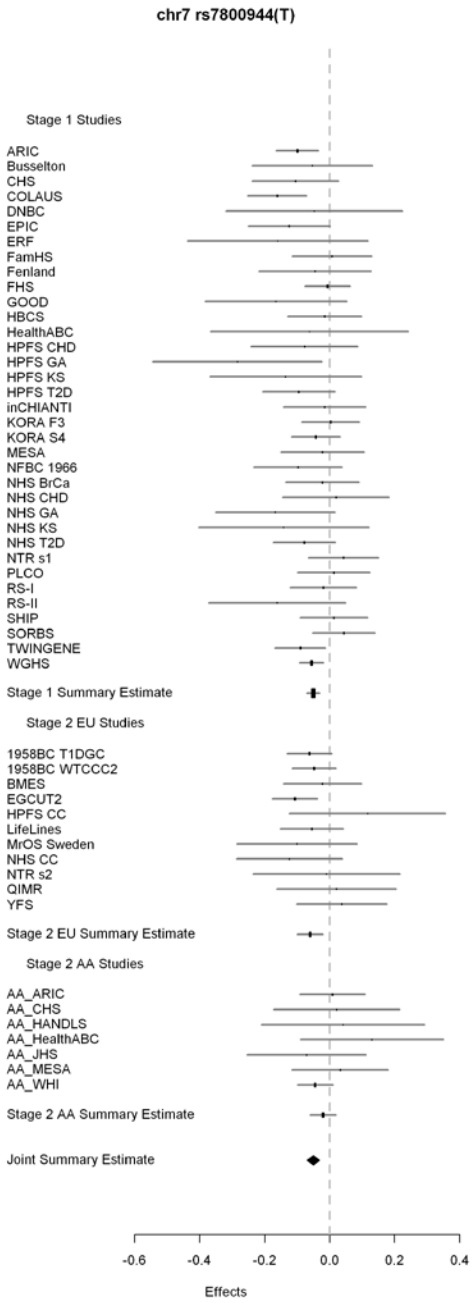
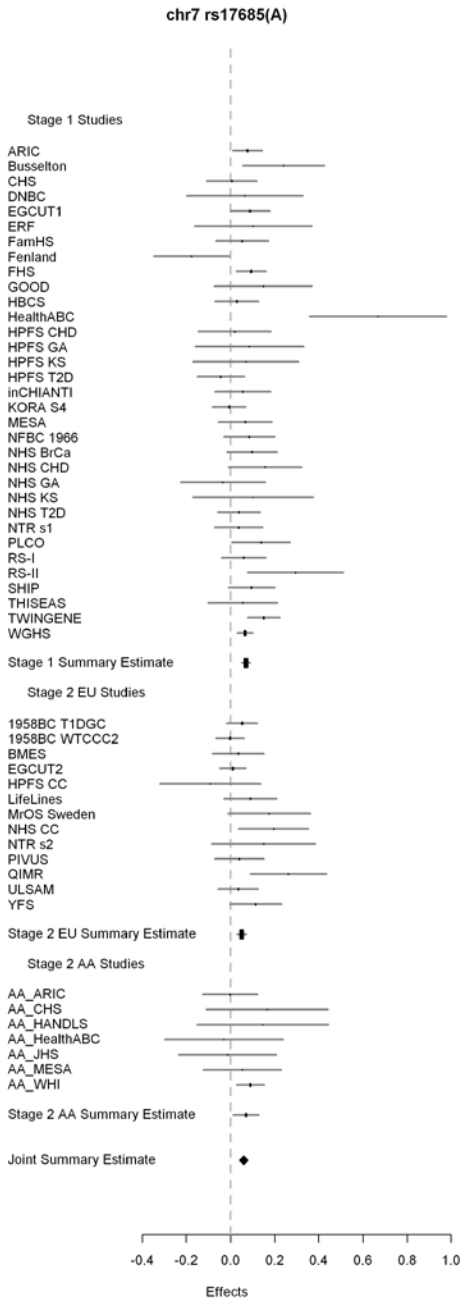

annotation key
framestop
splice
nonsyn
coding
utr
tfbscons
mcs44placental
no annotation
none

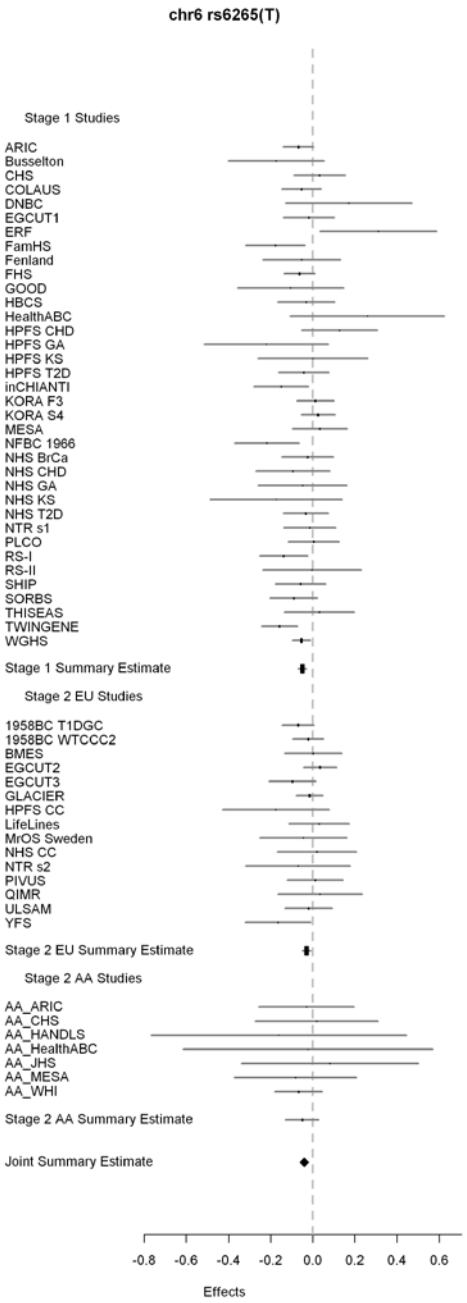
Figure S8. Forest Plots of Genome-Wide Significant Loci Associated With Coffee Consumption (Phenotype 1)

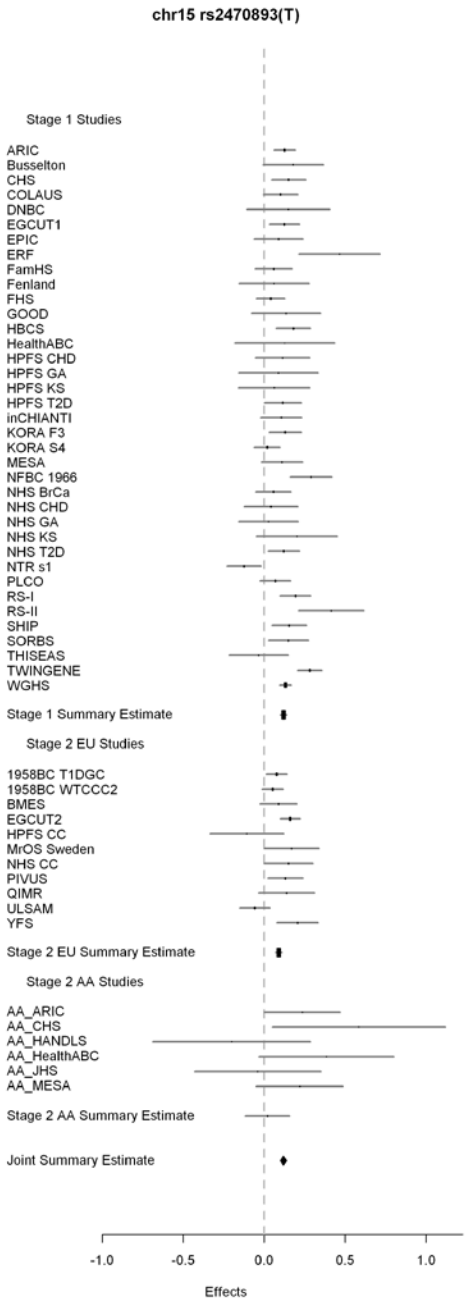


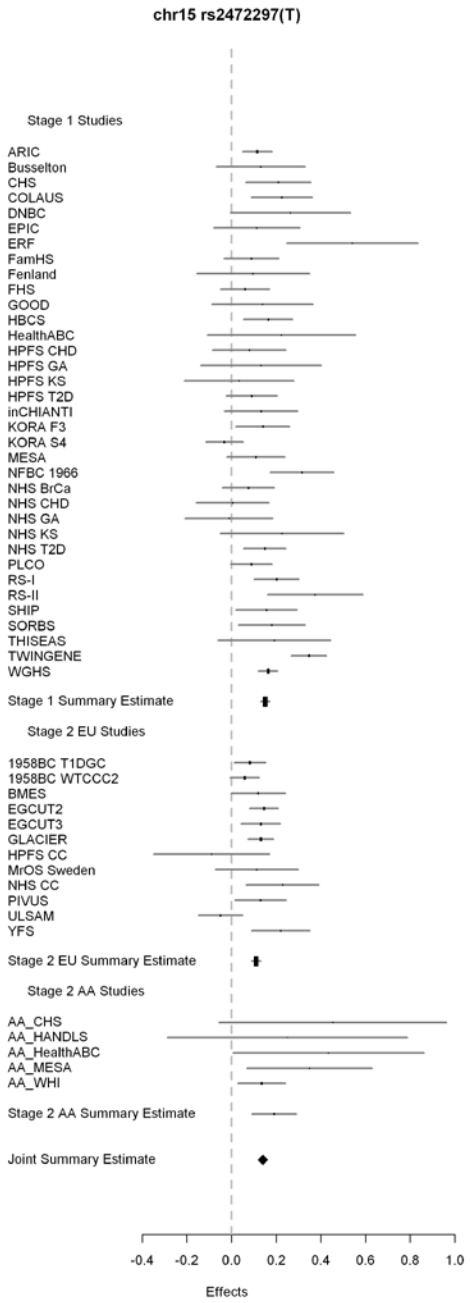

chr7 rs6968554(A)
Stage 1 Studies
ARIC
Busselton
CHS
COLAUS
DNBC
ERF
FamHS
Fenland
FHS
GOOD
HBCS
HealthABC
HPFS CHD
HPFS GA
HPFS KS
HPFS T2D
inCHIANTI
KORA F3
KORA S4
MESA
NFBC 1966
NHS BrCa
NHS CHD
NHS GA
NHS KS
NHS T2D
NTR s1
PLCO
RS-I
RS-II
SHIP
SORBS
THISEAS
TWINGENE
WGHS
Stage 1 Summary Estimate
Stage 2 EU Studies
1958BC T1DGC
1958BC WTCCC2
BMES
EGCUT2
EGCUT3
GLACIER
HPFS CC
LifeLines
MrOS Sweden
NHS CC
NTR s2
PIVUS
QIMR
ULSAM
YFS
Stage 2 EU Summary Estimate
Stage 2 AA Studies
AA_ARIC
AA_CHS
AA_HANDLS
AA_HealthABC
AA_JHS
AA_MESA
AA_WHI
Stage 2 AA Summary Estimate
Joint Summary Estimate
-0.8
-0.6
-0.4
-0.2
0.0
0.2
Effects


chr7 rs7800944(T)
Stage 1 Studies
ARIC
Busselton
CHS
COLAUS
DNBC
EPIC
ERF
FamHS
Fenland
FHS
GOOD
HBCS
HealthABC
HPFS CHD
HPFS GA
HPFS KS
HPFS T2D
inCHIANTI
KORA F3
KORA S4
MESA
NFBC 1966
NHS BrCa
NHS CHD
NHS GA
NHS KS
NHS T2D
NTR s1
PLCO
RS-I
RS-II
SHIP
SORBS
TWINGENE
WGHS
Stage 1 Summary Estimate
Stage 2 EU Studies
1958BC T1DGC
1958BC WTCCC2
BMES
EGCUT2
HPFS CC
LifeLines
MrOS Sweden
NHS CC
NTR s2
QIMR
YFS
Stage 2 EU Summary Estimate
Stage 2 AA Studies
AA_ARIC
AA_CHS
AA_HANDLS
AA_HealthABC
AA_JHS
AA_MESA
AA_WHI
Stage 2 AA Summary Estimate
Joint Summary Estimate
-0.6
-0.4
-0.2
0.0
0.2
0.4
Effects


chr7 rs17685(A)
Stage 1 Studies
ARIC
Busselton
CHS
DNBC
EGCUT1
ERF
FamHS
Fenland
FHS
GOOD
HBCS
HealthABC
HPFS CHD
HPFS GA
HPFS KS
HPFS T2D
inCHIANTI
KORA S4
MESA
NFBC 1966
NHS BrCa
NHS CHD
NHS GA
NHS KS
NHS T2D
NTR s1
PLCO
RS-I
RS-II
SHIP
THISEAS
TWINGENE
WGHS
Stage 1 Summary Estimate
Stage 2 EU Studies
1958BC T1DGC
1958BC WTCCC2
BMES
EGCUT2
HPFS CC
LifeLines
MrOS Sweden
NHS CC
NTR s2
PIVUS
QIMR
ULSAM
YFS
Stage 2 EU Summary Estimate
Stage 2 AA Studies
AA_ARIC
AA_CHS
AA_HANDLS
AA_HealthABC
AA_JHS
AA_MESA
AA_WHI
Stage 2 AA Summary Estimate
Joint Summary Estimate
-0.4
-0.2
0.0
0.2
0.4
0.6
0.8
1.0
Effects


chr6 rs6265(T)
Stage 1 Studies
ARIC
Busselton
CHS
COLAUS
DNBC
EGCUT1
ERF
FamHS
Fenland
FHS
GOOD
HBCS
HealthABC
HPFS CHD
HPFS GA
HPFS KS
HPFS T2D
inCHIANTI
KORA F3
KORA S4
MESA
NFBC 1966
NHS BrCa
NHS CHD
NHS GA
NHS KS
NHS T2D
NTR s1
PLCO
RS-I
RS-II
SHIP
SORBS
THISEAS
TWINGENE
WGHS
Stage 1 Summary Estimate
Stage 2 EU Studies
1958BC T1DGC
1958BC WTCCC2
BMES
EGCUT2
EGCUT3
GLACIER
HPFS CC
LifeLines
MrOS Sweden
NHS CC
NTR s2
PIVUS
QIMR
ULSAM
YFS
Stage 2 EU Summary Estimate
Stage 2 AA Studies
AA_ARIC
AA_CHS
AA_HANDLS
AA_HealthABC
AA_JHS
AA_MESA
AA_WHI
Stage 2 AA Summary Estimate
Joint Summary Estimate
-0.8 -0.6 -0.4 -0.2 0.0 0.2 0.4 0.6
Effects


chr15 rs2470893(T)
Stage 1 Studies
ARIC
Busselton
CHS
COLAUS
DNBC
EGCUT1
EPIC
ERF
FamHS
Fenland
FHS
GOOD
HBCS
HealthABC
HPFS CHD
HPFS GA
HPFS KS
HPFS T2D
inCHIANTI
KORA F3
KORA S4
MESA
NFBC 1966
NHS BrCa
NHS CHD
NHS GA
NHS KS
NHS T2D
NTR s1
PLCO
RS-I
RS-II
SHIP
SORBS
THISEAS
TWINGENE
WGHS
Stage 1 Summary Estimate
Stage 2 EU Studies
1958BC T1DGC
1958BC WTCCC2
BMES
EGCUT2
HPFS CC
MrOS Sweden
NHS CC
PIVUS
QIMR
ULSAM
YFS
Stage 2 EU Summary Estimate
Stage 2 AA Studies
AA_ARIC
AA_CHS
AA_HANDLS
AA_HealthABC
AA_JHS
AA_MESA
Stage 2 AA Summary Estimate
Joint Summary Estimate
-1.0 -0.5 0.0 0.5 1.0
Effects


chr15 rs2472297(T)
Stage 1 Studies
ARIC
Busselton
CHS
COLAUS
DNBC
EPIC
ERF
FamHS
Fenland
FHS
GOOD
HBCS
HealthABC
HPFS CHD
HPFS GA
HPFS KS
HPFS T2D
inCHIANTI
KORA F3
KORA S4
MESA
NFBC 1966
NHS BrCa
NHS CHD
NHS GA
NHS KS
NHS T2D
PLCO
RS-I
RS-II
SHIP
SORBS
THISEAS
TWINGENE
WGHS
Stage 1 Summary Estimate
Stage 2 EU Studies
1958BC T1DGC
1958BC WTCCC2
BMES
EGCUT2
EGCUT3
GLACIER
HPFS CC
MrOS Sweden
NHS CC
PIVUS
ULSAM
YFS
Stage 2 EU Summary Estimate
Stage 2 AA Studies
AA_CHS
AA_HANDLS
AA_HealthABC
AA_MESA
AA_WHI
Stage 2 AA Summary Estimate
Joint Summary Estimate
-0.4 -0.2 0.0 0.2 0.4 0.6 0.8 1.0
Effects

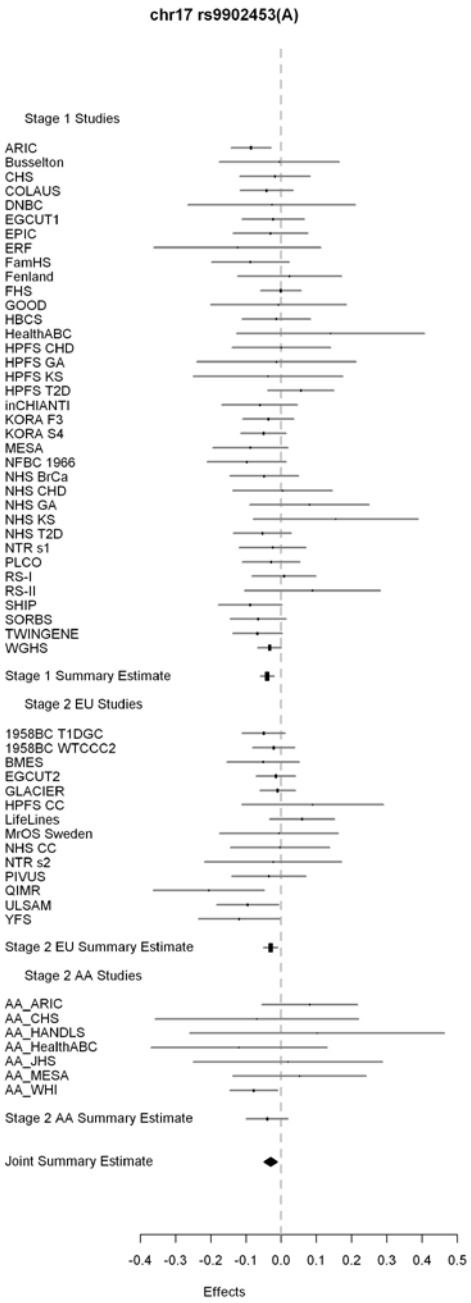
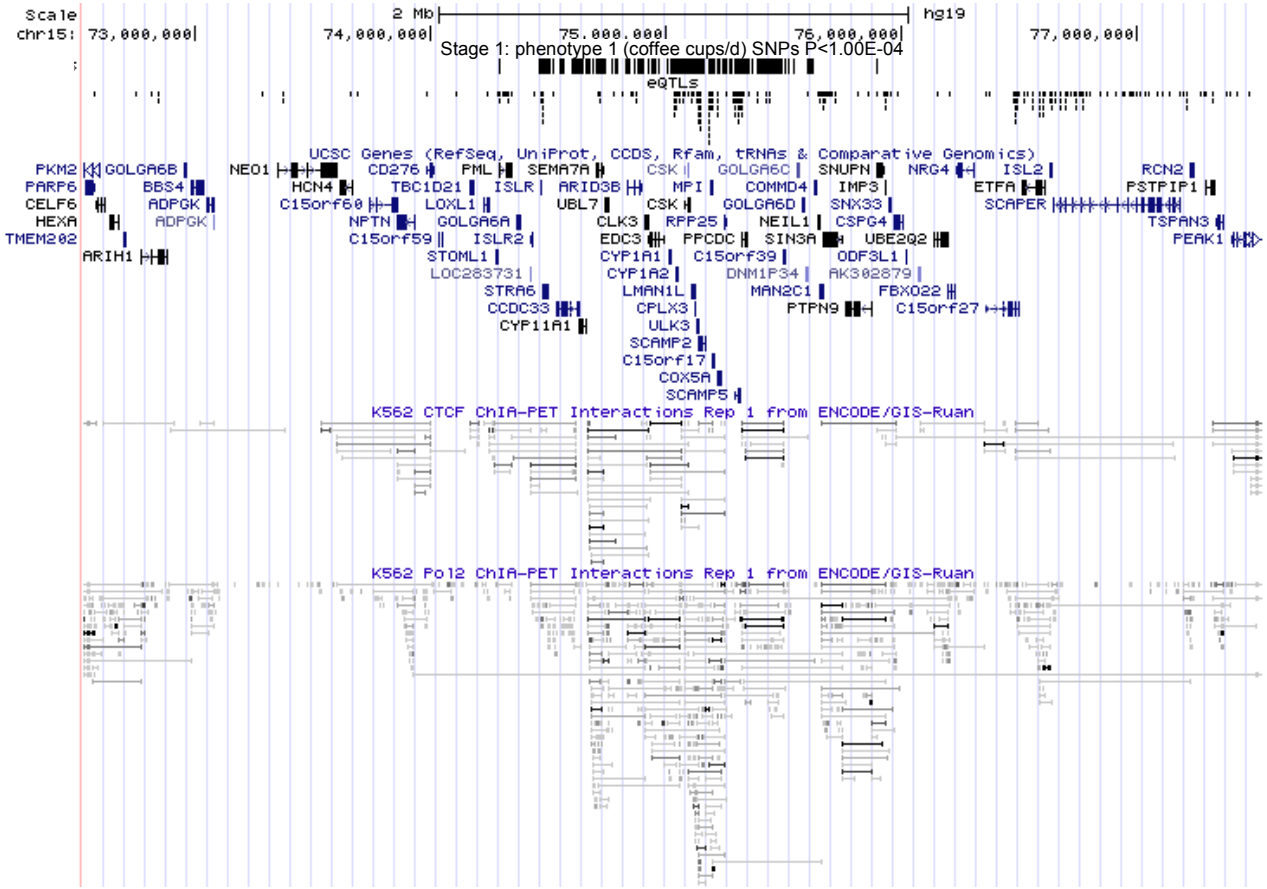

chr17 rs9902453(A)
Stage 1 Studies
ARIC
Busselton
CHS
COLAUS
DNBC
EGCUT1
EPIC
ERF
FamHS
Fenland
FHS
GOOD
HBCS
HealthABC
HPFS CHD
HPFS GA
HPFS KS
HPFS T2D
inCHIANTI
KORA F3
KORA S4
MESA
NFBC 1966
NHS BrCa
NHS CHD
NHS GA
NHS KS
NHS T2D
NTR s1
PLCO
RS-I
RS-II
SHIP
SORBS
TWINGENE
WGHS
Stage 1 Summary Estimate
Stage 2 EU Studies
1958BC T1DGC
1958BC WTCCC2
BMES
EGCUT2
GLACIER
HPFS CC
LifeLines
MrOS Sweden
NHS CC
NTR s2
PIVUS
QIMR
ULSAM
YFS
Stage 2 EU Summary Estimate
Stage 2 AA Studies
AA_ARIC
AA_CHS
AA_HANDLS
AA_HealthABC
AA_JHS
AA_MESA
AA_WHI
Stage 2 AA Summary Estimate
Joint Summary Estimate
-0.4 -0.3 -0.2 -0.1 0.0 0.1 0.2 0.3 0.4 0.5
Effects

Figure S9. Long-Range Chromatin Interactions Spanning 15q24 Locus

UCSC Genome browser displaying 15q24 locus (Hg19, track 1) and location of CTCF (track 5, top) and Pol12 (track 5, bottom) mediated chromatin interactions determined by Chromatin Interaction Analysis with Paired-End Tag (ChIA-PET) data[117] extracted from the K562 (chronic myeloid leukemia, tier 1) cell line. This track was produced as part of the ENCODE project. A chromatin interaction is defined as the association of two regions of the genome that are far apart in terms of genomic distance, but are spatially proximate to each other in the 3-dimensional cellular nucleus. PETs are represented by two blocks one for each end. These blocks are connected by a horizontal line if both ends are in the same chromosome. If the two ends are on different chromosomes, only one block will display. PET sequences that overlap at both ends form PET clusters. The number of PETs in a cluster reflects the strength of a chromatin interaction. PET clusters of more than 3 PETs could indicate genuine chromatin interactions. Track 2: packed view of stage 1 SNPs associated with coffee consumption (phenotype 1, p<1.0E-04). Track 3: packed view of eQTLs (eQTL Chicago). Track 4: packed view of UCSC genes.



Supplementary Tables S1-S33

Table S1. Study-specific design and self-reported measures of coffee consumption

Study	Study Abbreviation	Design	Country	Questionnaire Detail	Date of collection	Type of coffee[a]	Ref.
STAGE 1 / DISCOVERY							
Atherosclerosis Risk in Communities Study	ARIC	cohort	United States	Average coffee intake (1 cup) over the past year: (choice response) almost never 1-3 cups/month 1 cup/week 2-4 cups/week 5-6 cups/week 1 cup/day 2-3 cups/day 4-5 cups/day 6+ cups/day	1987-1989	regular coffee only	[1, 18, 94, 95]
Busselton Health Study	Busselton	nest case-control, asthma	Australia	How many of the following do you drink per day? (line items) cups of coffee/day (integer response)	1994-1995	all coffee	[3]
Cardiovascular Health Study	CHS	Cohort	United States.	Average coffee intake (8 oz) over the past year: (choice response) never 1-3 cups/month 1 cup/week 2-4 cups/week 5-6 cups/week 1 cup/day 2-3 cups/day 4-5 cups/day 6+ cups/day	1995-1996	Separate line items for regular and decaffeinated coffee	[18, 118]
Cohorte Lausannoise	COLAUS	Population-based	Switzerland	Average coffee intake over the past 4 weeks: (choice response)	2009-2011	Regular coffee only	[5]

				never 1 cup/month 2-3 cups/month 1-2 cups/week 3-4 cups/week 1 cup/day 2+ cups/day			
Danish National Birth Cohort: Preterm Birth Study	DNBC	Nested case-control	Denmark	Cups of regular coffee/day (integer response)	1998-2002	Regular coffee only	[6]
Estonian Genome Center of the University of Tartu	EGCUT1	Population-based	Estonia	Cups of coffee/day (integer response)	2003-2009	All coffee	[8]
European Prospective Investigation in Cancer and Nutrition Norfolk Study	EPIC-Norfolk	Case-cohort	United Kingdom	Average coffee intake (190 ml) over last 12 months (choice response) never or <1 cup/month 1-3 cups/month 1 cup/week 2-4 cups/week 5-6 cups/week 1 cup/day 2-3 cups/day 4-5 cups/day 6+ cups/day	1993-1997	Separate line items for regular and decaffeinated coffee	[119]
Erasmus Rucphen Family study	ERF	Family-based	Netherlands	Cups of coffee/day (integer response)	1995	All coffee	[10]
Family Heart Study	FamHS	Family-based	United States	In past year, how often consume coffee, not decaffeinated (1 cup) (choice response) never 1-3 cups/month 1 cup/week 2-4 cups/week 5-6 cups/week 1 cup/day 2-3 cups/day 4-5 cups/day 6+ cups/day	1992-1995	All coffee	[11, 18]
Fenland Study	Fenland	Population-based	United Kingdom	Average coffee intake (190 ml) over last 12 months (choice response) never or <1 cup/month	2005-2013	Separate line items for regular and decaffeinated coffee	[120, 121]

				1-3 cups/month 1 cup/week 2-4 cups/week 5-6 cups/week 1 cup/day 2-3 cups/day 4-5 cups/day 6+ cups/day			
Framingham Heart Study	FHS	cohort	United States	Average coffee intake (8oz) over last 6 months (choice response) never 1-3 cups/month 1 cup/week 2-4 cups/week 5-6 cups/week 1 cup/day 2-3 cups/day 4-5 cups/day 6+ cups/day	offspring exam5: 1991-1994 gen3 exam1: 2002 - 2005	Separate line items for regular and decaffeinated coffee	[18, 122-124]
Gothenburg Osteoporosis and Obesity Determinants Study	GOOD	Cohort	Sweden	Questionnaire: "Do you drink coffee?" "If yes, How many cups per day?"	2008-2009	All coffee	[125]
Helsinki Birth Cohort Study	HBCS	Cohort	Finland	Average coffee intake over the past year (choice response) never 1-3 cups/month 1 cup/week 2-4 cups/week 5-6 cups/week 1 cup/day 2-3 cups/day 4-5 cups/day 6+ cups/day	2001-2003	All coffee	[97, 98, 126]
Health, Aging and Body Composition Study	HealthABC	Cohort	United States	How many cups of coffee, regular or decaf? (usual intake over the past year) (choice response) Never or 1-11 per year 1-3 cups/month 1 cups/week	1998-1999	All coffee	[127]

				2-4 cups/week 5-6 cups/week 1 cup/day 2-3 cups/day 4 cups/day 5+ cups/day			
Health Professionals Follow-up Study	HPFS CHD	Nested case-control, heart disease	United States	Average coffee intake (1 cup) over the past year (choice response) never 1-3 cups/month 1 cup/week 2-4 cups/week 5-6 cups/week 1 cup/day 2-3 cups/day 4-5 cups/day 6+ cups/day	Mean of 1986 and 1990	Separate line items for regular and decaffeinated coffee	[18, 90]
	HPFS GA	Nested case-control, open-angle glaucoma					
	HPFS KS	Nested case-control, kidney stone disease					
	HPFS T2D	Nested case-control, type 2 diabetes					
Invecchiare in Chianti	inCHIANTI	Cohort	Italy	Cups of coffee/day (integer response)	1998-2000	Separate line items for regular and decaffeinated coffee	[12, 13]
Cooperative Health Research in the Augsburg Region	KORA_F3	Population-based, cross-sectional	Germany	Cups of coffee/day (integer response)	2004-2005	All coffee	[128]
	KORA_S4	Population-based, cross-sectional	Germany	Cups of coffee/day (integer response)	1999-2001	All coffee	[128]
The Multi-Ethnic Study of Atherosclerosis	MESA	Cohort	United States	Average coffee intake (not including latte, café au lait) Serving size captured as small, medium, large (weighted as 0.5, 1, 1.5) (choice response) Rare or never 1-3 cups/month 1 cup/week 2-4 cups/week 5-6 cups/week 1 cup/day 2-3 cups/day 4-5 cups/day 6+ cups/day	2000-2002	All coffee	[14, 129, 130]
Northern Finland Birth	NFBC 1966	Cohort	Finland	How many cups of coffee do you	1997	All coffee	[15]

Cohort 1966				usually drink in a day? filtered coffee: integer response boiled coffee: integer response			
Nurses' Health Study	NHS BrCa	Nested case-control, breast cancer	United States.	Average coffee intake (1 cup) over the past year (choice response) never 1-3 cups/month 1 cup/week 2-4 cups/week 5-6 cups/week 1 cup/day 2-3 cups/day 4-5 cups/day 6+ cups/day	Mean of 1984 and 1986	Separate line items for regular and decaffeinated coffee	[18, 90]
	NHS CHD	Nested case-control, heart disease					
	NHS GA	Nested case-control, open-angle glaucoma					
	NHS KS	Nested case-control, kidney stone disease					
	NHS T2D	Nested case-control, type 2 diabetes					
Netherlands Twin Register (Stage 1)	NTR s1	Population-based	Netherlands	Cups of regular coffee/day (integer response)	2000	Regular coffee only	[20, 131]
Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial	PLCO	Nested case-control	United States.	Cups of coffee/day (integer response)	1992-2001	Separate line items for regular and decaffeinated coffee	[132, 133]
Rotterdam Study	RS-I	Cohort	Netherlands	Cups of coffee/day (integer response)	1990 – 1993	All coffee	[24]
	RS-II	Cohort	Netherlands	Cups of coffee/day (integer response)	2000 – 2001	All coffee	
Study of Health in Pomerania	SHIP	Population-based	Germany/Poland	Cups of coffee/day (integer response)	1997-2001	Separate line items for regular and decaffeinated coffee	[134]
Sorbs	SORBS	Population based	Germany	Do you regularly drink coffee? Yes/no If yes, how many cups do you drink on average per day? (choice response) 1 cup/day 2 cups/day 3-4 cups/day > 4 cups/day	2005-2008	All coffee	[25, 91]
The Hellenic study of Interactions between SNPs & Eating in Atherosclerosis Susceptibility	THISEAS	Nested case-controls	Greece	Average cups of regular coffee over last year (choice response)	2006-2009	Regular coffee only	n/a

				never 1-3 cups/ month 1-2 cups/ week 3-4 cups/ week 5-6 cups/ week How many cups per day? (integer response)			
TwinGene	TWINGENE	Twin cohort	Sweden	How many cups of coffee do you usually drink a day? (choice response) never 1 cup/day 2 cups/day 3 cups/day 4 cups/day 5 cups/day (if more than 5 provide precise amount) don't know refuse	1998-2003	All coffee	[135]
Women's Genome Health Study	WGHS	Randomized trial of aspirin and vitamin E	United States.	Average coffee intake (8 oz) over the past year (choice response) never 1-3 cups/month 1 cup/week 2-4 cups/week 5-6 cups/week 1 cup/day 2-3 cups/day 4-5 cups/day 6+ cups/day	1993	Separate line items for regular and decaffeinated coffee	[18, 90]
STAGE 2 / REPLICATION							
Blue Mountain Eye Study	BMES	Cohort	Australia	Average coffee intake during the past year (choice response) never <1 cups/month 1-3 cup/month	1997-2000	Separate line items for regular and decaffeinated coffee	[18, 136, 137]

				1 cup/week 2-4 cups /week 5-6 cups /week 1 cup/day 2-3 cups/day 4+/day			
British Birth Cohort 1958	1958BC T1DGC 1958BC WTCCC2	Cohort	United Kingdom	How often do you drink coffee? (choice response) > 4 times a day 2-4 times a day once a day 3-6 days a week 1 or 2 days a week less than 1 day a week occasionally never	2002-2003	All coffee	[138, 139]
Estonian Genome Center of the University of Tartu	EGCUT2	Population-based	Estonia	Cups of coffee/day (integer response)	2003-2012	All coffee	[8]
	EGCUT3	Population-based	Estonia	Cups of coffee/day (integer response)	2003-2010	All coffee	[8]
Gene-Lifestyle interactions And Complex traits Involved in Elevated disease Risk Study	GLACIER	Population-based, prospective, cohort	Sweden	For boiled and filtered coffee respectively; Average consumption during the previous year: (choice response) Never Occasionally 1-3/month 1/week 2-3/week 4-6/week 1/day 2-3/day ≥4/day	1990-2007	Boiled and filtered coffee, with no distinction between regular and decaffeinated.	[140-142]
Health Professionals Follow-up Study	HPFS CC	Nested case-control, colon cancer	United States	Average coffee intake (1 cup) over the past year (choice response) never 1-3 cups/month 1 cup/week 2-4 cups/week 5-6 cups/week	Mean of 1986 and 1990	Separate line items for regular and decaffeinated coffee	[18, 90]

				1 cup/day 2-3 cups/day 4-5 cups/day 6+ cups/day			
Life Lines	LifeLines	Population-based	Netherlands	'How often did you drink coffee during the recent month? Please consider all types of coffee including decaffeinated coffee' 'How many cups did you drink on a regular day? (integer response)	2007-2011	All coffee	[30]
Osteoporotic Fractures in Men Sweden	MrOS Sweden	Cohort	Sweden	Questionnaire: "Do you drink coffee?" "If yes, How many cups per day?" (integer response)	2008-2009	All coffee	[143]
Netherlands Twin Register (Stage 2)	NTR s2	Population-based	Netherlands	Cups of regular coffee/day (integer response)	2000 and 2012	Regular coffee only	[20, 131]
Nurses' Health Study	NHS CC	Nested case-control, colon cancer	United States	Average coffee intake (1 cup) over the past year (choice response) never 1-3 cups/month 1 cup/week 2-4 cups/week 5-6 cups/week 1 cup/day 2-3 cups/day 4-5 cups/day 6+ cups/day	Mean of 1984 and 1986	Separate line items for regular and decaffeinated coffee	[18, 90]
Prospective Investigation of the Vasculature in Uppsala Seniors	PIVUS	Cohort	Sweden	7-day dietary records: Coffee intake (cups) was recorded 6 times daily (breakfast, lunch, supper, between meals, and in the evening). (integer response)	2001-2004	All coffee	[144, 145]
Queensland Institute of Medical Research	QIMR	Twin Cohorts	Australia	'On average, how many cups of coffee would you drink?' (integer response)	1980-1982	All coffee	[93, 146]
Uppsala Longitudinal Study of Adult Men	ULSAM	Cohort	Sweden	7-day dietary records: Coffee intake (cups) was recorded 6 times daily (breakfast, lunch, supper, between meals, and in the evening).	1990-1994	All coffee	[145, 147]

				(integer response)			
Cardiovascular Risk in Young Finns	YFS	Cohort	Finland	Average coffee intake (1 cup) over the past year (choice response) Never or less frequently 1-3 cups/month 1 cup/week 2-4 cups/week 5-6 cups/week 1 cup/day 2-3 cups/day 4-5 cups/day 6+ cups/day	2007	All coffee	[148]
Atherosclerosis Risk in Communities (African Americans) Study	AA_ARIC	Cohort	United States	Average coffee intake (1 cup) over the past year (choice response) almost never 1-3 cups/month 1 cup/week 2-4 cups/week 5-6 cups/week 1 cup/day 2-3 cups/day 4-5 cups/day 6+ cups/day	visit 1: 1987-1989 (data used)	Regular coffee only	[1, 18]
Cardiovascular Health Study (African Americans)	AA_CHS	Cohort	United States.	Average coffee intake (8 oz) over the past year (choice response) never 1-3 cups/month 1 cup/week 2-4 cups/week 5-6 cups/week 1 cup/day 2-3 cups/day 4-5 cups/day 6+ cups/day	1995-1996	Separate line items for regular and decaffeinated coffee	[18, 118]
Health, Aging in Neighborhoods of Diversity across the Life Span (African Americans)	AA_HANDLS	Cohort	United States	How many cups of coffee, regular or decaf? (usual intake over the past year) (choice response)	2010	All coffee	[127]

				never or 1-11 per year 1-3 cups/month 1 cups/week 2-4 cups/week 5-6 cups/week 1 cup/day 2-3 cups/day 4 cups/day 5+ cups/day			
Health Aging and Body Composition Study (African Americans)	AA_HealthABC	Cohort	United States	How many cups of coffee, regular or decaf? (usual intake over the past year) (choice response) Never or 1-11 per year 1-3 cups/month 1 cups/week 2-4 cups/week 5-6 cups/week 1 cup/day 2-3 cups/day 4 cups/day 5+ cups/day	1998-1999	All coffee	[127]
Jackson Heart Study (African Americans)	AA_JHS	Cohort	United States	How often, on the average do you drink coffee? (usual intake over the past year) (choice response) Never < once/month 1-3 times/month 1 time/week 2-4 times/week 5-6 times /week 1 time/day 2-3 times/day 4-5 times/day 6+ times/day When you drink coffee, your portion is usually closest to… (choice response) Small (8 fl. oz) Medium (16 fl oz)	2000-2004	Separate line items for regular and decaffeinated coffee	[149]

				Large (24 fl oz) X large (32 fl oz)			
The Multi-Ethnic Study of Atherosclerosis (African Americans)	AA_MESA	Cohort	United States	Average coffee intake (not including latte, café au lait) Serving size captured as small, medium, large (weighted as 0.5, 1, 1.5) (choice response) Rare or never 1-3 cups/month 1 cup/week 2-4 cups/week 5-6 cups/week 1 cup/day 2-3 cups/day 4-5 cups/day 6+ cups/day	2000-2002	All coffee	[14, 129, 130]
Women's Health Initiative SNP Health Association Resource (African Americans)	AA_WHI	Nested case-control	United States	Lifestyle Questionnaire Do you drink coffee each day, Y/N How many cups of regular coffee do you drink each day?	1991	Regular coffee only	[150]
EXPLORATORY							
Pakistan Risk of Myocardial Infarction Study	PROMIS	Case-control	Pakistan	Average coffee intake (1 cup) over the past year (choice response) >4 cups/day 2-4 cups/day 1cup/day >3 cups/week 2-3 cups/week 1cup/week <1 cup/week but ≥1 cup/month 1 cup/month <1 cup/month or occasionally in ramazan only none	2005-2011	All coffee	[89]

n/a, not available

[a]For most European and Australian populations, the majority of coffee consumed is of the regular type.

Table S2. Study characteristics for phenotype 1: cups of coffee consumed per day among coffee consumers

Study	N	Ancestry	Age, Years (SD)	Female, %	Current Smokers, %	Coffee Consumption, cups/day	
						Mean (SD)	range
STAGE 1							
ARIC	6982	EUR	54.3(5.7)	51	27	2.6 (1.9)	0.07-6
Busselton	839	EUR	53.1 (16.6)	57	14	2.8 (1.8)	1-10
CHS	1321	EUR	77.2 (4.6)	60	10	1.7 (1.3)	0.07-6.0
COLAUS	3135	EUR	58.0 (10.5)	54	21	2.2 (1.4)	0.03-6
DNBC	996	EUR	30.6 (4.2)	100	27	3.4 (2.9)	0.14-12.0
EGCUT1	2911	EUR	47.0 (21.8)	59	27	2.2 (1.5)	1.0-14.0
EPIC-Norfolk	1656	EUR	59.3 (0.0)	53	12	1.7 (1.7)	1.0-6.0
ERF	1920	EUR	49.5 (13.4)	53	42	5.8 (3.8)	1.0-40.0
FamHS	2103	EUR	53.3 (13.3)	50	21	2.4 (1.8)	0.07-6.0
Fenland	985	EUR	45.1 (7.2)	52	11	2.0 (1.7)	0.07-6.0
FHS	5576	EUR	47.9 (11.7)	51	19	2.0 (1.4)	0.07-6.0
GOOD	472	EUR	24.0 (0.62)	0	8	2.5 (1.6)	1-11
HBCS	1614	EUR	61.5 (2.9)	58	25	2.5 (1.5)	0.07-12.0
HealthABC	208	EUR	75.8 (2.9)	50	5	1.8 (1.5)	0.0-5.0
HPFS CHD	831	EUR	56.8 (8.7)	0	11	1.8 (1.5)	0.04-6.0
HPFS GA	337	EUR	58.0 (8.2)	0	5	1.7 (1.5)	0.04-6.0
HPFS KS	386	EUR	49.3 (6.8)	0	8	1.7 (1.5)	0.04-6.0
HPFS T2D	1848	EUR	55.5 (8.3)	0	9	1.8 (1.5)	0.04-6.0
inCHIANTI	1141	EUR	68.4 (15.3)	55	19	2.5 (1.3)	0.03-11.3
KORA_F3	1444	EUR	62.5 (10.1)	51	12	2.8 (2.4)	0.5-30.0
KORA_S4	1530	EUR	53.9 (8.9)	51	17	3.3 (2.9)	0.5-24.0
MESA	1952	EUR	63.1 (10.3)	52	12	2.1 (1.8)	0.04-9.0
NFBC 1966	4704	EUR	31.0 (0.0)	51	36	5.1 (3.0)	1.0-30.0
NHS BrCa	1640	EUR	52.8 (6.4)	100	17	2.0 (1.5)	0.04-6.0
NHS CHD	917	EUR	54.2 (6.5)	100	33	2.3 (1.6)	0.04-6.0
NHS GA	592	EUR	53.4 (6.2)	100	18	2.1 (1.5)	0.04-6.0
NHS KS	355	EUR	49.0 (6.7)	100	18	2.0 (1.6)	0.04-6.0
NHS T2D	2450	EUR	51.8 (6.7)	100	16	2.1 (1.5)	0.04-6.0
NTR s1	866	EUR	33.6 (11.8)	65	28	3.9 (2.5)	1.1-20.0
PLCO	3371	EUR	67.4 (5.4)	23	26	2.3 (1.7)	0.01-7.0
RS1	4478	EUR	67.5 (7.6)	58	11	4.0 (1.8)	0.3-20.0
RS2	1896	EUR	64.8 (8.0)	54	19	4.5 (2.5)	1.0-34.0
SHIP	3385	EUR	50.0 (15.7)	51	33	3.3 (2.0)	1.0-20.0
SORBS	745	EUR	49.2 (15.6)	60	13	1.8 (0.8)	1-4
THISEAS	555	EUR	57.4 (13.1)	38	34	1.8 (1.2)	0.07-10.0
TWINGENE	9028	EUR	59.5 (8.4)	53	18	4.1 (2.4)	0.05-25.0
WGHS	17332	EUR	54.8 (7.1)	100	13	2.4 (1.6)	0.07-6.0
STAGE 2							
BMES	1635	EUR	66.6 (8.8)	57	7	2.3 (1.5)	0.07-4.0
1958BC T1DGC	2086	EUR	45	51	24	2.2 (1.5)	0.03-4.0
1958BC WTCCC2	2249	EUR	45	47	25	2.2 (1.5)	0.04-4.0
EGCUT2	4434	EUR	56.6 (19.7)	52	28	2.0 (1.4)	1-30
EGCUT3	1947	EUR	58.1 (11.4)	64	26	2.1 (1.3)	1-12

GLACIER	5742	EUR	49.5 (8.6)	60	23	3.0 (1.4)	0.003-8.0
HPFS CC	342	EUR	58.4 (8.7)	0	5	1.7 (1.4)	0.04-6.0
LifeLines	4103	EUR	45.5 (8.3)	56	26	4.0 (2.2)	0.1-11.2
MrOS Sweden	881	EUR	75.3 (3.2)	0	8	3.0 (1.8)	0.5-15.0
NHS CC	896	EUR	54.0 (6.4)	100	18	2.2 (1.6)	0.04-6.0
NTR s2	909	EUR	37.9 (15.6)	61	21	3.8 (2.3)	1.0-15.0
PIVUS	761	EUR	70.2 (0.17)	50	11	3.2 (1.4)	0.07-6.0
QIMR	1445	EUR	32.4 (11.3)	69	50	3.2 (2.3)	1.0-22.0
ULSAM	1043	EUR	70.9 (0.63)	0	17	3.3 (1.4)	0.07-6.0
YFS	1599	EUR	37.7 (5.0)	55	23	3.3 (1.7)	0.07-10.5
AA_ARIC	1709	AA	53.3 (5.7)	62	33	1.6 (1.3)	0.07-6.0
AA_CHS	238	AA	75.3 (5.0)	61	19	1.2 (1.1)	0.07-6.0
AA_HANDLS	275	AA	50.4 (8.4)	57	47	1.7 (1.4)	0.25-14.0
AA_HealthABC	170	AA	75.4 (2.9)	50	12	0.86 (0.97)	1.0-2.0
AA_JHS	1212	AA	49.9 (12.1)	61	15	1.43 (2.13)	0.02-24.0
AA_MESA	916	AA	62.7 (9.8)	55	21	1.2 (1.4)	0.04-9.0
AA_WHI	3604	AA	61.5 (7.0)	100	17	1.9 (1.0)	1.0-6.0
EXPLORATORY							
PROMIS	365	IN	51.7 (10.3)	14	31	0.51 (0.87)	0.03-4.0

EUR, European; AA, African American; IN, Indian

Table S3. Study characteristics for phenotype 2: high vs no/low coffee consumption[a,b]

Study	Ancestry	Age, Years (SD)	Female, %	Current Smokers, %	Low Intake		High Intake	
					N	Cut-off	N	Cut-off
STAGE 1								
ARIC	EUR	54.0 (5.7)	54	27	3361	<1	1966	≥4
Busselton	EUR	52.5 (17.2)	58	13	463	<1	228	≥4
COLAUS	EUR	57.6 (10.7)	52	20	343	<1	112	≥6
DNBC	EUR	30.3 (4.2)	100	28	279	<2	211	≥6
EGCUT1	EUR	40.5 (19.9)	48	36	552	0	371	≥4
EPIC-Norfolk	EUR	58.6 (9.0)	53	13	831	<1	348	≥4
ERF	EUR	49.5 (12.5)	50	51	115	<2	855	≥6
FamHS	EUR	50.9 (13.7)	53	12	1920	<1	476	≥4
Fenland	EUR	45.1 (7.3)	59	12	591	<1	166	≥4
FHS	EUR	47.6 (12.0)	48	23	1171	<1	690	≥4
GOOD	EUR	24.0 (0.65)	0	8	141	≤1	96	≥4
HBCS	EUR	61.8 (3.0)	55	21	528	<2	71	≥6
HPFS CHD	EUR	56.7 (8.7)	0	8	570	<1	88	≥4
HPFS GA	EUR	58.5 (8.2)	0	4	214	<1	30	≥4
HPFS KS	EUR	48.7 (6.7)	0	6	307	<1	38	≥4
HPFS T2D	EUR	55.6 (8.7)	0	6	1158	<1	180	≥4
inCHIANTI	EUR	61.9 (18.2)	53	29	86	<1	179	≥4
KORA_F3	EUR	62.5 (10.1)	51	12	844	<2	158	≥6
KORA_S4	EUR	53.9 (8.9)	51	17	431	<2	325	≥6
MESA	EUR	62.5 (10.1)	52	11	890	<1	268	≥4
NFBC 1966	EUR	31.0 (0.0)	46	40	751	<2	1621	≥6
NHS BrCa	EUR	52.8 (6.5)	100	14	845	<1	212	≥4
NHS CHD	EUR	54.6 (6.3)	100	29	393	<1	161	≥4
NHS GA	EUR	53.1 (6.1)	100	16	269	<1	77	≥4
NHS KS	EUR	48.3 (6.7)	100	15	224	<1	54	≥4
NHS T2D	EUR	51.7 (6.8)	100	13	1321	<1	317	≥4
NTR s1	EUR	30.5 (11.1)	73	21	496	<2	231	≥6
PLCO	EUR	67.1 (5.6)	22	30	645	<1	638	≥4
RS1	EUR	64.8 (7.0)	38	42	99	<2	655	≥6
RS2	EUR	63.1 (7.3)	49	25	66	<2	569	≥6
SHIP	EUR	46.7 (17.4)	46	34	1049	<2	344	≥6
SORBS	EUR	44.3 (17.2)	56	13	162	0	371	≥4
THISEAS	EUR	55.7 (14.1)	42	25	215	<1	53	≥4
TWINGENE	EUR	58.8 (8.3)	46	24	1194	<2	1718	≥6
WGHS	EUR	54.4 (7.0)	100	13	8525	<1	3602	≥4
STAGE 2								
BMES	EUR	66.7 (8.8)	59	9	846	<1	602	≥4
1958BC_T1DGC	EUR	45	51	28	791	<1	439	≥4
1958BC_WTCCC2	EUR	45	48	27	810	<1	490	≥4
EGCUT2	EUR	60.0 (11.4)	42	36	864	0	521	≥4
EGCUT3	EUR	51.3 (21.0)	51	35	316	0	235	≥4
GLACIER	EUR	48.6 (8.4)	57	29	481	<1	1861	≥4
LifeLines	EUR	45.6 (8.2)	43	36	259	<1	636	≥6

MrOS Sweden	EUR	75.1 (3.2)	0	10	198	<2	78	≥6
NTR s2	EUR	30.8 (13.2)	71	19	536	<2	189	≥6
NHS CC	EUR	53.9 (6.9)	100	19	441	<1	134	≥4
PIVUS	EUR	70.2 (0.16)	46	8	167	<2	40	≥6
QIMR	EUR	31.7 (11.1)	68	49	732	<1	559	≥4
ULSAM	EUR	70.9 (0.71)	0	17	190	<2	63	≥6
YFS	EUR	37.6 (5.0)	52	30	341	<1	650	≥4
AA_ARIC	AA	53.1 (5.9)	63	25	1360	<1	129	≥4
AA_MESA	AA	62.5 (10.1)	52	15	952	<1	47	≥4
AA_WHI	AA	62.2 (6.9)	100	14	1101	<1	256	≥4
EXPLORATORY								
PROMIS	IN	53.5 (10.1)	19	31	11871	0	365	≥1

EUR, European; AA, African American; IN, Indian

[a]Different cut-offs were used to allow for geographical/cultural differences in the distribution of coffee intake. Generally, for populations where coffee consumption is the norm (i.e. European countries) higher cut-points were used.

[b]The following studies were unable to contribute to phenotype 2 due to unstable models: stage 1: CHS, HealthABC; stage 2: HPFS CC, AA_CHS, AA_JHS, AA_HealthABC, AA_HANDLS

Table S4. Stage 1 study characteristics of decaffeinated coffee consumers[a]

Study	N	Ancestry	Age, Years (SD)	Female, %	Current Smokers, %	Decaffeinated Coffee Consumption, cups/day	
						Mean (SD)	range
CHS	1172	EUR	77.2 (4.6)	63	7	1.3 (1.1)	0.07-6
Fenland	351	EUR	45.9 (6.9)	61	10	1.5 (1.5)	0.07-6
FHS	2444	EUR	50.9 (11.3)	61	13	1.0 (1.1)	0.07-6
HPFS CHD	700	EUR	57.0 (8.7)	0	9	1.1 (1.2)	0.04-6
HPFS GA	286	EUR	58.2 (8.1)	0	4	1.2 (1.2)	0.04-6
HPFS KS	316	EUR	49.1 (6.6)	0	7	1.0 (1.0)	0.04-6
HPFS T2D	1500	EUR	56.0 (8.3)	0	6	1.1 (1.2)	0.04-6
NHS BrCa	1332	EUR	53.1 (6.3)	100	13	1.2 (1.2)	0.04-6
NHS CHD	689	EUR	54.3 (6.4)	100	26	1.3 (1.2)	0.04-6
NHS GA	460	EUR	53.4 (6.1)	100	16	1.3 (1.2)	0.04-6
NHS KS	299	EUR	49.0 (6.7)	100	13	1.2 (1.2)	0.04-6
NHS T2D	1959	EUR	52.1 (6.6)	100	13	1.2 (1.2)	0.04-6
PLCO	2112	EUR	68.1 (5.4)	25	17	1.5 (1.3)	0.01-7
SHIP	405	EUR	57.9 (16.0)	56	19	2.3 (1.4)	1-9
WGHS	10659	EUR	55.0 (7.1)	100	9	1.3 (1.3)	0.07-6

EUR, European;

[a]Excludes non-decaffeinated coffee consumers but may include individuals reporting consumption of regular coffee.

Table S5. Study-specific genotyping and imputation

Study	Ref	Genotyping					Imputation		
		Platform	Exclusion criteria			SNPs met QC criteria	Software	Reference Panel	Quality filter[f]
			MAF	Call rate	P HWE				
STAGE 1 – DISCOVERY									
ARIC	[151]	Affymetrix 6.0	<.01	≤.98	<1e-5	669450	MACH	HM R22, CEU	0.3
Busselton	[4]	Ilumina 610Q	<.01	<.97	<5e-7	508061	MACH	HM R22, CEU	0.3
CHS	n/a	Illumina 370CNV	Heterozygotes=0	<.97	<1e-5	306655	BIMBAM	HM R22, CEU	0.3
COLAUS	[5]	Affymetrix 500K	<.01	<.90	<1e-7	390631	IMPUTE	HM R21, CEU	0.3[g]
DNBC	[151]	Illumina 660W-Q	<.005	≤.98	<1e-3	518097	MACH	HM R22, CEU	0.3
EGCUT1	[152]	Illumina 370K, Omni 770	<.01	≤.95	<1e-6	194589	IMPUTE	HM R22, CEU	0.4
EPIC-Norfolk	[153]	Affymetrix 500K	<.01	≤.90	<1e-6	386170	IMPUTE	HM R22, CEU	0.4
ERF	[154]	Illumina 6K, 318K, 370K, 610Q, Affymetrix 250K	<.01	<.95	-	n/a	MACH	HM R22, CEU	0.3
FamHS[a]	n/a	Illumina HapMap 550k Illumina 610Qv1 Illumina IM-Duov3	<.01	≤.98	<1e-6	493938 520193 849551	MACH	HM R22, CEU	0.3
Fenland	n/a	Affymetrix 500K	<.01	<.90	<1e-6	362055	IMPUTE	HM R22, CEU	0.4
FHS	n/a	Affymetrix 500K (250K Nsp & 250K Sty), MIPS 50K	--	≤.97	<1e-6	378163	MACH	HM R22, CEU	0.3
GOOD	n/a	Ilumina 610Q	<.01	<.98	<1e-6	521160	MACH	HM R22, CEU	0.3
HealthABC	n/a	Ilumina IM	<.01	<.95	<1e-7	914263	MACH	HM R22, CEU	0.3
HBCS	n/a	Ilumina 610Q	-	-	-	2543887	MACH	HM R22, CEU	0.3
HPFS CHD	[155]	Affymetrix 6.0	<.02	≤.98	<1e-4	724,881	MACH	HM R22, CEU	0.3
HPFS GA	[156]	Illumina 660Q	<.02	≤.98	<1e-4	495132	MACH	HM R22, CEU	0.3
HPFS KS	[90]	Illumina 610Q	<.01	<.95	<1e-5	2244671	MACH	HM R22, CEU	0.3
HPFS T2D	[151]	Affymetrix 6.0	<.02	≤.98	<1e-4	706,040	MACH	HM R22, CEU	0.3
inCHIANTI	[157]	Illumina HapMap 550k	<.01	≤.99	<1e-6	498838	MACH	HM R22, CEU	0.3
KORA-F3	n/a	Affymetrix 500K	<.01	<.95	<1e-6	2552925	IMPUTE	HM R22, CEU	0.4
KORA-S4	n/a	Affymetrix 1000K	<.01	<.95	<1e-6	2743205	IMPUTE	HM R22, CEU	0.4
MESA	[158]	Affymetrix 6.0	-	<.95	-	854578	IMPUTE	HM R24 (I & II), CEU	0.4
NFBC 1966	[159]	Illumina HumanCNV370DUO	<.01	<.95 (<.99 if MAF <.05)	<5.7e-7	324896	IMPUTE	HM R22, CEU	0.4
NHS BrCa	[160]	Illumina 550k	<.01	<.90	-	528173	MACH	HR R22, CEU	0.3
NHS CHD	[155]	Affymetrix 6.0	<.02	≤.98	<1e-4	721316	MACH	HR R22, CEU	0.3
NHS GA	[156]	Illumina 660Q	<.02	≤.98	<1e-4	495132	MACH	HR R22, CEU	0.3
NHS KS	[90]	Illumina 610Q	<.01	<.95	<1e-5	546344	MACH	HR R22, CEU	0.3
NHS T2D	[151]	Affymetrix 6.0	<.02	≤.98	<1e-4	704409	MACH	HR R22, CEU	0.3
NTR s1[b]	n/a	Affy/Perlegen	<.01	<.90	<1e-5	1443848	IMPUTE	HM R22, CEU	0.4

		Illumina 660Q Affymetrix 6.0 Illumina OmniExpress 1M							
PLCO	[90]	Illumina 550k Illumina 610Q	MAC< 10	<.95	--	515922	IMPUTE	HM R22, CEU	0.4
RS1	[161]	Illumina 550K	<.01	<.98	<1e-6	508333	MACH	HR R22, CEU	0.3
RS2	[161]	Illumina 550K	<.01	<.98	<1e-6	524337	MACH	HR R22, CEU	0.3
SHIP	n/a	Affymetrix 6.0	-	-	-	869224	IMPUTE	HM R22, CEU	0.4
SORBS	n/a	Affymetrix 500K (250K Nsp & 250K Sty), Affymetrix 6.0	<.01	<.95	<1e-4	378513	IMPUTE	HM R21, CEU	0.4
THISEAS	n/a	Illumina OmniExpress	--	≤.98	--	725582	n/a	n/a	n/a
TWINGENE	n/a	Illumina OmniExpress	<.03	≤.97	<1e-7	644556	IMPUTE	HM R22, CEU	0.4
WGHS		Illumina HumanHap300 Duo+ (some iSelect)	<.01	<.90	<1e-5	335603 (includes 32521 custom content)	MACH	HM R22, CEU	0.3
STAGE 2: (IN SILICO) REPLICATION									
BMES	[162]	Illumina Human670Q Custom chip v1	<.05	<.95	<1e-6	544802	IMPUTE	1000G (v1)	0.5
1958BC T1DGC	[29]	Illumina 550k	<.01	<.95	<1e-7	2446857	IMPUTE	HM R22, CEU	0.5
1958BC WTCCC2[c]	[28]	Affymetrix 6.0	<.01	≤.98	<1e-20	2541273	IMPUTE	HM R22, CEU	0.5
EGCUT2	[152]	Illumina OmniExpress	<.01	≤.95	<1e-6	630155	IMPUTE	HM R22, CEU	0.5
EGCUT3	[152]	Illumina Metabochip	<.01	≤.95	<1e-6	132363	-	-	-
GLACIER	n/a	Illumina Metabochip	<.01	≤95	<1e-6	189315	-	-	-
HPFS CC	[163]	Illumina HumanOmniExpress-12v1_B	<.02	≤.98	<1e-4	619089	MACH	HM R22, CEU	0.5
LifeLines	n/a	Illumina Cyto SNP 12 v2	<0.01	≤.95	<10e-4	257581	BEAGLE	HM R24, CEU	0.5
MrOS Sweden	n/a	Illumina HumanOmni1_Quad_v1-0 B	<.01	≤.98	<1e-4	739477	Minimach	HM Build 36, CEU, R22	0.5
NTR s2[d]	n/a	Affymetrix 6.0 Affymetrix-Perlegen 5.0 Illumina 370K Illumina 660Q Illumina 1M	<.01	<.95	<1e-5	>350000	IMPUTE	1000G (v1)	0.5
NHS CC	[163]	Illumina HumanOmniExpress-12v1_B	<.02	≤.98	<1e-4	619148	MACH	HM R22, CEU	0.5
PIVUS	n/a	Illumina HumanOmniExpress Metabochip	<.01	<.95	<1e-6	738879	IMPUTE	HM R22, CEU	0.5
QIMR	[146]	Illumina 317K Illumina 370K Illumina 610Q	<.01	<.95	<1e-5	274604	MACH	HR R22, CEU	0.5

ULSAM	n/a	Illumina Omni 2.5 M Metabochip	<.01	<.95	<1e-6	1621908	IMPUTE	HM R22, CEU	0.5
YFS	[164]	Illumina custom BeadChip, 670Q	<.01	<.95	<1e-6	546674	IMPUTE	HM R22, CEU	0.5
AA_ARIC	n/a	Affymetrix 6.0	<.01	≤.95	--	796384	MACH	HM R22, CEU & YRI	0.3
AA_CHS[e]	n/a	Illumina Omni1M	Heterozygotes=0	<.97	<1e-5	963248	BEAGLE	HM III, ASW, YRI,CEU	0.3
AA_HANDLS	n/a	Illumina Human IM	<.01	<.95	<1e-7	907763	MACH	HM R22, CEU & YRI	0.3
AA_HealthABC	n/a	Illumina Human IM	<.01	<.95	<1e-7	688867	MACH	HM R22, CEU & YRI	0.3
AA_JHS	n/a	Affymetrix 6.0	<.01	<.95	<1e-6	n/a	MACH	HM R22, CEU & YRI	0.3
AA_MESA	[158]	Affymetrix 6.0	-	<.95	-	854578	IMPUTE	HM 1 & 2 (R24), CEU+YRI+CHB+JPT	0.4
AA_WHI	n/a	Affymetrix 6.0	<.01	<.95	<1e-6	829370	MACH	HM R22, CEU & YRI	0.3
EXPLORATORY									
PROMIS	[89]	Illumina 660Q Illumina OmniExpress	<.01	<.95	<1e-7	541656 663455	IMPUTE	HM R22 & HM III, GIH	0.4

n/a, not available; HM, HapMap,

Study-specific exclusions:

[a]SNPs not in HapMap

[b]SNPs where Mendel and double error rate >.01, between platform MAF differences >.15

[c]SNPs with evidence of plate association (<.00001) and statistical info rate ≥ 0.975

[d]SNPs where Mendel and double error rate >.02, AT and GC SNPs MAF>0.35, MAF difference with ref > 0.20

[e] >2 duplicate errors

[f]SNP quality filter thresholds applied to stage 1 studies prior to current meta-analysis: MACH: <0.3, BIMBAM: <0.3, IMPUTE:<0.4, BEAGLE:<0.3. For stage 2 European studies, which provided results for only up to 44 SNPs, we requested genotyped or imputed SNPs with quality scores >0.5 (regardless of imputation software).

[g]The MACH-equivalent quality matrix (i.e. Rsq) was derived and used in place of IMPUTE's matrix (i.e. proper_info).

Table S6. Study-specific sample quality control

Study	Sample Quality Control	
	Call rate	Other exclusion criteria
STAGE 1		
ARIC	≤.95	-no DNA consent -first-degree relatives -PCA outliers - sex/gender phenotype and genotype mismatch -non-concordance between Affy/Birdsuite genotype and genotype assayed by TaqMan (47 SNPs) -missing phenotype & covariate information
Busselton	≤.95	-PCA outliers -sex mismatch -duplicates -related individuals (IBD threshold of 0.1875) -low heterozygosity >5 SD below mean heterozygosity
CHS	≤.95	-sex mismatch -discordance with prior genotyping -PCA outliers
COLAUS	<.90	-duplicates and first/second degree relatives (favoring the younger for inclusion) -PCA outliers
DNBC	≤.95	-sex discrepancy with genetic data from X-linked markers -MDS outliers
EGCUT1	≤.95	-sex discrepancy with genetic data from X-linked markers -duplicates and first/second degree relatives -PCA outliers -missing phenotype information
EPIC-Norfolk	<.94	-ethnic outliers -heterozygosity <23.0% or >30.0%
ERF	<.95	-sex mismatch -excess IBS -ethnic outliers
FamHS	<.98	-mendelian errors -sex discrepancies -missing phenotype & covariate information -PCA outliers
Fenland	<.95	-heterozygosity <27.3% or >28.8% -relatedness check -duplicate check
FHS	≤.97	-sex discrepancy with genetic data from X-linked markers -duplicates -heterozygosity > 5SD from mean(<25.758% or >29.958%) - >1000 Mendelian errors -missing phenotype & covariate information
GOOD	<.975	-heterozygosity > 33%; -related individuals and duplicates
HBCS	-	-sex discrepancy with genetic data from X-linked markers

		-relatedness check -duplicate check -ethnic outliers
Health ABC	<.95	-sex discrepancy with genetic data from X-linked markers -duplicates and first/second degree relatives -PCA outliers -heterozygosity -autosomal chromosome abberations -missing phenotype information
HPFS CHD	≤.98	-sex discrepancy with genetic data from X-linked markers -duplicates and first/second degree relatives -PCA outliers -heterozygosity -missing phenotype & covariate information
HPFS GA	≤.98	-sex discrepancy with genetic data from X-linked markers -duplicates and first/second degree relatives -PCA outliers -heterozygosity -autosomal chromosome abberations -missing phenotype & covariate information
HPFS KS	<.95	-duplicates and first/second degree relatives -PCA outliers -missing phenotype & covariate information
HPFS T2D	≤.98	-sex discrepancy with genetic data from X-linked markers -duplicates and first/second degree relatives -PCA outliers -heterozygosity -autosomal chromosome abberations -missing phenotype & covariate information
inCHIANTI	≤.97	-sex discrepancy with genetic data from X-linked markers -duplicates and first/second degree relatives -heterozygosity -missing phenotype information
KORA_F3	≤.90	-sex discrepancy with genetic data from X-linked markers -missing phenotype & covariate information
KORA_S4	≤.90	-sex discrepancy with genetic data from X-linked markers -missing phenotype & covariate information
MESA	<.95	-sex discrepancy with genetic data from X-linked markers -duplicates and first degree relatives -heterozygosity > 53% -missing phenotype & covariate information -PCA outliers
NFBC 1966	<.95	-Low mean heterozygosity [exclude if <0.29 & MDS outliers] -Duplicates: concordance with other DNA>0.99 -Contaminated samples: IBS pairwise with most other samples >0.99 -IBS pairwise sharing>0.20 -Withdrew consent -Gender mismatch: genotypic gender different from phenotypic -MDS outliers
NHS BrCa	≤.90	-duplicates and first/second degree relatives

		-PCA outliers -missing phenotype & covariate information
NHS CHD	≤.98	-sex discrepancy with genetic data from X-linked markers -duplicates and first/second degree relatives -PCA outliers -heterozygosity -missing phenotype & covariate information
NHS GA	≤.98	-sex discrepancy with genetic data from X-linked markers -duplicates and first/second degree relatives -PCA outliers -heterozygosity -autosomal chromosome abberations -missing phenotype & covariate information
NHS KS	<.95	-duplicates and first/second degree relatives -PCA outliers -missing phenotype & covariate information
NHS T2D	≤.98	-sex discrepancy with genetic data from X-linked markers -duplicates and first/second degree relatives -PCA outliers -heterozygosity -autosomal chromosome abberations -missing phenotype & covariate information
NTR s1	--	-sex discrepancy with genetic data from X-linked markers -heterozygosity -autosomal chromosome abberations -missing phenotype information -non European ancestry (outliers)
PLCO	≤.98	-sex discrepancy with genetic data from X-linked markers -duplicates and first/second degree relatives -PCA outliers -heterozygosity -autosomal chromosome abberations -missing phenotype & covariate information
RS-I	<.90	-sex mismatch -excess IBS -ethnic outliers
RS-II	<.90	-sex mismatch -excess IBS -ethnic outliers
SHIP	≤.92	-sex discrepancy with genetic data from X-linked markers -duplicates -missing phenotype information
SORBS	<.94	-ethnic outliers -duplicates -gender mismatch -IBS>0.2
THISEAS	≤.95	-sex discrepancy with genetic data from X-linked markers -heterozygosity -missing phenotype information
TWINGENE	≤.97	-sex discrepancy with genetic data from X-linked markers

		-heterozygosity -missing phenotype information -PCA outliers
WGHS	≤.98	-MDS outliers -missing phenotype & covariate information
STAGE 2		
BMES	≤.95	-outlier on distribution of genome-wide proportion of heterozygous genotypes discrepancies between clinical and genotypic gender -evidence of unintended sample duplication (IBD(2) > 0.95) -evidence of cryptic relatedness with one or more other samples (one member excluded for each pair with IBD(0) sharing proportion < 0.95) -concordance < 0.9 for ~30 SNPs genotyped in duplicate on the Sequenom platform -evidence of sample swap/contamination, based on comparison of 1356 samples independently genotyped in duplicate on the Illumina 610K array -PCA outliers
1958BC T1DGC	≤.97	-exceeding heterozygosity thresholds -MDS outliers -gender discrepancy
1958BC WTCCC2	--	-likely relatives -exceeding heterozygosity thresholds -MDS outliers -gender discrepancy -outlying allele intensities -discordance with external genotyping
EGCUT2	≤.95	-sex discrepancy with genetic data from X-linked markers -duplicates and first/second degree relatives -missing phenotype information
EGCUT3	≤.95	-sex discrepancy with genetic data from X-linked markers -duplicates and first/second degree relatives -missing phenotype information
GLACIER	≤.95	-sex discrepancy with genetic data from X-linked markers -missing phenotype & covariate information
HPFS CC	<.98	-duplicates and first/second degree relatives -PCA outliers -missing phenotype & covariate information
LifeLines	<.95	- duplicates and first degree relatives -PCA outliers - sex discrepancy - heterozygosity >4SD from mean - missing phenotype & covariate information
MrOS Sweden	<.97	-exclusion based on IBD clustering -identical twins -sample duplicates
NTR s2	≤.90	-non European ancestry (outliers) -in case of a monozygotic twin pair, one of the individuals was excluded -subjects (and their relatives) in NTR s1 -sex discrepancy with genetic data from X-linked markers

		-heterozygosity -large chromosomal genome abberations -unexpected IBD/ IBS sharing between samples in relation to pedigrees - mendelian error rate > 2% -missing phenotype information
NHS CC	<.98	-duplicates and first/second degree relatives -PCA outliers -missing phenotype & covariate information
PIVUS	≤.95	-sex discrepancy with genetic data from X-linked markers -heterozygosity -missing phenotype information - Large position disagreements and not mapping in the genome and/or mapping more than once in the genome (Will's list) and Bad probe assays -monomorphic SNPs -missing phenotype information -MDS outliers
QIMR	<.95	-sex discrepancy with genetic data from X-linked markers -duplicates and first/second degree relatives -PCA outliers -heterozygosity -autosomal chromosome abberations
ULSAM	≤.95	-sex discrepancy with genetic data from X-linked markers -heterozygosity (+/- 3 SD) -missing phenotype information -IBD check on a LD pruned (CEU: r^2<0.2) set of overlapping SNPs ~30K -PCA outliers
YFS	<.95	-missing gender -related individuals and duplicates -missing phenotype information
AA_ARIC	≤.95	<18 fingerprinting assays working >3 discordant fingerprinting assays -duplicates -extreme heterozygosity -low-level IBD/IBS sharing (PI_HAT>0.05) with lg # of samples -nearest neighbor analysis outliers -outliers from clustering based on missingness-missing phenotype & covariate information
AA_CHS	<.95	-sex mismatch -discordance with prior genotyping
AA_HANDLS	<.95	-sex discrepancy with genetic data from X-linked markers -duplicates and first/second degree relatives -PCA outliers -heterozygosity -autosomal chromosome abberations -missing phenotype information
AA_HealthABC	<.95	-sex discrepancy with genetic data from X-linked markers -duplicates and first/second degree relatives -PCA outliers -heterozygosity -autosomal chromosome abberations

		-missing phenotype information
AA_JHS	<.95	<18 fingerprinting assays working >3 discordant fingerprinting assays -duplicates -extreme heterozygosity -low-level IBD/IBS sharing (PI_HAT>0.05) with lg # of samples -nearest neighbor analysis outliers -outliers from clustering based on missingness-missing phenotype & covariate information
AA_MESA	<.95	-sex discrepancy with genetic data from X-linked markers -duplicates and first degree relatives -heterozygosity > 53% -missing phenotype & covariate information -PCA outliers
AA_WHI	<.95	-first degree relatives -duplicates -Y chromosome markers -missing phenotype information -PCA outliers
EXPLORATORY		
PROMIS	<.95	-sex discrepancy with genetic data from X-linked markers -duplicates and first/second degree relatives -missing phenotype information

PCA, principal component analysis; MDS, multidimensional scaling

Table S7. Stage 1 study-specific statistical analysis

Study	Linear regression, phenotype 1		Logistic regression, phenotype 2		Software	covariates
	SNPs in meta-analysis	λ_{GC}	SNPs in meta-analysis	λ_{GC}		
ARIC	2378426	1.022	2378420	1.026	probABEL	age, sex, center, smoking (never, former, current ≤20 cigs/d, current >21 cigs/d, 10 EVs
Busselton	2379720	1.045	2380024	1.032	probABEL	age, sex, asthma case-control status, smoking (never, former, current <15 cigs/d, current ≥15 cigs/d)
CHS	2237112	1.012	n/a	n/a	R	age, sex, smoking (never, former, current), site, 10 EVs
COLAUS	2364136	1.01	2358274	1.052	Matlab	age, sex, smoking (never, former, current <15 cigs/d, current ≥15 cigs/d), 4 PCs
DNBC	2378653	0.999	2378473	1.017	Mach2qtl, mach2dat	age, smoking (never, former, current <15 cigs/d, current ≥15 cigs/d)
EGCUT1	1164952	1.031	1167121	1.007	SNPTEST	age, sex, smoking (never, former, current <15 cigs/d, current ≥15 cigs/d), 2 EVs
EPIC-Norfolk	2265797	1.011	2291202	1.017	SNPTEST	age, sex, smoking (never, former, current <15 cigs/d, current ≥15 cigs/d)
ERF	2347504	0.999	2347472	1.057	probABEL (linear mixed model)	age, sex, smoking (never, former, current <15 cigs/d, current ≥15 cigs/d)
FamHS	2198784	1.002	2382970	1.021	SAS/R kinship matrix	age, sex, field center, smoking (never, former, current <20 cig/d, current ≥20 cig/d), platform
Fenland	2305059	1.01	2304877	1.01	SNPTEST	age, sex, smoking (never, former, current <15 cigs/d, current ≥15 cigs/d)
FHS	2401467	1.005	2401467	1.025	R (LMEKIN, GEE)	age, sex, smoking (never, former, current <15 cigs/d, current ≥15 cigs/d), PC1,2,6,9
GOOD	2376827	1.006	2376631	1.022	Mach2qtl	age, smoking (never, former, current <15 cig/d, current ≥15 cig/d)
HealthABC	2398505	1.013	n/a	n/a	Mach2qtl	age, smoking (never, former, current), sex, 2 EVs, study site
HBCS	2361930	1.01	2361855	1.041	PLINK, probABEL	age, sex, smoking (never, former, current <20 cig/d, current ≥20 cig/d)
HPFS CHD	2392485	1.005	2392005	1.007	probABEL	age, case-status, smoking (never, former, current <15 cig/d, current ≥15 cig/d), 3 EVs
HPFS GA	2385357	1.005	297747	2.934[a]	probABEL	age, case-status, DNA extraction method, smoking (never, former, current <15 cig/d, current ≥15 cig/d), 3 EVs
HPFS KS	2389689	1.007	2389956	1.039	probABEL	age, case-status, smoking (never, former, current <15 cig/d, current ≥15 cig/d), 4 EVs
HPFS T2D	2393735	1.012	2392220	1.021	probABEL	age, case-status, smoking (never, former, current <15 cig/d, current ≥15 cig/d), 4 EVs
inCHIANTI	2383633	0.994	2380771	1.04	Merlin (offline)	age, smoking (never, former, current <15 cigs/d, current ≥15 cigs/d), study site
KORA_F3	2316536	0.999	2315207	1.035	SNPTEST	age, smoking (never, former, current <15 cigs/d, current ≥15 cigs/d)
KORA_S4	2387459	0.995	2381636	1.005	SNPTEST	age, smoking (never, former, current <15 cigs/d, current ≥15 cigs/d)
MESA	2481815	0.994	2482577	1.017	SNPTEST	age, sex, smoking (former, current <20, current ≥20), 2EVs, study site
NFBC 1966	2462217	1.022	2462196	1.018	SNPTEST	Sex, smoking (never, former, current <15 cigs/d, current ≥15 cigs/d), 3 EVs
NHS BrCa	2388959	0.995	2363509	1.019	probABEL	age, case-status, smoking (never, former, current <15 cig/d, current ≥15 cig/d),4 EVs

NHS CHD	2388255	1.004	2386437	1.022	probABEL	age, case-status, smoking (never, former, current <15 cig/d, current ≥15 cig/d), 3 EVs
NHS GA	2384314	1.011	2383036	1.04	probABEL	age, case-status, DNA extraction method, DNA source, smoking (never, former, current <15 cig/d, current ≥15 cig/d), 3 EVs
NHS KS	2382433	0.994	2380043	1.048	probABEL	age, case-status, smoking (never, former, current <15 cig/d, current ≥15 cig/d), 4 EVs
NHS T2D	2389161	1.037	2388125	1.022	probABEL	age, case-status, smoking (never, former, current <15 cig/d, current ≥15 cig/d), 3 EVs
NTR s1	2318535	1.005	2321190	0.998	SNPTEST	age. sex, smoking (never, former, current ≤20 cigs/d, current >20 cigs/d), 10 EVs
PLCO	2291294	0.991	2291396	0.996	R	age, sex, smoking ((never, former, current <20 cigs/d, current ≥20 cigs/d), 2 EVs
RS-I	2377076	1.027	2376975	1	probABEL	age, sex, smoking (never, former, current <15 cigs/d, current ≥15 cigs/d)
RS-II	2375001	0.989	2374440	1.03	probABEL	age, sex, smoking (never, former, current <15 cigs/d, current ≥15 cigs/d)
SHIP	2456479	1.02	2455983	1.015	QUICKTEST	age, sex, smoking (never, former, current <15 cigs/d, current ≥15 cigs/d)
SORBS	2306278	1.028	2283493	1.183	SNPTEST	age, sex, smoking (never, former, current <15 cig/d, current ≥15 cig/d)
THISEAS	627993	0.999	627939	1.047	PLINK	age, case status, smoking (never, former, current <15 cigs/d, current ≥15 cigs/d)
TWINGENE	2323229	1.021	2322714	1.011	PLINK	age, sex, smoking (never, former, current <15 cigs/d, current ≥15 cigs/d), 3 EVs
WGHS	2424130	1.009	2424129	1.019	probABEL	age, smoking (never, former, current ≤20 cig/d, current >20 cig/d), 5 EVs

n/a, not available; EV, eigenvector; PC, principal component

[a]Relatively few study-level test statistics were non-missing due to sample size, contributing to an inflated λ_{GC} . The sample size and penalty imposed by GC for this study would, at most, reduce power of the meta-analysis.

Table S8. Stage 1 meta-analysis results for SNPs selected for follow-up[a]

Chr	Position (Hg18)	Best-SNP	EA/NEA	EAF	Closest genes[b]	Phenotype 1 cups of coffee consumed per day among coffee consumers					Phenotype 2 high vs. no/low coffee consumption				
						N	β (SE)	SNP-level		Gene-level	N	β (SE)	SNP-level		Gene-level
								P	I^2	Top 10 %P			*P*	I^2	Top 10 %P
1	171939422	rs6681766	A/G	0.21	*ANKRD45, KLHL20*	88069	-0.04 (0.01)	6.40E-05	12	8.8E-05	45992	-0.06 (0.02)	2.26E-03	0	2.7E-03
2	1001316	rs6548172	C/G	0.87	*SNTG2*	90885	-0.03 (0.01)	1.50E-02	0	0.28	46879	-0.12 (0.02)	1.38E-06	0	8.3E-05
2	2370732	rs713347	T/G	0.58	*MYT1L*	91455	0.03 (0.01)	4.42E-06	16	2.5E-04	47037	0.04 (0.02)	3.15E-02	45	0.02
2	27584444	rs1260326	T/C	0.41	*GCKR, FNDC4*	91407	-0.04 (0.01)	1.06E-07	0	1.0E-06	47020	-0.05 (0.02)	1.74E-03	0	1.5E-03
2	27693043	rs2068834	T/C	0.71	*ZNF512, CCDC121*	90823	0.03 (0.01)	3.18E-05	2	1.1E-04	46853	0.04 (0.02)	1.23E-02	9	4.5E-03
4	89258106	rs1481012	A/G	0.89	*ABCG2, PKD2, PPM1K*	87994	0.06 (0.01)	1.13E-06	0	1.8E-05	45985	0.08 (0.03)	2.01E-03	0	0.01
4	106768545	rs10007278	T/G	0.15	*FLJ20184, ATP5EP1*	88518	0.05 (0.01)	4.50E-06	2	1.1E-04	46116	0.13 (0.02)	6.54E-08	15	1.5E-05
4	109081710	rs10461142	T/C	0.67	*CYP2U1, SGMS2, HADH*	90935	0.01 (0.01)	1.96E-01	22	0.63	47152	0.08 (0.02)	4.12E-06	24	3.4E-05
6	119477071	rs649979	C/G	0.10	*MCM9, MAN1A1, C6orf60*	87987	0.06 (0.01)	1.82E-06	0	2.9E-04	45979	0.07 (0.03)	9.09E-03	0	0.1
6	145203694	rs4895657	A/G	0.64	*UTRN*	88523	0.02 (0.01)	4.69E-03	0	0.01	46121	0.08 (0.02)	6.53E-06	0	7.0E-05
6	159376363	rs4709267	A/G	0.90	*TAGAP, RSPH3*	88525	-0.05 (0.01)	3.63E-04	0	0.001	46122	-0.14 (0.03)	4.88E-07	0	1.0E-06
7	17163382	rs17137304	A/G	0.88	*AHR*	85528	0.07 (0.01)	1.09E-07	0	<1.0E-06	44939	0.13 (0.03)	6.50E-06	0	<1.0E-06
7	17251102	rs4410790	T/C	0.37	*AHR*	86338	-0.14 (0.01)	1.48E-57	54	<1.0E-06	44806	-0.24 (0.02)	1.82E-37	38	<1.0E-06
7	17253631	rs6968554	A/G	0.39	*AHR*	86867	-0.13 (0.01)	2.54E-57	57	<1.0E-06	45821	-0.23 (0.02)	3.13E-38	43	<1.0E-06
7	17409204	rs1077773	A/G	0.52	*AHR, SNX13*	89773	-0.04 (0.01)	4.70E-08	0	<1.0E-06	45847	-0.07 (0.02)	1.01E-05	20	<1.0E-06
7	31638661	rs10235961	A/T	0.53	*CCDC129, NEUROD6*	87892	0.04 (0.01)	5.90E-06	0	3.3E-04	45932	0.06 (0.02)	9.17E-04	10	1.9E-03
7	72673793	rs7800944	T/C	0.72	*MLXIPL, TBL2, VPS37D*	87998	-0.05 (0.01)	7.82E-09	6	<1.0E-06	n/a[3]	n/a	n/a	n/a	2.0E-06
7	75454041	rs17685	A/G	0.29	*POR, SNORA14A, TMEM120A*	85140	0.07 (0.01)	9.06E-14	37	<1.0E-06	44717	0.11 (0.02)	4.41E-09	0	<1.0E-06
7	144693673	rs10227393	C/G	0.75	*CNTNAP2*	87998	-0.02 (0.01)	2.01E-02	20	0.27	46230	-0.09 (0.02)	3.20E-06	0	0.25
8	20608550	rs2597398	T/G	0.46	*LZTS1*	90928	-0.04 (0.01)	2.56E-06	0	0.03	46905	-0.05 (0.02)	1.83E-03	0	0.04
8	72378104	rs12549065	T/C	0.63	*EYA1, XKR9, SEDLP2*	88521	-0.04 (0.01)	8.39E-06	0	0.02	46116	-0.05 (0.02)	8.44E-03	0	0.03
9	91405458	rs1571536	T/C	0.49	*SEMA4D, GADD45G*	89306	-0.04 (0.01)	2.32E-06	0	2.4E-05	47004	-0.06 (0.02)	2.72E-04	0	1.4E-03
9	139632376	rs1045777	A/G	0.21	*ARRDC1, C9orf37*	86420	0.04 (0.01)	1.70E-05	0	6.9E-05	46118	0.04 (0.02)	6.95E-02	41	0.09
10	16645488	rs11254079	A/G	0.88	*C1QL3, RSU1*	85156	-0.05 (0.01)	2.01E-04	23	0.01	44840	-0.13 (0.03)	2.04E-06	15	2.4E-04
11	27636492	rs6265	T/C	0.19	*BDNF*	89803	-0.05 (0.01)	3.40E-07	30	<1.0E-06	45866	-0.12 (0.02)	4.86E-08	16	2.0E-06
11	27650817	rs2049045	C/G	0.19	*BDNF*	89280	-0.05 (0.01)	2.08E-07	30	<1.0E-06	45731	-0.11 (0.02)	7.78E-08	16	2.0E-06
11	27704247	rs12288512	A/G	0.19	*BDNF*	87718	0.05 (0.01)	2.52E-07	1	<1.0E-06	44962	0.07 (0.02)	2.79E-04	8	2.0E-06
11	79095197	rs2264517	A/C	0.32	*ODZ4*	87983	0.04 (0.01)	2.85E-06	0	0.48	45979	0.06 (0.02)	6.97E-04	0	0.41
11	132761771	rs4245116	T/C	0.47	*OPCML*	87120	0.04 (0.01)	1.54E-06	11	0.01	45250	0.05 (0.02)	3.24E-03	0	0.07
12	117272386	rs17512574	T/C	0.17	*TAOK3, SUDS3*	91227	0.04 (0.01)	3.68E-05	35	9.0E-05	46028	0.08 (0.02)	6.45E-04	26	2.5E-03
12	117926875	rs11069228	A/G	0.15	*KIAA1853, SUDS3, HSPB8*	88525	0.02 (0.01)	5.77E-02	10	0.04	46121	0.10 (0.02)	9.85E-06	0	0.01
15	72806502	rs2470893	T/C	0.31	*CYP1A1, CYP1A2*	91462	0.12 (0.01)	6.89E-44	58	<1.0E-06	47040	0.21 (0.02)	1.33E-30	43	<1.0E-06

15	72814933	rs2472297	T/C	0.24	*CYP1A1, CYP1A2*	87622	0.15 (0.01)	6.45E-47	57	<1.0E-06	45213	0.23 (0.02)	1.63E-29	47	<1.0E-06
17	25373221	rs9902453	A/G	0.54	*EFCAB5, SSH2, CCDC55*	90770	-0.04 (0.01)	2.26E-06	0	1.0E-05	46841	-0.06 (0.02)	5.67E-05	0	4.5E-05
17	25555919	rs3794808	T/C	0.42	*SLC6A4, CCDC55, BLMH*	91422	0.03 (0.01)	1.57E-05	0	2.6E-05	47017	0.08 (0.02)	1.19E-06	0	1.1E-05
17	31939244	rs17560870	A/G	0.59	*MYO19, ZNHIT3, PIGW*	90851	0.03 (0.01)	9.59E-06	0	3.4E-05	46871	0.07 (0.02)	1.83E-05	0	5.6E-05
17	61762419	rs12450534	A/G	0.24	*PRKCA, APOH, CACNG5*	88514	-0.03 (0.01)	2.37E-03	21	0.03	46115	-0.09 (0.02)	7.04E-06	0	9.7E-03
19	4090773	rs164631	A/G	0.72	*MAP2K2, CREB3L3*	87519	-0.04 (0.01)	1.64E-05	0	5.6E-05	n/a[3]	n/a	n/a		0.05
19	5239875	rs17676218	T/C	0.23	*PTPRS, ZNRF4*	77003	0.04 (0.01)	8.65E-04	5	0.005	42049	0.10 (0.02)	5.62E-06	0	2.6E-05
22	23189667	rs9620388	A/C	0.91	*ADORA2A, UPB1*	91231	-0.06 (0.01)	3.85E-06	20	9.0E-05	46945	-0.10 (0.03)	2.45E-04	0	7.7E-04
22	23277452	rs738820	T/C	0.84	*UPB1, SNRPD3, C22orf13*	81820	0.06 (0.01)	1.02E-06	0	5.0E-06	n/a[3]	n/a	n/a		3.9E-04
22	39532082	rs138312	T/C	0.51	*SLC25A17, ST13*	91460	-0.03 (0.01)	6.54E-04	0	0.004	47042	-0.07 (0.02)	3.58E-05	0	2.6E-04
22	39745590	rs4821981	T/G	0.39	*RBX1, RPS9P2*	91332	-0.03 (0.01)	8.83E-05	22	3.6E-05	46974	-0.07 (0.02)	1.10E-05	0	1.4E-05
22	39985654	rs2235852	T/G	0.32	*RANGAP1, CHADL, ZC3H7B*	89000	-0.04 (0.01)	5.94E-07	0	1.5E-05	45513	-0.07 (0.02)	3.50E-05	0	6.9E-05

EA, effect allele; NEA, non-effect allele; EAF, effect allele frequency; n/a, not available

[a]Presented for each SNP-level regression model is sample size ('N'), beta coefficients and standard errors ('β (SE)'), P value (P, columns 9 and 14), and I^2 statistic for heterogeneity (columns 10 and 15). P values for gene-level analyses are presented in columns 11 and 16.

[b]Underlined genes are those yielding the lowest p-values in the gene-level meta-analysis (presented in column 11)

[c]These SNPs did not pass the sample size and/or imputation quality filters in-place for the meta-analysis

Table S9. Stage 2 study-specific statistical analysis[a]

Study	SNPs in Replication analysis				Software	covariates
	Linear regression, phenotype 1		Logistic regression, phenotype 2			
BMES	44		44		PLINK	age, sex, smoking (never, former, current <15 cig/d, current ≥15 cig/d)
1958BC T1DGC	44		44		SNPTEST	sex, smoking (never, former, current <20 cig/d, current ≥20 cig/d), region (Greater London, South of England, Middle England & Wales, North of England, Scotland), 2 PCs
1958BC WTCCC2	43		43		SNPTEST	sex, smoking (never, former, current <20 cig/d, current ≥20 cig/d), region (Greater London, South of England, Middle England & Wales, North of England, Scotland), 2 PCs
EGCUT2	43		43		SNPTEST	age, sex, smoking (never, former, current <15 cigs/d, current >15 cigs/d)
EGCUT3	13		13		SNPTEST	age, sex, smoking (never, former, current <15 cigs/d, current >15 cigs/d)
GLACIER	17[b]		17[b]		SAS 9.3	age, sex, smoking (never, former, current), top 4 PCs
HPFS CC	44		n/a		probABEL	age, case-status, smoking (never, former, current <15 cig/d, current ≥15 cig/d), 3 EVs
LifeLines	37[c]		37[c]		PLINK	age, sex, smoking (never, former, current <20 cigs/d, current ≥20 cigs/d) , 10 PCs
MrOS Sweden	44		44		PLINK	age, smoking (never, former, current <15 cig/d, current ≥15 cig/d), 4 EVs
NTR s2	39		39		PLINK	age. sex, smoking (never, former, current ≤20 cigs/d, current >20 cigs/d), 9 EVs
NHS CC	44		44		probABEL	age, case-status, smoking (never, former, current <15 cig/d, current ≥15 cig/d), 3 EVs
PIVUS	42		42		SNPTEST	age, sex, smoking (never, former, current <20 cig/d, current ≥20 cig/d)
QIMR	43		43		PLINK	age, sex, smoking (never, former, current <15 cig/d, current ≥15 cig/d
ULSAM	40		40		SNPTEST	age, smoking (never, former, current <20 cig/d, current ≥20 cig/d)
YFS	44		44		PLINK, probABEL	age, sex, smoking (never, former, current <15 cig/d, current ≥15 cig/d), 2 EVs
AA_ARIC	41		41		probABEL	age, sex, center, smoking (never, former, current ≤20 cigs/d, current >20 cigs/d), 10 EVs
AA_CHS	42		n/a		R	age, sex, smoking (never, former, current), site, top 10 EVs
AA_HANDLS	42		n/a		Mach2qtl	age, smoking (never, former, current), sex, study site, 5 EVs
AA_HealthABC	43		n/a		Mach2qtl	age, smoking (never, former, current), sex, study site, 5 EVs
AA_JHS	41		n/a		Mach2qtl	age, sex, smoking (non-current, current), 10 EVs
AA_MESA	43		43		SNPTEST	age, sex, smoking (never, former, current <20 cig/d, current ≥20 cig/d), study site, 2 EVs
AA_WHI	40		40		PLINK, probABEL	age, smoking(never, former, current <15 cig/d, current ≥15 cig/d), 4 EVs

[a] All stage 2 imputed SNPs have quality scores >0.5 (regardless of imputation software). For logistic models 'n/a' indicates no valid statistical tests available

[b] rs17309930 served as a proxy for rs12288512 (r^2=1); rs12150261 served as a proxy for rs9902453 (r^2=0.97)

[c] rs4449655 served as a proxy for rs4709267 (r^2=1); rs11819040 served as a proxy for rs11254079 (r^2=1); rs7285057 served as a proxy for rs9620388 (r^2=1)

Table S10. Stage 2 results from meta-analyses of coffee consumption (cups/d, phenotype 1) for SNPs selected for follow-up[a]

Chr	Position (Hg18)	Best-SNP	A1/A2	Stage 2 Europeans					Stage 2 African Americans				
				EAF	N	β (SE)	P	I^2	EAF	N	β (SE)	P	I^2
1	171939422	rs6681766	A/G	0.19	30059	-0.01 (0.01)	0.622	0	0.15	7962	-0.04 (0.03)	0.111	0
2	1001316	rs6548172	C/G	0.87	22306	-0.02 (0.02)	0.292	0	0.87	7937	0.04 (0.03)	0.163	0
2	2370732	rs713347	T/G	0.55	22373	0.02 (0.01)	0.097	29	0.48	7964	-0.01 (0.02)	0.602	19
2	27584444	rs1260326	T/C	0.36	30046	-0.03 (0.01)	0.020	38	0.17	7964	-0.01 (0.03)	0.768	0
2	27693043	rs2068834	T/C	0.73	30057	0.02 (0.01)	0.189	8	0.76	7961	0.01 (0.02)	0.679	0
4	89258106	rs1481012	A/G	0.89	30061	0.03 (0.02)	0.112	9	0.95	7963	0.16 (0.05)	1.27E-03	3
4	106768545	rs10007278	T/G	0.14	22281	0.04 (0.02)	0.033	29	0.56	7964	0.02 (0.02)	0.330	53
4	109081710	rs10461142	T/C	0.66	22370	0.01 (0.01)	0.402	0	0.61	7964	0.02 (0.02)	0.291	42
6	119477071	rs649979	C/G	0.12	18241	-0.02 (0.02)	0.340	0	0.04	7682	0.08 (0.06)	0.145	34
6	145203694	rs4895657	A/G	0.65	21417	-0.02 (0.01)	0.207	35	0.12	4466	-0.01 (0.05)	0.829	0
6	159376363	rs4709267	A/G	0.89	22339	0.02 (0.02)	0.314	55	0.84	7964	0.004 (0.03)	0.865	8
7	17163382	rs17137304	A/G	0.89	25833	0.04 (0.02)	0.046	0	0.94	7962	-0.03 (0.04)	0.427	6
7	17251102	rs4410790	T/C	0.35	22371	-0.05 (0.01)	1.66E-04	54	0.52	7964	-0.09 (0.02)	2.37E-06	17
7	17253631	rs6968554	A/G	0.35	30017	-0.07 (0.01)	2.78E-10	71	0.33	7964	-0.05 (0.02)	0.016	0
7	17409204	rs1077773	A/G	0.53	22371	-0.01 (0.01)	0.247	33	0.39	7964	0.01 (0.02)	0.531	0
7	31638661	rs10235961	A/T	0.56	22336	-0.01 (0.01)	0.649	0	0.47	7964	-0.03 (0.02)	0.173	53
7	72673793	rs7800944	T/C	0.72	20454	-0.06 (0.02)	4.20E-04	0	0.67	7964	-0.02 (0.02)	0.365	0
7	75454041	rs17685	A/G	0.30	22361	0.05 (0.01)	1.01E-03	37	0.19	7964	0.07 (0.03)	7.55E-03	0
7	144693673	rs10227393	C/G	0.77	18130	-0.01 (0.02)	0.544	0	0.88	7964	-0.01 (0.03)	0.767	57
8	20608550	rs2597398	T/G	0.46	22031	-0.02 (0.01)	0.130	37	0.49	7964	0.004 (0.02)	0.814	0
8	72378104	rs12549065	T/C	0.61	20997	-0.01 (0.02)	0.589	0	0.66	7964	-0.003 (0.02)	0.889	46
9	91405458	rs1571536	T/C	0.48	22374	0.002 (0.01)	0.899	14	0.57	7964	-0.02 (0.02)	0.253	0
9	139632376	rs1045777	A/G	0.21	19954	0.04 (0.02)	0.013	0	0.55	7964	-0.01 (0.02)	0.758	43
10	16645488	rs11254079	A/G	0.90	30014	-0.02 (0.02)	0.184	25	n/a[b]	n/a	n/a	n/a	n/a
11	27636492	rs6265	T/C	0.18	30062	-0.03 (0.01)	0.073	0	0.07	7963	-0.05 (0.04)	0.252	0
11	27650817	rs2049045	C/G	0.18	30046	-0.02 (0.01)	0.105	8	0.04	4228	-0.05 (0.08)	0.561	0
11	27704247	rs12288512	A/G	0.21	26289	0.03 (0.02)	0.093	0	0.11	7962	-0.0002 (0.03)	0.995	3
11	79095197	rs2264517	A/C	0.31	21360	-0.01 (0.01)	0.507	32	0.46	7954	0.04 (0.02)	0.059	56
11	132761771	rs4245116	T/C	0.47	21327	0.03 (0.01)	0.063	0	0.63	7964	-0.004 (0.02)	0.833	7
12	117272386	rs17512574	T/C	0.16	22362	0.02 (0.02)	0.249	0	0.15	7936	0.02 (0.03)	0.442	0
12	117926875	rs11069228	A/G	0.15	18261	-0.01 (0.02)	0.725	0	0.13	7964	-0.02 (0.03)	0.464	0
15	72806502	rs2470893	T/C	0.32	17344	0.09 (0.01)	9.92E-11	60	0.06	4466	0.20 (0.07)	4.23E-03	28
15	72814933	rs2472297	T/C	0.26	23602	0.11 (0.01)	3.26E-16	50	0.06	5048	0.19 (0.05)	8.62E-05	13
17	25373221	rs9902453	A/G	0.53	28089	-0.03 (0.01)	9.13E-03	19	0.80	7960	-0.04 (0.03)	0.174	6
17	25555919	rs3794808	T/C	0.43	30036	0.02 (0.01)	0.047	24	0.39	7964	0.04 (0.02)	0.032	5
17	31939244	rs17560870	A/G	0.59	22306	-0.0002 (0.01)	0.990	46	0.86	7959	-0.04 (0.03)	0.249	0
17	61762419	rs12450534	A/G	0.25	30056	-0.02 (0.01)	0.216	15	0.30	7964	0.002 (0.02)	0.916	5
19	4090773	rs164631	A/G	0.72	17564	-0.004 (0.02)	0.817	0	0.58	7964	-0.01 (0.02)	0.715	13
19	5239875	rs17676218	T/C	0.23	13786	-0.005 (0.02)	0.792	0	0.13	4989	-0.02 (0.03)	0.548	0
22	23189667	rs9620388	A/C	0.91	22219	0.004 (0.02)	0.861	37	0.87	7947	-0.02 (0.03)	0.522	12
22	23277452	rs738820	T/C	0.84	21317	0.01 (0.02)	0.780	16	0.51	7964	0.01 (0.02)	0.530	35
22	39532082	rs138312	T/C	0.50	22376	-0.02 (0.01)	0.159	0	0.43	7962	0.01 (0.02)	0.565	0
22	39745590	rs4821981	T/G	0.40	29906	-0.02 (0.01)	0.179	48	0.41	7947	0.005 (0.02)	0.810	0
22	39985654	rs2235852	T/G	0.31	22373	-0.01 (0.01)	0.389	26	0.29	7850	-0.01 (0.02)	0.669	56

EA, effect allele; NEA, non-effect allele; EAF, effect allele frequency; n/a, not available

[a]Presented are SNP-level sample size ('N'), beta coefficients and standard errors ('β (SE)', representing cups/day per effect allele), P values and I^2 statistic for heterogeneity from race-specific meta-analyses of stage 2 studies.

[b]rs11254079 was not available in YRI HapMap R22.

Table S11. Stage 2 results from meta-analyses of high vs. no/low coffee consumption (phenotype 2) for SNPs selected for follow-up[a]

Chr	Position (Hg18)	Best-SNP	EA/NEA	Stage 2 Europeans					Stage 2 African Americans				
				EAF	N	β (SE)	P	I^2	EAF	N	β (SE)	P	I^2
1	171939422	rs6681766	A/G	0.20	17655	0.003 (0.02)	0.879	0	0.15	3589	-0.06 (0.11)	0.596	31
2	1001316	rs6548172	C/G	0.87	14697	-0.02 (0.02)	0.491	0	0.88	3578	-0.06 (0.11)	0.596	0
2	2370732	rs713347	T/G	0.56	14762	0.02 (0.02)	0.201	10	0.48	3589	-0.01 (0.08)	0.921	0
2	27584444	rs1260326	T/C	0.36	17651	-0.01 (0.02)	0.470	35	0.17	3589	-0.08 (0.11)	0.460	0
2	27693043	rs2068834	T/C	0.76	17653	-0.02 (0.02)	0.401	38	0.76	3587	0.08 (0.10)	0.415	0
4	89258106	rs1481012	A/G	0.91	17656	0.06 (0.03)	0.070	0	0.95	3589	0.24 (0.21)	0.251	0
4	106768545	rs10007278	T/G	0.15	14685	0.03 (0.02)	0.164	0	0.57	3587	0.07 (0.08)	0.400	42
4	109081710	rs10461142	T/C	0.67	14761	-0.01 (0.02)	0.680	0	0.61	3589	-0.03 (0.08)	0.662	0
6	119477071	rs649979	C/G	0.12	13562	-0.06 (0.05)	0.211	35	0.04	3588	0.23 (0.21)	0.271	0
6	145203694	rs4895657	A/G	0.71	13987	-0.01 (0.02)	0.634	6	0.11	2488	-0.14 (0.19)	0.460	0
6	159376363	rs4709267	A/G	0.91	14737	0.06 (0.03)	0.037	0	0.84	3589	0.08 (0.11)	0.464	0
7	17163382	rs17137304	A/G	0.90	16368	0.10 (0.05)	0.043	17	0.94	3589	-0.18 (0.16)	0.266	0
7	17251102	rs4410790	T/C	0.41	14763	-0.07 (0.02)	6.80E-05	56	0.52	3589	-0.21 (0.09)	0.014	0
7	17253631	rs6968554	A/G	0.35	17623	-0.09 (0.02)	3.94E-08	61	0.33	3589	-0.19 (0.09)	0.048	62
7	17409204	rs1077773	A/G	0.57	14761	-0.01 (0.02)	0.763	0	0.39	3589	0.11 (0.08)	0.198	0
7	31638661	rs10235961	A/T	0.59	14728	-0.03 (0.02)	0.163	0	0.47	3589	-0.11 (0.08)	0.166	54
7	72673793	rs7800944	T/C	0.68	14209	-0.03 (0.02)	0.136	15	0.67	3589	-0.05 (0.09)	0.598	0
7	75454041	rs17685	A/G	0.27	14759	0.05 (0.02)	6.05E-03	30	0.19	3589	0.05 (0.10)	0.623	0
7	144693673	rs10227393	C/G	0.77	13444	0.003 (0.04)	0.949	0	0.88	3589	-0.24 (0.12)	0.039	0
8	20608550	rs2597398	T/G	0.44	14453	0.002 (0.02)	0.889	49	0.49	3589	0.08 (0.08)	0.299	0
8	72378104	rs12549065	T/C	0.65	13625	-0.003 (0.02)	0.866	0	0.66	3589	0.07 (0.09)	0.425	72
9	91405458	rs1571536	T/C	0.45	14766	-0.02 (0.02)	0.272	0	0.57	3589	-0.01 (0.08)	0.935	0
9	139632376	rs1045777	A/G	0.21	13319	0.01 (0.02)	0.729	57	0.56	3589	-0.04 (0.09)	0.627	59
10	16645488	rs11254079	A/G	0.88	17639	-0.05 (0.02)	0.058	42	n/a[b]	n/a	n/a	n/a	n/a
11	27636492	rs6265	T/C	0.17	17657	-0.03 (0.02)	0.158	0	0.06	3589	-0.24 (0.20)	0.222	66
11	27650817	rs2049045	C/G	0.16	17648	-0.03 (0.03)	0.174	0	0.04	2488	0.11 (0.26)	0.664	49
11	27704247	rs12288512	A/G	0.20	16624	0.002 (0.02)	0.938	0	0.11	3589	0.13 (0.13)	0.287	0
11	79095197	rs2264517	A/C	0.31	13936	0.01 (0.02)	0.672	0	0.47	3585	-0.04 (0.08)	0.593	80
11	132761771	rs4245116	T/C	0.42	14510	-0.01 (0.02)	0.455	1	0.63	3589	-0.16 (0.08)	0.050	0
12	117272386	rs17512574	T/C	0.16	14758	0.002 (0.02)	0.915	0	0.15	3569	0.11 (0.10)	0.298	0
12	117926875	rs11069228	A/G	0.15	13565	-0.04 (0.04)	0.411	0	0.13	3589	-0.11 (0.12)	0.366	0
15	72806502	rs2470893	T/C	0.32	12832	0.19 (0.03)	2.08E-08	0	0.05	2488	0.57 (0.27)	0.036	11
15	72814933	rs2472297	T/C	0.26	14453	0.22 (0.03)	3.10E-10	0	0.07	2100	0.34 (0.19)	0.080	82
17	25373221	rs9902453	A/G	0.53	17089	0.01 (0.02)	0.429	0	0.80	3589	-0.16 (0.11)	0.147	24
17	25555919	rs3794808	T/C	0.42	17642	0.003 (0.02)	0.858	0	0.39	3589	0.15 (0.08)	0.077	14
17	31939244	rs17560870	A/G	0.59	14707	0.02 (0.02)	0.353	0	0.87	3586	-0.08 (0.13)	0.518	0
17	61762419	rs12450534	A/G	0.25	17651	-0.03 (0.02)	0.092	0	0.30	3589	-0.08 (0.09)	0.354	0
19	4090773	rs164631	A/G	0.72	12979	-0.01 (0.04)	0.822	0	0.58	3589	0.04 (0.08)	0.648	0
19	5239875	rs17676218	T/C	0.23	8230	0.04 (0.04)	0.328	0	0.13	2069	0.20 (0.13)	0.141	0
22	23189667	rs9620388	A/C	0.89	14629	0.004 (0.02)	0.868	11	0.86	3579	-0.09 (0.10)	0.375	0
22	23277452	rs738820	T/C	0.85	14497	0.01 (0.02)	0.592	0	0.50	3589	0.02 (0.08)	0.830	4
22	39532082	rs138312	T/C	0.51	14766	-0.04 (0.02)	0.012	0	0.43	3589	-0.02 (0.08)	0.823	14
22	39745590	rs4821981	T/G	0.37	17523	-0.03 (0.02)	0.095	5	0.40	3582	-0.04 (0.08)	0.653	0
22	39985654	rs2235852	T/G	0.29	14764	-0.02 (0.02)	0.169	0	0.29	3554	-0.01 (0.09)	0.930	79

EA, effect allele; NEA, non-effect allele; EAF, effect allele frequency; n/a, not available

[a]Presented are SNP-level sample size ('N'), beta coefficients and standard errors ('β (SE)'), P values and I^2 statistic for heterogeneity from race-specific meta-analyses of stage 2 studies.

[b]rs11254079 was not available in YRI HapMap R22.

Table S12. African American gene-region meta-analysis of coffee consumption (cups/d, phenotype 1)[a]

Chr	Gene	Gene-level P	Best-SNP	Position (Hg18)	Best-SNP P
A priori genes tested[b]					
1	*KLHL20*	0.13	rs2068871	172062792	2.75E-03
2	*ZNF512*	0.86	rs11901534	27874494	0.072
4	*CYP2U1*	0.16	rs1313649	109063396	0.025
12	*TAOK3*	0.35	rs4767669	117312838	6.13E-03
19	*CREB3L3*	0.74	rs8109965	4147014	0.044
22	*RBX1*	0.42	rs5758179	39736865	9.93E-03
GRAIL's significant candidate genes tested					
7	*AHR*	3.760E-04	rs17137472	17262597	5.49E-05
7	*POR*	0.031	rs6965343	75430861	7.10E-03
15	*CYP1A2*	0.005	rs2472297	72814933	8.62E-05
All genes in region tested (results from most significant gene shown)					
2	*SNTGT*	0.42	rs4854428	904015	0.018
2	*MYT1L*	0.27	rs1978703	2006354	4.19E-03
2	*SLC5A6*	0.51	rs934986	27258350	0.014
4	*ABCG2*	0.008	rs1481012	89258106	1.27E-03
4	*INTS12*	0.031	rs4698950	106813617	1.31E-03
6	*FAM184A*	0.60	rs808031	119502415	0.016
6	*UTRN*	0.12	rs1544157	145198244	4.30E-04
6	*RSPH3,TAGAP*	0.19	rs182429	159389562	0.062
7	*CCDC129*	0.23	rs12701087	31481376	0.010
7	*BCL7B*	0.12	rs10275549	72552747	0.047
7	*CNTNAP2*	0.26	rs10244661	146941232	2.14E-03
8	*EYA1*	0.058	rs17785083	72411771	1.45E-04
9	*SEMA4D,GADD45G*	0.054	rs17054918	91272997	5.03E-04
9	*ZMYND19,ARRDC*	0.009	rs1328921	139653450	7.66E-04
10	*RSU1*	0.012	rs17333493	16859862	4.49E-04
11	*BDNF*	0.24	rs7933739	27748599	9.56E-03
11	*ODZ*	0.13	rs7482532	78557348	2.92E-03
11	*OPCML*	0.58	rs10791242	132001509	3.19E-03
12	*KIAA1853*	0.26	rs9651910	118016537	8.82E-03
17	*SLC6A4*	0.024	rs2020936	25574940	2.94E-03
17	*MRM*	0.059	rs17138364	31984652	0.015
17	*APOH*	0.19	rs12452188	61610860	0.022
19	*PTPRS*	0.11	rs11085118	5165158	0.016
22	*ADORA2A*	0.004	rs11704465	23192597	1.01E-03
22	*ADORA2A*	0.010	rs11704465	23192597	1.01E-03

[a]Results from gene-based tests of all SNPs in regions selected for follow-up from stage 1.
[b]Significant gene-level tests from stage 1 gene-level meta-analysis.

Table S13. African American gene-region meta-analysis of high vs. no/low coffee consumption (phenotype 2)[a]

Chr	*Gene*	Gene-level P	Best-SNP	Position	Best-SNP P
A priori genes tested[b]					
1	*KLHL20*	0.22	rs2068871	172062792	0.017
2	*ZNF512*	0.42	rs1728918	27488967	0.22
4	*CYP2U1*	0.57	rs4373225	109109572	0.18
12	*TAOK3*	0.071	rs1051470	117067615	0.013
19	*CREB3L3*	0.79	rs352509	4145090	0.42
22	*RBX1*	0.57	rs1109003	39725163	0.22
GRAIL's significant candidate genes tested					
7	*AHR*	0.024	rs2282885	17312139	0.011
7	*POR*	0.57	rs757589	75333460	0.16
15	*CYP1A2*	0.19	rs2472297	72814933	0.080
All genes in region tested (results from most significant gene shown)					
2	*SNTGT*	0.22	rs7606829	1086540	6.57E-03
2	*MYT1L*	0.14	rs7586598	2260315	8.20E-03
2	*TRIM54*	0.070	rs13404327	27372657	0.037
4	*PKD2*	0.097	rs10516798	89102561	2.41E-03
4	*FLJ20184*	0.024	rs3960769	106671248	0.010
6	*FAM184A*	0.29	rs4946399	119470457	0.034
6	*UTRN*	0.50	rs12664200	144786537	0.034
6	*RSPH3*	0.16	rs4709252	159300372	0.062
7	*CCDC129*	0.12	rs6973479	31674522	3.11E-03
7	*BCL7B ,MLXIPL,,TBL2*	0.13	rs2286276	72625290	0.087
7	*CNTNAP2*	8.370E-03	rs826802	146666000	9.02E-04
8	*EYA1*	0.081	rs2890502	72337071	2.46E-03
9	*GADD45G*	0.14	rs11265853	91453169	0.015
9	*EHMT*	0.16	rs3123501	139723379	0.024
10	*RSU1*	0.044	rs11254088	16654041	3.47E-03
11	*LIN7C*	0.096	rs16917069	27482700	0.040
11	*ODZ*	0.072	rs490195	78716905	2.36E-03
11	*OPCML*	0.090	rs11223060	131778763	5.20E-03
12	*KIAA1853*	0.29	rs1917879	117934996	0.027
17	*SLC6A4*	0.025	rs7223821	25603446	9.25E-03
17	*MGC4172, MRM1,GGNBP2*	0.042	rs10468612	32056847	0.011
17	*PRKCA*	0.34	rs8068129	62085771	0.012
19	*FUT6*	0.035	rs10426709	5816645	0.012
22	*ADORA2A*	0.010	rs5760444	23208218	2.09E-03
22	*ADORA2A*	0.018	rs5760444	23208218	2.09E-03

[a]Results from gene-based tests of all SNPs in regions selected for follow-up from Stage 1.
[b]Significant gene-level tests from Stage 1 gene-level meta-analysis.

Table S14. Stage 1 and stage 2 trans-ethnic meta-analysis of coffee consumption (cups/d, phenotype 1)

Chr	Position (Hg18)	Best-SNP	EA/NEA	Trans-ethnic meta-analysis[a]			Random effects meta-analysis[b]		
				N	Log10BF	Post Prob	β (SE)	P	I^2
1	171939422	rs6681766	A/G	126090	2.92	0.07	-0.03 (0.01)	1.01E-04	2
2	1001316	rs6548172	C/G	121128	0.45	0.17	-0.02 (0.01)	0.04	0
2	2370732	rs713347	T/G	121792	3.77	0.18	0.03 (0.01)	5.12E-04	23
2	27584444	rs1260326	T/C	129417	6.48	0.07	-0.04 (0.01)	7.14E-08	5
2	27693043	rs2068834	T/C	128841	3.37	0.06	0.03 (0.01)	3.46E-05	0
4	89258106	rs1481012	A/G	126019	6.08	0.23	0.06 (0.01)	8.93E-08	5
4	106768545	rs10007278	T/G	118764	4.69	0.14	0.04 (0.01)	2.52E-05	19
4	109081710	rs10461142	T/C	121270	0.25	0.08	0.01 (0.01)	0.18	18
6	119477071	rs649979	C/G	113910	2.58	0.23	0.04 (0.01)	1.85E-03	13
6	145203694	rs4895657	A/G	114407	0.08	0.07	0.01 (0.01)	0.09	6
6	159376363	rs4709267	A/G	118828	0.82	0.46	-0.02 (0.01)	0.15	28
7	17163382	rs17137304	A/G	119323	5.41	0.28	0.05 (0.01)	5.25E-07	2
7	17251102	rs4410790	T/C	116674	58.87	0.96	-0.10 (0.01)	3.08E-17	60
7	17253631	rs6968554	A/G	124849	69.69	1.00	-0.10 (0.01)	5.23E-17	65
7	17409204	rs1077773	A/G	120108	4.75	0.16	-0.03 (0.01)	1.69E-05	12
7	31638661	rs10235961	A/T	118192	2.83	0.14	0.03 (0.01)	1.86E-04	8
7	72673793	rs7800944	T/C	116417	8.83	0.09	-0.05 (0.01)	2.29E-11	0
7	75454041	rs17685	A/G	115465	15.12	0.08	0.07 (0.01)	4.26E-11	31
7	144693673	rs10227393	C/G	114093	0.80	0.07	-0.02 (0.01)	0.11	18
8	20608550	rs2597398	T/G	120923	3.87	0.09	-0.03 (0.01)	2.72E-05	5
8	72378104	rs12549065	T/C	117481	3.03	0.11	-0.03 (0.01)	6.77E-05	0
9	91405458	rs1571536	T/C	119644	3.37	0.09	-0.03 (0.01)	3.05E-05	0
9	139632376	rs1045777	A/G	114338	3.95	0.21	0.04 (0.01)	2.43E-05	9
10	16645488	rs11254079[c]	A/G	115171	2.78	0.11	-0.04 (0.01)	5.58E-03	24
11	27636492	rs6265	T/C	127828	5.76	0.10	-0.04 (0.01)	2.69E-06	12
11	27650817	rs2049045	C/G	123553	5.54	0.12	-0.04 (0.01)	2.03E-05	20
11	27704247	rs12288512	A/G	121969	5.24	0.07	0.04 (0.01)	5.86E-07	2
11	79095197	rs2264517	A/C	117297	3.27	0.10	0.03 (0.01)	9.05E-04	19
11	132761771	rs4245116	T/C	116411	4.58	0.09	0.03 (0.01)	1.76E-06	0
12	117272386	rs17512574	T/C	121525	2.84	0.11	0.03 (0.01)	7.39E-04	21
12	117926875	rs11069228	A/G	114749	-0.10	0.11	0.01 (0.01)	0.34	5
15	72806502	rs2470893	T/C	113273	57.79	1.00	0.12 (0.01)	2.72E-19	57
15	72814933	rs2472297	T/C	116272	62.77	0.97	0.14 (0.01)	2.47E-24	56
17	25373221	rs9902453	A/G	126819	6.29	0.05	-0.03 (0.01)	2.44E-08	0
17	25555919	rs3794808	T/C	129422	4.77	0.05	0.03 (0.01)	2.73E-07	0
17	31939244	rs17560870	A/G	121116	2.12	0.12	0.02 (0.01)	3.98E-03	20
17	61762419	rs12450534	A/G	126535	1.41	0.08	-0.02 (0.01)	0.05	18
19	4090773	rs164631	A/G	113047	2.55	0.10	-0.03 (0.01)	1.81E-04	0
19	5239875	rs17676218	T/C	95778	0.85	0.25	0.02 (0.01)	0.02	0
22	23189667	rs9620388	A/C	121397	2.87	0.19	-0.04 (0.01)	5.04E-03	29
22	23277452	rs738820	T/C	111101	2.69	0.15	0.04 (0.01)	5.15E-04	17
22	39532082	rs138312	T/C	121798	1.55	0.09	-0.02 (0.01)	9.94E-04	0
22	39745590	rs4821981	T/G	129185	2.90	0.06	-0.02 (0.01)	2.93E-03	27
22	39985654	rs2235852	T/G	119223	4.29	0.07	-0.03 (0.01)	1.28E-04	17

EA, effect allele; NEA, non-effect allele; EAF, effect allele frequency; BF, Bayes factor, Post Prob, posterior probability

[a]Results from trans-ethnic meta-analysis of all stage 1 and 2 studies. Presented is SNP-level sample size ('N'), Log10 Bayes-factor ('Log10BF') and posterior probabilities ('Post Prob').

[b]Results from random-effects meta-analysis of all stage 1 and 2 studies. Presented are SNP-level beta coefficients and standard errors ('β (SE)', representing cups/day per effect allele), P values and I^2 statistic for heterogeneity.

[c]rs11254079 was not available in YRI HapMap R22, thus results are based on European studies only.

Table S15. Stage 1 and stage 2 trans-ethnic meta-analysis of high vs. no/low coffee consumption (phenotype 2)

Chr	Position (Hg18)	Best-SNP	EA/EAF	Transethnic Meta-Analysis[a]			Random Effects Meta-analysis[b]		
				N	Log10BF	Post Prob	β (SE)	P^1	I^2
1	171939422	rs6681766	A/G	67236	0.46	0.12	-0.04 (0.02)	0.02	3
2	1001316	rs6548172	C/G	65155	2.81	0.20	-0.07 (0.02)	7.53E-05	0
2	2370732	rs713347	T/G	65389	0.50	0.15	0.03 (0.02)	0.09	36
2	27584444	rs1260326	T/C	68261	1.26	0.31	-0.04 (0.01)	2.48E-03	13
2	27693043	rs2068834	T/C	68093	-0.18	0.27	0.04 (0.02)	3.34E-02	24
4	89258106	rs1481012	A/G	67231	2.40	0.17	0.07 (0.02)	2.64E-04	0
4	106768545	rs10007278	T/G	64389	4.63	0.56	0.10 (0.02)	1.06E-05	20
4	109081710	rs10461142	T/C	65502	1.69	0.64	0.06 (0.02)	1.40E-03	34
6	119477071	rs649979	C/G	63129	0.24	0.19	0.04 (0.03)	0.17	9
6	145203694	rs4895657	A/G	62597	1.55	0.17	0.05 (0.02)	4.57E-03	21
6	159376363	rs4709267	A/G	64449	2.24	0.98	-0.06 (0.03)	1.30E-02	16
7	17163382	rs17137304	A/G	64896	4.60	0.45	0.11 (0.02)	2.47E-06	0
7	17251102	rs4410790	T/C	63158	41.58	1.00	-0.20 (0.03)	3.36E-14	61
7	17253631	rs6968554	A/G	67033	45.78	1.00	-0.20 (0.03)	7.41E-15	63
7	17409204	rs1077773	A/G	64197	1.73	0.24	-0.05 (0.02)	7.96E-03	25
7	31638661	rs10235961	A/T	64249	0.15	0.29	0.03 (0.02)	0.07	28
7	72673793	rs7800944	T/C	63782	2.65	0.38	-0.06 (0.02)	2.55E-04	17
7	75454041	rs17685	A/G	63066	7.96	0.19	0.08 (0.01)	1.13E-09	1
7	144693673	rs10227393	C/G	63265	3.82	0.13	-0.08 (0.02)	9.07E-06	0
8	20608550	rs2597398	T/G	64946	0.35	0.15	-0.03 (0.02)	0.03	21
8	72378104	rs12549065	T/C	63331	0.47	0.14	-0.03 (0.01)	0.05	0
9	91405458	rs1571536	T/C	65359	1.96	0.12	-0.04 (0.01)	8.32E-04	0
9	139632376	rs1045777	A/G	63026	0.01	0.23	0.03 (0.03)	0.27	46
10	16645488	rs11254079[c]	A/G	62479	4.48	0.55	-0.10 (0.03)	4.78E-04	31
11	27636492	rs6265	T/C	67112	5.26	0.39	-0.08 (0.02)	1.35E-04	21
11	27650817	rs2049045	C/G	65867	4.99	0.35	-0.08 (0.02)	3.12E-04	20
11	27704247	rs12288512	A/G	65175	0.84	0.22	0.05 (0.02)	4.28E-03	10
11	79095197	rs2264517	A/C	63500	1.01	0.15	0.03 (0.01)	0.02	4
11	132761771	rs4245116	T/C	63350	0.12	0.27	0.02 (0.02)	0.12	14
12	117272386	rs17512574	T/C	64355	1.12	0.29	0.05 (0.02)	0.01	20
12	117926875	rs11069228	A/G	63275	1.78	0.15	0.07 (0.02)	7.54E-04	0
15	72806502	rs2470893	T/C	62360	36.88	0.83	0.20 (0.02)	5.05E-19	31
15	72814933	rs2472297	T/C	61767	38.30	0.96	0.23 (0.03)	3.06E-18	37
17	25373221	rs9902453	A/G	67519	2.16	0.95	-0.05 (0.01)	1.64E-03	17
17	25555919	rs3794808	T/C	68249	3.34	0.78	0.05 (0.01)	3.53E-05	5
17	31939244	rs17560870	A/G	65164	2.53	0.16	0.04 (0.01)	2.35E-04	0
17	61762419	rs12450534	A/G	67355	3.44	0.13	-0.06 (0.01)	1.28E-05	0
19	4090773	rs164631	A/G	62175	0.89	0.17	-0.04 (0.02)	0.02	2
19	5239875	rs17676218	T/C	52349	4.09	0.12	0.09 (0.02)	2.93E-06	0
22	23189667	rs9620388	A/C	65153	1.03	0.60	-0.07 (0.02)	3.30E-03	13
22	23277452	rs738820	T/C	61784	0.34	0.22	0.05 (0.02)	2.40E-03	0
22	39532082	rs138312	T/C	65398	4.03	0.10	-0.05 (0.01)	2.76E-06	0
22	39745590	rs4821981	T/G	68079	2.16	0.12	-0.05 (0.01)	1.78E-05	0
22	39985654	rs2235852	T/G	63831	2.81	0.14	-0.05 (0.01)	1.04E-04	0

EA, effect allele; NEA, non-effect allele; EAF, effect allele frequency; BF, Bayes-factor, Post Prob, posterior probability

[a]Results from trans-ethnic meta-analysis of all stage 1 and 2 studies. Presented is SNP-level sample size ('N'), Log10 Bayes-factor ('Log10BF') and posterior probabilities ('Post Prob').

[b]Results from random-effects meta-analysis of all stage 1 and 2 studies. Presented are SNP-level beta coefficients and standard errors ('β (SE)'), P values and I^2 statistic for heterogeneity.

[c]rs11254079 was not available in YRI HapMap R22, thus results are based on European studies only.

Table S16. Fine mapping of coffee consumption (cups/d, phenotype 1) loci using MANTRA[a]

Lead SNP[b]	Chr	Position (Hg18)	European ancestry only (N≤91,464)						European ancestry and African Americans (N≤99,421)					
			Log10 BF	Post prob	95% set, #SNPs	95% interval, #bp	99% set, #SNPs	99% interval, #bp	Log10 BF	Post prob	95% set, #SNPs	95% interval, #bp	99% set, #SNPs	99% interval, #bp
rs1260326[c]	2	27584444	5.63	0.07	3	11663	5	17684	5.38	0.07	3	11663	5	17684
rs1481012	4	89258106	4.82	0.08	10	189102	29	204945	6.31	0.17	2	6249	7	188989
rs10007278	4	106768545	4.32	0.27	37	371922	69	686322	4.03	0.27	39	371922	67	686322
rs6968554[d]	7	17253631	57.14	0.97	1	0	1	0	59.91	1	2	2529	2	2529
rs7800944[e]	7	72673793	6.95	0.16	9	179588	13	179588	6.54	0.16	9	179588	13	179588
rs17685	7	75454041	11.91	0.09	3	20135	5	161729	13.23	0.11	1	0	2	20135
rs6265	11	27636492	5.33	0.22	13	110786	19	110786	5.69	0.14	9	90882	17	110786
rs2472297	15	72814933	48.24	0.99	1	0	1	0	50.98	0.99	1	0	1	0
rs9902453	17	25373221	4.42	0.07	57	615044	85	615044	4.85	0.07	54	603027	78	615044

Post prob, posterior probability. BF, Bayes-factor

[a]Applying the fine mapping approach described by Franceschini et al [38] and additionally summarized in Methods, we defined 95 (columns 6 and 12) and 99% (columns 8 and 14) ‘‘credible sets’’ of SNPs with the strongest signals of association on the basis of the European-only trans-ethnic meta-analysis and then after inclusion of the African American studies. For each locus, high quality SNPs ±500kb of lead SNP with results available from Stage 1 European *and* Stage 2 African American studies were considered for fine-mapping analysis.

[b]From European Ancestry Stage 1. The same lead SNPs were identified in trans-ethnic analysis unless noted below.

Lead SNPs in trans-ethnic analysis:

[c]rs780094: log10BF=5.62, post prob=0.05

[d]rs4410790: log10BF=60.77, post prob=0.88

[e]rs2074755: log10BF=7.09, post prob=0.51

Table S17. Significant[a] correlations between coffee-consumption associated SNPs and tissue-specific gene expression

Locus	SNP[b]	EA[c]	Tissue	eQTL	Probe	Effect	P value[a]	Reference
2p24	**rs1260326**	C	blood	*SNX17*	3360468	-	1.27E-11	63
			blood	*EIF2B4*	5960546	+	1.12E-04	63
			blood	*NRBP1*	430239	-	1.12E-04	63
			LCL	*EIF2B4*	ILMN_2356672	+	3.63E-04	54
7q11.23	**rs7800944**	C	blood	*WBSCR22*	6370538	+	3.47E-12	63
			adipose	*MLXIPL*	ILMN_2399919	+	1.58E-06	54
			adipose	*MLXIPL*	ILMN_1722073	+	4.93E-06	54
7q11.23	**rs17685**	A	blood	*RHBDD2*	510373	+	1.42E-44	63
			blood	*POR*	1230754	+	5.55E-32	63
			blood	*RHBDD2*	6650746	+	7.00E-32	63
			blood	*MDH2*	6420369	-	2.99E-10	63
			LCL	*TMEM120A*[d]	n/a	n/a	3.97E-05	52
			LCL	*STYXL1*	ILMN_22107290	+	7.00E-06	53
			LCL	*TMEM120A*	ILMN_1654516	+	3.89E-13	54
15q24	**rs2470893**	T	adipose	*SNUPN*	ILMN_2364535	-	3.38E-05	54
			adipose	*SNUPN*	ILMN_1733932	-	1.28E-04	54
			adipose	*RPP25*	ILMN_1695271	-	4.14E-04	54
			LCL	*ULK3*	ILMN_1679495	-	1.06E-05	54
			LCL	*CSK*	ILMN_1754121	-	8.73E-04	54
			blood	*CSK*	3170239	-	2.42E-23	63
			blood	*SCAMP2*	50341	+	1.43E-11	63
			blood	*ULK3*	4480132	-	8.35E-08	63
			blood	*MPI*	4010041	+	2.60E-05	63
	rs2472297	T	adipose	*ISLR*	ILMN_1747593	-	7.69E-04	54
			adipose	*SNUPN*	ILMN_2364535	-	9.90E-05	54
			adipose	*RPP25*	ILMN_1695271	-	4.02E-04	54
			LCL	*ULK3*	ILMN_1679495	-	1.52E-04	54
			blood	*CSK*	3170239	-	1.97E-21	63
			blood	*SCAMP2*	50341	+	3.42E-11	63
			blood	*ULK3*	4480132	-	2.13E-05	63
17q11.2	**rs9902453**	G	LCL	*GIT1*[d]	n/a	n/a	1.83E-04	52
	rs4328498	G	LCL-LWK	*ATAD5*	n/a	+	7.00E-04	55
	rs4465650	G	LCL-GIH	*SLC6A4*	n/a	-	8.00E-04	55

EA, effect allele; LCL, lymphoblastoid cell lines; n/a not available

[a]Statistical significance as defined by original publication (reported in column 8). All results reported by Westra *et al* [63] additionally met the false-discovery rate threshold of 0.05.

[b]Lead SNPs (bold) and perfect proxies (r^2=1)

[c]Variant associated with *increased* coffee consumption

[d]Exon QTL

Table S18. Significant[a] correlations between coffee-consumption associated SNPs and gene methylation in cerebellum and frontal cortex [64]

Locus	SNP	EA[b]	Tissue	Probe	mQTL	Effect	P value[a]
2p24	rs1260326	C	Cbm	Cg11618577	*KRTCAP3*	-	4.6E-03
			FrCtx	Cg11618577	*KRTCAP3*	-	1.8E-03
			FrCtx	Cg15296858	*PPM1G*	-	2.2E-03
7p21	rs6968554	G	Cbm	Cg13676215	*AHR*	+	4.2E-03
	rs4410790	C	Cbm	Cg13676215	*AHR*	+	2.0E-03
7q11.23	rs7800944	C	FrCtx	Cg18438300	*FZD9*	-	1.4E-03
7q11.23	rs17685	A	FrCtx	Cg06772202	*STYXL1*	-	1.1E-03
15q24	rs2470893	T	Cbm	Cg10253484	*SCAMP2*	-	1.4E-04
	rs2472297	T	Cbm	Cg10253484	*SCAMP2*	-	1.1E-03
17q11.2	rs9902453	G	Cbm	Cg26813908	*CCDC55*	+	1.6E-04
			Cbm	Cg06038133	*CORO6*	-	3.4E-03
			FrCtx	Cg02717570	*ANKRD13B*	+	1.9E-03
			FrCtx	Cg10394139	*CRLF3*	+	4.0E-03

EA, effect allele; Cbm, cerebellum; FrCtx, frontal cortex

[a]Significance was defined as P<0.00625 after correcting for number of SNPs tested (0.05/8 confirmed loci)

[b]Variant associated with increased coffee consumption

Table S19. Non-synonymous variants in linkage disequilibrium ($r^2 > 0.80$) with lead SNPs[48]

Chr	Lead SNP	Coding SNP	r^2 with lead SNP (HapMap CEU)	CEU MAF CEU/YRI/CHB	Gene	Coding change
2p24	rs1260326	rs1260326	1	.42/.10/.58	*GCKR*	Leu446Pro
4q22	rs1481012	rs2231142	0.92	.12/.00/.29	*ABCG2*	Gln141Lys
7p21	rs4410790 rs6968554	—	—	—	—	—
7q11.23	rs7800944	—	—	—	—	—
7q11.23	rs17685	rs1057868	0.93	.32/.10/.42	*POR*	Ala503Val
11p13	rs6265	rs6265	1	.20/.00/.42	*BDNF*	Val66Met
15q24	rs2470893 rs2472297	—	—	—	—	—
17q11.2	rs9902453	rs9897794	0.90	.48/.43/.71	*EFCAB5*	Leu237Val

MAF, minor allele frequency

Table S20. Candidate genes considered for biological inferences

Locus, Lead SNP	Gene ID	*Official Symbol* and Other Aliases	Gene Name
2p24 rs1260326	790	*CAD*	carbamoyl-phosphate synthetase 2, aspartate transcarbamylase, and dihydroorotase
	2498	*FTH1P3* FTHL3, FTHL3P	ferritin, heavy polypeptide-like 3 pseudogene
	2646	*GCKR* FGQTL5, GKRP	glucokinase (hexokinase 4) regulator
	2976	*GTF3C2* TFIIIC-BETA, TFIIIC110	general transcription factor IIIC, polypeptide 2, beta 110kDa
	4358	*MPV17* SYM1, MTDPS6	MpV17 mitochondrial inner membrane protein
	5496	*PPM1G* PP2CG, PPP2CG, PP2CGAMMA	protein phosphatase 1G (formerly 2C), magnesium-dependent, gamma isoform
	7349	*UCN* UI, UROC	Urocortin
	7781	*SLC30A3* ZNT3	solute carrier family 30 (zinc transporter), member 3
	8884	*SLC5A6* SMVT	solute carrier family 5 (sodium-dependent vitamin transporter), member 6
	8890	*EIF2B4* EIF2Bdelta, EIF2B, EIF-2B	eukaryotic translation initiation factor 2B, subunit 4 delta, 67kDa
	9784	*SNX17*	sorting nexin 17
	9913	*SUPT7L* STAF65G, SUPT7H, STAF65(gamma), STAF65, SPT7L	suppressor of Ty 7 (S. cerevisiae)-like
	10113	*PREB* *SEC12*	prolactin regulatory element binding
	11321	*GPN1* NTPBP, RPAP4, MBDIN, ATPBD1A, XAB1	GPN-loop GTPase 1
	22950	*SLC4A1AP* HLC3	solute carrier family 4 (anion exchanger), member 1, adaptor protein
	26160	*IFT172* wim, osm-1, SLB	intraflagellar transport 172 homolog (Chlamydomonas)
	29959	*NRBP1* NRBP, MADM, BCON3, MUDPNP	nuclear receptor binding protein 1
	51374	*ATRAID* C2orf28, HSPC013, APR-3, APR, p18, PRO240	all-trans retinoic acid-induced differentiation factor
	57159	*TRIM54* muRF3, MURF-3, RNF30, MURF	tripartite motif-containing 54
	64838	*FNDC4* FRCP1	fibronectin type III domain containing 4
	79635	*CCDC121*	coiled-coil domain containing 121
	84226	*C2orf16*	chromosome 2 open reading frame 16
	84450	*ZNF512* KIAA1805	zinc finger protein 512
	84696	*ABHD1* LABH1	abhydrolase domain containing 1
	130557	*ZNF513* RP58, HMFT0656	zinc finger protein 513
	150921	*TCF23* TCF-23, bHLHa24, OUT	transcription factor 23
	200634	*KRTCAP3* KCP3, PRO9898, MRV222	keratinocyte associated protein 3
	285126	*DNAJC5G* CSP-gamma	DnaJ (Hsp40) homolog, subfamily C, member 5 gamma
	339779	*C2orf53*	chromosome 2 open reading frame 53
4q22 rs1481012	5311	*PKD2* PC2, APKD2, Pc-2, TRPP2, PKD4	polycystic kidney disease 2 (autosomal dominant)
	6696	*SPP1* BSPI, OPN, ETA-1, BNSP	secreted phosphoprotein 1
	9429	*ABCG2* BCRP1, BCRP, MRX, EST157481, GOUT1, MXR, CDw338, CD338, ABC15, ABCP, UAQTL1, MXR1,	ATP-binding cassette, sub-family G (WHITE), member 2

		BMDP	
7p21 rs6968554	196	*AHR* bHLHe76	aryl hydrocarbon receptor
7q11.23.i rs7800944	8326	*FZD9* FZD3, CD349	frizzled homolog 9 (Drosophila)
	8468	*FKBP6* PPIase, FKBP36	FK506 binding protein 6, 36kDa
	9031	*BAZ1B* WSTF, WBSCR10, WBSCR9	bromodomain adjacent to zinc finger domain, 1B
	9275	*BCL7B*	B-cell CLL/lymphoma 7B
	26608	*TBL2* WBSCR13, WS-betaTRP	transducin (beta)-like 2
	51085	*MLXIPL* MIO, WS-bHLH, MONDOB, CHREBP, bHLHd14, WBSCR14	MLX interacting protein-like
	55695	*NSUN5* WBSCR20A, WBSCR20, p120, NOL1, NSUN5A, NOL1R	NOL1/NOP2/Sun domain family, member 5
	84277	*DNAJC30* WBSCR18	DnaJ (Hsp40) homolog, subfamily C, member 30
	114049	*WBSCR22* WBMT, HUSSY-3, HASJ4442, PP3381, MERM1	Williams Beuren syndrome chromosome region 22
	155382	*VPS37D* WBSCR24	vacuolar protein sorting 37 homolog D (S. cerevisiae)
7q11.23.ii rs17685	3315	*HSPB1* HMN2B, HS.76067, SRP27, CMT2F, HSP27, Hsp25, HSP28	heat shock 27kDa protein-like 2 pseudogene; heat shock 27kDa protein 1
	4191	*MDH2* MDH, MOR1, M-MDH, MGC:3559	malate dehydrogenase 2, NAD (mitochondrial)
	5447	*POR* CPR, CYPOR, P450R	P450 (cytochrome) oxidoreductase
	7532	*YWHAG* 14-3-3GAMMA	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, gamma polypeptide
	7784	*ZP3* Zp-3, ZPC, ZP3B, ZP3A	zona pellucida glycoprotein 3 (sperm receptor)
	51657	*STYXL1* MK-STYX, DUSP24	serine/threonine/tyrosine interacting-like 1
	57414	*RHBDD2* RHBDL7, NPD007	rhomboid domain containing 2
	83862	*TMEM120A* TMPIT, NET29	transmembrane protein 120A
	222183	*SRRM3* FLJ37078	serine/arginine repetitive matrix 3
	677801	*SNORA14A* ACA14a	small nucleolar RNA, H/ACA box 14A; small nucleolar RNA, H/ACA box 14B
11p13 rs6265	627	*BDNF* BULN2, ANON2	brain-derived neurotrophic factor
	55327	*LIN7C* VELI3, LIN-7C, LIN-7-C, MALS3, MALS-3	lin-7 homolog C (C. elegans)
	91057	*CCDC34* NY-REN-41, L15, RAMA3	coiled-coil domain containing 34
	497258	*BDNF-AS* ANTI-BDNF, BDNF, BDNF-AS1, BDNFOS, NCRNA00049	BDNF opposite strand (non-protein coding)
	196074	*METTL15* METT5D1	methyltransferase like 15
15q24 rs2472297	1198	*CLK3* PHCLK3/152, PHCLK3	CDC-like kinase 3
	1445	*CSK*	c-src tyrosine kinase
	1543	*CYP1A1* P450DX, P1-450, AHH, AHRR, P450-C, CYP1, CP11	cytochrome P450, family 1, subfamily A, polypeptide 1
	1544	*CYP1A2* CP12, P3-450, P450(PA)	cytochrome P450, family 1, subfamily A, polypeptide 2
	1583	*CYP11A1* CYPXIA1, CYP11A, P450SCC	cytochrome P450, family 11, subfamily A, polypeptide 1
	3671	*ISLR* HsT17563	immunoglobulin superfamily containing leucine-rich repeat
	4351	*MPI*	mannose phosphate isomerase

		CDG1B, PMI, PMI1	
	8482	*SEMA7A* H-Sema-L, CDw108, SEMAK1, CD108, JMH, H-SEMA-K1, SEMAL	semaphorin 7A, GPI membrane anchor (John Milton Hagen blood group)
	9377	*COX5A* VA, COX-VA, COX	cytochrome c oxidase subunit Va
	10066	*SCAMP2*	secretory carrier membrane protein 2
	10073	*SNUPN* Snurportin1, RNUT1, KPNBL	snurportin 1
	10620	*ARID3B* DRIL2, BDP	AT rich interactive domain 3B (BRIGHT-like)
	25989	*ULK3*	unc-51-like kinase 3 (C. elegans)
	54913	*RPP25*	ribonuclease P/MRP 25kDa subunit
	57184	*FAM219B* C15orf17	family with sequence similarity 219, member B
	60490	*PPCDC* MDS018	phosphopantothenoylcysteine decarboxylase
	79748	*LMAN1L* ERGL, ERGIC-53L	lectin, mannose-binding, 1 like
	80153	*EDC3* LSM16, YJDC, YJEFN2	enhancer of mRNA decapping 3 homolog (S. cerevisiae)
	84993	*UBL7* TCBA1, BMSC-UbP	ubiquitin-like 7 (bone marrow stromal cell-derived)
	192683	*SCAMP5*	secretory carrier membrane protein 5
	594855	*CPLX3* CPX-III, Nbla11589, CPXIII	complexin 3
17q11.2 rs9902453	642	*BLMH* BMH, BH	bleomycin hydrolase
	6532	*SLC6A4* SERT, 5HTT, SERT1, hSERT, 5-HTTLPR, HTT, 5-HTT, OCD1	solute carrier family 6 (neurotransmitter transporter, serotonin), member 4
	28964	*GIT1*	G protein-coupled receptor kinase interacting ArfGAP 1
	57551	*TAOK1* hKFC-B, hTAOK1, MAP3K16, KFC-B, MARKK, PSK2, PSK-2, TAO1	TAO kinase 1
	79915	*ATAD5* FRAG1, C17orf41, ELG1	ATPase family, AAA domain containing 5
	84081	*NSRP1* CCDC55, HSPC095, NSrp70	nuclear speckle splicing regulatory protein 1
	84940	*CORO6*	coronin 6
	85464	*SSH2* SSH-2	slingshot homolog 2 (Drosophila)
	90313	*TP53I13* DSCP1	tumor protein p53 inducible protein 13
	116236	*ABHD15*	abhydrolase domain containing 15
	124930	*ANKRD13B*	ankyrin repeat domain 13B
	374786	*EFCAB5*	EF-hand calcium binding domain 5

Table S21. Experimentally defined relationships among novel and a priori candidate[a] genes associated with coffee consumption

Node1 (source)	Edge	Node2 (target)	Mechanism	Pathway tools used to identify relationship	Reference
ADORA1[a]	inhibits	*YWHAG*	EM	MetaCore	[165]
ADORA1[a]	activates	*HSPB1*	EM	MetaCore	[166, 167]
ADORA2A[a]	inhibits	*SPP1*	EM	MetaCore	[168]
ADORA2A[a]	inhibits	*HSPB1*	EM	STITCH	[169]
ADORA2A[a]	unspecified	*PPP1R1B*[a]	Ph	STITCH	[170, 171]
AHR	activates	*CYP1A1*	TR	MetaCore, IPA	[172-194]
AHR	activates	*ABCG2*	TR	MetaCore, IPA	[195-199]
AHR	unspecified	*SLC5A6*	TR	MetaCore	[189, 200]
AHR	unspecified	*PPM1G*	TR	MetaCore	[201]
AHR	activates	*SPP1*	EM	MetaCore, IPA	[202]
AHR	unspecified	*LIN7C*	TR	MetaCore	[203]
AHR	unspecified	*ZNF513*	TR	MetaCore	[201]
AHR	unspecified	*LSM16*	TR	MetaCore	[189, 191]
AHR	activates	*CYP1A2*	TR	MetaCore, IPA	[180, 184, 186-189, 191, 192, 204-219]
AHR	unspecified	*PYGL*[a]	PPI	MetaCore	[189]
ATAD5	unspecified	*BAZ1B*	PPI	MetaCore, IPA	[220] (not in Fig 1.)
BDNF	activates	*DRD2*[a]	EM	MetaCore	[221]
BDNF	activates	*SLC6A4*	EM	MetaCore, IPA, STITCH	[222-224]
BDNF	activates	*PPP1R1B*[a]	EM	MetaCore, STITCH	[225-229]
BDNF	inhibits	*HSPB1*	EM	STITCH	[230]
BLMH	unspecified	*CYP2C8*[a]	PPI	MetaCore, IPA	[231] (not in Fig 1.)
CSK	unspecified	*PDE4A*[a]	PPI	MetaCore	[232]
CSK	unspecified	*MDH2*	Ph	MetaCore	[233]
CSK	unspecified	*GIT1*	PPI	MetaCore	[234]
PPP1R1B[a]	inhibits	*PPM1G*	PPI	MetaCore	[235]
DRD2[a]	activates	*BDNF*	EM	MetaCore	[236]
DRD2[a]	inhibits	*PPP1R1B*[a]	Ph	STITCH	[237]
TAOK1	unspecified	*YWHAG*	PPI	MetaCore	[238]
MLXIPL	activates	*GCKR*	TR	MetaCore	[239]
POR	unspecified	*CYP3A4*[a]	PPI	MetaCore, IPA	[240, 241]
POR	unspecified	*CYP1A2*	PPI	MetaCore, IPA, STRING	[242, 243]
POR	activates	*CYP2C9*[a]	CM	MetaCore, STITCH, IPA	[242, 244]
POR	activates	*CYP2E1*[a]	CM	MetaCore, IPA	[242, 245]
PYGL[a]	unspecified	*CSK*	PPI	MetaCore	[234]
SCAMP2	unspecified	*SLC6A4*	PPI	MetaCore, STITCH	[246]
YWHAG	unspecified	*MLXIPL*	PPI	MetaCore	[247]
YWHAG	unspecified	*GIT1*	PPI	MetaCore, IPA	[248]
YWHAG	unspecified	*PPM1G*	PPI	MetaCore	[249]
YWHAG	unspecified	*EDC3*	PPI	MetaCore, IPA	[238, 248]
YWHAG	unspecified	*CLK3*	PPI	MetaCore, IPA	[238, 248]

EM, expression modification; TR, transcription regulation; PPI, protein-protein interaction; Ph, phosphorylation; CM, covalent modification

[a]Genes encoding proteins known to be involved in caffeine metabolism/response [73-77, 250, 251].

Table S22. Liver (human), brain (human) and taste bud (primate) expression of candidate genes in confirmed regions associated with coffee consumption

Locus	Gene ID	Gene Symbol	**Liver Human RNA-seq[a]**	**Brain Human RNA-seq[a]**	**Brain Atlas Human Microarray[b]**	**Taste Buds Primate Microarray[c]**
2p24	790	*CAD*			CbCx, WM	
	2498	*FTH1P3*				
	2646	*GCKR*			OL	
	2976	*GTF3C2*			CbCx, WM	
	4358	*MPV17*			VT	
	5496	*PPM1G*				
	7349	*UCN*			MES	
	7781	*SLC30A3*				
	8884	*SLC5A6*				
	8890	*EIF2B4*				
	9784	*SNX17*				
	9913	*SUPT7L*				
	10113	*PREB*				
	11321	*GPN1*				
	22950	*SLC4A1AP*				
	26160	*IFT172*				
	29959	*NRBP1*			CbCx	
	51374	*ATRAID*			WM, GP	
	57159	*TRIM54*			OL	
	64838	*FNDC4*				
	79635	*CCDC121*			GP	
	84226	*C2orf16*				
	84450	*ZNF512*				
	84696	*ABHD1*			GP, WM	
	130557	*ZNF513*				
	150921	*TCF23*				
	200634	*KRTCAP3*			CbCx	
	285126	*DNAJC5G*			OL, PL	
	339779	*C2orf53*				
4q22	5311	*PKD2*			WM, VT	
	6696	*SPP1*			WM, CbN	
	9429	*ABCG2*			SbT, CbN	
7p21	196	*AHR*			WM	
7q11.23(i)	8326	*FZD9*			DT	
	8468	*FKBP6*				
	9031	*BAZ1B*			CbCx, GP	
	9275	*BCL7B*				
	26608	*TBL2*				
	51085	*MLXIPL*				
	55695	*NSUN5*				
	84277	*DNAJC30*				
	114049	*WBSCR22*			SbT, VT	
	155382	*VPS37D*				
7q11.23(ii)	3315	*HSPB1*				
	4191	*MDH2*				
	5447	*POR*				
	7532	*YWHAG*			Bpons	
	7784	*ZP3*				
	51657	*STYXL1*			WM, VT	
	57414	*RHBDD2*				
	83862	*TMEM120A*			Bpons	
	222183	*SRRM3*				

	677801	*SNORA14A*				
11p13	627	*BDNF*			CI, HiF	
	55327	*LIN7C*			CbCx	
	91057	*CCDC34*				
	497258	*BDNF-AS*				
	196074	*METTL15*			Bpons	
15q24	1198	*CLK3*			GP, WM	
	1445	*CSK*			CbCx, Bpons	
	1543	*CYP1A1*				
	1544	*CYP1A2*				
	1583	*CYP11A1*			Bpons, ET	
	3671	*ISLR*			CI	
	4351	*MPI*				
	8482	*SEMA7A*			Bpons	
	9377	*COX5A*			SbT	
	10066	*SCAMP2*			WM, VT	
	10073	*SNUPN*				
	10620	*ARID3B*			CbCx	
	25989	*ULK3*			CbCx, Bpons	
	54913	*RPP25*			SbT	
	57184	*FAM219B*				
	60490	*PPCDC*				
	79748	*LMAN1L*			PTg	
	80153	*EDC3*				
	84993	*UBL7*				
	192683	*SCAMP5*				
	594855	*CPLX3*			PHG, DT	
17q11.2	642	*BLMH*			CbCx, WM	
	6532	*SLC6A4*			MES, PTg	
	28964	*GIT1*			HiF	
	57551	*TAOK1*				
	79915	*ATAD5*			CbCx	
	84081	*NSRP1*			GP, WM	
	84940	*CORO6*				
	85464	*SSH2*			WM	
	90313	*TP53I13*				
	116236	*ABHD15*				
	124930	*ANKRD13B*				
	374786	*EFCAB5*				

[a]Expression level with respects to normalized and absolute log values of Illumina Human Body Map 2.0 data. Green (low): <log 500, Red (high): >log 500.
[b]Regions of the brain that highly (z-score >1.5) express the gene of interest. Source: Allen Human Brain Atlas[69]. FL, Frontal Lobe; Ins, Insula; CgG, Cingulate gyrus; HiF, hippocampal formation; PHG, parahippocampal gyrus; OL, Occipital Lobe; PL, Parietal Lobe; TL, Temporal Lobe; Amg, Amygdala; BF, Basal Forebrain; GP, Globus Pallidus; Str, Striatum; Cl, Claustrum; ET, Epithalamus; Hy, Hypothalamus; SbT, Subthalamus; DT, Dorsal Thalamus; VT, Ventral Thalamus; MES, Mesencephalon; CbCx, Cerebellar Cortex; CbN, Cerebellar Nuclei; Bpons, Basal Part of Pons; PTg, Pontine Tegmentum; MY, Myelencephalon; WM, White Matter;
[c]Red cells indicate 'taste bud gene' of the Rhesus Macaque as originally defined by Hevezi et al [70](GEO:GSE16485).

Table S23. Phenotyped mouse orthologs[a]

Human Gene Symbol[b]	Mouse Ortholog Gene Symbol	Mouse phenotype[c]	Ref.[4]
GCKR	Gckr	abnormal enzyme/ coenzyme level abnormal glucose homeostasis	
MPV17	Mpv17	hypoactivity (disease-related) abnormal blood homeostasis abnormal circulating protein level abnormal coat/hair pigmentation abnormal cochlea morphology abnormal glomerular capillary morphology abnormal hepatocyte morphology abnormal hypodermis muscle layer morphology abnormal liver lobule morphology abnormal liver sinusoid morphology abnormal mitochondrial crista morphology abnormal portal triad morphology abnormal renal glomerulus morphology abnormal renal tubule morphology abnormal renal/urinary system physiology abnormal scala media morphology abnormal sebaceous gland morphology abnormal skin morphology cachexia decreased body weight decreased brainstem auditory evoked potential decreased subcutaneous adipose tissue amount decreased urine osmolality expanded mesangial matrix fused podocyte foot processes glomerulosclerosis hypertension increased blood urea nitrogen level increased circulating alanine transaminase level increased circulating aspartate transaminase level increased circulating cholesterol level increased circulating creatine kinase level increased circulating creatinine level increased heart rate increased or absent threshold for auditory brainstem response podocyte foot process effacement premature death sensorineural hearing loss spiral ligament degeneration stria vascularis degeneration	
PPM1G	Ppm1g	decreased prepulse inhibition partial perinatal lethality	
UCN	Ucn	**abnormal emotion/affect behavior** **decreased brainstem auditory evoked potential** **abnormal startle reflex** **increased anxiety-related response** abnormal circulating corticosterone level abnormal distortion product otoacoustic emission increased thigmotaxis short cochlear outer hair cells	252-254
SLC30A3	Slc30a3	**abnormal hippocampal mossy fiber morphology** **abnormal synaptic vesicle morphology** **abnormal zinc homeostasis** **amyloid beta deposits**	255, 256
SLC5A6	Slc5a6	embryonic lethality	
SUPT7L	Supt7l	embryonic lethality preweaning lethality	

IFT172	Ift172	**abnormal brain morphology/development** **decreased motor neuron number** abnormal cardiovascular system morphology abnormal cell morphology abnormal craniofacial morphology abnormal diencephalon morphology abnormal direction of heart looping abnormal embryogenesis/ development abnormal embryonic cilium morphology abnormal embryonic neuroepithelium morphology abnormal floor plate morphology abnormal kidney morphology abnormal left-right axis patterning abnormal limb morphology abnormal mesendoderm development abnormal neural tube morphology/development abnormal outflow tract development abnormal renal glomerulus morphology abnormal rhombomere morphology abnormal spinal cord morphology abnormal telencephalon morphology abnormal trachea morphology anophthalmia atrioventricular septal defect complete atrioventricular septal defect complete embryonic lethality during organogenesis edema esophageal atresia esophagus hypoplasia esophagus stenosis exencephaly hemorrhage holoprosencephaly hydroencephaly perinatal lethality polydactyly preaxial polydactyly	[257-259]
NRBP1	Nrbp1	abnormal crypts of Lieberkuhn morphology abnormal enterocyte proliferation abnormal intestinal enteroendocrine cell morphology abnormal intestinal goblet cell morphology abnormal intestinal mucosa morphology abnormal intestine morphology abnormal liver morphology abnormal Paneth cell morphology complete embryonic lethality between implantation and somite formation distended stomach increased gastrointestinal tumor incidence increased lymphoma incidence increased tumor incidence leukemia lung carcinoma premature death	
ATRAID	Atraid	cataracts	
TRIM54	Trim54	abnormal cardiac muscle contractility abnormal heart left ventricle morphology abnormal heart morphology abnormal sarcomere morphology lethargy premature death	

PKD2	Pkd2	**abnormal liver morphology/development/function** abnormal digestive organ placement abnormal direction of embryo turning abnormal heart development abnormal hepatic vein morphology abnormal intrahepatic bile duct morphology abnormal left-right axis patterning abnormal kidney morphology/development/function abnormal lung development abnormal mitotic spindle morphology abnormal pancreas morphology abnormal papillary duct morphology abnormal placenta vasculature morphology abnormal primitive node morphology embryonic lethality during organogenesis lethality throughout fetal growth and development dextrocardia edema hemorrhage hydrops fetalis increased apoptosis increased blood urea nitrogen level mesocardia pericardial effusion polyhydramnios	[260-263]
SPP1	Spp1	**abnormal striatum morphology** **abnormal substantia nigra morphology** **decreased susceptibility to dopaminergic neuron neurotoxicity** abnormal liver morphology (not significant) abnormal physiological sensitivity to bleomycin decreased incidence of chemically-induced tumors hyporesponsive to tactile stimuli abnormal aorta morphology abnormal blood cell morphology/development abnormal blood flow velocity abnormal blood vessel morphology/physiology abnormal body weight abnormal bone mineralization abnormal cardiovascular system physiology abnormal cell chemotaxis abnormal chemokine level abnormal hematopoietic system physiology abnormal immune system physiology abnormal inflammatory response abnormal kidney morphology abnormal leukocyte migration/physiology abnormal macrophage chemotaxis/physiology abnormal microglial cell morphology abnormal muscle physiology abnormal osteoblast/osteoclast physiology abnormal physical strength abnormal physiological neovascularization abnormal response to infection abnormal trabecular bone morphology abnormal urine nucleoside level abnormal vasodilation abnormal wound healing improved glucose homeostasis complete postnatal lethality decreased circulating triglyceride level decreased dendritic cell number increased anti-double stranded DNA antibody level increased anti-single stranded DNA antibody level increased blood urea nitrogen level increased circulating cholesterol level increased circulating triglyceride level increased lung weight increased respiratory quotient increased sensitivity to induced morbidity/mortality internal hemorrhage lung cysts nephrocalcinosis postnatal lethality premature death	[264, 265]

ABCG2	Abcg2	**anxiety** **amyloid beta deposits** **i**ncreased physiological sensitivity to protoporphyrin IX abnormal bile color abnormal hematopoietic system morphology/development increased circulating bilirubin level phototoxicity porphyria	[266, 267]
AHR	Ahr	**abnormal hepatocyte morphology** **abnormal liver morphology/physiology** **abnormal xenobiotic pharmacokinetics** abnormal auchene hair morphology abnormal blood vessel morphology abnormal branching of the mammary ductal tree abnormal chemokine level abnormal circulating alanine transaminase level abnormal circulating hormone level abnormal coat/ hair morphology abnormal colon morphology abnormal coronary artery morphology abnormal cytokine secretion abnormal enzyme/coenzyme activity abnormal epidermal layer morphology abnormal eye development abnormal heart morphology abnormal immune system physiology abnormal interleukin level abnormal keratinocyte physiology abnormal kidney blood vessel morphology abnormal Langerhans cell morphology/physiology abnormal macrophage physiology abnormal ovary morphology/physiology abnormal skin condition abnormal spleen morphology abnormal stomach pyloric region morphology abnormal superovulation abnormal urinary bladder morphology abnormal uterus morphology abnormal vascular regression decreased body size decreased body weight decreased susceptibility to type IV hypersensitivity reaction increased body weight increased heart weight increased hepatoma incidence increased liver adenoma incidence increased susceptibility to bacterial infection increased susceptibility to injury increased urine uric acid level partial postnatal lethality postnatal growth retardation premature death weight loss	[218, 268-276]
FZD9	Fzd9	**abnormal dentate gyrus morphology** **abnormal forebrain development** **abnormal spatial learning** **increased susceptibility to pharmacologically induced seizures** abnormal bone mineralization/ossification/structure abnormal lymph node morphology abnormal osteoblast differentiation/physiology abnormal skeleton development/physiology abnormal spleen morphology abnormal vertebral body morphology decreased B cell number decreased body length decreased body weight increased eosinophil cell number increased leukocyte cell number increased monocyte cell number increased neutrophil cell number postnatal growth retardation premature death thymus atrophy	[277]
FKBP6	Fkbp6	abnormal spermatocyte morphology	

BAZ1B	Baz1b	**hyperactivity** abnormal double-strand DNA break repair abnormal fourth branchial arch morphology abnormal heart morphology abnormal palatine bone morphology abnormal tooth morphology decreased body length/size/weight decreased lean body mass / total body fat increased carbon dioxide production increased circulating calcium level increased energy expenditure increased oxygen consumption increased susceptibility to bacterial infection malocclusion micrognathia postnatal lethality	WTSI/MGP [278]
MLXIPL	Mlxipl	**abnormal liver physiology** abnormal glucose homeostasis abnormal lipid homeostasis decreased body temperature decreased brown adipose tissue amount decreased epididymal fat pad weight decreased white adipose tissue amount	[279]
POR	Por	**abnormal liver morphology** **abnormal xenobiotic pharmacokinetics** **enhanced behavioral response to xenobiotic** abnormal bile salt level abnormal branchial arch morphology abnormal cell adhesion abnormal craniofacial development abnormal embryogenesis/ development abnormal enzyme/coenzyme activity abnormal eye morphology abnormal heart ventricle morphology abnormal limb development abnormal myocardium layer morphology abnormal neural fold elevation formation abnormal outflow tract development absent trabeculae carneae decreased body weight abnormal progesterone/testosterone level abnormal lipid levels decreased kidney weight decreased litter size decreased lung weight embryonic growth retardation enhanced behavioral response to anesthetic increased apoptosis increased circulating alanine transaminase level intracranial hemorrhage prenatal/postnatal lethality pericardial edema reduced fertility	[280-283]
ZP3	Zp3	abnormal cumulus oophorus abnormal embryogenesis/ development abnormal oocyte morphology abnormal ovarian morphology abnormal ovulation impaired fertility	
BDNF	Bdnf	**abnormal action potential** **abnormal adrenergic neuron morphology** **abnormal axon pruning** **abnormal barrel cortex morphology** **abnormal brain interneuron morphology** **abnormal cerebellum external granule cell layer morphology** **abnormal cerebral cortex morphology** **abnormal CNS synaptic transmission** **abnormal cochlea morphology** **abnormal cochlear ganglion morphology** **abnormal cochlear OHC afferent innervation pattern** **abnormal contextual conditioning behavior** **abnormal cranial ganglia morphology** **abnormal crista ampullaris morphology** **abnormal cued conditioning behavior** **abnormal dendrite morphology**	[284-315]

		abnormal dendritic spine morphology **abnormal dentate gyrus morphology** **abnormal depression-related behavior** **abnormal dorsal root ganglion morphology** **abnormal excitatory postsynaptic potential** **abnormal hippocampus CA2 region morphology** **abnormal hippocampus morphology** **abnormal inhibitory postsynaptic currents** **abnormal innervation** **abnormal liquid preference** **abnormal locomotor activation** **abnormal medium spiny neuron morphology** **abnormal motor capabilities/coordination/movement** **abnormal motor coordination/ balance** **abnormal neuron apoptosis** **abnormal neuron differentiation** **abnormal neuron morphology** **abnormal neuron physiology** **abnormal sensory ganglion morphology** **abnormal sensory neuron innervation pattern** **abnormal sensory neuron morphology** **abnormal serotonergic neuron morphology** **abnormal serotonin level** **abnormal somatic nervous system morphology** **abnormal somatic sensory system morphology** **abnormal striatum morphology** **abnormal substantia nigra morphology** **abnormal sympathetic system morphology** **abnormal synaptic plasticity** **impaired behavioral response to xenobiotic** **increased aggression** **increased anxiety-related response** **abnormal olfactory bulb morphology** **abnormal olfactory nerve morphology** **abnormal fungiform papillae morphology** **abnormal gustatory papillae taste bud morphology** **abnormal gustatory system morphology** **abnormal circumvallate papillae morphology** abnormal glucose homeostasis abnormal homeostasis abnormal hair cycle abnormal GABAergic neuron morphology abnormal geniculate ganglion morphology abnormal colon morphology abnormal intestine morphology abnormal keratinocyte morphology abnormal kindling response abnormal nervous system electrophysiology abnormal nervous system physiology abnormal nest building behavior abnormal optic nerve morphology abnormal trigeminal nerve morphology abnormal type I vestibular cell abnormal vestibular ganglion morphology abnormal vestibular nerve morphology abnormal vestibulocochlear ganglion morphology abnormal visual cortex morphology ataxia neonatal/postnatal lethality/growth retardation abnormal body size/weight decreased pulmonary respiratory rate impaired fertility increased cholesterol level increased circulating leptin level increased susceptibility to age related obesity lethargy obese polyphagia weight loss	
LIN7C	Lin7c	abnormal breathing pattern abnormal excitatory postsynaptic currents abnormal kidney morphology/physiology neonatal/postnatal lethality decreased body weight	

CSK	Csk	**abnormal cranial ganglia morphology** **abnormal food preference** **absent olfactory bulb** **decreased anxiety-related response** abnormal allantois morphology abnormal T cell differentiation abnormal embryonic neuroepithelium morphology abnormal granulocyte physiology abnormal lens induction abnormal myelopoiesis abnormal neural tube morphology/development abnormal object recognition memory abnormal otic vesicle development abnormal social investigation absent visceral yolk sac blood islands absent vitelline blood vessels embryonic lethality increased acute inflammation increased susceptibility to bacterial infection necrosis poor circulation	[316, 317]
CYP1A1	Cyp1a1	**abnormal liver physiology** **abnormal physiological response to xenobiotic** **abnormal xenobiotic pharmacokinetics** abnormal hemoglobin content increased circulating alanine transaminase level increased circulating aspartate transaminase level increased susceptibility to weight loss small spleen small thymus	[318-321]
CYP1A2	Cyp1a2	**abnormal liver physiology** **abnormal physiological/behavioral response to xenobiotic** **abnormal xenobiotic pharmacokinetics** abnormal lung development abnormal production of surfactant abnormal pulmonary alveolar duct morphology abnormal type II pneumocyte morphology absent gastric milk in neonates atelectasis cyanosis decreased body temperature decreased leukocyte cell number increased circulating alanine transaminase level increased circulating aspartate transaminase level neonatal/postnatal lethality small spleen small thymus	[318, 319, 322-326]
CYP11A1	Cyp11a1	abnormal food intake (anorexia) hypoactivity (lethargy) abnormal adrenal gland morphology/physiology/function abnormal circulating hormone level abnormal circulating potassium level abnormal circulating sodium level abnormal corticosterone level abnormal epididymis morphology abnormal lipid level abnormal mitochondrion morphology abnormal noradrenaline level abnormal seminiferous tubule morphology abnormal spermatogenesis abnormal testis morphology neonatal/postnatal lethality/growth retardation decreased T cell apoptosis enlarged adrenal glands enlarged adrenocortical cells muscular atrophy	[327]

MPI	Mpi	abnormal angiogenesis abnormal cell physiology abnormal cell proliferation abnormal chorioallantoic fusion abnormal placenta morphology abnormal pupil morphology abnormal vascular morphology/development decreased total body fat amount embryonic lethality/growth retardation increased apoptosis increased erythrocyte cell number increased prepulse inhibition	
SEMA7A	Sema7a	**abnormal olfactory tract morphology** **abnormal axon outgrowth**	[328]
ARID3B	Arid3b	abnormal head morphology abnormal heart development abnormal vascular regression embryonic lethality/growth retardation increased apoptosis small branchial arch wavy neural tube	
CPLX3	Cplx3	**abnormal excitatory postsynaptic currents** **abnormal inhibitory postsynaptic currents** **abnormal miniature excitatory postsynaptic currents** **abnormal neuron physiology** **abnormal neurotransmitter secretion** abnormal eye electrophysiology abnormal cone electrophysiology abnormal pre-Botzinger complex physiology abnormal retinal photoreceptor morphology abnormal vision neonatal lethality	[329]
BLMH	Blmh	increased physiological sensitivity to bleomycin abnormal tail morphology decreased body size dermatitis neonatal lethality	
SLC6A4	Slc6a4	**abnormal action potential** **abnormal active avoidance behavior** **abnormal anxiety-related response** **abnormal brain morphology** **abnormal fear/anxiety-related behavior** **abnormal neuron physiology** **abnormal response to new environment** **abnormal response to novel object** **abnormal serotonergic neuron morphology** **abnormal serotonin level** **abnormal sleep pattern** **abnormal social investigation** **hyperactivity** **hypoactivity** **abnormal physiological/behavioral response to xenobiotic** **abnormal conditioned place preference behavior** abnormal pulmonary artery morphology abnormal adenohypophysis morphology abnormal autonomic nervous system physiology abnormal barrel cortex morphology abnormal body temperature homeostasis abnormal heart valve morphology abnormal heart ventricle morphology postnatal lethality/growth retardation	[330-341]
GIT1	Git1	**abnormal brain wave pattern** **abnormal CNS synaptic transmission** **abnormal dendrite morphology** **abnormal object recognition memory** **abnormal operant conditioning behavior** **abnormal spatial learning** **decreased anxiety-related response** **decreased fear-related response** **hyperactivity** abnormal vascular endothelial cell physiology abnormal lung development/morphology decreased body weight decreased vascular endothelial cell number neonatal/ postnatal lethality	[342-344]

ATAD5	Atad5	abnormal cell physiology abnormal DNA repair adenocarcinoma aneuploidy chromosomal instability prenatal lethality hemangiosarcoma increased lymphoma incidence increased spindle cell carcinoma incidence increased tumor incidence lung carcinoma reticulocytosis T cell derived lymphoma uterus tumor	
CCDC55	Ccdc55	embryonic lethality	

[a]Data presented were retrieved from the Mouse Genome Database, Mouse Genome Informatics (MGI), The Jackson Laboratory, Bar Harbor, Maine (URL: http://www.informatics.jax.org, March 2013). Data from the Wellcome Trust Sanger Institute (WTSI) Mouse Resources Portal (http://www.sanger.ac.uk/mouseportal) are also downloaded and integrated into MGI.
[b]Genes in cells with the same color are in close proximately in the human genome.
[c]Mouse models queried: transgenic, targeted (knock-out, knock-in, reporter, floxed/frt, others), gene-trapped, chemically-induced, spontaneous. Phenotypes in bold-face are relevant to coffee consumption behavior and have been additionally curated for accuracy. Notes in brackets accompany suspect entries subsequently confirmed to be 'not relevant' to coffee consumption behavior.
[d]References are provided for phenotypes in bold-face (column 3) only.

Table S24. Fold-change (FC) values of human hepatocyte gene expression in response to 1500 or 7500 μM caffeine vs. vehicle exposure[a]

LOCUS	GENE	PROBE	1500 μM				7500 μM			
			8 hours		24 hours		8 hours		24 hours	
			FC	FDR	FC	FDR	FC	FDR	FC	FDR
2p24	*SNX17*	200991_s_at	0.95	0.87	0.94	0.62	0.75	**0.01**	0.63	**1.4E-03**
	GCKR	206867_at	1.03	0.93	1.12	0.29	1.01	0.97	1.71	**1.9E-03**
	PPM1G	200913_at	0.96	0.90	0.79	0.08	0.90	0.22	0.61	**2.1E-03**
	CCDC121	220321_s_at	1.20	0.52	1.22	0.26	0.90	0.29	1.57	**3.3E-03**
	GPN1	209313_at	0.99	0.98	0.88	0.24	0.93	0.38	0.71	**0.01**
	SUPT7L	201836_s_at	0.83	0.53	0.95	0.83	0.61	**4.0E-03**	0.61	**0.01**
	SUPT7L	201838_s_at	1.10	0.81	0.90	0.35	0.98	0.92	0.74	**0.01**
	FNDC4	218843_at	1.05	0.87	1.18	0.17	0.98	0.84	1.29	**0.01**
	MPV17	203466_at	0.91	0.70	1.03	0.85	0.80	**0.03**	0.80	**0.02**
	SLC4A1AP	1558201_s_at	0.98	0.95	0.95	0.66	0.82	0.05	0.81	**0.02**
	SLC4A1AP	218682_s_at	1.02	0.97	0.95	0.72	0.90	0.31	0.81	**0.03**
	ZNF513	225753_at	1.13	0.77	1.22	0.16	0.92	0.40	1.26	**0.03**
	PREB	217861_s_at	0.97	0.93	1.08	0.50	0.74	**0.01**	1.07	0.44
	ATRAID	219329_s_at	1.00	1.00	1.09	0.44	0.81	**0.05**	1.05	0.56
	GTF3C2	204366_s_at	0.92	0.76	1.02	0.94	0.78	**0.02**	1.00	1.00
4q22	*SPP1*	209875_s_at	0.82	0.52	0.71	**0.04**	0.62	**0.01**	0.43	**2.3E-04**
	ABCG2	209735_at	1.15	0.68	1.11	0.37	1.14	0.27	1.29	**0.02**
7p21	*AHR*	202820_at	1.11	0.74	0.99	0.94	1.92	**3.4E-03**	0.85	0.06
7q23i	*DNAJC30*	223367_at	0.85	0.55	0.96	0.80	0.54	**1.7E-03**	0.70	**4.1E-03**
	MLXIPL	221163_s_at	0.93	0.82	1.12	0.32	1.08	0.43	1.58	**0.01**
	TBL2	212685_s_at	1.01	1.00	0.93	0.52	0.92	0.36	0.83	**0.04**
	WBSCR22	207628_s_at	1.00	1.00	0.93	0.52	0.95	0.57	0.83	**0.04**
	BCL7B	202518_at	1.14	0.65	1.05	0.71	1.25	**0.04**	1.17	0.07
7q23ii	*RHBDD2*	222995_s_at	1.14	0.67	1.28	0.07	1.19	0.21	2.10	**4.5E-04**
	POR	208928_at	1.02	0.95	1.42	**0.04**	1.12	0.20	3.33	**4.6E-04**
	RHBDD2	232053_x_at	1.24	0.44	1.32	0.06	1.16	0.16	1.91	**6.3E-04**
	MDH2	213333_at	1.02	0.96	1.06	0.72	0.96	0.71	0.67	**2.6E-03**
	STYXL1	232353_s_at	1.07	0.83	1.06	0.65	0.98	0.88	1.38	**0.01**
	STYXL1	230370_x_at	1.01	0.99	1.05	0.67	0.93	0.52	1.32	**0.01**
	HSPB1	201841_s_at	0.99	0.98	0.85	0.16	0.97	0.79	0.76	**0.01**
	STYXL1	218321_x_at	1.04	0.92	1.10	0.39	0.98	0.87	1.33	**0.01**
	MDH2	209036_s_at	0.96	0.89	0.95	0.64	0.88	0.15	0.76	**0.01**
	STYXL1	233982_x_at	1.06	0.88	1.06	0.65	0.95	0.68	1.30	**0.01**
	TMEM120A	223482_at	1.12	0.70	1.43	**0.04**	1.03	0.76	1.38	**0.01**
11p13	*CCDC34*	226287_at	1.10	0.76	0.77	0.38	0.87	0.14	0.52	**3.5E-03**
	LIN7C	219399_at	0.95	0.87	0.87	0.22	0.76	**0.02**	0.66	**0.01**
	METTL15	242247_at	0.83	0.70	1.02	0.89	0.53	**0.03**	0.68	**0.02**
	LIN7C	221568_s_at	0.87	0.70	0.96	0.77	0.90	0.48	0.77	**0.03**
15q24	*PPCDC*	219066_at	1.66	0.18	1.40	0.05	1.26	**0.03**	2.00	**4.4E-04**
	ARID3B	218964_at	1.07	0.84	1.14	0.23	1.61	**0.01**	1.55	**1.7E-03**
	SEMA7A	230345_at	0.97	0.94	0.64	**0.03**	0.72	**0.01**	0.55	**2.0E-03**
	EDC3	219207_at	0.97	0.94	0.99	0.96	0.64	**0.02**	0.70	**3.9E-03**
	COX5A	229426_at	1.05	0.88	0.96	0.79	0.94	0.61	0.69	**4.4E-03**
	CSK	202329_at	1.01	0.99	0.90	0.34	0.72	**0.02**	0.79	**0.02**
	ULK3	225067_at	0.96	0.92	1.11	0.44	1.01	0.94	1.21	**0.03**
	RPP25	219143_s_at	0.81	0.53	0.84	0.27	0.75	**0.02**	0.92	0.28
	MPI	202472_at	0.94	0.89	0.97	0.89	0.78	**0.03**	1.00	0.98
17q11.2	*BLMH*	202179_at	0.92	0.73	0.96	0.85	0.71	**0.03**	0.52	**2.2E-03**
	TAOK1	238420_at	1.16	0.81	1.03	0.86	1.66	**0.03**	1.23	**0.04**
	SLC6A4	207519_at	0.90	0.75	0.93	0.55	1.07	0.67	1.19	**0.04**
	NSRP1	223236_at	1.05	0.87	1.08	0.47	0.80	0.14	1.22	**0.04**
	TAOK1	224778_s_at	1.16	0.65	0.98	0.91	1.46	**0.02**	1.10	0.58
	TAOK1	224769_at	1.22	0.57	0.96	0.75	1.41	**0.01**	1.05	0.62
	TAOK1	227454_at	1.06	0.83	0.94	0.57	1.35	**0.01**	1.07	0.63

[a]Shown are results for all candidate gene probes associated with differential expression in response to caffeine at false discovery rate (FDR) <0.05 under one or more experimental conditions.

Table S25. Between-study heterogeneity at 7p21 and 15q24

A. Subgroup analysis according to study characteristics considered as potential sources of heterogeneity at 15q24 and 7p21[a]

Study characteristic	*CYP1A2* (rs2472297, EA=T)					*AHR* (rs6968554, EA=G)				
	# studies	Summary β (SE)	Measure of Heterogeneity Q-value	P-value	I^2	# studies	Summary β (SE)	Measure of Heterogeneity Q-value	P-value	I^2
Overall	47	0.13 (0.01)	108.82	<.0001	58	50	0.11 (0.01)	147.0	<.0001	67
Mean age of sample										
<60 years	34	0.13 (0.02)	86.07	<.0001	62	37	0.12 (0.01)	105.8	<.0001	66
≥60 years	13	0.13 (0.03)	22.54	0.032	47	13	0.08 (0.03)	35.0	<.0001	66
Proportion of sample female										
≤50%	13	0.07 (0.02)	10.42	0.58	0	13	0.08 (0.02)	11.9	0.46	0
>50%	34	0.15 (0.02)	81.93	<.0001	60	37	0.11(0.02)	130.9	<.0001	73
Proportion of sample currently smoking										
<20%	30	0.13 (0.02)	78.3	<.0001	63	29	0.10 (0.02)	74.7	<.0001	63
≥20%	17	0.14 (0.02)	29.75	0.02	46	21	0.11 (0.02)	68.8	<.0001	71
Mean coffee intake of sample										
< 2 cups/d	10	0.12 (0.03)	6.3	0.71	0	9	0.11(0.02)	5.9	0.66	0
2 to 3 cups/d	24	0.12 (0.01)	21.1	0.57	0	24	0.09(0.02)	61.4	<.0001	63
> 3 cups/d	13	0.19 (0.04)	78.2	<.0001	85	17	0.14 (0.03)	78.5	<.0001	80
Geographic residence of sample										
North America	19	0.12 (0.01)	18.1	0.45	1	19	0.14 (0.01)	17.8	0.47	0
Europe /Australia	26/2	0.15 (0.02)	90.1	<.0001	70	28/13	0.10 (0.02)	117.5	<.0001	75
rs2472297 quality										
genotyped	32	0.11 (0.02)	80.3	<.0001	61	-	-	-	-	-
imputed	15	0.18 (0.02)	17.6	0.23	20	-	-	-	-	-
rs6968554 quality										
genotyped	-	-	-	-	-	13	0.08 (0.03)	54.1	<.0001	78
imputed	-	-	-	-	-	37	0.11 (0.02)	91.5	<.0001	61

EA, effect allele

[a]Results from random effects meta-analysis of rs2472297 and rs6968554 with coffee consumption (cups/day, phenotype 1) stratified by pre-specified study characteristics. Only studies of European ancestry were included for this post-hoc analysis.

B. Univariate and multivariate meta-regression analysis of polymorphism beta coefficients for coffee consumption and study characteristics[a]

Study characteristic, unit	*CYP1A2* (rs2472297, EA=T)				*AHR* (rs6968554, EA=G)			
	Univariate β (SE),	P-value	Multivariate β (SE)	P-value	Univariate β (SE)	P-value	Multivariate β (SE)	P-value
Mean age, years	-0.0008 (0.001)	0.56	----	----	-0.0009 (0.001)	0.47	----	----
Female, %	0.0008 (0.0005)	0.09	0.0008 (0.0005)	0.07	0.0002 (0.0005)	0.66	0.0002 (0.0005)	0.66
Current smokers, %[b]	0.002 (0.002)	0.18	----	----	0.0009 (0.002)	0.56	----	----
Mean coffee intake, cups/d	0.06 (0.02)	<0.0001	0.05 (0.02)	0.001	0.04 (0.02)	0.005	0.04 (0.02)	0.006
North American (yes/no)[b]	-0.04 (0.03)	0.14	----	----	-0.02 (0.03)	0.48	----	----
Imputation Quality index	-0.07 (0.08)	0.38	----	----	0.03 (0.10)	0.76	----	----
Imputed (yes/no)	0.06 (0.03)	0.03	0.04 (0.03)	0.14	0.01 (0.03)	0.69	----	----

EA, effect allele

[a]Results from random effects meta-regressions using the restricted maximum likelihood method. Characteristics yielding differential heterogeneity in stratified analysis (Supplementary Table S25A) or moderate associations (P<0.1) with effect sizes according to univariate meta-regressions (column 2 and 6) were considered for multivariate meta-regressions (columns 4 and 8). Only studies of European ancestry were included for this post-hoc analysis.

[b]These covariates was highly correlated (r>0.4) with study mean coffee consumption and thus excluded from all multivariate analysis.

Table S26. Candidate gene literature mining

Gene	Top Pubmed articles linking gene to 'caffeine', 'coffee' , 'psychostimulant' , 'addiction' , 'taste', 'smell' ,'gustation' or 'olfaction'	Ref.
PPM1G	•Kimura H et al. "A novel histone exchange factor, protein phosphatase 2Cgamma, mediates the exchange and dephosphorylation of H2A-H2B." J Cell Biol. 2006 Nov 6;175(3):389-400. PMID 17074886 *'The disruption of PP2Cgamma in chicken DT40 cells increased the sensitivity to **caffeine**, a reagent that disturbs DNA replication and damage checkpoints, suggesting the involvement of PP2Cgamma-mediated histone dephosphorylation and exchange in damage response or checkpoint recovery in higher eukaryotes '*	[345]
SPP1	•Vízdalová M et al. "The role of the HCR system in the repair of lethal lesions of Bacillus subtilis phages and their transfecting DNA damaged by radiation and alkylating agents." Folia Microbiol (Praha). 1980;25(5):369-80. PMID 6776018 *'**caffeine** lowers the survival of UV-irradiated phage SPP1 in exponentially growing hcr+ cells but has no effect on its survival in competent hcr+ cells'* •Kunii Y et al. "The immunohistochemical expression profile of osteopontin in normal human tissues using two site-specific antibodies reveals a wide distribution of positive cells and extensive expression in the central and peripheral nervous systems." Med Mol Morphol. 2009 Sep;42(3):155-61. PMID 19784742 *'To elucidate the cellular distribution of osteopontin (OPN) in normal human tissues, we undertook immunohistochemistry using two site-specific OPN antibodies* [10A16, O-17]'. *' lutein cells and **taste** bud cells exhibited O-17 reactivity alone.'*	[346, 347]
PKD2	•Harris PC et al. "Polycystic Kidney Disease, Autosomal Dominant." GeneReviews™ 1993;. PMID 20301424 *'Agents/circumstances to avoid: Long-term administration of nephrotoxic agents, **caffeine** (which may promote renal cyst growth), use of estrogens by individuals with severe polycystic liver disease, and smoking'* •Morel N et al. "PKD1 haploinsufficiency is associated with altered vascular reactivity and abnormal calcium signaling in the mouse aorta." Pflugers Arch. 2009 Feb;457(4):845-56. PMID 18679710 *'Basal cytosolic calcium, KCl, and phenylephrine-evoked calcium signals were significantly lower in the Pkd1+/- aortas, whereas calcium release evoked by **caffeine** or thapsigargin was significantly larger* [then Pkd1+/+]*'* . •Anyatonwu GI et al. "Regulation of ryanodine receptor-dependent calcium signaling by polycystin-2." Proc Natl Acad Sci U S A. 2007 Apr 10;104(15):6454-9. PMID 17404231 *'In the presence of **caffeine*** [an RYR agonist]*, Pkd2(-/-) cardiomyocytes exhibited decreased peak fluorescence, a slower rate of rise, and a longer duration of Ca(2+) transients compared with Pkd2(+/+). These data suggest that PC2 is important for regulation of RyR2 function and that loss of this regulation of RyR2, as occurs when PC2 is mutated, results in altered Ca(2+) signaling in the heart'* •Qian Q et al. "Pkd2 haploinsufficiency alters intracellular calcium regulation in vascular smooth muscle cells." Hum Mol Genet. 2003 Aug 1;12(15):1875-80. PMID 12874107 *'The resting [Ca(2+)](i) is 17.1% lower in Pkd2 (+/-) compared with wild-type cells (P=0.0003) and the total sarcoplasmic reticulum Ca(2+) store (emptied by **caffeine** plus thapsigargin) is decreased (P<0.0001).'* •Volk T et al. "A polycystin-2-like large conductance cation channel in rat left ventricular myocytes." Cardiovasc Res. 2003 Apr	[348-354]

	1;58(1):76-88. PMID 12667948 *'Application of 10 mM **caffeine** to the bath solution to increase the intracellular Ca(2+) concentration led to activation of* [large conductance nonselective cation channels] *LCC in 56% of the myocytes investigated (total n=651), in approximately 10%, more than three LCCs were detected'* •Nauli SM et al. "Polycystins 1 and 2 mediate mechanosensation in the primary cilium of kidney cells." Nat Genet. 2003 Feb;33(2):129-37. PMID 12514735 *'Cells isolated from transgenic mice that lack functional PC1 formed cilia but did not increase Ca(2+) influx in response to physiological fluid flow. Blocking antibodies directed against* [polycystin] *PC2 similarly abolished the flow response in wild-type cells as did inhibitors of the ryanodine receptor* [**caffeine**]*, whereas inhibitors of G-proteins, phospholipase C and InsP(3) receptors had no effect'* •Molland KL et al. "Identification of the structural motif responsible for trimeric assembly of the C-terminal regulatory domains of polycystin channels PKD2L1 and PKD2." Biochem J. 2010 Jul 1;429(1):171-83. PMID 20408813 *'Polycystin 2-type cation channels PKD2 and PKD2L1 interact with polycystin 1-type proteins PKD1 and PKD1L3 respectively, to form receptor-cation-channel complexes. The PKD2L1-PKD1L3 complex perceives sour **taste**, whereas disruption of the PKD2-PKD1 complex, responsible for mechanosensation, leads to development of ADPKD (autosomal-dominant polycystic kidney disease).'*	
ABCG2	•Ding R et al. "Xanthines down-regulate the drug transporter ABCG2 and reverse multidrug resistance." Mol Pharmacol. 2012 Mar;81(3):328-37. PMID 22113078 *'In this study, we found that a group of xanthines including **caffeine**, theophylline, and dyphylline can dramatically decrease ABCG2 protein in cells that have either moderate (BeWo, a placental choriocarcinoma cell line) or high (MCF-7/MX100, a breast cancer drug-resistant cell subline) levels of ABCG2 expression. This down-regulation is time-dependent, dose-dependent, and reversible.'* •Isshiki M et al. "**Coffee** induces breast cancer resistance protein expression in Caco-2 cells." Biol Pharm Bull. 2011;34(10):1624-7. PMID 21963506 ***'Coffee** induced BCRP gene expression in Caco-2 cells in a **coffee**-dose dependent manner. **Coffee** treatment of Caco-2 cells also increased the level of BCRP protein, which corresponded to induction of gene expression, and also increased cellular efflux activity, as judged by Hoechst33342 accumulation. None of the major constituents of **coffee** tested could induce BCRP gene expression. The constituent of **coffee** that mediated this induction was extractable with ethyl acetate and was produced during the roasting process. Dehydromethylepoxyquinomicin (DHMEQ), an inhibitor of nuclear factor (NF)-κB, inhibited **coffee**-mediated induction of BCRP gene expression, suggesting involvement of NF-κB in this induction.'* •Tournier N et al. "Interaction of drugs of abuse and maintenance treatments with human P-glycoprotein (ABCB1) and breast cancer resistance protein (ABCG2)." Int J Neuropsychopharmacol. 2010 Aug;13(7):905-15. PMID 19887017 *'We used in vitro P-gp and BCRP inhibition flow cytometric assays with hMDR1- and hBCRP-transfected HEK293 cells to test 14 compounds or metabolites frequently involved in **addiction**'* *'BCRP was inhibited by buprenorphine>norbuprenorphine>ibogaine and THC'* *'BCRP did not transport any of the tested compounds.'.*	355-357
AHR	•Dobrinas M et al. "Pharmacogenetics of CYP1A2 activity and inducibility in smokers and exsmokers." Pharmacogenet Genomics. 2013 May;23(5):286-92. PMID 23492909 *'A significant influence on CYP1A2 inducibility was observed for the NR1I3 rs2502815 (P=0.0026), rs4073054 (P=0.029), NR2B1*	358-364

rs3818740 (P=0.0045), rs3132297 (P=0.036), AhR rs2282885 (P=0.040), rs2066853 (P=0.019), NR1I1 rs2228570 (P=0.037), and NR1I2 rs1523130 (P=0.044) polymorphisms.'

•Hung WT et al. "Genetic susceptibility to dioxin-like chemicals' induction of cytochrome P4501A2 in the human adult linked to specific AhRR polymorphism." Chemosphere. 2013 Mar;90(9):2358-64. PMID 23168330

'The goal of this study was to determine the relationship between inducibility of CYP1A2 [**caffeine** breath test] *and genetic polymorphisms of AhR, ARNT, and AhRR in human AhRR (rs2292596) genotypes predict the inducibility of CYP1A2 in people highly exposed to toxic dioxin-like chemicals'*

•Legendre A et al. "Metabolic characterization of primary rat hepatocytes cultivated in parallel microfluidic biochips." J Pharm Sci. 2013 Feb 19;. PMID 23423727

*'The functionality of primary rat hepatocytes was assessed in an Integrated Dynamic Cell Cultures in Microsystem (IDCCM) device.We characterized the hepatocytes over 96 h of culture and evaluated the impact of dynamic cell culture on their viability, inducibility, and metabolic activity. Reverse Transcription quantitative Polymerase Chain Reaction (RTqPCR) was performed on selected genes: liver transcription factors (HNF4α and CEBP), nuclear receptors sensitive to xenobiotics (AhR, PXR, CAR, and FXR), cytochromes P450 (CYPs) (1A2,3A2, 3A23/3A1, 7A1, 2B1, 2C6, 2C, 2D1, 2D2, and 2E1), phase II metabolism enzymes (GSTA2,SULT1A1, and UGT1A6), ABC transporters (ABCB1b and ABCC2), and oxidative stress related enzymes (HMOX1 and NQO1). Metabolic activities were also confirmed with the detection of metabolism rate and induced mRNAs after exposure to several inducers: 3-methylcholanthrene, **caffeine**, phenacetin, paracetamol, and midazolam.'*

•Vaynshteyn D et al. "**Caffeine** induces CYP1A2 expression in rat hepatocytes but not in human hepatocytes." Drug Metab Lett. 2012 Jun 1;6(2):116-9. PMID 23167901

*'Our results from luciferase assays performed in HepG2 cells showed that **caffeine** is not an activator of the aromatic hydrocarbon receptor (AhR), a major transcription factor involved in upregulation of CYP1A2.'*

•Kalthoff S et al. "**Coffee** induces expression of glucuronosyltransferases by the aryl hydrocarbon receptor and Nrf2 in liver and stomach." Gastroenterology. 2010 Nov;139(5):1699-710, 1710.e1-2. PMID 20600030

*'Incubation of cells with **coffee** induced transcription of UGT1A1 (5.4-fold), UGT1A3 (5.2-fold), UGT1A4 (4.8-fold), UGT1A7 (6.2-fold), UGT1A8 (5.2-fold), UGT1A9 (3.5-fold), and UGT1A10 (6.1-fold). Induction was independent of **caffeine**, methylxanthines, or the diterpenes cafestol and kahweol. Mutagenesis and short interfering RNA knockdown studies showed that UGT1A is regulated by the aryl hydrocarbon receptor (AhR) and the nuclear factor erythroid-related factor 2 (Nrf2) by cis-acting antioxidant and xenobiotic response elements (ARE/XRE). '*

•de Waard PW et al. "A human intervention study with foods containing natural Ah-receptor agonists does not significantly show AhR-mediated effects as measured in blood cells and urine." Chem Biol Interact. 2008 Oct 22;176(1):19-29. PMID 18762178

' we performed a human intervention study with [natural Ah-receptor agonists] *NAhRA-containing cruciferous vegetables and grapefruit juice. The expression of the prototypical AhR-responsive genes CYP1A1, CYP1B1 and NQO1 in whole blood cells and in freshly isolated lymphocytes was not significantly affected. Also enzyme activities of CYP1A2, CYP2A6, N-acetyltransferase 2 (NAT2) and xanthine oxidase (XO), as judged by **caffeine** metabolites in urine, were unaffected, except for a small down-regulation of NAT2 activity by grapefruit juice.'*

•Long JR et al. "Population-based case-control study of AhR (aryl hydrocarbon receptor) and CYP1A2 polymorphisms and breast cancer risk." Pharmacogenet Genomics. 2006 Apr;16(4):237-43. PMID 16538170

'For the AhR gene, the A (Lys) allele was associated with a decreased risk of breast cancer'.

	'There was no significant association between the AhR gene and CYP1A2 activity [**caffeine** metabolite ratio] *in either cases or controls'*	
UCN	•Fatima A et al. "Urocortin 1 administered into the hypothalamic supraoptic nucleus inhibits food intake in freely fed and food-deprived rats." Amino Acids. 2013 Mar;44(3):879-85. PMID 23076252 *'We first established the dose-related effects of Ucn1 injected into the* [supraoptic nucleus] *SON on the feeding response in both freely fed and 24-h food-deprived rats. A conditioned **taste** avoidance paradigm was performed to investigate possible generalised effects of local Ucn1 treatment.'* *Administration of Ucn1 into the SON at doses equal to or higher than 0.5 µg significantly decreased food intake in both freely fed and food-deprived rats. 'The Ucn1-mediated suppression of food intake was delayed in freely fed as compared to food-deprived animals. Conditioning for **taste** aversion to saccharine appeared at 0.5 and 1 µg of Ucn1.'* •Fekete EM et al. "Systemic urocortin 2, but not urocortin 1 or stressin 1-A, suppresses feeding via CRF2 receptors without malaise and stress." Br J Pharmacol. 2011 Dec;164(8):1959-75. PMID 21627635 *'We examined whether rats show peripheral CRF/Ucn-induced anorexia and determined its behavioural and pharmacological bases.'* *'Stressin(1) -A and Ucn 1, but not Ucn 2, produced a conditioned **taste** aversion, reduced feeding efficiency and weight regain and elicited diarrhoea.'* •Fekete EM et al. "Delayed satiety-like actions and altered feeding microstructure by a selective type 2 corticotropin-releasing factor agonist in rats: intra-hypothalamic urocortin 3 administration reduces food intake by prolonging the post-meal interval." Neuropsychopharmacology. 2007 May;32(5):1052-68. PMID 17019404 '*The present study sought to identify the receptor subtype, brain site, and behavioral mode of action through which Ucn 3 reduces nocturnal food intake in rats.'* *'Ucn 3 effects were behaviorally specific, because minimal effective anorectic Ucn 3 doses did not alter drinking rate or promote a conditioned **taste** aversion, and site-specific, because intra-MeA Ucn 3 produced a nibbling pattern of more, but smaller meals without altering total intake.'* •Wang C et al. "Urocortin in the lateral septal area modulates feeding induced by orexin A in the lateral hypothalamus." Am J Physiol Regul Integr Comp Physiol. 2002 Aug;283(2):R358-67. PMID 12121849 *'Injection of 10 or 30 pmol UCN into LSi [lateral septum[significantly decreased feeding in food-deprived rats for 24 h without producing conditioned **taste** aversion (CTA).'* •Wang C et al. "Feeding inhibition by urocortin in the rat hypothalamic paraventricular nucleus." Am J Physiol Regul Integr Comp Physiol. 2001 Feb;280(2):R473-80. PMID 11208577 *'In the current studies we examined the effect of UCN in the hypothalamic paraventricular nucleus (PVN) on feeding.'* *'Ten and thirty picomoles UCN did not induce a [conditioned **taste** aversion] CTA, whereas 100 pmol UCN produced a CTA.'* •Benoit SC et al. "Comparison of central administration of corticotropin-releasing hormone and urocortin on food intake, conditioned **taste** aversion, and c-Fos expression." Peptides. 2000 Mar;21(3):345-51. PMID 10793215 *'CRH but not Ucn promoted robust and reliable [conditioned **taste** aversion] CTA learning'* •Giardino WJ et al. "Characterization of Genetic Differences within the Centrally Projecting Edinger-Westphal Nucleus of C57BL/6J and DBA/2J Mice by Expression Profiling." Front Neuroanat. 2012;6:5. PMID 22347848	365-375

'Detailed examination of the midbrain Edinger-Westphal (EW) nucleus revealed the existence of two distinct nuclei. One population of EW preganglionic (EWpg) neurons was found to control oculomotor functions, and a separate population of EW centrally projecting (EWcp) neurons was found to contain stress- and feeding-related neuropeptides.'
'To identify genetic differences in the EWcp of inbred mouse strains that differ in behaviors relevant to EWcp function, we used publicly available tools from the Allen Brain Atlas to identify 68 transcripts that were selectively expressed in the EWcp, and examined their expression within tissue punch microdissection samples containing the EWcp of adult male C57BL/6J (B6) and DBA/2J (D2) mice. Using 96-well quantitative real-time PCR (qPCR) arrays that included the EWcp-specific genes, several other genes of interest, and five housekeeping genes, we identified strain differences in expression of 11 EWcp-specific genes (BC023892, Btg3, Bves, Cart, Cck, Ghsr, Neto1, Postn, Ptprn, Rcn1, and Ucn), two immediate early genes (Egr1 and Fos), and one dopamine-related gene (Drd5).'
*'our identification of differentially expressed EWcp-specific genes between B6 and D2 mice may hold powerful insight into the neurogenetic contributions of the EWcp to stress- and **addiction**-related behaviors'*

•Fonareva I et al. "Increased perioculomotor urocortin 1 immunoreactivity in genetically selected alcohol preferring rats." Alcohol Clin Exp Res. 2009 Nov;33(11):1956-65. PMID 19673740

*'These findings extend previous reports of increased Ucn 1-positive cell distribution in preferring lines of animals. They indicate that Ucn1 contributes to increased alcohol consumption across different species and that this contribution could be gender specific. The results also suggest that Ucn1 regulates **positive reinforcing** rather than aversive properties of alcohol and that these effects could be mediated by CRF(2) receptors, independent of direct actions of DA'*

•Sharpe AL et al. "Central urocortin 3 administration decreases limited-access ethanol intake in nondependent mice." Behav Pharmacol. 2009 Jul;20(4):346-51. PMID 19581799

*'The CRF family of endogenous ligands includes urocortin 3 (Ucn 3), which binds selectively to the CRF type 2 receptor and has been implicated in ethanol consumption in **dependent and withdrawing** rats. The objective of this study was to examine the effect of Ucn 3, delivered centrally to nondependent mice, on limited-access ethanol consumption.'*
'There was a significant decrease in ethanol (both ml and g/kg), but not water, intake following Ucn 3 treatment, explained by a change in size of the largest lick run. Food intake at both 2 h and 24 h after injection was statistically unaffected by Ucn 3 administration.'

•Tao J et al. "Activation of corticotropin-releasing factor receptor 1 selectively inhibits CaV3.2 T-type calcium channels." Mol Pharmacol. 2008 Jun;73(6):1596-609. PMID 18292205

*'The corticotropin-releasing factor (CRF) peptides CRF and uro-cortins 1 to 3 are crucial regulators of mammalian stress and inflammatory responses, and they are also implicated in disorders such as anxiety, depression, and **drug addiction**.'*
' here we report that the native CRF receptor 1 (CRFR1) endogenous to the human embryonic kidney 293 cells can functionally couple to mammalian Ca(V)3.2 T-type calcium channels.'
'Activation of CRFR1 by either CRF or urocortin (UCN) 1 reversibly inhibits Ca(V)3.2 currents (IC(50) of approximately 30 nM), but it does not affect Ca(V)3.1 or Ca(V)3.3 channels.'

•Zoumakis E et al. "Potential uses of corticotropin-releasing hormone antagonists." Ann N Y Acad Sci. 2006 Nov;1083:239-51. PMID 17148743
'Corticotropin-releasing hormone (CRH), its natural homologs urocortins (UCN) 1, 2, and 3, and several types of CRH receptors (R), coordinate the behavioral, endocrine, autonomic, and immune responses to stress. The potential use of CRH antagonists is currently under intense investigation. Selective antagonists have been used experimentally to clarify the role of CRH-related peptides in anxiety and depression, **addictive behavior**, inflammatory disorders, acute and chronic neurodegeneration, and sleep disorders, as well as preterm labor' [Review]

POR	•Dobrinas M et al. "Influence of cytochrome P450 oxidoreductase genetic polymorphisms on CYP1A2 activity and inducibility by smoking." Pharmacogenet Genomics. 2012 Feb;22(2):143-51. PMID 22246422 *While smoking, none of the tested POR polymorphisms showed a significant influence on CYP1A2 activity. After smoking cessation, significantly higher CYP1A2 activity was found in POR rs2302429A carriers (P=0.038) and in carriers of rs17148944G-rs10239977C-rs3815455T-rs2286823G-rs2302429A-rs1057868T haplotype (P=0.038), whereas carriers of POR rs2286823A (P=0.031) and of the rs17148944G-rs10239977C-rs3815455C-rs2286823A-rs2302429G-rs1057868C haplotype (P=0.031) had decreased CYP1A2 activity. In the complete regression model, only POR rs2302429G>A showed a significant effect (P=0.017). No influence of POR genotypes or haplotypes was observed on the inducibility of CYP1A2* •Tomalik-Scharte D et al. "Impaired hepatic drug and steroid metabolism in congenital adrenal hyperplasia due to P450 oxidoreductase deficiency." Eur J Endocrinol. 2010 Dec;163(6):919-24. PMID 20844025 *We studied an adult patient with ORD due to homozygous POR A287P, the most frequent POR mutation in Caucasians, and her clinically unaffected, heterozygous mother.* *Though CYP enzyme genotyping predicted normal or high enzymatic activities* [cocktail with **caffeine**] *in both subjects, in vivo assessment showed subnormal activities of CYP1A2, CYP2C9, CYP2D6 and CYP3A4 in the patient and of CYP1A2 and CYP2C9 in her mother*	376, 377
HSPB1	•Das T et al. "Potential role of Hsp25 in calcium-modulated cardiomyocytes." Proteomics. 2012 Feb;12(3):411-20. PMID 22140065 *'In this study, we investigated the effect of **caffeine**, an inducer of intracellular Ca^{2+} accumulation, on HL-1 cardiomyocytes by using a proteomic approach.'* *' we identified 24* [i.e. Hsp25] *differentially expressed protein spots in the **caffeine**-treated group as compared with the controls '* *'Depletion of Hsp25 transcripts by siRNA increased **caffeine**-mediated signaling, including ERK activation, and decreased the Ca^{2+} transient peak and expression of calsequestrin 2 in HL-1 cardiomyocytes.'* •Xu JW et al. "Downregulation of Rac1 activation by caffeic acid in aortic smooth muscle cells." Life Sci. 2005 Apr 29;76(24):2861-72. PMID 15808886 *'Caffeic acid, a dietary phenol from **coffee**, fruits and vegetables, is an efficient antioxidant'.* *'Our results showed that caffeic acid decrease Rac1 protein level under basal conditions and incubation with angiotensin II (ANG II) in vascular smooth muscle cells'.* *' pretreatment with caffeic acid for 24 hours was able to prevent phosphorylation of MLC and HSP27, when cells were challenged with ANG II through the redox sensitive pathway.'* •Carr VM et al. "Odorants as cell-type specific activators of a heat shock response in the rat olfactory mucosa." J Comp Neurol. 2001 Apr 16;432(4):425-39. PMID 11268007 *'Heat shock, or stress, proteins (HSPs) are induced in response to conditions that cause protein denaturation. Activation of cellular stress responses as a protective and survival mechanism is often associated with chemical exposure.'* *'To determine whether environmental odorants affect [**olfactory** epithelium] OE HSP expression, rats were exposed to a variety of odorants added to the cage bedding. Odorant exposure led to transient, selective induction of HSP70, HSC70, HSP25, and ubiquitin immunoreactivities (IRs) in supporting cells and subepithelial Bowman's gland acinar cells '* *'HSP25 IR occurred throughout the entire supporting cell cytoplasm'* •Mjahed H et al. "Heat shock proteins in hematopoietic malignancies." Exp Cell Res. 2012 Sep 10;318(15):1946-58. PMID 22652452	378-381

	*'This cancer cell **addiction** for HSPs is the basis for the use of HSP inhibitors in cancer therapy'*	
BDNF	•Moy GA et al. "**Caffeine** prevents weight gain and cognitive impairment caused by a high-fat diet while elevating hippocampal BDNF." Physiol Behav. 2013 Jan 17;109:69-74. PMID 23220362 *'Here we investigated the impact of **caffeine** administration on metabolism and cognitive performance, both in control rats and in rats placed on a high-fat diet.'* *'**Caffeine** did not alter hippocampal metabolism or insulin signaling, likely because the high-fat-fed animals did not develop full-blown diabetes; however, **caffeine** did prevent or reverse a decrease in hippocampal brain-derived neurotrophic factor (BDNF) seen in high-fat-fed animals.'* •Reyes-Izquierdo T et al. "Modulatory effect of **coffee** fruit extract on plasma levels of brain-derived neurotrophic factor in healthy subjects." Br J Nutr. 2013 Jan 14;:1-6. PMID 23312069 *'The present single-dose study was performed to assess the effect of whole **coffee** fruit concentrate powder (WCFC), green **coffee** **caffeine** powder (N677), grape seed extract powder (N31) and green **coffee** bean extract powder (N625) on blood levels of brain-derived neurotrophic factor (BDNF).'* *'The collected data revealed that treatments with N31 and N677 increased levels of plasma BDNF by about 31 % under these experimental conditions, whereas treatment with WCFC increased it by 143 % (n 10), compared with baseline'* •Alzoubi KH et al. "Chronic **caffeine** treatment prevents stress-induced LTP impairment: the critical role of phosphorylated CaMKII and BDNF." J Mol Neurosci. 2013 Jan;49(1):11-20. PMID 22706686 *'**Caffeine** prevented stress-induced [long-term potentiation] LTP impairment. Western blot analysis showed reduction of the basal levels of the phosphorylated calcium calmodulin kinase II (P-CAMKII), total CaMKII, and brain-derived neurotrophic factor (BDNF) in area CA1 of stressed rats. These reductions were prevented by chronic **caffeine** treatment'* •Sallaberry C et al. "Chronic **caffeine** prevents changes in inhibitory avoidance memory and hippocampal BDNF immunocontent in middle-aged rats." Neuropharmacology. 2013 Jan;64:153-9. PMID 22841916 *' we here compare the effects of chronic **caffeine** (1 mg/mL drinking solution for 30 days) on short- and long term memory and on levels of hippocampal proBDNF, mature BDNF, TrkB and CREB in young (3 month old) and middle-aged (12 month old) rats. **Caffeine** treatment substantially reduced i) age-related impairments in the two types of memory in an inhibitory avoidance paradigm, and ii) parallel increases in hippocampal BDNF levels. In addition, chronic **caffeine** increased proBDNF and CREB concentrations, and decreased TrkB levels, in hippocampus regardless of age.'* •Capiotti KM et al. "Early exposure to **caffeine** affects gene expression of adenosine receptors, DARPP-32 and BDNF without affecting sensibility and morphology of developing zebrafish (Danio rerio)." Neurotoxicol Teratol. 2011 Nov-Dec;33(6):680-5. PMID 21914471 *'BDNF was also expressed since 24 hpf* [hours postfertilization] *and **caffeine** treatment increased its expression at 48 and 72 hpf'.* •Alhaider IA et al. "Sleep deprivation prevents stimulation-induced increases of levels of P-CREB and BDNF: protection by **caffeine**." Mol Cell Neurosci. 2011 Apr;46(4):742-51. PMID 21338685	382-417

	*'Sleep deprivation prevents the high frequency stimulation-induced increases in the levels of phosphorylated-cAMP response element binding protein (P-CREB) and brain-derived neurotrophic factor (BDNF) seen during the expression of late phase long-term potentiation (L-LTP). However, chronic **caffeine** treatment prevented the effect of sleep-deprivation on the stimulated levels of P-CREB and BDNF'.* •Connolly S et al. "**Caffeine** modulates CREB-dependent gene expression in developing cortical neurons." Biochem Biophys Res Commun. 2010 Jun 25;397(2):152-6. PMID 20493822 *'Quantitative real-time PCR analysis demonstrated that transcripts derived from endogenous CREB target genes, such as the gene encoding brain-derived neurotrophic factor BDNF, are increased following **caffeine** treatment*.' •Alhaider IA et al. "**Caffeine** prevents sleep loss-induced deficits in long-term potentiation and related signaling molecules in the dentate gyrus." Eur J Neurosci. 2010 Apr;31(8):1368-76. PMID 20384774 *' chronic **caffeine** treatment prevented the sleep deprivation-associated decreases in the basal levels of the phosphorylated calcium/calmodulin-dependent protein kinase II (P-CaMKII) and brain derived neurotrophic factor (BDNF) as well as in the stimulated levels of P-CaMKII in the [dentate gyrus] DG area.'* •Bairam A et al. "Neonatal **caffeine** treatment does not induce long-term consequences on TrkB receptors or BDNF expression in chemosensory organs of adult rats." Neurosci Lett. 2010 Jan 14;468(3):292-6. PMID 19914342 *'In male rats, when* [neonatal **caffeine** treatment] *NCT tended to decrease TrkBR mRNA transcript levels by about 32% in the CB and to reduce BDNF transcripts in the NTS by 22%, western blot analyses showed no parallel changes in final protein expression. NCT had no effects on TrkBR or BDNF mRNA and protein levels in the CB and NTS of female rats. Neither gene was altered by NCT in the superior cervical ganglion of male and female rats.'* •Prakash YS et al. "Neurotrophin effects on intracellular Ca2+ and force in airway smooth muscle." Am J Physiol Lung Cell Mol Physiol. 2006 Sep;291(3):L447-56. PMID 16648236 *'Basal [Ca(2+)](i), peak responses to all agonists* [i.e. **caffeine**]*, SOCE, and force responses to ACh and histamine were all significantly enhanced by both acute and prolonged BDNF exposure (smaller effect of* [neurotrophin] *NT4) but decreased by NT3. Inhibition of the BDNF/NT4 receptor trkB by K252a prevented enhancement of [Ca(2+)](i) responses.'* •Wang Y et al. "Differential involvement of brain-derived neurotrophic factor in reconsolidation and consolidation of conditioned **taste** aversion memory." PLoS One. 2012;7(11):e49942. PMID 23185492 *'We have recently observed that BDNF signaling in the central nuclei of the amygdala (CeA) and insular cortex (IC) was involved in the consolidation of conditioned **taste** aversion (CTA) memory.'* *'In the present study, using a CTA memory paradigm, we observed increased BDNF expression in the IC but not in the CeA during CTA reconsolidation. We further determined that BDNF synthesis and signaling in the IC but not in the CeA was required for memory reconsolidation. '* •Runge EM et al. "Neurotrophin-4 is more potent than brain-derived neurotrophic factor in promoting, attracting and suppressing geniculate ganglion neurite outgrowth." Dev Neurosci. 2012;34(5):389-401. PMID 23151843 *'The geniculate ganglion, which provides innervation to **taste** buds in the anterior tongue and palate, is unique among sensory ganglia in that its neurons depend on both neurotrophin-4 (NT4) and brain-derived neurotrophic factor (BDNF) for survival. Whereas BDNF is additionally implicated in **taste** axon guidance at targeting stages, much less is known about the guidance role of NT4 during targeting, or about either neurotrophin during initial pathfinding.'* *'During early pathfinding to the tongue (embryonic days 12-13; E12-13), NT4 and BDNF promote significantly longer outgrowth than*	

during intralingual targeting (E15-18). NT4 is more potent than BDNF at stimulating neurite outgrowth and both factors exhibit concentration optima, i.e. intermediate concentrations (0.25 ng/ml NT4 or 25 ng/ml BDNF) promote maximal neurite extension and high concentrations (10 ng/ml NT4 or 200 ng/ml BDNF) suppress it.'

•Adaikkan C et al. "The role of protein phosphorylation in the gustatory cortex and amygdala during **taste** learning." Exp Neurobiol. 2012 Jun;21(2):37-51. PMID 22792024

*'In the present review we focus on the roles of several families of kinases, phosphatases, and other synaptic-plasticity-related proteins, which activate membrane receptors and various intracellular signals to promote transcription, translation and protein degradation, and to regulate the appropriate cellular proteomes required for **taste** memory acquisition, consolidation and maintenance.'*
*'The various temporal phases of **taste** learning require the activation of appropriate waves of biochemical signals. These include: extracellular signal regulated kinase I and II (ERKI/II) signal transduction pathways; Ca(2+)-dependent pathways; tyrosine kinase/phosphatase-dependent pathways; brain-derived neurotrophicfactor (BDNF)-dependent pathways; cAMP-responsive element bindingprotein (CREB); and translation-regulation factors, such as initiation and elongation factors, and the mammalian target of rapamycin (mTOR).'*

•de Souza FT et al. "Burning mouth syndrome: a therapeutic approach involving mechanical salivary stimulation." Headache. 2012 Jun;52(6):1026-34. PMID 22084903

'Twenty-six BMS patients underwent treatment with salivary mechanical stimulation. Resting and stimulated saliva were collected before and after therapy. Salivary levels of total protein, brain-derived neurotrophic factor, interleukin-10, tumor necrosis factor-α, interleukin-6, and nerve growth factor were assessed '
*'A significant reduction in the burning sensation and number of burning sites as well as an improvement of **taste** disturbances and xerostomia were observed after therapy'.*
' therapy resulted in a significant decrease in salivary levels of total protein and an increase of tumor necrosis factor-α.'

•Nosrat IV et al. "Targeted **taste** cell-specific overexpression of brain-derived neurotrophic factor in adult **taste** buds elevates phosphorylated TrkB protein levels in **taste** cells, increases **taste** bud size, and promotes gustatory innervation." J Biol Chem. 2012 May 11;287(20):16791-800. PMID 22442142

*'Brain-derived neurotrophic factor (BDNF) is the most potent neurotrophic factor in the peripheral **taste** system during embryonic development. It is also expressed in adult **taste** buds.'*
*'We show that **taste** buds in these mice are significantly larger and have a larger number of **taste** cells compared with controls.'*
*'Up-regulation of TrkB transcripts in **taste** buds and elevated **taste** cell-specific TrkB phosphorylation in response to increased BDNF levels indicate that BDNF controls the expression and activation of its high affinity receptor in **taste** cells.'*

•Patel AV et al. "Neurotrophin-4 regulates the survival of gustatory neurons earlier in development using a different mechanism than brain-derived neurotrophic factor." Dev Biol. 2012 May 1;365(1):50-60. PMID 22353733

*'The number of neurons in the geniculate ganglion that are available to innervate **taste** buds is regulated by neurotrophin-4 (NT-4) and brain-derived neurotrophic factor (BDNF).'*
' there was an increase in TUNEL-labeling, indicating an increase in cell death in Ntf4(-/-) mice compared with wild types. However, activated caspase-3, which is up-regulated in the absence of BDNF, was not increased. This finding indicates that cell death initiated by NT-4-removal occurs through a different cell death pathway than BDNF-removal.'

•Kranjac D et al. "Peripheral bacterial endotoxin administration triggers both memory consolidation and reconsolidation deficits in mice." Brain Behav Immun. 2012 Jan;26(1):109-21. PMID 21889586

	*'We used a similar, but partially **olfaction**-based, contextual fear conditioning paradigm to examine the effects of LPS on memory consolidation and reconsolidation in mice.'* *'LPS administered immediately or 2 h, but not 12 h, post-training impaired memory consolidation processes that support the storage of the conditioned contextual fear memory.'* *'Four hours post-injection, both central cytokine and peripheral cytokine and chemokine levels were heightened in LPS-treated animals, with a simultaneous decrease in BDNF, but not Zif-268, mRNA.'* •Martínez-Moreno A et al. "Late Protein Synthesis-Dependent Phases in CTA Long-Term Memory: BDNF Requirement." Front Behav Neurosci. 2011;5:61. PMID 21960964 *'Our previous studies on the insular cortex (IC), a region of the temporal cortex implicated in the acquisition and storage of conditioned **taste** aversion (CTA), have demonstrated that intracortical delivery of BDNF reverses the deficit in CTA memory caused by the inhibition of IC protein synthesis due to anisomycin administration during early acquisition.'* *"We observed that CTA memory become sensible to protein synthesis inhibition 5 and 7h after acquisition.'* *'Our results show that BDNF reverses the CTA memory deficit produced by protein synthesis inhibition in both phases'* •Zimmerberg B et al. "Olfactory association learning and brain-derived neurotrophic factor in an animal model of early deprivation." Dev Psychobiol. 2009 May;51(4):333-44. PMID 19308959 *'Neonatal rats primarily use olfaction for attachment, and Brain-Derived Neurotrophic Factor (BDNF) may be a key transcription target in olfactory association learning.'* *'Learning the odor association, as revealed in a position preference for the novel odor, was accompanied by an increase in hippocampal BDNF in O/M subjects from undisturbed Control litters. BDNF levels were also positively related to degree of preference for the odor in the O/M Control group. ED subjects did not make the classically conditioned odor association and did not show an increase in hippocampal BDNF.'* •Cao L et al. "Genetic modulation of BDNF signaling affects the outcome of axonal competition in vivo." Curr Biol. 2007 Jun 5;17(11):911-21. PMID 17493809 *'We establish an in vivo axonal-competition paradigm in the mouse olfactory system by employing a genetic strategy that permits suppression of neurosecretory activity in random subsets of olfactory sensory neurons (OSNs). Long-term follow up confirmed that this genetic manipulation triggers competition by revealing a bias toward selective stabilization of active arbors and local degeneration of synaptically silent ones. By using a battery of genetically modified mouse models, we demonstrate that a decrease either in the total levels or the levels of activity-dependent secreted BDNF (due to a val66met substitution), rescues silent arbors from withering.'* •Nosrat CA et al. "Lingual deficits in BDNF and NT3 mutant mice leading to gustatory and somatosensory disturbances, respectively." Development. 1997 Apr;124(7):1333-42. PMID 9118804 *'A combination of anatomical, histological and physiological data from wild-type and null-mutated mice have established crucial roles for BDNF and NT3 in **gustatory** and somatosensory innervation of the tongue, and indeed for proper development of the papillary surface of the tongue'* [Review] •Depoy LM et al. "Developmentally divergent effects of Rho-kinase inhibition on cocaine- and BDNF-induced behavioral plasticity." Behav Brain Res. 2013 Apr 15;243:171-5. PMID 23327740 *'We administered the Rho-kinase inhibitor HA-1077 during three adolescent periods in mice to destabilize dendritic spines. In adulthood, cocaine-induced locomotor activity was exaggerated. By contrast, when administered in adulthood, HA-1077 had no psychomotor consequences and normalized food-reinforced instrumental responding after orbitofrontal-selective knockdown of Brain-derived neurotrophic factor, a potential factor in **addiction**.'*	

	•Legastelois R et al. “Blockade of Ethanol-Induced Behavioral Sensitization by Sodium Butyrate: Descriptive Analysis of Gene Regulations in the Striatum.” Alcohol Clin Exp Res. 2013 Mar 12;. PMID 23488934 *‘Because recent data demonstrate that histone deacetylase inhibitor (HDACi) may be of interest in the treatment of **addiction**, we explored the effect of the HDACi sodium butyrate (NaB) on EtOH-induced behavioral sensitization (EIBS) in DBA/2J mice.’* *‘Among the 168 studied genes, EIBS blockade was associated with specific gene regulations (bcl-2, bdnf, hdac4, pak1, penk, tacr1, vip) and changes in brain-derived neurotrophic factor in both striatum and prefrontal cortex.’* •Schmidt HD et al. “Epigenetics and **psychostimulant** addiction.” Cold Spring Harb Perspect Med. 2013 Mar 1;3(3):a012047. PMID 23359110 *‘Here we review how alterations in histone modifications, DNA methylation, and microRNAs regulate gene expression and contribute to **psychostimulant** **addiction** with a focus on the epigenetic mechanisms that regulate brain-derived neurotrophic factor (BDNF) expression following chronic cocaine exposure.’* •Barish PA et al. “Design and functional evaluation of an optically active μ-opioid receptor.” Eur J Pharmacol. 2013 Feb 27;705(1-3):42-48. PMID 23454521 *The use of opioids, which achieve therapeutic analgesia through activation of μ-opioid receptors, are limited in the management of chronic pain by adverse effects including tolerance and **addiction**.’* *‘A prototype optoactive μ-opioid receptor (optoMOR) was designed by replacing the intracellular domains from rhodopsin with those of the native μ-opioid receptor and was transiently expressed in human embryonic kidney (HEK293) cells.’* *‘Photoactivation of optoMOR decreased the Ca2+ influx and inhibited the forskolin-induced cAMP generation, activation of CREB, and BDNF levels in optoMOR-expressing cells similar to the activation of native μ-opioid receptor by DAMGO.’* •Kaczmarczyk MM et al. “Methylphenidate prevents high-fat diet (HFD)-induced learning/memory impairment in juvenile mice.” Psychoneuroendocrinology. 2013 Feb 11;. PMID 23411461 *‘ a HFD rapidly impacts dopamine metabolism in the brain appearing to trigger anxiety-like behaviors and learning/memory impairments prior to the onset of weight gain and/or pre-diabetes.’* *‘Examination of mouse cortex, hippocampus and hypothalamus for dopamine and its metabolites demonstrated increased homovanillic acid (HVA) concentrations in the hippocampus and cortex that were associated with decreased cortical BDNF gene expression.’* *‘Administration to mice of the **psychostimulant** methylphenidate prevented HFD-dependent impairment of learning/memory.’* •Geisel O et al. “Serum levels of brain-derived neurotrophic factor in patients with internet use disorder.” Psychiatry Res. 2013 Jan 30;. PMID 23375675 *‘Internet use disorder (IUD) is characterised by excessive internet gaming use and has temporarily been conceptualised as a behavioural **addiction**.’* *‘Serum levels of BDNF were not correlated with severity of IUD or clinical and demographic variables in our study.’* •Li X et al. “Different roles of BDNF in nucleus accumbens core versus shell during the incubation of cue-induced cocaine craving and its long-term maintenance.” J Neurosci. 2013 Jan 16;33(3):1130-42. PMID 23325250 *‘Brain-derived neurotrophic factor (BDNF) contributes to diverse types of plasticity, including cocaine **addiction**’.* *‘These results suggest that basal levels of BDNF transmission in the NAc core exert a suppressive effect on cocaine seeking in early withdrawal (WD1), whereas the late elevation of BDNF protein in NAc shell contributes to incubation in late withdrawal*	

	(WD90).' •Harte-Hargrove LC et al. "Brain-derived neurotrophic factor-estrogen interactions in the hippocampal mossy fiber pathway: Implications for normal brain function and disease." Neuroscience. 2012 Dec 29;. PMID 23276673 'We suggest that the interactions between BDNF and 17β-estradiol in the MFs are potentially important in the normal function of the hippocampus, and have implications for sex differences in functions that depend on the MFs and in diseases where MF plasticity has been suggested to play an important role, Alzheimer's disease, epilepsy and addiction.' [Review] •McCarthy DM et al. "Regulation of BDNF expression by cocaine." Yale J Biol Med. 2012 Dec;85(4):437-46. PMID 23239946 *'Exposure to drugs of abuse is known to modulate epigenetic regulation of BDNF gene expression. This review will discuss how exposure to cocaine, one of the most **addictive** drugs known to mankind, can produce alterations in BDNF gene expression, especially in the mesolimbic dopaminergic system, which lead to alterations in the reward-mediated behaviors involved in **addiction**.'* •Tsai SJ et al. "Increased central brain-derived neurotrophic factor activity could be a risk factor for substance abuse: Implications for treatment." Med Hypotheses. 2007;68(2):410-4. PMID 16824691 *'The increased central BDNF activity hypothesis of drug **addiction** may provide new insights for improved therapeutic strategies for the prevention and treatment of drug addiction. Several strategies to decrease central BDNF activity that have potential use in the treatment of drug addiction are proposed'* •Kim DJ et al. "High concentrations of plasma brain-derived neurotrophic factor in methamphetamine users." Neurosci Lett. 2005 Nov 11;388(2):112-5. PMID 16039058 *'Methamphetamine is a highly **addictive** drug that has a neurotoxic effect on the brain'* *'The plasma BDNF concentrations of methamphetamine users were significantly higher compared with those of controls (2536.3 pg/ml versus 1352.6 pg/ml).'*	
CYP11A1	•Xu D et al. "**Caffeine**-induced activated glucocorticoid metabolism in the hippocampus causes hypothalamic-pituitary-adrenal axis inhibition in fetal rats." PLoS One. 2012;7(9):e44497. PMID 22970234 *'Pregnant Wistar rats were intragastrically administered 20, 60, and 180 mg/kg · d **caffeine** from gestational days 11-20.'* *'The fetal adrenal cortex changed into slight and the expression of fetal adrenal steroid acute regulatory protein (StAR) and cholesterol side-chain cleavage enzyme (P450scc), as well as the level of fetal adrenal endogenous corticosterone (CORT), were all significantly decreased after **caffeine** treatment.'* •Toyoshima K et al. "Immunohistochemical identification of cells expressing steroidogenic enzymes cytochrome P450scc and P450 aromatase in **taste** buds of rat circumvallate papillae." Arch Histol Cytol. 2007 Nov;70(4):215-24. PMID 18296822 *'The present study demonstrated for the first time the localizations and patterns of expression of key enzymes for steroidogenesis, cytochrome P450 side-chain-cleavage (P450scc), and P450 aromatase in the **taste** buds of rat circumvallate papillae, using immunoblot analyses and immunohistochemistry.'*	418, 419
CYP1A1	•Kravchenko LV et al. "[Effects of green tea extract and its components on antioxidant status and activities of xenobiotic metabolizing enzymes of rats]." Vopr Pitan. 2011;80(2):9-15. PMID 21692342	420-426

	'There were significant differences in the effects of EGCG [epigallocatechin gallate]*, Qu* [quercetine] *and GTE* [green tea extract] *on the activities and expression of mRNA for CYP1A1, CYP1A2 and CYP3A1. But feeding both GTE and Cf* [**caffeine**] *to rats results in similar elevated activities of CYP1A1, CYP1A2, UDP-glucuronosyl transferase and glutathion transferase'* [Article in Russian]. •Uno Y et al. "CYP1D1, pseudogenized in human, is expressed and encodes a functional drug-metabolizing enzyme in cynomolgus monkey." Biochem Pharmacol. 2011 Feb 1;81(3):442-50. PMID 21070747 *'The amino acid sequence deduced from cynomolgus monkey CYP1D1 cDNA shared the high sequence identity (91%) with human CYP1D1P (postulated from the gene sequence), and the highest sequence identity (44-45%) with CYP1A1 and CYP1A2 among cynomolgus monkey P450s.'* *'Cynomolgus monkey CYP1D1 protein heterologously expressed in Escherichia coli catalyzed ethoxyresorufin O-deethylation and **caffeine** 8-hydroxylation, which CYP1As also catalyze'* •Mills BM et al. "Current cytochrome P450 phenotyping methods applied to metabolic drug-drug interaction prediction in dogs." Drug Metab Dispos. 2010 Mar;38(3):396-404. PMID 20007294 *'The minor **caffeine**-fluvoxamine interaction (1.78-fold) was slightly higher than predicted values based on determination of a moderate f(m) value for CYP1A1, although CYP1A2 may also be involved in **caffeine** metabolism.'* •Eugster HP et al. "**Caffeine**, estradiol, and progesterone interact with human CYP1A1 and CYP1A2. Evidence from cDNA-directed expression in Saccharomyces cerevisiae." Drug Metab Dispos. 1993 Jan-Feb;21(1):43-9. PMID 8095225 *'**Caffeine** was shown to be metabolized by CYP1A2 and CYP1A1. Both enzymes formed paraxanthine and minor amounts of theobromine; however, trimethyluric acid was exclusively formed by CYP1A1.'* •Wiercinska P et al. "The roles of different porcine cytochrome P450 enzymes and cytochrome b5A in skatole metabolism." Animal. 2012 May;6(5):834-45. PMID 22558931 *'Boar taint is the unfavourable odour and **taste** from pork fat, which results in part from the accumulation of skatole (3-methylindole, 3MI).'* *'The results show that pig CYP1A1, CYP2A19, CYP2C33v4, CYP2C49, CYP2E1 and CYP3A and human CYP2E1 (hCYP2E1) are all capable of producing the major skatole metabolite 3-methyloxyindole (3MOI), as well as indole-3-carbinol (I3C), 5-hydroxy-3-methylindole (5-OH-3MI), 6-OH-3MI, 2-aminoacetophenone (2AAP) and 3-hydroxy-3-methyloxindole.'* •Smolowitz RM et al. "Cytochrome P4501A induction in tissues, including olfactory epithelium, of topminnows (Poeciliopsis spp.) by waterborne benzo[a]pyrene." Carcinogenesis. 1992 Dec;13(12):2395-402. PMID 1473249 *'We examined induction of cytochrome P4501A (CYP1A) in liver and other organs of the species P. monacha and P. lucida exposed to benzo[a]pyrene (B[a]P) in water (added in acetone carrier) at 1 mg/l for 48 and 90 h.'* *'There was a very strong specific induction by B[a]P in **olfactory** epithelium and epidermal **taste** bud epithelium of P. monacha, the first demonstration of strong CYP1A induction in chemosensory epithelia exposed to inducer in a physiologically relevant way.* •Bromek E et al. "The ability of cytochrome P450 2D isoforms to synthesize dopamine in the brain: An in vitro study." Eur J Pharmacol. 2010 Jan 25;626(2-3):171-8. PMID 19818757 *'The study was conducted with cDNA-expressed CYP isoforms (rat CYP1A1, 2A2, 2B1, 2C6/11/13, 2D1/2/4/18, 2E1, 3A2 and human CYP2D6) and with rat brain microsomes. Of the rat CYP isoforms tested, only CYP2D2, 2D4 and 2D18 (but not CYP2D1) were capable of forming dopamine from tyramine.'* *'The results are discussed in the context of the likelihood of CYP2D-mediated dopamine synthesis in vivo, the implications for*	

	*Parkinson's disease and the **addiction** process'*	
CYP1A2	•Dobrinas M et al. "Pharmacogenetics of CYP1A2 activity and inducibility in smokers and exsmokers." Pharmacogenet Genomics. 2013 May;23(5):286-92. PMID 23492909 *'A significant influence on CYP1A2 inducibility* [**caffeine** metabolite ratio] *was observed for the NR1I3 rs2502815 (P=0.0026), rs4073054 (P=0.029), NR2B1 rs3818740 (P=0.0045), rs3132297 (P=0.036), AhR rs2282885 (P=0.040), rs2066853 (P=0.019), NR1I1 rs2228570 (P=0.037), and NR1I2 rs1523130 (P=0.044) polymorphisms. Among these, the NR1I3 rs2502815 (P=0.0045), rs4073054 (P=0.048), and NR2B1 rs3818740 (P=0.031) also influenced CYP1A2 basal activity'* •Lin W et al. "Evaluation of the effect of TM208 on the activity of five cytochrome P450 enzymes using on-line solid-phase extraction HPLC-DAD: A cocktail approach." J Chromatogr B Analyt Technol Biomed Life Sci. 2013 Apr 1;923-924:29-36. PMID 23466445 *'A rapid, simple, and sensitive on-line solid-phase extraction HPLC-DAD method for simultaneous evaluation of the activity of five CYP450 isoforms (CYP1A2, CYP2C19, CYP2D6, CYP2E1 and CYP3A4) in vivo has been developed and validated. The five specific probe substrates include **caffeine** (1A2), metoprolol (2D6), dapsone (3A4), omeprazole (2C19) and chlorzoxazone (2E1).'* •Ke AB et al. "A Physiologically Based Pharmacokinetic Model to Predict Disposition of CYP2D6 and CYP1A2 Metabolized Drugs in Pregnant Women." Drug Metab Dispos. 2013 Apr;41(4):801-13. PMID 23355638 *We refined and verified our previously published pregnancy* [physiologically based pharmacokinetic] *PBPK model by incorporating cytochrome P450 CYP1A2 suppression (based on **caffeine** PK) and CYP2D6 induction (based on metoprolol PK) into the model.* •Hung WT et al. "Genetic susceptibility to dioxin-like chemicals' induction of cytochrome P4501A2 in the human adult linked to specific AhRR polymorphism." Chemosphere. 2013 Mar;90(9):2358-64. PMID 23168330 *'The goal of this study was to determine the relationship between inducibility of CYP1A2* [**caffeine** breath test] *and genetic polymorphisms of AhR, ARNT, and AhRR in human AhRR (rs2292596) genotypes predict the inducibility of CYP1A2 in people highly exposed to toxic dioxin-like chemicals'* •Doroshyenko O et al. "Drug Cocktail Interaction Study on the Effect of the Orally Administered Lavender Oil Preparation Silexan on Cytochrome P-450 Enzymes in Healthy Volunteers." Drug Metab Dispos. 2013 Feb 11;. PMID 23401474 *Repeated Silexan (160 mg/day) administration has no clinically relevant inhibitory or inducing effects on the CYP 1A2* [activity, **caffeine** metabolite ratio*], 2C9, 2C19, 2D6 and 3A4 enzymes in vivo.* •Perera V et al. "Diurnal variation in CYP1A2 enzyme activity in South Asians and Europeans." J Pharm Pharmacol. 2013 Feb;65(2):264-70. PMID 23278694 *This study observed diurnal variation of CYP1A2 activity* [**caffeine** metabolite ratio] *in South Asians, resulting in lower enzyme activity in the evening* •Li Q et al. "Assessment of effects of IR and IPC on activities of cytochrome P450 isozymes in rats by a five-drug cocktail approach." Drug Dev Ind Pharm. 2013 Jan 23;. PMID 23339682 *Cocktail approach was used to evaluate the influence of [ischemia and reperfusion] IR and [ischemic preconditioning] IPC on the*	358, 361, 427-437

*activities of CYP1A2, CYP2C9, CYP2E1, CYP2D6 and CYP3A4, which were reflected by the changes of pharmacokinetic parameters of five specific probe drugs: **caffeine**, chlorzoxazone, tolbutamide, metoprolol and midazolam, respectively. IR can variably decrease the activities of CYP isozymes in rats and this decrease can be attenuated by IPC*

•Vaynshteyn D et al. "**Caffeine** induces CYP1A2 expression in rat hepatocytes but not in human hepatocytes." Drug Metab Lett. 2012 Jun 1;6(2):116-9. PMID 23167901

*Our results from luciferase assays performed in HepG2 cells showed that **caffeine** is not an activator of the aromatic hydrocarbon receptor (AhR), a major transcription factor involved in upregulation of CYP1A2.*

•Dumond JB et al. "A phenotype-genotype approach to predicting CYP450 and P-glycoprotein drug interactions with the mixed inhibitor/inducer tipranavir/ritonavir." Clin Pharmacol Ther. 2010 Jun;87(6):735-42. PMID 20147896

*'Multiple dosing produced weak induction of CYP1A2 [**caffeine** metabolite ratio], moderate induction of CYP2C19, potent induction of intestinal P-gp, and potent inhibition of CYP2D6 and CYP3A '*
*'CYP1A2, NAT-2, and XO genetic polymorphisms showed no effects on **caffeine** pharmacokinetics'*

•Blake MJ et al. "Effect of diet on the development of drug metabolism by cytochrome P-450 enzymes in healthy infants." Pediatr Res. 2006 Dec;60(6):717-23. PMID 17065585

*Formula feeding appears to accelerate maturation of **caffeine** and* [dextromethorphan] *DM metabolism by increasing the activity of CYP1A2 and CYP3A4, respectively*

•Ma JD et al. "Maribavir pharmacokinetics and the effects of multiple-dose maribavir on cytochrome P450 (CYP) 1A2, CYP 2C9, CYP 2C19, CYP 2D6, CYP 3A, N-acetyltransferase-2, and xanthine oxidase activities in healthy adults." Antimicrob Agents Chemother. 2006 Apr;50(4):1130-5. PMID 16569820

'Maribavir (1263W94, VP-41263) is an oral anticytomegalovirus agent under clinical development. The pharmacokinetics and safety of maribavir and the effects of maribavir on the activities of cytochrome P450 (CYP) 1A2, CYP 2C9, CYP 2C19, CYP 2D6, CYP 3A, N-acetyltransferase-2 (NAT-2), and xanthine oxidase (XO) were evaluated in a randomized, double-blind, placebo-controlled study.'
'Maribavir did not affect the CYP 1A2, CYP 2C9, CYP 3A, NAT-2, or XO activities.'
*'**Taste** disturbance was the most frequently reported adverse event.'*

•Kot M et al. "**Caffeine** as a marker substrate for testing cytochrome P450 activity in human and rat." Pharmacol Rep. 2008 Nov-Dec;60(6):789-97. PMID 19211970 [Review]

'The current knowledge on the involvement of cytochrome P450 (P450, CYP) isoforms in the metabolism of **caffeine** in rat and human liver is reviewed. Attention is also paid to species- and concentration-dependent metabolism of **caffeine**. Finally, we discuss the P450-mediated metabolism of **caffeine** in relation to **coffee addiction** and drug interactions.'

•Pardo Lozano R et al. "[**Caffeine**: a nutrient, a drug or a drug of abuse]." Adicciones. 2007;19(3):225-38. PMID 17724925

*'Basically, **caffeine** is metabolized by the hepatic cytochrome P-450 1A2 enzymes (CYP1A2).'*
*'Finally, **caffeine** can be considered a **drug of abuse**. It has positive reinforcing actions, produces tolerance, and a withdrawal syndrome after stopping its consumption. **Caffeine** can cause different mental disorders such as **dependence**, which is not included in the DSM-IV-R, withdrawal syndrome and intoxication. Depending on its use, **caffeine** can be considered a nutrient, a drug or a drug of abuse'*

SLC6A4	•Keating E et al. "The effect of a series of organic cations upon the plasmalemmal serotonin transporter, SERT." Life Sci. 2004 Nov 19;76(1):103-19. PMID 15501483 *Initial rates of (3)H-serotonin ((3)H-5HT; 200 nM) uptake were not changed by some of the organic cations tested (guanidine, N-methylnicotinamide, choline, atenolol,* ***caffeine*** *and theophylline),* •Alaux-Cantin S et al. "Alcohol intoxications during adolescence increase motivation for alcohol in adult rats and induce neuroadaptations in the nucleus accumbens." Neuropharmacology. 2013 Apr;67:521-31. PMID 23287538 *'We evaluated the effect of adolescent binge-like ethanol intoxication on vulnerability to alcohol abuse in Sprague-Dawley rats.'* *'In young adult animals, we measured free ethanol consumption in the two-bottle choice paradigm, motivation for ethanol in the operant self-administration task and both ethanol's rewarding and aversive properties in the conditioned place preference (CPP) and* ***taste*** *aversion (CTA) paradigms.'* *'This vulnerability to ethanol abuse was associated with a lower c-Fos immunoreactivity in the Nac and enduring alterations of the expression of Penk and Slc6a4, 2 neurotransmission-related genes that have been shown to play critical roles in the behavioral effects of ethanol and alcoholism.'* •Wheeler JM et al. "Genetically correlated effects of selective breeding for high and low methamphetamine consumption." Genes Brain Behav. 2009 Nov;8(8):758-71. PMID 19689456 *'We produced mouse lines that orally self-administer high (MAHDR) or low (MALDR) amounts of methamphetamine, representing the first demonstration of selective breeding for self-administration of any* ***psychostimulant*** *drug. Conditioned place preference and* ***taste*** *aversion results indicate that MAHDR mice are relatively more sensitive to the rewarding effects and less sensitive to the aversive effects of methamphetamine, compared to MALDR mice.'* *'Genes differentially expressed in the drug-naï ve state, including Slc6a4 (serotonin transporter), Htr3a (serotonin receptor 3A), Rela [nuclear factor kappaB (NFkappaB)] and Fos (cFos), represent candidates whose expression levels may predict methamphetamine consumption and susceptibility to methamphetamine reward and aversion'* •Moy SS et al. "Social approach in genetically engineered mouse lines relevant to autism." Genes Brain Behav. 2009 Mar;8(2):129-42. PMID 19016890 *'In this study, a three-chambered choice task was used to evaluate sociability and social novelty preference in five lines of mice with mutations in genes implicated in autism spectrum disorders.'* *'Slc6a4, Igf-1, En2 and Dhcr7 mice: (1) neurobehavioral screen and home cage observation, (2) activity in an open field, (3) rotarod, (4) social approach test, (5) buried food test for* ***olfactory*** *ability and (6) elevated plus maze.'* *'there were no differences between wild-type and mutant mice within each study for body weight, anxiety-like behavior on the elevated plus maze, motor coordination on an accelerating rotarod or performance in the buried food test for* ***olfactory*** *ability.'* *'Male mice with targeted disruption of Slc6a4 displayed significantly less sociability than wild-type controls.'* •Gómez C et al. "Heterogeneous targeting of centrifugal inputs to the glomerular layer of the main olfactory bulb." J Chem Neuroanat. 2005 Jun;29(4):238-54. PMID 15927786 *'The centrifugal systems innervating the olfactory bulb are important elements in the functional regulation of the* ***olfactory*** *pathway. In this study, the selective innervation of specific glomeruli by serotonergic, noradrenergic and cholinergic centrifugal axons was analyzed.'* *'Serotonin-, serotonin transporter-immunostaining and acetylcholinesterase-staining revealed a higher heterogeneity in the glomerular layer of the main olfactory bulb than previously reported.'*	333, 438-446

•Elkins RL et al. “Serotonin reuptake is less efficient in **taste** aversion resistant than in **taste** aversion-prone rats.” Pharmacol Biochem Behav. 2000 Jul;66(3):609-14. PMID 10899378
*‘Earlier studies demonstrated that the **taste** aversion resistant (TAR) animals exhibited lower concentrations of brain serotonin and consumed greater amounts of ethanol than their **taste** aversion prone (TAP) counterparts. In the present study, TAR rats demonstrated significantly less efficient brain serotonin transport compared to TAP rats, but the rat lines demonstrated similar levels of serotonin transporter or V(max) and similar whole brain paroxetine (a specific serotonin reuptake inhibitor) binding (B(max)).’*

•Ren Y et al. “Immunocytochemical localization of serotonin and serotonin transporter (SET) in **taste** buds of rat.” Brain Res Mol Brain Res. 1999 Dec 10;74(1-2):221-4. PMID 10640694

*‘We used an immunocytochemical approach to study the localization of serotonin and its termination system, serotonin transporter (SET), in the **taste** buds of rats using specific antibodies against serotonin and SET. Under confocal laser scanning microscopy, both serotonin and SET immunoreactivity were detected in the **taste** buds of rat vallate papillae.’*
*‘SET-immunoreactivity was mainly localized in the periphery or interfaces between the **taste** cells’.*

•Hansson SR et al. “Serotonin transporter messenger RNA expression in neural crest-derived structures and sensory pathways of the developing rat embryo.” Neuroscience. 1999 Mar;89(1):243-65. PMID 10051233

*‘Several sensory organs (cochlear and retinal ganglionic cells, **taste** buds, whisker and hair follicles) contained serotonin transporter messenger RNA by late gestation. The expression of serotonin transporter messenger RNA throughout the sensory pathways from central nervous system relay stations [Hansson S. R. et al. (1997) Neuroscience 83, 1185-1201; Lebrand C. et al. (1996) Neuron 17, 823-835] to sensory nerves and target organs as shown in this study suggests that serotonin may regulate peripheral synaptogenesis, and thereby influence later processing of sensory stimuli.’*

•Alaux-Cantin S et al. “Alcohol intoxications during adolescence increase motivation for alcohol in adult rats and induce neuroadaptations in the nucleus accumbens.” Neuropharmacology. 2013 Apr;67:521-31. PMID 23287538

‘We evaluated the effect of adolescent binge-like ethanol intoxication on vulnerability to alcohol abuse in Sprague-Dawley rats.’
*‘As the nucleus accumbens (Nac) is particularly involved in **addictive** behavior, we analyzed IEI-induced long-term neuroadaptations in the Nac using c-Fos immunohistochemistry and an array of neurotransmission-related genes. This vulnerability to ethanol abuse was associated with a lower c-Fos immunoreactivity in the Nac and enduring alterations of the expression of Penk and Slc6a4, 2 neurotransmission-related genes that have been shown to play critical roles in the behavioral effects of ethanol and alcoholism.’*

•Sørensen L et al. “Interaction of antidepressants with the serotonin and norepinephrine transporters: mutational studies of the S1 substrate binding pocket.” J Biol Chem. 2012 Dec 21;287(52):43694-707. PMID 23086945

*‘A wide range of inhibitors of SERT and NET are used as treatment of depression and anxiety disorders or as **psychostimulant** drugs of abuse.’*
‘In this study, we determine the effect of mutating six key S1 residues in human SERT (hSERT) and NET (hNET) on the potency of 15 prototypical SERT/NET inhibitors belonging to different drug classes. Analysis of the resulting drug sensitivity profiles provides novel information on drug binding modes in hSERT and hNET and identifies specific S1 residues as important molecular determinants for inhibitor potency and hSERT/hNET selectivity’

•Naumenko VS et al. “[Implication of 5-HT2A receptors in the genetic mechanisms of the brain 5-HT system autoregulation].” Mol Biol (Mosk). 2012 May-Jun;46(3):416-22. PMID 22888631

*’ Brain serotonin (5-HT) system has been implicated in pathophysiology of anxiety, depression, drug **addiction**, and schizophrenia’*
‘Here, we investigated the role of 5-HT2A receptor in the autoregulation of the brain 5-HT system.’
‘The chronic treatment with agonist of 5-HT2A receptor DOI (1.0 mg/kg, i.p./14 days) produced considerable decrease of 5-HT2A

	receptor-mediated "head-twitches" in AKR/J mice indicating desensitization of 5-HT2A receptors. Chronic DOI treatment failed to alter 5-HT2A receptor gene expression in the midbrain, hippocampus and frontal cortex.'	
COX5A	•Mulligan MK et al. "Toward understanding the genetics of alcohol drinking through transcriptome meta-analysis." Proc Natl Acad Sci U S A. 2006 Apr 18;103(16):6368-73. PMID 16618939 *'Much evidence from studies in humans and animals supports the hypothesis that alcohol **addiction** is a complex disease with both hereditary and environmental influences.'* *'Microarray analyses of brain gene expression in three selected lines, and six isogenic strains of mice known to differ markedly in voluntary alcohol consumption provided >4.5 million data points for a meta-analysis.'* *'cis-regulated candidate genes for an alcohol preference quantitative trait locus on chromosome 9 were identified: Arhgef12, Carm1, Cryab, Cox5a, Dlat, Fxyd6, Limd1, Nicn1, Nmnat3, Pknox2, Rbp1, Sc5d, Scn4b, Tcf12, Vps11, and Zfp291 and four ESTs.'*	[447]
GIT1	•Won H et al. "GIT1 is associated with ADHD in humans and ADHD-like behaviors in mice." Nat Med. 2011 May;17(5):566-72. PMID 21499268 *'An intronic single-nucleotide polymorphism in GIT1, the minor allele of which causes reduced GIT1 expression, shows a strong association with ADHD susceptibility in humans. Git1-deficient mice show ADHD-like phenotypes, with traits including hyperactivity, enhanced electroencephalogram theta rhythms and impaired learning and memory. Hyperactivity in Git1(-/-) mice is reversed by amphetamine and methylphenidate, **psychostimulant**s commonly used to treat ADHD.'* *'Our study identifies a previously unknown involvement of GIT1 in human ADHD and shows that GIT1 deficiency in mice causes **psychostimulant**-responsive ADHD-like phenotypes'*	[344]

Table S27. Associations between coffee consumption loci and other traits (from GWAS catalogue)[a]

Locus Lead SNP	SNP	r^2 with Lead SNP CEU/YRI	EA	Trait	Ref
2p24 rs1260326	rs1260326	1	C	↓serum albumin	38
	rs1260326	1	C	? 5-HDL-P	448
	rs1260326	1	C	↑non-albumin protein	449
	rs1260326	1	C	↓Ala/Gln ratio	450
	rs1260326	1	C	↓platelet counts	451
	rs1260326	1	C	↓GGT	452
	rs1260326	1	C	↓TGs	116, 159, 453-455
	rs1260326	1	C	↓CRP	456
	rs1260326	1	C	↓tCHOL	116
	rs1260326	1	C	↓hypertriglyceridemia	457
	rs1260326	1	C	↓CKD	458
	rs1260326	1	C	↓2-hr glucose challenge	104
	rs1260326	1	C	↑mannose	85
	rs1260326	1	C	↓Ala	85
	rs1260333	.81/.09	G	↓TGs	459
	rs780093	.93/.64	C	↓SHBG	460
	rs780093	.93/.64	C	↓Crohn's disease	461
	rs780094	.93/.43	C	↓MetS	462
	rs780094	.93/.43	C	↓glucose/mannose ratio	463
	rs780094	.93/.43	C	↓docosapentaenoic acid	464
	rs780094	.93/.43	C	↑fasting glucose	103
	rs780094	.93/.43	C	↑HOMA-IR	103
	rs780094	.93/.43	C	↑fasting insulin	103
	rs780094	.93/.43	C	↓uric acid	465
	rs780094	.93/.43	C	↓TGs	466
	rs780094	.93/.43	C	↓CRP	467
	rs780094	.93/.43	C	↓TGs	468, 469
4q22 rs1481012	rs1481012	1	A	↑response to statins (LDL)	470
	rs1481012	1	A	↓uric acid	471
	rs2199936	.84/na	G	↓uric acid	453, 472
	rs2231142	.92/na	G	↓uric acid	465, 473
7p21 rs6968554	rs6968554	1	G	↓caffeine	463
7q11.23 rs7800944	rs2286276	.81/.44	T	↓TGs	474
11p13 rs6265	rs6265	1	C	↑BMI	475
	rs6265	1	C	↑smoking behavior	476
15q24 rs2472297	rs2472297	1	T	↓caffeine[b]	85
	rs1378942[c]	.10/na	A	↓diastolic blood pressure	477

EA, effect allele; na, not available

[a]Traits associated with lead SNP (or close proxies: CEU: $r^2>0.80$) according to previous GWAS. Data were obtained from the National Human Genome Research Institute GWAS catalogue[84] ($P<5\times10^{-8}$) and the Metabolomics GWAS server ($P<1.03\times10^{-10}$)[85]. Any ambiguities were resolved by reviewing the original publications.

[b]Borderline significant ($P<1.50\times10^{-10}$) accordinging to Shin et al [85].

[c]rs1378942 A, in low LD with rs2472297, was also associated with higher coffee consumption ($P<1.46\times10^{-17}$) in stage 1 of the current report.

Table S28. Association between confirmed loci and coffee consumption (cups/d, phenotype 1) with further adjustment for plasma lipids in the Women's Genome Health Study

CHR	SNP	EA/NEA	Model 1[a] Original Model n=17,332		Model 2[bd] Original Model + triglycerides[e] n=17,252		Model 3[cd] Original Model + cholesterol[e] n=17,252	
			β (SE)	P	β (SE)	P	β (SE)	P
2	rs1260326	T/C	-0.028 (0.02)	0.10	-0.011 (.02)	0.53	-0.027(0.02)	0.12
4	rs1481012	G/A	-0.023 (0.03)	0.38	-0.019 (0.03)	0.48	-0.023 (0.03)	0.40
7	rs4410790	T/C	-0.183 (0.02)	2.41E-27	-0.183(0.02)	1.75E-27	-0.183 (0.02)	4.68E-27
7	rs6968554	A/G	-0.179 (0.02)	3.22E-26	-0.179 (0.02)	2.18E-26	-0.178 (0.02)	6.37E-26
7	rs7800944	C/T	0.056 (0.02)	2.89E-03	0.046 (0.02)	1.48E-02	0.055 (0.02)	2.47E-03
7	rs17685	A/G	0.067 (0.02)	2.69E-05	0.071 (0.02)	1.13E-04	0.068 (0.02)	2.38E-04
11	rs6265	T/C	-0.055 (0.02)	0.01	-0.054 (0.02)	0.01	-0.053 (0.02)	0.01
15	rs2470893	T/C	0.132 (0.02)	9.73E-14	0.132 (0.02)	8.89E-14	0.132 (0.02)	1.13E-13
15	rs2472297	T/C	0.163 (0.02)	3.66E-14	0.162 (0.02)	4.06E-14	0.163 (0.02)	4.08E-14
17	rs9902453	G/A	0.033 (0.02)	0.05	0.032 (0.02)	5.09E-02	0.032 (0.02)	5.18E-02

EA, effect allele; NEA, non effect allele

[a]Model 1: adjusted for age, smoking, randomization, EVs (Supplementary Table S7).

[b]Model 2: Model 1 further adjusted for natural log transformed plasma triglycerides.

[c]Model 3: Model 1 further adjusted for plasma cholesterol.

[d]Model 2 and Model 3 results are not significantly different from Model 1 results (P≥0.48).

[e]Methods and assays for lipids have been described in detail previously[478].

Table S29. Association between confirmed loci and coffee consumption (cups/d, phenotype 1) with further adjustment for fasting plasma glucose in the TwinGene Study

CHR	SNP	EA/NEA	Model 1[a] Original Model n=8952		Model 2[bc] Original Model + glucose[d] n≤8846[e]	
			β (SE)	P	β (SE)	P
2	rs1260326	T/C	-0.079 (0.04)	0.03	-0.071 (0.04)	0.05
4	rs1481012	G/A	-0.086 (0.06)	0.13	-0.067 (0.06)	0.25
7	rs4410790	T/C	-0.194(0.04)	2.09E-07	-0.198 (0.04)	2.34E-07
7	rs6968554	A/G	-0.222 (0.04)	4.03E-10	-0.218 (0.04)	1.27E-09
7	rs7800944	C/T	0.0901 (0.04)	0.02	0.097 (0.04)	0.02
7	rs17685	A/G	0.150 (0.04)	5.84E-5	0.152 (0.04)	5.61E-05
11	rs6265	T/C	-0.159 (0.04)	2.19E-04	-0.164 (0.04)	1.46E-04
15	rs2470893	T/C	0.282 (0.04)	9.60E-14	0.277 (0.04)	3.61E-13
15	rs2472297	T/C	0.347 (0.04)	6.23E-18	0.345 (0.04)	1.96E-17
17	rs9902453	G/A	0.068 (0.04)	0.06	0.071 (0.04)	0.05

EA, effect allele; NEA, non effect allele

[a]Model 1: adjusted for age, sex, smoking, PCs (Supplementary Table S7).

[b]Model 2: Model 1 further adjusted for fasting plasma glucose.

[c]Model 2 and Model 3 results are not significantly different from Model 1 results (P≥0.82).

[d]Glucose measures were performed by the Karolinska Hospital laboratory according to their standard for clinical blood samples using the Synchron Lx System (Beckman Coulter).

[e]Results based on samples (n range: 8183 to 8846) with available phenotype 1, glucose measures and genotypes.

Table S30. Results of GRAIL analysis[a]

Locus	SNP	GENE	GRAIL P-value	Similar genes (overall rank)
8 query regions, 0 seed genes				
2p24	rs1260326	*GCKR*	0.03	*MLXIPL* (28), *MPI* (48)
4q22	rs1481012	*ABCG2*	0.38	*SLC6A4* (134)
7p21	rs6968554	*AHR*	0.08	*CYP1A1* (6), *CYP1A2* (7)
7q11.23	rs7800944	*TBL2*	0.29	*BCL7B* (4), *BAZ1B (10), FZD9 (32), MLXIPL (88)*
7q11.23	rs17685	*POR*	0.044	*CYP1A2 (29), COX5A (43), CYP1A1 (48)*
11p13	rs6265	*BDNFOS*	0.088	*BDNF* (1), *EFCAB5* (14)
15q24	rs2472297	*CYP1A1*	0.00016	*CYP1A2 (1), AHR (3), POR (14), COX5A(57)*
17q11.2	rs9902453	*SLC6A4*	0.10	*SLC5A6 (130), CYP1A1 (175)*
2 query regions, 6 seed genes				
2p24	rs1260326	*SLC5A6*	0.015	*SLC6A4* (126)
4q22	rs1481012	*seed ABCG2*	n/a	n/a
7p21	rs6968554	*seed AHR*	n/a	n/a
7q11.23	rs7800944	*MLXIPL*	0.32	
7q11.23	rs17685	*seed POR*	n/a	n/a
11p13	rs6265	*seed BDNF*	n/a	n/a
15q24	rs2472297	*seed CYP1A2*	n/a	n/a
17q11.2	rs9902453	*seed SLC6A4*	n/a	n/a

[a]Results of GRAIL analysis, detailing the genes (column 3) at each confirmed locus with the most significantly associated other genes (column 5) based on text-mining of literature abstracts. Shown are P values (column 4) for co-occurrence in published abstracts and a list of similar genes within the top 200 highest rankings among 18,875 genes tested.

Table S31. Genes highlighted by DAPPLE analysis due to connectivity

Locus	SNP	GENE	DAPPLE corrected P-value
8 query regions, 0 genes specified			
2p24	rs1260326	*SNX17*	0.56
4q22	rs1481012	*ABCG2*	0.99
7p21	rs6968554	*AHR*	0.54
7q11.23	rs7800944	*MLXIPL*	0.27
7q11.23	rs17685	*POR*	0.63
11p13	rs6265	*BDNF*	0.99
15q24	rs2472297	*CYP1A2*	0.06
17q11.2	rs9902453	*CCDC55*	0.58
2 query regions, 6 genes specified			
2p24	rs1260326	*PPM1G*	0.91
4q22	rs1481012	*specified ABCG2*	0.71
7p21	rs6968554	*specified AHR*	0.36
7q11.23	rs7800944	*MLXIPL*	0.91
7q11.23	rs17685	*specified POR*	*0.06*
11p13	rs6265	*specified BDNF*	*0.52*
15q24	rs2472297	*specified CYP1A2*	*0.002*
17q11.2	rs9902453	*specified SLC6A4*	*0.62*

Table S32. Association between coffee consumption-loci and coffee-implicated diseases and traits

A. Type 2 diabetes[a] and related traits[b]

Locus	SNP	EA[c]/NEA	EAF	Type 2 diabetes		HbA1c		Fasting glucose		2 hr glucose challenge (adjusted for BMI)		Fasting insulin		HOMA-B		HOMA-IR		Proinsulin	
				Effect	P	Effect	P	Effect	P	Effect	P	Effect	P	Effect	P	Effect	P	Effect	P
2p24	rs1260326	C/T	0.59	+	0.06	+	0.31	+	**4.25E-13**	-	**1.53E-06**	+	**1.22E-04**	+	0.33	+	**9.16E-07**	+	0.06
4q22	rs1481012	A/G	0.89	+	0.73	-	0.53	-	0.91	+	0.30	-	0.49	-	0.18	-	0.56	+	0.83
7p21	rs4410790	C/T	0.63	+	0.36	+	0.56	-	0.82	-	0.34	+	0.15	+	0.30	+	0.44	+	0.45
	rs6968554	G/A	0.61	+	0.26	+	0.36	-	0.90	-	0.24	+	**0.05**	+	0.20	+	0.22	+	0.58
7q11.23	rs7800944	C/T	0.28	+	0.53	+	0.15	+	0.19	+	0.38	+	0.31	+	0.83	+	0.30	+	0.78
7q11.23	rs17685	A/G	0.29	n/a	n/a	-	0.42	+	0.64	n/a	n/a	-	0.26	-	0.24	-	0.19	-	0.53
11p13	rs6265	C/T	0.81	+	0.42	+	**0.04**	+	0.79	-	0.17	+	0.39	+	0.49	+	0.35	+	**0.05**
15q24	rs2470893	T/C	0.31	+	0.49	+	0.41	+	0.74	+	0.82	+	0.45	+	0.07	+	0.23	-	0.20
	rs2472297	T/C	0.24	-	0.74	+	0.82	+	0.62	+	0.54	+	0.86	+	0.25	+	0.45	-	0.12
17q11.2	rs9902453	G/A	0.46	-	**0.05**	+	0.54	+	0.10	-	0.20	-	0.89	-	0.23	+	0.97	-	0.42

EA, effect allele; NEA, non-effect allele; EAF, effect allele frequency
[a]Summary-level results for type 2 diabetes[107] were contributed by DIAGRAM+ investigators and were downloaded from http://www.well.ox.ac.uk/DIAGRAM.
[b]Summary-level results for glycaemic traits[103-106] including HbAlc, fasting glucose, 2 hr glucose challenge, fasting insulin, HOMA-B, HOMA-IR and proinsulin were contributed by MAGIC investigators and were downloaded from www.magicinvestigators.org.
[c]Effect allele corresponds to the allele associated with increased coffee consumption.

B. Blood lipid traits[a]

Locus	SNP	EA[b]/NEA	EAF	LDL		Total Cholesterol		Triglycerides		HDL	
				Effect	P	Effect	P	Effect	P	Effect	P
2p24	rs1260326	C/T	0.59	-	**2.33E-04**	-	**7.31E-27**	-	**5.68E-133**	+	0.08
4q22	rs1481012	A/G	0.89	-	0.07	-	0.09	-	0.13	+	**0.05**
7p21	rs4410790	C/T	0.63	+	0.24	+	0.36	+	**0.04**	-	**2.94E-03**
	rs6968554	G/A	0.61	+	0.36	+	0.58	+	0.06	-	**1.18E-03**
7q11.23	rs7800944	C/T	0.28	+	0.59	-	**0.01**	-	**1.02E-41**	+	**2.24E-03**
7q11.23	rs17685	A/G	0.29	-	0.87	+	0.86	+	**0.04**	-	0.28
11p13	rs6265	C/T	0.81	+	0.44	+	0.32	+	0.22	-	0.57
15q24	rs2470893	T/C	0.31	+	0.27	+	0.47	-	0.12	+	0.80
	rs2472297	T/C	0.24	+	0.71	+	0.93	-	0.09	+	0.49
17q11.2	rs9902453	G/A	0.46	-	0.44	-	0.77	+	0.35	+	0.17

EA, effect allele; NEA, non-effect allele; EAF, effect allele frequency
[a]Summary-level results for all blood lipid traits were contributed by Teslovich et al, *Nature* (2010)[116] and were downloaded from http://www.sph.umich.edu/csg/abecasis/public/lipids2010 .
[b]Effect allele corresponds to the allele associated with increased coffee consumption.

C. Birth weight[a] and anthropometric[b] and blood pressure traits[c]

Locus	SNP	EA[d]/NEA	EAF	Birth weight		Waist-to-Hip		Body mass index		Systolic blood pressure		Diastolic blood pressure	
				Effect	P	Effect	P	Effect	P	Effect	P	Effect	P
2p24	rs1260326	C/T	0.59	-	0.61	-	**3.40E-04**	+	0.13	-	0.30	-	0.76
4q22	rs1481012	A/G	0.89	+	0.27	-	0.50	+	**4.85E-03**	+	0.06	+	0.28
7p21	rs4410790	C/T	0.63	+	0.13	+	0.04	+	0.34	-	0.43	-	0.42
	rs6968554	G/A	0.61	+	0.16	+	**0.03**	+	0.32	-	0.20	-	0.23
7q11.23	rs7800944	C/T	0.28	+	**2.1E-03**	-	0.57	+	0.68	-	0.06	-	0.11
7q11.23	rs17685	A/G	0.29	+	0.92	n/a	n/a	+	0.35	n/a	n/a	n/a	n/a
11p13	rs6265	C/T	0.81	+	0.51	+	0.66	+	**1.88E-12**	-	0.06	-	**6.58E-04**
15q24	rs2470893	T/C	0.31	+	**0.03**	+	0.35	+	0.07	-	**2.14E-04**	-	**1.26E-05**
	rs2472297	T/C	0.24	+	0.21	+	0.78	+	0.25	-	**6.81E-05**	-	**6.75E-06**
17q11.2	rs9902453	G/A	0.46	-	0.79	-	0.83	-	0.50	-	**6.05E-03**	-	**0.02**

EA, effect allele; NEA, non-effect allele; EAF, effect allele frequency; n/a, not available

[a]Summary-level results for birth weight[110] were contributed by the EGG Consortium and were downloaded from www.egg-consortium.org

[b]Summary-level results for waist-to-hip ratio[108] and body-mass index[109] were contributed by GIANT investigators and were downloaded from http://www.broadinstitute.org/collaboration/giant/index.php/GIANT_consortium_data_files

[c]Summary-level results for systolic and diastolic blood pressure were contributed by ICBP investigators and were downloaded from http://www.georgehretlab.org/icbp_088023401234-9812599.html. Beta-coefficients were not made public and were thus obtained via personal communication.

[d]Effect allele corresponds to the allele associated with increased coffee consumption.

D. Smoking behavior traits[a]

Locus	SNP	EA[b]/NEA	EAF	Cigarettes per day		Age of initiation		Ever vs. Never		Current vs. Former	
				Effect	P	Effect	P	Effect	P	Effect	P
2p24	rs1260326	C/T	0.59	-	0.32	-	0.88	-	**0.04**	-	0.57
4q22	rs1481012	A/G	0.89	+	0.83	+	0.35	+	0.36	-	0.20
7p21	rs4410790	C/T	0.63	-	0.45	+	0.14	-	0.29	-	0.16
	rs6968554	G/A	0.61	-	0.53	+	0.23	-	0.35	-	0.13
7q11.23	rs7800944	C/T	0.28	-	0.37	-	0.14	+	0.91	-	0.06
7q11.23	rs17685	A/G	0.29	+	0.31	-	0.36	-	0.75	-	0.87
11p13	rs6265	C/T	0.81	+	0.66	-	0.89	+	**1.72E-05**	-	0.52
15q24	rs2470893	T/C	0.31	-	**0.03**	+	0.28	+	0.52	-	0.47
	rs2472297	T/C	0.24	-	0.23	+	0.61	-	0.75	-	0.87
17q11.2	rs9902453	G/A	0.46	-	0.43	+	0.13	-	0.92	-	0.90

EA, effect allele; NEA, non-effect allele; EAF, effect allele frequency

[a]Summary-level results for smoking behavior traits[115] were contributed by TAG investigators and were downloaded from https://pgc.unc.edu

[b]Effect allele corresponds to the allele associated with increased coffee consumption.

E. Neurological[a] and psychiatric disorders[b]

Locus	SNP	EA[c]/NEA	EAF	Parkinson's Disease		Schizophrenia		Major Depression Disorder		Attention Deficit Hyperactivity Disorder		Bipolar Disorder	
				Effect	P	Effect	P	Effect	P	Effect	P	Effect	P
2p24	rs1260326	C/T	0.59	+	0.43	+	**0.02**	+	0.94	-	0.40	-	**2.31E-03**
4q22	rs1481012	A/G	0.89	+	**7.11E-03**	+	0.53	+	0.21	-	0.46	+	0.93
7p21	rs4410790	C/T	0.63	n/a	n/a	-	0.47	-	0.33	-	0.52	-	0.99
	rs6968554	G/A	0.61	+	0.71	-	0.56	-	0.34	-	0.29	+	0.96
7q11.23	rs7800944	C/T	0.28	-	0.08	+	0.41	+	0.29	-	0.39	+	0.16
7q11.23	rs17685	A/G	0.29	-	0.68	+	0.61	-	**6.98E-03**	-	0.50	-	**0.05**
11p13	rs6265	C/T	0.81	-	0.96	+	0.13	-	0.14	+	0.84	+	0.64
15q24	rs2470893	T/C	0.31	n/a	n/a	-	0.12	-	0.66	-	0.85	+	0.91
	rs2472297	T/C	0.24	+	0.51	-	0.18	-	0.94	-	0.56	+	0.71
q11.2	rs9902453	G/A	0.46	+	0.86	+	**0.04**	+	0.67	-	0.45	+	**7.30E-03**

EA, effect allele; NEA, non-effect allele; EAF, effect allele frequency

[a]Summary-level result for Parkinson' Disease[64] was contributed by IPDGC.

[b]Summary-level results for Schizophrenia[111], Major Depression Disorder[112], Attention Deficit Hyperactivity Disorder[113] and Bipolar Disorder[114] were contributed by PGC investigators and were downloaded from https://pgc.unc.edu .

[c]Effect allele corresponds to the allele associated with increased coffee consumption.

Table S33. Association between confirmed loci and coffee consumption using linear and ordinal regression in the Women's Genome Health Study[a]

CHR	SNP	EA/NEA	Linear Regression cups/d among consumers n=17,332		Ordinal Regression[b] categories of cups/d among consumers n=17,332		Ordinal Regression[b] categories of cups/d + non-consumers n=23,250	
			β (SE)	P	β (SE)	P	β (SE)	P
2	rs1260326	T/C	-0.028 (0.02)	0.10	-0.032 (0.02)	0.12	-0.032 (0.02)	0.06
4	rs1481012	G/A	-0.023 (0.03)	0.38	-0.026 (0.03)	0.42	0.007 (0.03)	0.80
7	rs4410790	T/C	-0.183 (0.02)	2.41E-27	-0.218 (0.02)	7.98E-27	-0.150 (0.02)	1.68E-18
7	rs6968554	A/G	-0.179 (0.02)	3.22E-26	-0.213 (0.02	7.53E-26	-0.151 (0.02)	9.93E-19
7	rs7800944	C/T	0.056 (0.02)	2.89E-03	0.060 (0.02)	7.18E-03	0.049 (0.02)	9.42E-03
7	rs17685	A/G	0.067 (0.02)	2.69E-05	0.080 (0.02)	2.74E-04	0.050 (0.02)	7.46E-03
11	rs6265	T/C	-0.055 (0.02)	0.01	-0.053 (0.03)	0.03	-0.051 (0.02)	0.016
15	rs2470893	T/C	0.132 (0.02)	9.73E-14	0.151 (0.02)	1.78E-12	0.099 (0.02)	4.86E-08
15	rs2472297	T/C	0.163 (0.02)	3.66E-14	0.187 (0.03)	6.63E-13	0.116 (0.02)	1.29E-07
17	rs9902453	G/A	0.033 (0.02)	0.05	0.032 (0.02)	0.10	0.019 (0.02)	0.25

EA, effect allele; NEA, non effect allele

[a]Results from linear regression models of coffee consumption used in the current GWAS (cups/day, columns 4 and 5) and from two ordinal regression models (categories of consumption, columns 6 to 9). Models are adjusted for age, smoking, randomization, and EVs (Supplementary Table S7).

[b]Results from the clm() function in the R package 'ordinal' using a logit link function.

References

1. The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. The ARIC investigators. *Am J Epidemiol* 1989; **129**(4)**:** 687-702.

2. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* 2006; **38**(8)**:** 904-909.

3. Busselton Health Study. http://www.busseltonhealthstudy.com/.

4. Moffatt MF, Gut IG, Demenais F, Strachan DP, Bouzigon E, Heath S *et al.* A large-scale, consortium-based genomewide association study of asthma. *N Engl J Med* 2010; **363**(13)**:** 1211-1221.

5. Firmann M, Mayor V, Vidal PM, Bochud M, Pecoud A, Hayoz D *et al.* The CoLaus study: a population-based study to investigate the epidemiology and genetic determinants of cardiovascular risk factors and metabolic syndrome. *BMC cardiovascular disorders* 2008; **8:** 6.

6. Olsen J, Melbye M, Olsen SF, Sorensen TI, Aaby P, Andersen AM *et al.* The Danish National Birth Cohort--its background, structure and aim. *Scandinavian journal of public health* 2001; **29**(4)**:** 300-307.

7. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007; **81**(3)**:** 559-575.

8. Estonian Genome Project of University of Tartu. www.biobank.ee.

9. McGinnis RE, Deloukas P, McLaren WM, Inouye M. Visualizing chromosome mosaicism and detecting ethnic outliers by the method of "rare" heterozygotes and homozygotes (RHH). *Hum Mol Genet* 2010; **19**(13)**:** 2539-2553.

10. Aulchenko YS, Heutink P, Mackay I, Bertoli-Avella AM, Pullen J, Vaessen N *et al.* Linkage disequilibrium in young genetically isolated Dutch population. *Eur J Hum Genet* 2004; **12**(7)**:** 527-534.

11. Higgins M, Province M, Heiss G, Eckfeldt J, Ellison RC, Folsom AR *et al.* NHLBI Family Heart Study: objectives and design. *Am J Epidemiol* 1996; **143**(12)**:** 1219-1228.

12. Ferrucci L, Bandinelli S, Benvenuti E, Di Iorio A, Macchi C, Harris TB *et al.* Subsystems contributing to the decline in ability to walk: bridging the gap between epidemiology and geriatric practice in the InCHIANTI study. *Journal of the American Geriatrics Society* 2000; **48**(12)**:** 1618-1625.

13. Bartali B, Turrini A, Salvini S, Lauretani F, Russo CR, Corsi AM *et al.* Dietary intake estimated using different methods in two Italian older populations. *Archives of gerontology and geriatrics* 2004; **38**(1)**:** 51-60.

14. Bild DE, Bluemke DA, Burke GL, Detrano R, Diez Roux AV, Folsom AR *et al.* Multi-ethnic study of atherosclerosis: objectives and design. *Am J Epidemiol* 2002; **156**(9)**:** 871-881.

15. Rantakallio P. Groups at risk in low birth weight infants and perinatal mortality. *Acta paediatrica Scandinavica* 1969; **193:** Suppl 193:191+.

16. R Development Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing: Vienna, Austria, 2010.

17. Colditz GA, Hankinson SE. The Nurses' Health Study: lifestyle and health among women. *Nat Rev Cancer* 2005; **5**(5)**:** 388-396.

18. Willett WC. *Nutritional Epidemiology*. Oxford University Press: New York, 1998.

19. Willemsen G, Vink JM, Abdellaoui A, den Braber A, van Beek JH, Draisma HH *et al.* The adult Netherlands twin register: twenty-five years of survey and biological data collection. *Twin Res Hum Genet* 2013; **16**(1)**:** 271-281.

20. Vink JM, Staphorsius AS, Boomsma DI. A genetic analysis of coffee consumption in a sample of Dutch twins. *Twin Res Hum Genet* 2009; **12**(2)**:** 127-131.

21. Willemsen G, de Geus EJ, Bartels M, van Beijsterveldt CE, Brooks AI, Estourgie-van Burk GF *et al.* The Netherlands Twin Register biobank: a resource for genetic epidemiological studies. *Twin Res Hum Genet* 2010; **13**(3)**:** 231-245.

22. Abdellaoui A, Hottenga JJ, Knijff PD, Nivard MG, Xiao X, Scheet P *et al.* Population structure, migration, and diversifying selection in the Netherlands. *Eur J Hum Genet* 2013.

23. Pritchard JK, Stephens M, Rosenberg NA, Donnelly P. Association mapping in structured populations. *Am J Hum Genet* 2000; **67**(1)**:** 170-181.

24. Hofman A, van Duijn CM, Franco OH, Ikram MA, Janssen HL, Klaver CC *et al.* The Rotterdam Study: 2012 objectives and design update. *Eur J Epidemiol* 2011; **26**(8)**:** 657-686.

25. Veeramah KR, Tonjes A, Kovacs P, Gross A, Wegmann D, Geary P *et al.* Genetic variation in the Sorbs of eastern Germany in the context of broader European genetic diversity. *Eur J Hum Genet* 2011; **19**(9)**:** 995-1001.

26. Theodoraki EV, Nikopensius T, Suhorutsenko J, Peppes V, Fili P, Kolovou G *et al.* Fibrinogen beta variants confer protection against coronary artery disease in a Greek case-control study. *BMC Med Genet* 2010; **11:** 28.

27. Ridker PM, Chasman DI, Zee RY, Parker A, Rose L, Cook NR *et al.* Rationale, design, and methodology of the Women's Genome Health Study: a genome-wide association study of more than 25,000 initially healthy american women. *Clin Chem* 2008; **54**(2)**:** 249-255.

28. Sawcer S, Hellenthal G, Pirinen M, Spencer CC, Patsopoulos NA, Moutsianas L *et al.* Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. *Nature* 2011; **476**(7359)**:** 214-219.

29. Barrett JC, Clayton DG, Concannon P, Akolkar B, Cooper JD, Erlich HA *et al.* Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. *Nat Genet* 2009; **41**(6)**:** 703-707.

30. Stolk RP, Rosmalen JG, Postma DS, de Boer RA, Navis G, Slaets JP *et al.* Universal risk factors for multifactorial diseases: LifeLines: a three-generation population-based study. *Eur J Epidemiol* 2008; **23**(1)**:** 67-74.

31. Taylor HA, Jr., Wilson JG, Jones DW, Sarpong DF, Srinivasan A, Garrison RJ *et al.* Toward resolution of cardiovascular health disparities in African Americans: design and methods of the Jackson Heart Study. *Ethn Dis* 2005; **15**(4 Suppl 6)**:** S6-4-17.

32. Design of the Women's Health Initiative clinical trial and observational study. The Women's Health Initiative Study Group. *Controlled clinical trials* 1998; **19**(1)**:** 61-109.

33. Segre AV, Groop L, Mootha VK, Daly MJ, Altshuler D. Common inherited variation in mitochondrial genes is not enriched for associations with type 2 diabetes or related glycemic traits. *PLoS Genet* 2010; **6**(8)**:** e1001058.

34. Liu JZ, McRae AF, Nyholt DR, Medland SE, Wray NR, Brown KM *et al.* A versatile gene-based test for genome-wide association studies. *Am J Hum Genet* 2010; **87**(1)**:** 139-145.

35. Yang J, Ferreira T, Morris AP, Medland SE, Madden PA, Heath AC *et al.* Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nat Genet* 2012; **44**(4)**:** 369-375, S361-363.

36. Raychaudhuri S, Plenge RM, Rossin EJ, Ng AC, Purcell SM, Sklar P *et al.* Identifying relationships among genomic disease regions: predicting genes at pathogenic SNP associations and rare deletions. *PLoS Genet* 2009; **5**(6)**:** e1000534.

37. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *Bmj* 2003; **327**(7414)**:** 557-560.

38. Franceschini N, van Rooij FJ, Prins BP, Feitosa MF, Karakas M, Eckfeldt JH *et al.* Discovery and fine mapping of serum protein loci through transethnic meta-analysis. *Am J Hum Genet* 2012; **91**(4)**:** 744-753.

39. ENCODE Project Consortium. A user's guide to the encyclopedia of DNA elements (ENCODE). *PLoS Biol* 2011; **9**(4)**:** e1001046.

40. Ward LD, Kellis M. HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic Acids Res* 2012; **40**(Database issue)**:** D930-934.

41. Boyle AP, Hong EL, Hariharan M, Cheng Y, Schaub MA, Kasowski M *et al.* Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res* 2012; **22**(9)**:** 1790-1797.

42. Matys V, Kel-Margoulis OV, Fricke E, Liebich I, Land S, Barre-Dirrie A *et al.* TRANSFAC and its module TRANSCompel: transcriptional gene regulation in eukaryotes. *Nucleic Acids Res* 2006; **34**(Database issue)**:** D108-110.

43. Portales-Casamar E, Thongjuea S, Kwon AT, Arenillas D, Zhao X, Valen E *et al.* JASPAR 2010: the greatly expanded open-access database of transcription factor binding profiles. *Nucleic Acids Res* 2010; **38**(Database issue)**:** D105-110.

44. Ernst J, Kheradpour P, Mikkelsen TS, Shoresh N, Ward LD, Epstein CB *et al.* Mapping and analysis of chromatin state dynamics in nine human cell types. *Nature* 2011; **473**(7345)**:** 43-49.

45. Berger MF, Badis G, Gehrke AR, Talukder S, Philippakis AA, Pena-Castillo L *et al.* Variation in homeodomain DNA binding revealed by high-resolution analysis of sequence preferences. *Cell* 2008; **133**(7)**:** 1266-1276.

46. Badis G, Berger MF, Philippakis AA, Talukder S, Gehrke AR, Jaeger SA *et al.* Diversity and complexity in DNA recognition by transcription factors. *Science* 2009; **324**(5935)**:** 1720-1723.

47. Berger MF, Philippakis AA, Qureshi AM, He FS, Estep PW, 3rd, Bulyk ML. Compact, universal DNA microarrays to comprehensively determine transcription-factor binding site specificities. *Nat Biotechnol* 2006; **24**(11)**:** 1429-1435.

48. Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM, Handsaker RE *et al.* An integrated map of genetic variation from 1,092 human genomes. *Nature* 2012; **491**(7422)**:** 56-65.

49. Davydov EV, Goode DL, Sirota M, Cooper GM, Sidow A, Batzoglou S. Identifying a high fraction of the human genome to be under selective constraint using GERP++. *PLoS computational biology* 2010; **6**(12)**:** e1001025.

50. Garber M, Guttman M, Clamp M, Zody MC, Friedman N, Xie X. Identifying novel constrained elements by exploiting biased substitution patterns. *Bioinformatics* 2009; **25**(12)**:** i54-62.

51. Lindblad-Toh K, Garber M, Zuk O, Lin MF, Parker BJ, Washietl S *et al.* A high-resolution map of human evolutionary constraint using 29 mammals. *Nature* 2011; **478**(7370)**:** 476-482.

52. Montgomery SB, Sammeth M, Gutierrez-Arcelus M, Lach RP, Ingle C, Nisbett J *et al.* Transcriptome genetics using second generation sequencing in a Caucasian population. *Nature* 2010; **464**(7289)**:** 773-777.

53. Nica AC, Parts L, Glass D, Nisbet J, Barrett A, Sekowska M *et al.* The architecture of gene regulatory variation across multiple human tissues: the MuTHER study. *PLoS Genet* 2011; **7**(2)**:** e1002003.

54. Grundberg E, Small KS, Hedman AK, Nica AC, Buil A, Keildson S *et al.* Mapping cis- and trans-regulatory effects across multiple tissues in twins. *Nat Genet* 2012; **44**(10)**:** 1084-1089.

55. Stranger BE, Montgomery SB, Dimas AS, Parts L, Stegle O, Ingle CE *et al.* Patterns of cis regulatory variation in diverse human populations. *PLoS Genet* 2012; **8**(4)**:** e1002639.

56. Veyrieras JB, Kudaravalli S, Kim SY, Dermitzakis ET, Gilad Y, Stephens M *et al.* High-resolution mapping of expression-QTLs yields insight into human gene regulation. *PLoS Genet* 2008; **4**(10)**:** e1000214.

57. Gaffney DJ, Veyrieras JB, Degner JF, Pique-Regi R, Pai AA, Crawford GE *et al.* Dissecting the regulatory architecture of gene expression QTLs. *Genome Biol* 2012; **13**(1)**:** R7.

58. Dimas AS, Deutsch S, Stranger BE, Montgomery SB, Borel C, Attar-Cohen H *et al.* Common regulatory variation impacts gene expression in a cell type-dependent manner. *Science* 2009; **325**(5945)**:** 1246-1250.

59. Schadt EE, Molony C, Chudin E, Hao K, Yang X, Lum PY *et al.* Mapping the genetic architecture of gene expression in human liver. *PLoS Biol* 2008; **6**(5)**:** e107.

60. Innocenti F, Cooper GM, Stanaway IB, Gamazon ER, Smith JD, Mirkov S *et al.* Identification, replication, and functional fine-mapping of expression quantitative trait loci in primary human liver tissue. *PLoS Genet* 2011; **7**(5)**:** e1002078.

61. Myers AJ, Gibbs JR, Webster JA, Rohrer K, Zhao A, Marlowe L *et al.* A survey of genetic human cortical gene expression. *Nat Genet* 2007; **39**(12)**:** 1494-1499.

62. Yang TP, Beazley C, Montgomery SB, Dimas AS, Gutierrez-Arcelus M, Stranger BE *et al.* Genevar: a database and Java application for the analysis and visualization of SNP-gene associations in eQTL studies. *Bioinformatics* 2010; **26**(19)**:** 2474-2476.

63. Westra HJ, Peters MJ, Esko T, Yaghootkar H, Schurmann C, Kettunen J *et al.* Systematic identification of trans eQTLs as putative drivers of known disease associations. *Nat Genet* 2013; **45**(10)**:** 1238-1243.

64. Nalls MA, Plagnol V, Hernandez DG, Sharma M, Sheerin UM, Saad M *et al.* Imputation of sequence variants for identification of genetic risks for Parkinson's disease: a meta-analysis of genome-wide association studies. *Lancet* 2011; **377**(9766)**:** 641-649.

65. Gibbs JR, van der Brug MP, Hernandez DG, Traynor BJ, Nalls MA, Lai SL *et al.* Abundant quantitative trait loci exist for DNA methylation and gene expression in human brain. *PLoS Genet* 2010; **6**(5)**:** e1000952.

66. Rossin EJ, Lage K, Raychaudhuri S, Xavier RJ, Tatar D, Benita Y *et al.* Proteins encoded in genomic regions associated with immune-mediated disease physically interact and suggest underlying biology. *PLoS Genet* 2012; **7**(1)**:** e1001273.

67. Thorvaldsdottir H, Robinson JT, Mesirov JP. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Briefings in bioinformatics* 2013; **14**(2)**:** 178-192.

68. Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G *et al.* Integrative genomics viewer. *Nat Biotechnol* 2011; **29**(1)**:** 24-26.

69. Hawrylycz MJ, Lein ES, Guillozet-Bongaarts AL, Shen EH, Ng L, Miller JA *et al.* An anatomically comprehensive atlas of the adult human brain transcriptome. *Nature* 2012; **489**(7416)**:** 391-399.

70. Hevezi P, Moyer BD, Lu M, Gao N, White E, Echeverri F *et al.* Genome-wide analysis of gene expression in primate taste buds reveals links to diverse processes. *PLoS One* 2009; **4**(7)**:** e6395.

71. Jensen LJ, Kuhn M, Stark M, Chaffron S, Creevey C, Muller J *et al.* STRING 8--a global view on proteins and their functional interactions in 630 organisms. *Nucleic Acids Res* 2009; **37**(Database issue)**:** D412-416.

72. Kuhn M, Szklarczyk D, Franceschini A, von Mering C, Jensen LJ, Bork P. STITCH 3: zooming in on protein-chemical interactions. *Nucleic Acids Res* 2012; **40**(Database issue)**:** D876-880.

73. Fredholm BB, Battig K, Holmen J, Nehlig A, Zvartau EE. Actions of caffeine in the brain with special reference to factors that contribute to its widespread use. *Pharmacol Rev* 1999; **51**(1)**:** 83-133.

74. Thorn CF, Aklillu E, Klein TE, Altman RB. PharmGKB summary: very important pharmacogene information for CYP1A2. *Pharmacogenet Genomics* 2012; **22**(1)**:** 73-77.

75. Xie X, Ramkumar V, Toth LA. Adenosine and dopamine receptor interactions in striatum and caffeine-induced behavioral activation. *Comparative medicine* 2007; **57**(6)**:** 538-545.

76. Ward S, Sotsios Y, Dowden J, Bruce I, Finan P. Therapeutic potential of phosphoinositide 3-kinase inhibitors. *Chemistry & biology* 2003; **10**(3)**:** 207-213.

77. Fisone G, Borgkvist A, Usiello A. Caffeine as a psychomotor stimulant: mechanism of action. *Cell Mol Life Sci* 2004; **61**(7-8)**:** 857-872.

78. Saito R, Smoot ME, Ono K, Ruscheinski J, Wang PL, Lotia S *et al.* A travel guide to Cytoscape plugins. *Nat Methods* 2012; **9**(11)**:** 1069-1076.

79. Lage K, Karlberg EO, Storling ZM, Olason PI, Pedersen AG, Rigina O *et al.* A human phenome-interactome network of protein complexes implicated in genetic disorders. *Nat Biotechnol* 2007; **25**(3)**:** 309-316.

80. Lage K, Hansen NT, Karlberg EO, Eklund AC, Roque FS, Donahoe PK *et al.* A large-scale analysis of tissue-specific pathology and gene expression of human disease genes and complexes. *Proc Natl Acad Sci U S A* 2008; **105**(52)**:** 20870-20875.

81. Smyth GK. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Statistical applications in genetics and molecular biology* 2004; **3:** Article3.

82. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* 2005; **102**(43)**:** 15545-15550.

83. Mootha VK, Lindgren CM, Eriksson KF, Subramanian A, Sihag S, Lehar J *et al.* PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat Genet* 2003; **34**(3)**:** 267-273.

84. Hindorf L, MacArthur J, Morales J, Junkins H, Hall P, Klemm A *et al.* Catalogue of Published Genome-Wide Association Studies. accessed January 1, 2013.

85. Shin SY, Fauman EB, Petersen AK, Krumsiek J, Santos R, Huang J *et al.* An atlas of genetic influences on human blood metabolites. *Nat Genet* 2014; **46**(6)**:** 543-550.

86. Reed DR, Zhu G, Breslin PA, Duke FF, Henders AK, Campbell MJ *et al.* The perception of quinine taste intensity is associated with common genetic variants in a bitter receptor cluster on chromosome 12. *Hum Mol Genet* 2010; **19**(21)**:** 4278-4285.

87. Ledda M, Kutalik Z, Souza Destito MC, Souza MM, Cirillo CA, Zamboni A *et al.* GWAS of Human Bitter Taste Perception Identifies New Loci and Reveals Additional Complexity of Bitter Taste Genetics. *Hum Mol Genet* 2013; **23**(1)**:** 259-267.

88. Byrne EM, Johnson J, McRae AF, Nyholt DR, Medland SE, Gehrman PR *et al.* A genome-wide association study of caffeine-related sleep disturbance: confirmation of a role for a common variant in the adenosine receptor. *Sleep* 2012; **35**(7)**:** 967-975.

89. Saleheen D, Zaidi M, Rasheed A, Ahmad U, Hakeem A, Murtaza M *et al.* The Pakistan Risk of Myocardial Infarction Study: a resource for the study of genetic, lifestyle and other determinants of myocardial infarction in South Asia. *Eur J Epidemiol* 2009; **24**(6)**:** 329-338.

90. Cornelis MC, Monda KL, Yu K, Paynter N, Azzato EM, Bennett SN *et al.* Genome-wide meta-analysis identifies regions on 7p21 (AHR) and 15q24 (CYP1A2) as determinants of habitual caffeine consumption. *PLoS Genet* 2011; **7**(4)**:** e1002033.

91. Sulem P, Gudbjartsson DF, Geller F, Prokopenko I, Feenstra B, Aben KK *et al.* Sequence variants at CYP1A1-CYP1A2 and AHR associate with coffee consumption. *Hum Mol Genet* 2011; **20**(10)**:** 2071-2077.

92. Amin N, Byrne E, Johnson J, Chenevix-Trench G, Walter S, Nolte IM *et al.* Genome-wide association analysis of coffee drinking suggests association with CYP1A1/CYP1A2 and NRCAM. *Molecular psychiatry* 2011; **17**(11)**:** 1116-1129.

93. Luciano M, Kirk KM, Heath AC, Martin NG. The genetics of tea and coffee drinking and preference for source of caffeine in a large community sample of Australian twins. *Addiction* 2005; **100**(10)**:** 1510-1517.

94. Willett WC, Sampson L, Stampfer MJ, Rosner B, Bain C, Witschi J *et al.* Reproducibility and validity of a semiquantitative food frequency questionnaire. *Am J Epidemiol* 1985; **122**(1)**:** 51-65.

95. Stevens J, Metcalf P, Dennis B, Tell G, Shimakawa T, Folsom A. Reliability of a food frequency questionnaire by ethnicity, gender, age and education. *Nutrition Research* 1996; **16:** 735-745.

96. Feskanich D, Rimm EB, Giovannucci EL, Colditz GA, Stampfer MJ, Litin LB *et al.* Reproducibility and validity of food intake measurements from a semiquantitative food frequency questionnaire. *J Am Diet Assoc* 1993; **93**(7)**:** 790-796.

97. Mannisto S, Virtanen M, Mikkonen T, Pietinen P. Reproducibility and validity of a food frequency questionnaire in a case-control study on breast cancer. *J Clin Epidemiol* 1996; **49**(4)**:** 401-409.

98. Paalanen L, Mannisto S, Virtanen MJ, Knekt P, Rasanen L, Montonen J *et al.* Validity of a food frequency questionnaire varied by age and body mass index. *J Clin Epidemiol* 2006; **59**(9)**:** 994-1001.

99. Han B, Eskin E. Random-effects model aimed at discovering associations in meta-analysis of genome-wide association studies. *Am J Hum Genet* 2011; **88**(5)**:** 586-598.

100. Wang X, Chua HX, Chen P, Ong RT, Sim X, Zhang W *et al.* Comparing methods for performing trans-ethnic meta-analysis of genome-wide association studies. *Hum Mol Genet* 2013; **22**(11)**:** 2303-2311.

101. Stephens M, Balding DJ. Bayesian statistical methods for genetic association studies. *Nat Rev Genet* 2009; **10**(10)**:** 681-690.

102. Sellke T, Bayarri M, Berger J. Calibration of p values for testing precise null hypotheses. *Am Stat* 2001; **55**(1)**:** 62-71.

103. Dupuis J, Langenberg C, Prokopenko I, Saxena R, Soranzo N, Jackson AU *et al.* New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. *Nat Genet* 2010; **42**(2)**:** 105-116.

104. Saxena R, Hivert MF, Langenberg C, Tanaka T, Pankow JS, Vollenweider P *et al.* Genetic variation in GIPR influences the glucose and insulin responses to an oral glucose challenge. *Nat Genet* 2010; **42**(2)**:** 142-148.

105. Soranzo N, Sanna S, Wheeler E, Gieger C, Radke D, Dupuis J *et al.* Common variants at 10 genomic loci influence hemoglobin A(1)(C) levels via glycemic and nonglycemic pathways. *Diabetes* 2010; **59**(12)**:** 3229-3239.

106. Strawbridge RJ, Dupuis J, Prokopenko I, Barker A, Ahlqvist E, Rybin D *et al.* Genome-wide association identifies nine common variants associated with fasting proinsulin levels and provides new insights into the pathophysiology of type 2 diabetes. *Diabetes* 2011; **60**(10)**:** 2624-2634.

107. Morris AP, Voight BF, Teslovich TM, Ferreira T, Segre AV, Steinthorsdottir V *et al.* Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. *Nat Genet* 2012; **44**(9)**:** 981-990.

108. Heid IM, Jackson AU, Randall JC, Winkler TW, Qi L, Steinthorsdottir V *et al.* Meta-analysis identifies 13 new loci associated with waist-hip ratio and reveals sexual dimorphism in the genetic basis of fat distribution. *Nat Genet* 2010; **42**(11)**:** 949-960.

109. Speliotes EK, Willer CJ, Berndt SI, Monda KL, Thorleifsson G, Jackson AU *et al.* Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nat Genet* 2010; **42**(11)**:** 937-948.

110. Horikoshi M, Yaghootkar H, Mook-Kanamori DO, Sovio U, Taal HR, Hennig BJ *et al.* New loci associated with birth weight identify genetic links between intrauterine growth and adult height and metabolism. *Nat Genet* 2012; **45**(1)**:** 76-82.

111. Schizophrenia Psychiatric Genome-Wide Association Study (GWAS) Consortium. Genome-wide association study identifies five new schizophrenia loci. *Nat Genet* 2011; **43**(10)**:** 969-976.

112. Ripke S, Wray NR, Lewis CM, Hamilton SP, Weissman MM, Breen G *et al.* A mega-analysis of genome-wide association studies for major depressive disorder. *Molecular psychiatry* 2012; **18**(4)**:** 497-511.

113. Neale BM, Medland SE, Ripke S, Asherson P, Franke B, Lesch KP *et al.* Meta-analysis of genome-wide association studies of attention-deficit/hyperactivity disorder. *Journal of the American Academy of Child and Adolescent Psychiatry* 2010; **49**(9)**:** 884-897.

114. Psychiatric GWAS Consortium Bipolar Disorder Working Group (172 contributors). Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near ODZ4. *Nat Genet* 2012; **43**(10)**:** 977-983.

115. Tobacco and Genetics Consortium. Genome-wide meta-analyses identify multiple loci associated with smoking behavior. *Nat Genet* 2010; **42**(5)**:** 441-447.

116. Teslovich TM, Musunuru K, Smith AV, Edmondson AC, Stylianou IM, Koseki M *et al.* Biological, clinical and population relevance of 95 loci for blood lipids. *Nature* 2010; **466**(7307)**:** 707-713.

117. Fullwood MJ, Han Y, Wei CL, Ruan X, Ruan Y. Chromatin interaction analysis using paired-end tag sequencing. *Current protocols in molecular biology / edited by Frederick M Ausubel [et al* 2010; **Chapter 21:** Unit 21 15 21-25.

118. Fried LP, Borhani NO, Enright P, Furberg CD, Gardin JM, Kronmal RA *et al.* The Cardiovascular Health Study: design and rationale. *Ann Epidemiol* 1991; **1**(3)**:** 263-276.

119. Day N, Oakes S, Luben R, Khaw KT, Bingham S, Welch A *et al.* EPIC-Norfolk: study design and characteristics of the cohort. European Prospective Investigation of Cancer. *British journal of cancer* 1999; **80 Suppl 1:** 95-103.

120. Rolfe Ede L, Loos RJ, Druet C, Stolk RP, Ekelund U, Griffin SJ *et al.* Association between birth weight and visceral fat in adults. *Am J Clin Nutr* 2010; **92**(2)**:** 347-352.

121. Bingham SA, Welch AA, McTaggart A, Mulligan AA, Runswick SA, Luben R *et al.* Nutritional methods in the European Prospective Investigation of Cancer in Norfolk. *Public health nutrition* 2001; **4**(3)**:** 847-858.

122. Splansky GL, Corey D, Yang Q, Atwood LD, Cupples LA, Benjamin EJ *et al.* The Third Generation Cohort of the National Heart, Lung, and Blood Institute's Framingham Heart Study: design, recruitment, and initial examination. *Am J Epidemiol* 2007; **165**(11)**:** 1328-1335.

123. Kannel WB, Feinleib M, McNamara PM, Garrison RJ, Castelli WP. An investigation of coronary heart disease in families. The Framingham offspring study. *Am J Epidemiol* 1979; **110**(3)**:** 281-290.

124. Dawber TR, Kannel WB, Lyell LP. An approach to longitudinal studies in a community: the Framingham Study. *Ann N Y Acad Sci* 1963; **107:** 539-556.

125. Lorentzon M, Swanson C, Andersson N, Mellstrom D, Ohlsson C. Free testosterone is a positive, whereas free estradiol is a negative, predictor of cortical bone size in young Swedish men: the GOOD study. *J Bone Miner Res* 2005; **20**(8)**:** 1334-1341.

126. Barker DJ, Osmond C, Forsen TJ, Kajantie E, Eriksson JG. Trajectories of growth among children who have coronary events as adults. *N Engl J Med* 2005; **353**(17)**:** 1802-1809.

127. Houston DK, Nicklas BJ, Ding J, Harris TB, Tylavsky FA, Newman AB *et al.* Dietary protein intake is associated with lean mass change in older, community-dwelling adults: the Health, Aging, and Body Composition (Health ABC) Study. *Am J Clin Nutr* 2008; **87**(1)**:** 150-155.

128. Wichmann HE, Gieger C, Illig T. KORA-gen--resource for population genetics, controls and a broad spectrum of disease phenotypes. *Gesundheitswesen (Bundesverband der Arzte des Offentlichen Gesundheitsdienstes (Germany))* 2005; **67 Suppl 1:** S26-30.

129. Mayer-Davis EJ, Vitolins MZ, Carmichael SL, Hemphill S, Tsaroucha G, Rushing J *et al.* Validity and reproducibility of a food frequency interview in a Multi-Cultural Epidemiology Study. *Ann Epidemiol* 1999; **9**(5)**:** 314-324.

130. Nettleton JA, Rock CL, Wang Y, Jenny NS, Jacobs DR. Associations between dietary macronutrient intake and plasma lipids demonstrate criterion performance of the Multi-Ethnic Study of Atherosclerosis (MESA) food-frequency questionnaire. *Br J Nutr* 2009; **102**(8)**:** 1220-1227.

131. Boomsma DI, de Geus EJ, Vink JM, Stubbe JH, Distel MA, Hottenga JJ *et al.* Netherlands Twin Register: from twins to twin families. *Twin Res Hum Genet* 2006; **9**(6)**:** 849-857.

132. Prorok PC, Andriole GL, Bresalier RS, Buys SS, Chia D, Crawford ED *et al.* Design of the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial. *Controlled clinical trials* 2000; **21**(6 Suppl)**:** 273S-309S.

133. Subar AF, Thompson FE, Kipnis V, Midthune D, Hurwitz P, McNutt S *et al.* Comparative validation of the Block, Willett, and National Cancer Institute food frequency questionnaires : the Eating at America's Table Study. *Am J Epidemiol* 2001; **154**(12)**:** 1089-1099.

134. Volzke H, Alte D, Schmidt CO, Radke D, Lorbeer R, Friedrich N *et al.* Cohort profile: the study of health in Pomerania. *Int J Epidemiol* 2011; **40**(2)**:** 294-307.

135. Lichtenstein P, Sullivan PF, Cnattingius S, Gatz M, Johansson S, Carlstrom E *et al.* The Swedish Twin Registry in the third millennium: an update. *Twin Res Hum Genet* 2006; **9**(6)**:** 875-882.

136. Mitchell P, Smith W, Attebo K, Wang JJ. Prevalence of age-related maculopathy in Australia. The Blue Mountains Eye Study. *Ophthalmology* 1995; **102**(10)**:** 1450-1460.

137. Attebo K, Mitchell P, Smith W. Visual acuity and the causes of visual loss in Australia. The Blue Mountains Eye Study. *Ophthalmology* 1996; **103**(3)**:** 357-364.

138. Power C, Elliott J. Cohort profile: 1958 British birth cohort (National Child Development Study). *Int J Epidemiol* 2006; **35**(1)**:** 34-41.

139. Fuller E, Power C, Shepherd P, Strachan D. Technical report on the National Child Development Study biomedical survey 2002-2004. London: Centre for Longitudinal Studies; 2006.

140. Norberg M, Wall S, Boman K, Weinehall L. The Vasterbotten Intervention Programme: background, design and implications. *Global health action* 2010; **3**.

141. Hallmans G, Agren A, Johansson G, Johansson A, Stegmayr B, Jansson JH *et al.* Cardiovascular disease and diabetes in the Northern Sweden Health and Disease Study Cohort - evaluation of risk factors and their interactions. *Scand J Public Health Suppl* 2003; **61:** 18-24.

142. Renstrom F, Shungin D, Johansson I, Florez JC, Hallmans G, Hu FB *et al.* Genetic predisposition to long-term nondiabetic deteriorations in glucose homeostasis: Ten-year follow-up of the GLACIER study. *Diabetes* 2011; **60**(1)**:** 345-354.

143. Mellstrom D, Johnell O, Ljunggren O, Eriksson AL, Lorentzon M, Mallmin H *et al.* Free testosterone is an independent predictor of BMD and prevalent fractures in elderly men: MrOS Sweden. *J Bone Miner Res* 2006; **21**(4)**:** 529-535.

144. Ingelsson E, Hulthe J, Lind L. Inflammatory markers in relation to insulin resistance and the metabolic syndrome. *Eur J Clin Invest* 2008; **38**(7)**:** 502-509.

145. Becker W. Befolkningens kostvanor och näringsintag i Sverige 1989. Metod och resultatanalys. (Dietary habits and nutrient intake in Sweden 1989. Methods and analysis): Nutritional Unit, Swedish National Food Administration; 1994.

146. Medland SE, Nyholt DR, Painter JN, McEvoy BP, McRae AF, Zhu G *et al.* Common variants in the trichohyalin gene are associated with straight hair in Europeans. *Am J Hum Genet* 2009; **85**(5)**:** 750-755.

147. Zethelius B, Byberg L, Hales CN, Lithell H, Berne C. Proinsulin and acute insulin response independently predict Type 2 diabetes mellitus in men--report from 27 years of follow-up study. *Diabetologia* 2003; **46**(1)**:** 20-26.

148. Raitakari OT, Juonala M, Ronnemaa T, Keltikangas-Jarvinen L, Rasanen L, Pietikainen M *et al.* Cohort profile: the cardiovascular risk in Young Finns Study. *Int J Epidemiol* 2008; **37**(6)**:** 1220-1226.

149. Carithers TC, Talegawkar SA, Rowser ML, Henry OR, Dubbert PM, Bogle ML *et al.* Validity and calibration of food frequency questionnaires used with African-American adults in the Jackson Heart Study. *J Am Diet Assoc* 2009; **109**(7)**:** 1184-1193.

150. Anderson GL, Manson J, Wallace R, Lund B, Hall D, Davis S *et al.* Implementation of the Women's Health Initiative study design. *Ann Epidemiol* 2003; **13**(9 Suppl)**:** S5-17.

151. Laurie CC, Doheny KF, Mirel DB, Pugh EW, Bierut LJ, Bhangale T *et al.* Quality control and quality assurance in genotypic data for genome-wide association studies. *Genet Epidemiol* 2010; **34**(6)**:** 591-602.

152. Nelis M, Esko T, Magi R, Zimprich F, Zimprich A, Toncheva D *et al.* Genetic structure of Europeans: a view from the North-East. *PLoS One* 2009; **4**(5)**:** e5472.

153. Loos RJ, Lindgren CM, Li S, Wheeler E, Zhao JH, Prokopenko I *et al.* Common variants near MC4R are associated with fat mass, weight and risk of obesity. *Nat Genet* 2008; **40**(6)**:** 768-775.

154. Pardo LM, MacKay I, Oostra B, van Duijn CM, Aulchenko YS. The effect of genetic drift in a young genetically isolated population. *Annals of human genetics* 2005; **69**(Pt 3)**:** 288-295.

155. Jensen MK, Pers TH, Dworzynski P, Girman CJ, Brunak S, Rimm EB. Protein interaction-based genome-wide analysis of incident coronary heart disease. *Circ Cardiovasc Genet* 2011; **4**(5)**:** 549-556.

156. Wiggs JL, Hee Kang J, Yaspan BL, Mirel DB, Laurie C, Crenshaw A *et al.* Common variants near CAV1 and CAV2 are associated with primary open-angle glaucoma in Caucasians from the USA. *Hum Mol Genet* 2011; **20**(23)**:** 4701-4713.

157. Tanaka T, Scheet P, Giusti B, Bandinelli S, Piras MG, Usala G *et al.* Genome-wide association study of vitamin B6, vitamin B12, folate, and homocysteine blood concentrations. *Am J Hum Genet* 2009; **84**(4)**:** 477-482.

158. Manichaikul A, Naj AC, Herrington D, Post W, Rich SS, Rodriguez A. Association of SCARB1 variants with subclinical atherosclerosis and incident cardiovascular disease: the multi-ethnic study of atherosclerosis. *Arterioscler Thromb Vasc Biol* 2012; **32**(8)**:** 1991-1999.

159. Sabatti C, Service SK, Hartikainen AL, Pouta A, Ripatti S, Brodsky J *et al.* Genome-wide association analysis of metabolic traits in a birth cohort from a founder population. *Nat Genet* 2009; **41**(1)**:** 35-46.

160. Hunter DJ, Kraft P, Jacobs KB, Cox DG, Yeager M, Hankinson SE *et al.* A genome-wide association study identifies alleles in FGFR2 associated with risk of sporadic postmenopausal breast cancer. *Nat Genet* 2007; **39**(7)**:** 870-874.

161. Hofman A, Breteler MM, van Duijn CM, Janssen HL, Krestin GP, Kuipers EJ *et al.* The Rotterdam Study: 2010 objectives and design update. *Eur J Epidemiol* 2009; **24**(9)**:** 553-572.

162. Holliday EG, Smith AV, Cornes BK, Buitendijk GH, Jensen RA, Sim X *et al.* Insights into the genetic architecture of early stage age-related macular degeneration: a genome-wide association study meta-analysis. *PLoS One* 2013; **8**(1)**:** e53830.

163. Peters U, Jiao S, Schumacher FR, Hutter CM, Aragaki AK, Baron JA *et al.* Identification of Genetic Susceptibility Loci for Colorectal Tumors in a Genome-Wide Meta-analysis. *Gastroenterology* 2013; **144**(4)**:** 799-807 e724.

164. Hernesniemi JA, Seppala I, Lyytikainen LP, Mononen N, Oksala N, Hutri-Kahonen N *et al.* Genetic profiling using genome-wide significant coronary artery disease risk variants does not improve the prediction of subclinical atherosclerosis: the Cardiovascular Risk in Young Finns Study, the Bogalusa Heart Study and the Health 2000 Survey--a meta-analysis of three independent studies. *PLoS One* 2012; **7**(1)**:** e28931.

165. Ashton KJ, Holmgren K, Peart J, Lankford AR, Paul Matherne G, Grimmond S *et al.* Effects of A1 adenosine receptor overexpression on normoxic and post-ischemic gene expression. *Cardiovascular research* 2003; **57**(3)**:** 715-726.

166. Joo JD, Kim M, Horst P, Kim J, D'Agati VD, Emala CW, Sr. *et al.* Acute and delayed renal protection against renal ischemia and reperfusion injury with A1 adenosine receptors. *Am J Physiol Renal Physiol* 2007; **293**(6)**:** F1847-1857.

167. Lee HT, Kim M, Jan M, Penn RB, Emala CW. Renal tubule necrosis and apoptosis modulation by A1 adenosine receptor expression. *Kidney international* 2007; **71**(12)**:** 1249-1261.

168. Mediero A, Kara FM, Wilder T, Cronstein BN. Adenosine A(2A) receptor ligation inhibits osteoclast formation. *The American journal of pathology* 2012; **180**(2)**:** 775-786.

169. Chiang MC, Chen HM, Lai HL, Chen HW, Chou SY, Chen CM *et al.* The A2A adenosine receptor rescues the urea cycle deficiency of Huntington's disease by enhancing the activity of the ubiquitin-proteasome system. *Hum Mol Genet* 2009; **18**(16)**:** 2929-2942.

170. Andersson M, Usiello A, Borgkvist A, Pozzi L, Dominguez C, Fienberg AA *et al.* Cannabinoid action depends on phosphorylation of dopamine- and cAMP-regulated phosphoprotein of 32 kDa at the protein kinase A site in striatal projection neurons. *J Neurosci* 2005; **25**(37)**:** 8432-8438.

171. Botsakis K, Pavlou O, Poulou PD, Matsokis N, Angelatou F. Blockade of adenosine A2A receptors downregulates DARPP-32 but increases ERK1/2 activity in striatum of dopamine deficient "weaver" mouse. *Neurochem Int* 2010; **56**(2)**:** 245-249.

172. Carrier F, Chang CY, Duh JL, Nebert DW, Puga A. Interaction of the regulatory domains of the murine Cyp1a1 gene with two DNA-binding proteins in addition to the Ah receptor and the Ah receptor nuclear translocator (ARNT). *Biochemical pharmacology* 1994; **48**(9)**:** 1767-1778.

173. Fallone F, Villard PH, Seree E, Rimet O, Nguyen QB, Bourgarel-Rey V *et al.* Retinoids repress Ah receptor CYP1A1 induction pathway through the SMRT corepressor. *Biochem Biophys Res Commun* 2004; **322**(2)**:** 551-556.

174. Gharavi N, El-Kadi AO. tert-Butylhydroquinone is a novel aryl hydrocarbon receptor ligand. *Drug Metab Dispos* 2005; **33**(3)**:** 365-372.

175. Hankinson O. Role of coactivators in transcriptional activation by the aryl hydrocarbon receptor. *Archives of biochemistry and biophysics* 2005; **433**(2)**:** 379-386.

176. Hestermann EV, Brown M. Agonist and chemopreventative ligands induce differential transcriptional cofactor recruitment by aryl hydrocarbon receptor. *Mol Cell Biol* 2003; **23**(21)**:** 7920-7925.

177. Ko HP, Okino ST, Ma Q, Whitlock JP, Jr. Transactivation domains facilitate promoter occupancy for the dioxin-inducible CYP1A1 gene in vivo. *Mol Cell Biol* 1997; **17**(7)**:** 3497-3507.

178. Kobayashi A, Sogawa K, Fujii-Kuriyama Y. Cooperative interaction between AhR.Arnt and Sp1 for the drug-inducible expression of CYP1A1 gene. *J Biol Chem* 1996; **271**(21)**:** 12310-12316.

179. Kumar MB, Ramadoss P, Reen RK, Vanden Heuvel JP, Perdew GH. The Q-rich subdomain of the human Ah receptor transactivation domain is required for dioxin-mediated transcriptional activity. *J Biol Chem* 2001; **276**(45)**:** 42302-42310.

180. Shimada T, Inoue K, Suzuki Y, Kawai T, Azuma E, Nakajima T *et al.* Arylhydrocarbon receptor-dependent induction of liver and lung cytochromes P450 1A1, 1A2, and 1B1 by polycyclic aromatic hydrocarbons and polychlorinated biphenyls in genetically engineered C57BL/6J mice. *Carcinogenesis* 2002; **23**(7)**:** 1199-1207.

181. Tian Y, Ke S, Chen M, Sheng T. Interactions between the aryl hydrocarbon receptor and P-TEFb. Sequential recruitment of transcription factors and differential phosphorylation of C-terminal domain of RNA polymerase II at cyp1a1 promoter. *J Biol Chem* 2003; **278**(45)**:** 44041-44048.

182. Wang S, Ge K, Roeder RG, Hankinson O. Role of mediator in transcriptional activation by the aryl hydrocarbon receptor. *J Biol Chem* 2004; **279**(14)**:** 13593-13600.

183. Watson AJ, Hankinson O. Dioxin- and Ah receptor-dependent protein binding to xenobiotic responsive elements and G-rich DNA studied by in vivo footprinting. *J Biol Chem* 1992; **267**(10)**:** 6874-6878.

184. Caron E, Rioux N, Nicolas O, Lebel-Talbot H, Hamelin BA. Quantification of the expression and inducibility of 12 rat cytochrome P450 isoforms by quantitative RT-PCR. *Journal of biochemical and molecular toxicology* 2005; **19**(6)**:** 368-378.

185. Han EH, Kim JY, Jeong HG. Effect of biochanin A on the aryl hydrocarbon receptor and cytochrome P450 1A1 in MCF-7 human breast carcinoma cells. *Archives of pharmacal research* 2006; **29**(7)**:** 570-576.

186. Kawajiri K, Fujii-Kuriyama Y. Cytochrome P450 gene regulation and physiological functions mediated by the aryl hydrocarbon receptor. *Archives of biochemistry and biophysics* 2007; **464**(2)**:** 207-212.

187. Murray IA, Reen RK, Leathery N, Ramadoss P, Bonati L, Gonzalez FJ *et al.* Evidence that ligand binding is a key determinant of Ah receptor-mediated transcriptional activity. *Archives of biochemistry and biophysics* 2005; **442**(1)**:** 59-71.

188. Okino ST, Quattrochi LC, Pookot D, Iwahashi M, Dahiya R. A dioxin-responsive enhancer 3' of the human CYP1A2 gene. *Mol Pharmacol* 2007; **72**(6)**:** 1457-1465.

189. Dere E, Lo R, Celius T, Matthews J, Zacharewski TR. Integration of genome-wide computation DRE search, AhR ChIP-chip and gene expression analyses of TCDD-elicited responses in the mouse liver. *BMC Genomics* 2011; **12:** 365.

190. Kinehara M, Fukuda I, Yoshida K, Ashida H. High-throughput evaluation of aryl hydrocarbon receptor-binding sites selected via chromatin immunoprecipitation-based screening in Hepa-1c1c7 cells stimulated with 2,3,7,8-tetrachlorodibenzo-p-dioxin. *Genes & genetic systems* 2008; **83**(6)**:** 455-468.

191. Lo R, Celius T, Forgacs AL, Dere E, MacPherson L, Harper P *et al.* Identification of aryl hydrocarbon receptor binding targets in mouse hepatic tissue treated with 2,3,7,8-tetrachlorodibenzo-p-dioxin. *Toxicol Appl Pharmacol* 2011; **257**(1)**:** 38-47.

192. Tanos R, Patel RD, Murray IA, Smith PB, Patterson AD, Perdew GH. Aryl hydrocarbon receptor regulates the cholesterol biosynthetic pathway in a dioxin response element-independent manner. *Hepatology* 2012; **55**(6)**:** 1994-2004.

193. Vikstrom Bergander L, Cai W, Klocke B, Seifert M, Pongratz I. Tryptamine serves as a proligand of the AhR transcriptional pathway whose activation is dependent of monoamine oxidases. *Mol Endocrinol* 2012; **26**(9)**:** 1542-1551.

194. Nebert DW, Dalton TP, Okey AB, Gonzalez FJ. Role of aryl hydrocarbon receptor-mediated induction of the CYP1 enzymes in environmental toxicity and cancer. *J Biol Chem* 2004; **279**(23)**:** 23847-23850.

195. Ebert B, Seidel A, Lampen A. Identification of BCRP as transporter of benzo[a]pyrene conjugates metabolically formed in Caco-2 cells and its induction by Ah-receptor agonists. *Carcinogenesis* 2005; **26**(10)**:** 1754-1763.

196. Tan KP, Wang B, Yang M, Boutros PC, Macaulay J, Xu H *et al.* Aryl hydrocarbon receptor is a transcriptional activator of the human breast cancer resistance protein (BCRP/ABCG2). *Mol Pharmacol* 2010; **78**(2)**:** 175-185.

197. To KK, Robey R, Zhan Z, Bangiolo L, Bates SE. Upregulation of ABCG2 by romidepsin via the aryl hydrocarbon receptor pathway. *Mol Cancer Res* 2011; **9**(4)**:** 516-527.

198. To KK, Yu L, Liu S, Fu J, Cho CH. Constitutive AhR activation leads to concomitant ABCG2-mediated multidrug resistance in cisplatin-resistant esophageal carcinoma cells. *Mol Carcinog* 2012; **51**(6)**:** 449-464.

199. de Boussac H, Orban TI, Varady G, Tihanyi B, Bacquet C, Brozik A *et al.* Stimulus-induced expression of the ABCG2 multidrug transporter in HepG2 hepatocarcinoma model cells involves the ERK1/2 cascade and alternative promoters. *Biochem Biophys Res Commun* 2012; **426**(2)**:** 172-176.

200. Dey S, Subramanian VS, Chatterjee NS, Rubin SA, Said HM. Characterization of the 5' regulatory region of the human sodium-dependent multivitamin transporter, hSMVT. *Biochim Biophys Acta* 2002; **1574**(2)**:** 187-192.

201. Reymann S, Borlak J. Transcriptome profiling of human hepatocytes treated with Aroclor 1254 reveals transcription factor regulatory networks and clusters of regulated genes. *BMC Genomics* 2006; **7:** 217.

202. Chuang CY, Chang H, Lin P, Sun SJ, Chen PH, Lin YY *et al.* Up-regulation of osteopontin expression by aryl hydrocarbon receptor via both ligand-dependent and ligand-independent pathways in lung cancer. *Gene* 2012; **492**(1)**:** 262-269.

203. Boutros PC, Moffat ID, Franc MA, Tijet N, Tuomisto J, Pohjanvirta R *et al.* Dioxin-responsive AHRE-II gene battery: identification by phylogenetic footprinting. *Biochem Biophys Res Commun* 2004; **321**(3)**:** 707-715.

204. Arpiainen S, Raffalli-Mathieu F, Lang MA, Pelkonen O, Hakkola J. Regulation of the Cyp2a5 gene involves an aryl hydrocarbon receptor-dependent pathway. *Mol Pharmacol* 2005; **67**(4)**:** 1325-1333.

205. Du L, Neis MM, Ladd PA, Keeney DS. Differentiation-specific factors modulate epidermal CYP1-4 gene expression in human skin in response to retinoic acid and classic aryl hydrocarbon receptor ligands. *J Pharmacol Exp Ther* 2006; **319**(3)**:** 1162-1171.

206. Gonzalez FJ, Fernandez-Salguero P. The aryl hydrocarbon receptor: studies using the AHR-null mice. *Drug Metab Dispos* 1998; **26**(12)**:** 1194-1198.

207. Jin B, Ryu DY. Regulation of CYP1A2 by histone deacetylase inhibitors in mouse hepatocytes. *Journal of biochemical and molecular toxicology* 2004; **18**(3)**:** 131-132.

208. Krusekopf S, Roots I. St. John's wort and its constituent hyperforin concordantly regulate expression of genes encoding enzymes involved in basic cellular pathways. *Pharmacogenet Genomics* 2005; **15**(11)**:** 817-829.

209. Li W, Harper PA, Tang BK, Okey AB. Regulation of cytochrome P450 enzymes by aryl hydrocarbon receptor in human cells: CYP1A2 expression in the LS180 colon carcinoma cell line after treatment with 2,3,7,8-tetrachlorodibenzo-p-dioxin or 3-methylcholanthrene. *Biochemical pharmacology* 1998; **56**(5)**:** 599-612.

210. Nakajima M, Iwanari M, Yokoi T. Effects of histone deacetylation and DNA methylation on the constitutive and TCDD-inducible expressions of the human CYP1 family in MCF-7 and HeLa cells. *Toxicol Lett* 2003; **144**(2)**:** 247-256.

211. Nakata K, Tanaka Y, Nakano T, Adachi T, Tanaka H, Kaminuma T *et al.* Nuclear receptor-mediated transcriptional regulation in Phase I, II, and III xenobiotic metabolizing systems. *Drug metabolism and pharmacokinetics* 2006; **21**(6)**:** 437-457.

212. Ovando BJ, Ellison CA, Vezina CM, Olson JR. Toxicogenomic analysis of exposure to TCDD, PCB126 and PCB153: identification of genomic biomarkers of exposure to AhR ligands. *BMC Genomics* 2010; **11:** 583.

213. Sienkiewicz P, Ciolino HP, Leslie BJ, Hergenrother PJ, Singletary K, Yeh GC. A novel synthetic analogue of a constituent of Isodon excisus inhibits transcription of CYP1A1, -1A2 and -1B1 by preventing activation of the aryl hydrocarbon receptor. *Carcinogenesis* 2007; **28**(5)**:** 1052-1057.

214. Tan Z, Chang X, Puga A, Xia Y. Activation of mitogen-activated protein kinases (MAPKs) by aromatic hydrocarbons: role in the regulation of aryl hydrocarbon receptor (AHR) function. *Biochemical pharmacology* 2002; **64**(5-6)**:** 771-780.

215. Williams SN, Shih H, Guenette DK, Brackney W, Denison MS, Pickwell GV *et al.* Comparative studies on the effects of green tea extracts and individual tea catechins on human CYP1A gene expression. *Chemico-biological interactions* 2000; **128**(3)**:** 211-229.

216. Shimizu Y, Nakatsuru Y, Ichinose M, Takahashi Y, Kume H, Mimura J *et al.* Benzo[a]pyrene carcinogenicity is lost in mice lacking the aryl hydrocarbon receptor. *Proc Natl Acad Sci U S A* 2000; **97**(2)**:** 779-782.

217. Carlson DB, Perdew GH. A dynamic role for the Ah receptor in cell signaling? Insights from a diverse group of Ah receptor interacting proteins. *Journal of biochemical and molecular toxicology* 2002; **16**(6)**:** 317-325.

218. Moriguchi T, Motohashi H, Hosoya T, Nakajima O, Takahashi S, Ohsako S *et al.* Distinct response to dioxin in an arylhydrocarbon receptor (AHR)-humanized mouse. *Proc Natl Acad Sci U S A* 2003; **100**(10)**:** 5652-5657.

219. Tijet N, Boutros PC, Moffat ID, Okey AB, Tuomisto J, Pohjanvirta R. Aryl hydrocarbon receptor regulates distinct dioxin-dependent and dioxin-independent gene batteries. *Mol Pharmacol* 2006; **69**(1)**:** 140-153.

220. Rahman S, Sowa ME, Ottinger M, Smith JA, Shi Y, Harper JW *et al.* The Brd4 extraterminal domain confers transcription activation independent of pTEFb by recruiting multiple proteins, including NSD3. *Mol Cell Biol* 2011; **31**(13)**:** 2641-2652.

221. Pineda JR, Canals JM, Bosch M, Adell A, Mengod G, Artigas F *et al.* Brain-derived neurotrophic factor modulates dopaminergic deficits in a transgenic mouse model of Huntington's disease. *Journal of neurochemistry* 2005; **93**(5)**:** 1057-1068.

222. Rumajogee P, Madeira A, Verge D, Hamon M, Miquel MC. Up-regulation of the neuronal serotoninergic phenotype in vitro: BDNF and cAMP share Trk B-dependent mechanisms. *Journal of neurochemistry* 2002; **83**(6)**:** 1525-1528.

223. Guiard BP, David DJ, Deltheil T, Chenu F, Le Maitre E, Renoir T *et al.* Brain-derived neurotrophic factor-deficient mice exhibit a hippocampal hyperserotonergic phenotype. *The international journal of neuropsychopharmacology / official scientific journal of the Collegium Internationale Neuropsychopharmacologicum (CINP)* 2008; **11**(1)**:** 79-92.

224. Outhred T, Das P, Dobson-Stone C, Griffiths K, Felmingham KL, Bryant RA *et al.* The functional epistasis of 5-HTTLPR and BDNF Val66Met on emotion processing: a preliminary study. *Brain and behavior* 2012; **2**(6)**:** 778-788.

225. Bogush A, Pedrini S, Pelta-Heller J, Chan T, Yang Q, Mao Z *et al.* AKT and CDK5/p35 mediate brain-derived neurotrophic factor induction of DARPP-32 in medium size spiny neurons in vitro. *J Biol Chem* 2007; **282**(10)**:** 7352-7359.

226. Stroppolo A, Guinea B, Tian C, Sommer J, Ehrlich ME. Role of phosphatidylinositide 3-kinase in brain-derived neurotrophic factor-induced DARPP-32 expression in medium size spiny neurons in vitro. *Journal of neurochemistry* 2001; **79**(5)**:** 1027-1032.

227. Gharami K, Xie Y, An JJ, Tonegawa S, Xu B. Brain-derived neurotrophic factor over-expression in the forebrain ameliorates Huntington's disease phenotypes in mice. *Journal of neurochemistry* 2008; **105**(2)**:** 369-379.

228. Ivkovic S, Polonskaia O, Farinas I, Ehrlich ME. Brain-derived neurotrophic factor regulates maturation of the DARPP-32 phenotype in striatal medium spiny neurons: studies in vivo and in vitro. *Neuroscience* 1997; **79**(2)**:** 509-516.

229. Petersen AA, Larsen KE, Behr GG, Romero N, Przedborski S, Brundin P *et al.* Brain-derived neurotrophic factor inhibits apoptosis and dopamine-induced free radical production in striatal neurons but does not prevent cell death. *Brain research bulletin* 2001; **56**(3-4)**:** 331-335.

230. Krueger-Naug AM, Emsley JG, Myers TL, Currie RW, Clarke DB. Administration of brain-derived neurotrophic factor suppresses the expression of heat shock protein 27 in rat retinal ganglion cells following axotomy. *Neuroscience* 2003; **116**(1)**:** 49-58.

231. Soler-Lopez M, Zanzoni A, Lluis R, Stelzl U, Aloy P. Interactome mapping suggests new mechanistic details underlying Alzheimer's disease. *Genome Res* 2011; **21**(3)**:** 364-376.

232. O'Connell JC, McCallum JF, McPhee I, Wakefield J, Houslay ES, Wishart W *et al.* The SH3 domain of Src tyrosyl protein kinase interacts with the N-terminal splice region of the PDE4A cAMP-specific phosphodiesterase RPDE-6 (RNPDE4A5). *Biochem J* 1996; **318 (Pt 1):** 255-261.

233. Xue L, Wang WH, Iliuk A, Hu L, Galan JA, Yu S *et al.* Sensitive kinase assay linked with phosphoproteomics for identifying direct kinase substrates. *Proc Natl Acad Sci U S A* 2012; **109**(15)**:** 5615-5620.

234. Yang G, Li Q, Ren S, Lu X, Fang L, Zhou W *et al.* Proteomic, functional and motif-based analysis of C-terminal Src kinase-interacting proteins. *Proteomics* 2009; **9**(21)**:** 4944-4961.

235. Hemmings HC, Jr., Nairn AC, Elliott JI, Greengard P. Synthetic peptide analogs of DARPP-32 (Mr 32,000 dopamine- and cAMP-regulated phosphoprotein), an inhibitor of protein phosphatase-1. Phosphorylation, dephosphorylation, and inhibitory activity. *J Biol Chem* 1990; **265**(33)**:** 20369-20376.

236. Takeuchi Y, Fukunaga K, Miyamoto E. Activation of nuclear Ca(2+)/calmodulin-dependent protein kinase II and brain-derived neurotrophic factor gene expression by stimulation of dopamine D2 receptor in transfected NG108-15 cells. *Journal of neurochemistry* 2002; **82**(2)**:** 316-328.

237. Lindskog M, Svenningsson P, Fredholm BB, Greengard P, Fisone G. Activation of dopamine D2 receptors decreases DARPP-32 phosphorylation in striatonigral and striatopallidal projection neurons via different mechanisms. *Neuroscience* 1999; **88**(4)**:** 1005-1008.

238. Ewing RM, Chu P, Elisma F, Li H, Taylor P, Climie S *et al.* Large-scale mapping of human protein-protein interactions by mass spectrometry. *Mol Syst Biol* 2007; **3:** 89.

239. Ma L, Robinson LN, Towle HC. ChREBP*Mlx is the principal mediator of glucose-induced gene expression in the liver. *J Biol Chem* 2006; **281**(39)**:** 28721-28730.

240. Taguchi M, Imaoka S, Yoshii K, Kobayashi K, Hosokawa M, Shimada N *et al.* Kinetics of testosterone 6beta-hydroxylation in the reconstituted system with similar ratios of purified CYP3A4, NADPH-cytochrome p450 oxidoreductase and cytochrome B5 to human liver microsomes. *Research communications in molecular pathology and pharmacology* 2001; **109**(1-2)**:** 53-63.

241. Grinkova YV, Denisov IG, Sligar SG. Functional reconstitution of monomeric CYP3A4 with multiple cytochrome P450 reductase molecules in Nanodiscs. *Biochem Biophys Res Commun* 2010; **398**(2)**:** 194-198.

242. Guengerich FP, Johnson WW. Kinetics of ferric cytochrome P450 reduction by NADPH-cytochrome P450 reductase: rapid reduction in the absence of substrate and variations among cytochrome P450 systems. *Biochemistry* 1997; **36**(48)**:** 14741-14750.

243. Dohr O, Paine MJ, Friedberg T, Roberts GC, Wolf CR. Engineering of a functional human NADH-dependent cytochrome P450 system. *Proc Natl Acad Sci U S A* 2001; **98**(1)**:** 81-86.

244. Locuson CW, Wienkers LC, Jones JP, Tracy TS. CYP2C9 protein interactions with cytochrome b(5): effects on the coupling of catalysis. *Drug Metab Dispos* 2007; **35**(7)**:** 1174-1181.

245. Ozalp C, Szczesna-Skorupa E, Kemper B. Bimolecular fluorescence complementation analysis of cytochrome p450 2c2, 2e1, and NADPH-cytochrome p450 reductase molecular interactions in living cells. *Drug Metab Dispos* 2005; **33**(9)**:** 1382-1390.

246. Muller HK, Wiborg O, Haase J. Subcellular redistribution of the serotonin transporter by secretory carrier membrane protein 2. *J Biol Chem* 2006; **281**(39)**:** 28901-28909.

247. Merla G, Howald C, Antonarakis SE, Reymond A. The subcellular localization of the ChoRE-binding protein, encoded by the Williams-Beuren syndrome critical region gene 14, is regulated by 14-3-3. *Hum Mol Genet* 2004; **13**(14)**:** 1505-1514.

248. Jin J, Smith FD, Stark C, Wells CD, Fawcett JP, Kulkarni S *et al.* Proteomic, functional, and domain-based analysis of in vivo 14-3-3 binding proteins involved in cytoskeletal regulation and cellular organization. *Curr Biol* 2004; **14**(16)**:** 1436-1450.

249. Bandyopadhyay S, Chiang CY, Srivastava J, Gersten M, White S, Bell R *et al.* A human MAP kinase interactome. *Nat Methods* 2010; **7**(10)**:** 801-805.

250. Rao FV, Andersen OA, Vora KA, Demartino JA, van Aalten DM. Methylxanthine drugs are chitinase inhibitors: investigation of inhibition and binding modes. *Chemistry & biology* 2005; **12**(9)**:** 973-980.

251. Kasvinsky PJ, Fletterick RJ, Madsen NB. Regulation of the dephosphorylation of glycogen phosphorylase a and synthase b by glucose and caffeine in isolated hepatocytes. *Canadian journal of biochemistry* 1981; **59**(6)**:** 387-395.

252. Vetter DE, Li C, Zhao L, Contarino A, Liberman MC, Smith GW *et al.* Urocortin-deficient mice show hearing impairment and increased anxiety-like behavior. *Nat Genet* 2002; **31**(4)**:** 363-369.

253. Wang X, Su H, Copenhagen LD, Vaishnav S, Pieri F, Shope CD *et al.* Urocortin-deficient mice display normal stress-induced anxiety behavior and autonomic control but an impaired acoustic startle response. *Mol Cell Biol* 2002; **22**(18)**:** 6605-6610.

254. Zalutskaya AA, Arai M, Bounoutas GS, Abou-Samra AB. Impaired adaptation to repeated restraint and decreased response to cold in urocortin 1 knockout mice. *Am J Physiol Endocrinol Metab* 2007; **293**(1)**:** E259-263.

255. Cole TB, Wenzel HJ, Kafer KE, Schwartzkroin PA, Palmiter RD. Elimination of zinc from synaptic vesicles in the intact mouse brain by disruption of the ZnT3 gene. *Proc Natl Acad Sci U S A* 1999; **96**(4)**:** 1716-1721.

256. Lee JY, Cole TB, Palmiter RD, Suh SW, Koh JY. Contribution by synaptic zinc to the gender-disparate plaque formation in human Swedish mutant APP transgenic mice. *Proc Natl Acad Sci U S A* 2002; **99**(11)**:** 7705-7710.

257. Friedland-Little JM, Hoffmann AD, Ocbina PJ, Peterson MA, Bosman JD, Chen Y *et al.* A novel murine allele of Intraflagellar Transport Protein 172 causes a syndrome including VACTERL-like features with hydrocephalus. *Hum Mol Genet* 2011; **20**(19)**:** 3725-3737.

258. Gorivodsky M, Mukhopadhyay M, Wilsch-Braeuninger M, Phillips M, Teufel A, Kim C *et al.* Intraflagellar transport protein 172 is essential for primary cilia formation and plays a vital role in patterning the mammalian brain. *Dev Biol* 2009; **325**(1)**:** 24-32.

259. Huangfu D, Liu A, Rakeman AS, Murcia NS, Niswander L, Anderson KV. Hedgehog signalling in the mouse requires intraflagellar transport proteins. *Nature* 2003; **426**(6962)**:** 83-87.

260. Kim I, Li C, Liang D, Chen XZ, Coffy RJ, Ma J *et al.* Polycystin-2 expression is regulated by a PC2-binding domain in the intracellular portion of fibrocystin. *J Biol Chem* 2008; **283**(46)**:** 31559-31566.

261. Pennekamp P, Karcher C, Fischer A, Schweickert A, Skryabin B, Horst J *et al.* The ion channel polycystin-2 is required for left-right axis determination in mice. *Curr Biol* 2002; **12**(11)**:** 938-943.

262. Stroope A, Radtke B, Huang B, Masyuk T, Torres V, Ritman E *et al.* Hepato-renal pathology in pkd2ws25/- mice, an animal model of autosomal dominant polycystic kidney disease. *The American journal of pathology* 2010; **176**(3)**:** 1282-1291.

263. Wu G, Tian X, Nishimura S, Markowitz GS, D'Agati V, Park JH *et al.* Trans-heterozygous Pkd1 and Pkd2 mutations modify expression of polycystic kidney disease. *Hum Mol Genet* 2002; **11**(16)**:** 1845-1854.

264. Berman JS, Serlin D, Li X, Whitley G, Hayes J, Rishikof DC *et al.* Altered bleomycin-induced lung fibrosis in osteopontin-deficient mice. *Am J Physiol Lung Cell Mol Physiol* 2004; **286**(6)**:** L1311-1318.

265. Maetzler W, Berg D, Schalamberidze N, Melms A, Schott K, Mueller JC *et al.* Osteopontin is elevated in Parkinson's disease and its absence leads to reduced neurodegeneration in the MPTP model. *Neurobiology of disease* 2007; **25**(3)**:** 473-482.

266. Schumacher T, Krohn M, Hofrichter J, Lange C, Stenzel J, Steffen J *et al.* ABC transporters B1, C1 and G2 differentially regulate neuroregeneration in mice. *PLoS One* 2012; **7**(4)**:** e35613.

267. Krohn M, Lange C, Hofrichter J, Scheffler K, Stenzel J, Steffen J *et al.* Cerebral amyloid-beta proteostasis is regulated by the membrane transport protein ABCC1 in mice. *J Clin Invest* 2011; **121**(10)**:** 3924-3931.

268. Curran CP, Miller KA, Dalton TP, Vorhees CV, Miller ML, Shertzer HG *et al.* Genetic differences in lethality of newborn mice treated in utero with coplanar versus non-coplanar hexabromobiphenyl. *Toxicol Sci* 2006; **89**(2)**:** 454-464.

269. Schmid FA, Pena RC, Robinson W, Tarnowski GS. Toxicity of intraperitoneal injections of 7, 12-dimethylbenz[a]anthracene in inbred mice. *Cancer Res* 1967; **27**(3)**:** 558-562.

270. Lin TM, Ko K, Moore RW, Simanainen U, Oberley TD, Peterson RE. Effects of aryl hydrocarbon receptor null mutation and in utero and lactational 2,3,7,8-tetrachlorodibenzo-p-dioxin exposure on prostate and seminal vesicle development in C57BL/6 mice. *Toxicol Sci* 2002; **68**(2)**:** 479-487.

271. Nukaya M, Lin BC, Glover E, Moran SM, Kennedy GD, Bradfield CA. The aryl hydrocarbon receptor-interacting protein (AIP) is required for dioxin-induced hepatotoxicity but not for the induction of the Cyp1a1 and Cyp1a2 genes. *J Biol Chem* 2010; **285**(46)**:** 35599-35605.

272. Yu Z, Mahadevan B, Lohr CV, Fischer KA, Louderback MA, Krueger SK *et al.* Indole-3-carbinol in the maternal diet provides chemoprotection for the fetus against transplacental carcinogenesis by the polycyclic aromatic hydrocarbon dibenzo[a,l]pyrene. *Carcinogenesis* 2006; **27**(10)**:** 2116-2123.

273. Bunger MK, Moran SM, Glover E, Thomae TL, Lahvis GP, Lin BC *et al.* Resistance to 2,3,7,8-tetrachlorodibenzo-p-dioxin toxicity and abnormal liver development in mice carrying a mutation in the nuclear localization sequence of the aryl hydrocarbon receptor. *J Biol Chem* 2003; **278**(20)**:** 17767-17774.

274. Thatcher TH, Maggirwar SB, Baglole CJ, Lakatos HF, Gasiewicz TA, Phipps RP *et al.* Aryl hydrocarbon receptor-deficient mice develop heightened inflammatory responses to cigarette smoke and endotoxin associated with rapid loss of the nuclear factor-kappaB component RelB. *The American journal of pathology* 2007; **170**(3)**:** 855-864.

275. Lahvis GP, Lindell SL, Thomas RS, McCuskey RS, Murphy C, Glover E *et al.* Portosystemic shunting and persistent fetal vascular structures in aryl hydrocarbon receptor-deficient mice. *Proc Natl Acad Sci U S A* 2000; **97**(19)**:** 10442-10447.

276. Schmidt JV, Su GH, Reddy JK, Simon MC, Bradfield CA. Characterization of a murine Ahr null allele: involvement of the Ah receptor in hepatic growth and development. *Proc Natl Acad Sci U S A* 1996; **93**(13)**:** 6731-6736.

277. Zhao C, Aviles C, Abel RA, Almli CR, McQuillen P, Pleasure SJ. Hippocampal and visuospatial learning defects in mice with a deletion of frizzled 9, a gene in the Williams syndrome deletion interval. *Development* 2005; **132**(12)**:** 2917-2927.

278. Ayadi A, Birling MC, Bottomley J, Bussell J, Fuchs H, Fray M *et al.* Mouse large-scale phenotyping initiatives: overview of the European Mouse Disease Clinic (EUMODIC)

and of the Wellcome Trust Sanger Institute Mouse Genetics Project. *Mamm Genome* 2012; **23**(9-10)**:** 600-610.

279. Iizuka K, Bruick RK, Liang G, Horton JD, Uyeda K. Deficiency of carbohydrate response element-binding protein (ChREBP) reduces lipogenesis as well as glycolysis. *Proc Natl Acad Sci U S A* 2004; **101**(19)**:** 7281-7286.

280. Gu J, Weng Y, Zhang QY, Cui H, Behr M, Wu L *et al.* Liver-specific deletion of the NADPH-cytochrome P450 reductase gene: impact on plasma cholesterol homeostasis and the function and regulation of microsomal cytochrome P450 and heme oxygenase. *J Biol Chem* 2003; **278**(28)**:** 25895-25901.

281. Henderson CJ, Otto DM, Carrie D, Magnuson MA, McLaren AW, Rosewell I *et al.* Inactivation of the hepatic cytochrome P450 system by conditional deletion of hepatic cytochrome P450 reductase. *J Biol Chem* 2003; **278**(15)**:** 13480-13486.

282. Wei Y, Zhou X, Fang C, Li L, Kluetzman K, Yang W *et al.* Generation of a mouse model with a reversible hypomorphic cytochrome P450 reductase gene: utility for tissue-specific rescue of the reductase expression, and insights from a resultant mouse model with global suppression of P450 reductase expression in extrahepatic tissues. *J Pharmacol Exp Ther* 2010; **334**(1)**:** 69-77.

283. Wu L, Gu J, Cui H, Zhang QY, Behr M, Fang C *et al.* Transgenic mice with a hypomorphic NADPH-cytochrome P450 reductase gene: effects on development, reproduction, and microsomal cytochrome P450. *J Pharmacol Exp Ther* 2005; **312**(1)**:** 35-43.

284. Chen ZY, Jing D, Bath KG, Ieraci A, Khan T, Siao CJ *et al.* Genetic variant BDNF (Val66Met) polymorphism alters anxiety-related behavior. *Science* 2006; **314**(5796)**:** 140-143.

285. Kaneko M, Xie Y, An JJ, Stryker MP, Xu B. Dendritic BDNF synthesis is required for late-phase spine maturation and recovery of cortical responses following sensory deprivation. *J Neurosci* 2012; **32**(14)**:** 4790-4802.

286. Sakata K, Woo NH, Martinowich K, Greene JS, Schloesser RJ, Shen L *et al.* Critical role of promoter IV-driven BDNF transcription in GABAergic transmission and synaptic plasticity in the prefrontal cortex. *Proc Natl Acad Sci U S A* 2009; **106**(14)**:** 5942-5947.

287. Bianchi LM, Conover JC, Fritzsch B, DeChiara T, Lindsay RM, Yancopoulos GD. Degeneration of vestibular neurons in late embryogenesis of both heterozygous and homozygous BDNF null mutant mice. *Development* 1996; **122**(6)**:** 1965-1973.

288. Conover JC, Erickson JT, Katz DM, Bianchi LM, Poueymirou WT, McClain J *et al.* Neuronal deficits, not involving motor neurons, in mice lacking BDNF and/or NT4. *Nature* 1995; **375**(6528)**:** 235-238.

289. Ernfors P, Lee KF, Jaenisch R. Mice lacking brain-derived neurotrophic factor develop with sensory deficits. *Nature* 1994; **368**(6467)**:** 147-150.

290. Agerman K, Baudet C, Fundin B, Willson C, Ernfors P. Attenuation of a caspase-3 dependent cell death in NT4- and p75-deficient embryonic sensory neurons. *Molecular and cellular neurosciences* 2000; **16**(3)**:** 258-268.

291. Agerman K, Hjerling-Leffler J, Blanchard MP, Scarfone E, Canlon B, Nosrat C *et al.* BDNF gene replacement reveals multiple mechanisms for establishing neurotrophin specificity during sensory nervous system development. *Development* 2003; **130**(8)**:** 1479-1491.

292. Liu X, Ernfors P, Wu H, Jaenisch R. Sensory but not motor neuron deficits in mice lacking NT4 and BDNF. *Nature* 1995; **375**(6528)**:** 238-241.

293. Bamji SX, Majdan M, Pozniak CD, Belliveau DJ, Aloyz R, Kohn J *et al.* The p75 neurotrophin receptor mediates neuronal apoptosis and is essential for naturally occurring sympathetic neuron death. *The Journal of cell biology* 1998; **140**(4)**:** 911-923.

294. Baquet ZC, Bickford PC, Jones KR. Brain-derived neurotrophic factor is required for the establishment of the proper number of dopaminergic neurons in the substantia nigra pars compacta. *J Neurosci* 2005; **25**(26)**:** 6251-6259.

295. Baquet ZC, Gorski JA, Jones KR. Early striatal dendrite deficits followed by neuron loss with advanced age in the absence of anterograde cortical brain-derived neurotrophic factor. *J Neurosci* 2004; **24**(17)**:** 4250-4258.

296. Fan G, Egles C, Sun Y, Minichiello L, Renger JJ, Klein R *et al.* Knocking the NT4 gene into the BDNF locus rescues BDNF deficient mice and reveals distinct NT4 and BDNF activities. *Nature neuroscience* 2000; **3**(4)**:** 350-357.

297. Gorski JA, Zeiler SR, Tamowski S, Jones KR. Brain-derived neurotrophic factor is required for the maintenance of cortical dendrites. *J Neurosci* 2003; **23**(17)**:** 6856-6865.

298. Saylor AJ, Meredith GE, Vercillo MS, Zahm DS, McGinty JF. BDNF heterozygous mice demonstrate age-related changes in striatal and nigral gene expression. *Experimental neurology* 2006; **199**(2)**:** 362-372.

299. Singh KK, Park KJ, Hong EJ, Kramer BM, Greenberg ME, Kaplan DR *et al.* Developmental axon pruning mediated by BDNF-p75NTR-dependent axon degeneration. *Nature neuroscience* 2008; **11**(6)**:** 649-658.

300. Matsumoto T, Rauskolb S, Polack M, Klose J, Kolbeck R, Korte M *et al.* Biosynthesis and processing of endogenous BDNF: CNS neurons store and secrete BDNF, not pro-BDNF. *Nature neuroscience* 2008; **11**(2)**:** 131-133.

301. Monteggia LM, Luikart B, Barrot M, Theobold D, Malkovska I, Nef S *et al.* Brain-derived neurotrophic factor conditional knockouts show gender differences in depression-related behaviors. *Biological psychiatry* 2007; **61**(2)**:** 187-197.

302. Lush ME, Ma L, Parada LF. TrkB signaling regulates the developmental maturation of the somatosensory cortex. *Int J Dev Neurosci* 2005; **23**(6)**:** 523-536.

303. Liebl DJ, Klesse LJ, Tessarollo L, Wohlman T, Parada LF. Loss of brain-derived neurotrophic factor-dependent neural crest-derived sensory neurons in neurotrophin-4 mutant mice. *Proc Natl Acad Sci U S A* 2000; **97**(5)**:** 2297-2302.

304. Lyons WE, Mamounas LA, Ricaurte GA, Coppola V, Reid SW, Bora SH *et al.* Brain-derived neurotrophic factor-deficient mice develop aggressiveness and hyperphagia in conjunction with brain serotonergic abnormalities. *Proc Natl Acad Sci U S A* 1999; **96**(26)**:** 15239-15244.

305. Mistretta CM, Goosens KA, Farinas I, Reichardt LF. Alterations in size, number, and morphology of gustatory papillae and taste buds in BDNF null mutant mice demonstrate neural dependence of developing taste organs. *The Journal of comparative neurology* 1999; **409**(1)**:** 13-24.

306. Liebl DJ, Tessarollo L, Palko ME, Parada LF. Absence of sensory neurons before target innervation in brain-derived neurotrophic factor-, neurotrophin 3-, and TrkC-deficient embryonic mice. *J Neurosci* 1997; **17**(23)**:** 9113-9121.

307. Jones KR, Farinas I, Backus C, Reichardt LF. Targeted disruption of the BDNF gene perturbs brain and sensory neuron development but not motor neuron development. *Cell* 1994; **76**(6)**:** 989-999.

308. Botchkarev VA, Botchkareva NV, Lommatzsch M, Peters EM, Lewin GR, Subramaniam A *et al.* BDNF overexpression induces differential increases among subsets of sympathetic innervation in murine back skin. *The European journal of neuroscience* 1998; **10**(10)**:** 3276-3283.

309. Chang Q, Khare G, Dani V, Nelson S, Jaenisch R. The disease progression of Mecp2 mutant mice is affected by the level of BDNF expression. *Neuron* 2006; **49**(3)**:** 341-348.

310. Choi DC, Maguschak KA, Ye K, Jang SW, Myers KM, Ressler KJ. Prelimbic cortical BDNF is required for memory of learned fear but not extinction or innate fear. *Proc Natl Acad Sci U S A* 2010; **107**(6)**:** 2675-2680.

311. Korte M, Carroll P, Wolf E, Brem G, Thoenen H, Bonhoeffer T. Hippocampal long-term potentiation is impaired in mice lacking brain-derived neurotrophic factor. *Proc Natl Acad Sci U S A* 1995; **92**(19)**:** 8856-8860.

312. Monteggia LM, Barrot M, Powell CM, Berton O, Galanis V, Gemelli T *et al.* Essential role of brain-derived neurotrophic factor in adult hippocampal function. *Proc Natl Acad Sci U S A* 2004; **101**(29)**:** 10827-10832.

313. Rauskolb S, Zagrebelsky M, Dreznjak A, Deogracias R, Matsumoto T, Wiese S *et al.* Global deprivation of brain-derived neurotrophic factor in the CNS reveals an area-specific requirement for dendritic growth. *J Neurosci* 2010; **30**(5)**:** 1739-1749.

314. Rios M, Fan G, Fekete C, Kelly J, Bates B, Kuehn R *et al.* Conditional deletion of brain-derived neurotrophic factor in the postnatal brain leads to obesity and hyperactivity. *Mol Endocrinol* 2001; **15**(10)**:** 1748-1757.

315. Zakharenko SS, Patterson SL, Dragatsis I, Zeitlin SO, Siegelbaum SA, Kandel ER *et al.* Presynaptic BDNF required for a presynaptic but not postsynaptic component of LTP at hippocampal CA1-CA3 synapses. *Neuron* 2003; **39**(6)**:** 975-990.

316. Nada S, Yagi T, Takeda H, Tokunaga T, Nakagawa H, Ikawa Y *et al.* Constitutive activation of Src family kinases in mouse embryos that lack Csk. *Cell* 1993; **73**(6)**:** 1125-1135.

317. Sinai L, Mathew R, Roder JC. Impaired social memories in 129P2 inbred mice are rescued by reduced Csk expression. *Genes, brain, and behavior* 2012; **11**(5)**:** 559-567.

318. Nukaya M, Moran S, Bradfield CA. The role of the dioxin-responsive element cluster between the Cyp1a1 and Cyp1a2 loci in aryl hydrocarbon receptor biology. *Proc Natl Acad Sci U S A* 2009; **106**(12)**:** 4923-4928.

319. Uno S, Dalton TP, Sinclair PR, Gorman N, Wang B, Smith AG *et al.* Cyp1a1(-/-) male mice: protection against high-dose TCDD-induced lethality and wasting syndrome, and resistance to intrahepatocyte lipid accumulation and uroporphyria. *Toxicol Appl Pharmacol* 2004; **196**(3)**:** 410-421.

320. Dong H, Dalton TP, Miller ML, Chen Y, Uno S, Shi Z *et al.* Knock-in mouse lines expressing either mitochondrial or microsomal CYP1A1: differing responses to dietary benzo[a]pyrene as proof of principle. *Mol Pharmacol* 2009; **75**(3)**:** 555-567.

321. Dragin N, Uno S, Wang B, Dalton TP, Nebert DW. Generation of 'humanized' hCYP1A1_1A2_Cyp1a1/1a2(-/-) mouse line. *Biochem Biophys Res Commun* 2007; **359**(3)**:** 635-642.

322. Greaves P, Clothier B, Davies R, Higginson FM, Edwards RE, Dalton TP *et al.* Uroporphyria and hepatic carcinogenesis induced by polychlorinated biphenyls-iron interaction: absence in the Cyp1a2(-/-) knockout mouse. *Biochem Biophys Res Commun* 2005; **331**(1)**:** 147-152.

323. Liang HC, Li H, McKinnon RA, Duffy JJ, Potter SS, Puga A *et al.* Cyp1a2(-/-) null mutant mice develop normally but show deficient drug metabolism. *Proc Natl Acad Sci U S A* 1996; **93**(4)**:** 1671-1676.

324. Shertzer HG, Clay CD, Genter MB, Schneider SN, Nebert DW, Dalton TP. Cyp1a2 protects against reactive oxygen production in mouse liver microsomes. *Free Radic Biol Med* 2004; **36**(5)**:** 605-617.

325. Sinclair PR, Gorman N, Walton HS, Bement WJ, Dalton TP, Sinclair JF *et al.* CYP1A2 is essential in murine uroporphyria caused by hexachlorobenzene and iron. *Toxicol Appl Pharmacol* 2000; **162**(1)**:** 60-67.

326. Zaher H, Buters JT, Ward JM, Bruno MK, Lucas AM, Stern ST *et al.* Protection against acetaminophen toxicity in CYP1A2 and CYP2E1 double-null mice. *Toxicol Appl Pharmacol* 1998; **152**(1)**:** 193-199.

327. Hu MC, Hsu NC, El Hadj NB, Pai CI, Chu HP, Wang CK *et al.* Steroid deficiency syndromes in mice with targeted disruption of Cyp11a1. *Mol Endocrinol* 2002; **16**(8)**:** 1943-1950.

328. Pasterkamp RJ, Peschon JJ, Spriggs MK, Kolodkin AL. Semaphorin 7A promotes axon outgrowth through integrins and MAPKs. *Nature* 2003; **424**(6947)**:** 398-405.

329. Xue M, Stradomska A, Chen H, Brose N, Zhang W, Rosenmund C *et al.* Complexins facilitate neurotransmitter release at excitatory and inhibitory synapses in mammalian central nervous system. *Proc Natl Acad Sci U S A* 2008; **105**(22)**:** 7875-7880.

330. Bengel D, Murphy DL, Andrews AM, Wichems CH, Feltner D, Heils A *et al.* Altered brain serotonin homeostasis and locomotor insensitivity to 3, 4-methylenedioxymethamphetamine ("Ecstasy") in serotonin transporter-deficient mice. *Mol Pharmacol* 1998; **53**(4)**:** 649-655.

331. Eddahibi S, Hanoun N, Lanfumey L, Lesch KP, Raffestin B, Hamon M *et al.* Attenuated hypoxic pulmonary hypertension in mice lacking the 5-hydroxytryptamine transporter gene. *J Clin Invest* 2000; **105**(11)**:** 1555-1562.

332. Lira A, Zhou M, Castanon N, Ansorge MS, Gordon JA, Francis JH *et al.* Altered depression-related behaviors and functional changes in the dorsal raphe nucleus of serotonin transporter-deficient mice. *Biological psychiatry* 2003; **54**(10)**:** 960-971.

333. Moy SS, Nadler JJ, Young NB, Nonneman RJ, Grossman AW, Murphy DL *et al.* Social approach in genetically engineered mouse lines relevant to autism. *Genes, brain, and behavior* 2009; **8**(2)**:** 129-142.

334. Page DT, Kuti OJ, Prestia C, Sur M. Haploinsufficiency for Pten and Serotonin transporter cooperatively influences brain size and social behavior. *Proc Natl Acad Sci U S A* 2009; **106**(6)**:** 1989-1994.

335. Persico AM, Baldi A, Dell'Acqua ML, Moessner R, Murphy DL, Lesch KP *et al.* Reduced programmed cell death in brains of serotonin transporter knockout mice. *Neuroreport* 2003; **14**(3)**:** 341-344.

336. Persico AM, Mengual E, Moessner R, Hall FS, Revay RS, Sora I *et al.* Barrel pattern formation requires serotonin uptake by thalamocortical afferents, and not vesicular monoamine release. *J Neurosci* 2001; **21**(17)**:** 6862-6873.

337. Sora I, Hall FS, Andrews AM, Itokawa M, Li XF, Wei HB *et al.* Molecular mechanisms of cocaine reward: combined dopamine and serotonin transporter knockouts eliminate cocaine place preference. *Proc Natl Acad Sci U S A* 2001; **98**(9)**:** 5300-5305.

338. Sora I, Wichems C, Takahashi N, Li XF, Zeng Z, Revay R *et al.* Cocaine reward models: conditioned place preference can be established in dopamine- and in serotonin-transporter knockout mice. *Proc Natl Acad Sci U S A* 1998; **95**(13)**:** 7699-7704.

339. Thompson BJ, Jessen T, Henry LK, Field JR, Gamble KL, Gresch PJ *et al.* Transgenic elimination of high-affinity antidepressant and cocaine sensitivity in the presynaptic serotonin transporter. *Proc Natl Acad Sci U S A* 2011; **108**(9)**:** 3785-3790.

340. Weisstaub NV, Zhou M, Lira A, Lambe E, Gonzalez-Maeso J, Hornung JP *et al.* Cortical 5-HT2A receptor signaling modulates anxiety-like behaviors in mice. *Science* 2006; **313**(5786)**:** 536-540.

341. Zhao S, Edwards J, Carroll J, Wiedholz L, Millstein RA, Jaing C *et al.* Insertion mutation at the C-terminus of the serotonin transporter disrupts brain serotonin function and emotion-related behaviors in mice. *Neuroscience* 2006; **140**(1)**:** 321-334.

342. Pang J, Hoefen R, Pryhuber GS, Wang J, Yin G, White RJ *et al.* G-protein-coupled receptor kinase interacting protein-1 is required for pulmonary vascular development. *Circulation* 2009; **119**(11)**:** 1524-1532.

343. Schmalzigaug R, Rodriguiz RM, Bonner PE, Davidson CE, Wetsel WC, Premont RT. Impaired fear response in mice lacking GIT1. *Neurosci Lett* 2009; **458**(2)**:** 79-83.

344. Won H, Mah W, Kim E, Kim JW, Hahm EK, Kim MH *et al.* GIT1 is associated with ADHD in humans and ADHD-like behaviors in mice. *Nat Med* 2011; **17**(5)**:** 566-572.

345. Kimura H, Takizawa N, Allemand E, Hori T, Iborra FJ, Nozaki N *et al.* A novel histone exchange factor, protein phosphatase 2Cgamma, mediates the exchange and dephosphorylation of H2A-H2B. *The Journal of cell biology* 2006; **175**(3)**:** 389-400.

346. Vizdalova M, Janovska E, Zhestyanikov VD. The role of the HCR system in the repair of lethal lesions of Bacillus subtilis phages and their transfecting DNA damaged by radiation and alkylating agents. *Folia microbiologica* 1980; **25**(5)**:** 369-380.

347. Kunii Y, Niwa S, Hagiwara Y, Maeda M, Seitoh T, Suzuki T. The immunohistochemical expression profile of osteopontin in normal human tissues using two site-specific antibodies reveals a wide distribution of positive cells and extensive expression in the central and peripheral nervous systems. *Medical molecular morphology* 2009; **42**(3)**:** 155-161.

348. Harris PC, Torres VE. Polycystic Kidney Disease, Autosomal Dominant. *GeneReviews [Internet]* 1993.

349. Morel N, Vandenberg G, Ahrabi AK, Caron N, Desjardins F, Balligand JL *et al.* PKD1 haploinsufficiency is associated with altered vascular reactivity and abnormal calcium signaling in the mouse aorta. *Pflugers Arch* 2009; **457**(4)**:** 845-856.

350. Anyatonwu GI, Estrada M, Tian X, Somlo S, Ehrlich BE. Regulation of ryanodine receptor-dependent calcium signaling by polycystin-2. *Proc Natl Acad Sci U S A* 2007; **104**(15)**:** 6454-6459.

351. Qian Q, Hunter LW, Li M, Marin-Padilla M, Prakash YS, Somlo S *et al.* Pkd2 haploinsufficiency alters intracellular calcium regulation in vascular smooth muscle cells. *Hum Mol Genet* 2003; **12**(15)**:** 1875-1880.

352. Volk T, Schwoerer AP, Thiessen S, Schultz JH, Ehmke H. A polycystin-2-like large conductance cation channel in rat left ventricular myocytes. *Cardiovascular research* 2003; **58**(1)**:** 76-88.

353. Nauli SM, Alenghat FJ, Luo Y, Williams E, Vassilev P, Li X *et al.* Polycystins 1 and 2 mediate mechanosensation in the primary cilium of kidney cells. *Nat Genet* 2003; **33**(2)**:** 129-137.

354. Molland KL, Narayanan A, Burgner JW, Yernool DA. Identification of the structural motif responsible for trimeric assembly of the C-terminal regulatory domains of polycystin channels PKD2L1 and PKD2. *Biochem J* 2010; **429**(1)**:** 171-183.

355. Ding R, Shi J, Pabon K, Scotto KW. Xanthines down-regulate the drug transporter ABCG2 and reverse multidrug resistance. *Mol Pharmacol* 2012; **81**(3)**:** 328-337.

356. Isshiki M, Umezawa K, Tamura H. Coffee induces breast cancer resistance protein expression in Caco-2 cells. *Biological & pharmaceutical bulletin* 2011; **34**(10)**:** 1624-1627.

357. Tournier N, Chevillard L, Megarbane B, Pirnay S, Scherrmann JM, Decleves X. Interaction of drugs of abuse and maintenance treatments with human P-glycoprotein

(ABCB1) and breast cancer resistance protein (ABCG2). *The international journal of neuropsychopharmacology / official scientific journal of the Collegium Internationale Neuropsychopharmacologicum (CINP)* 2010; **13**(7)**:** 905-915.

358. Dobrinas M, Cornuz J, Eap CB. Pharmacogenetics of CYP1A2 activity and inducibility in smokers and exsmokers. *Pharmacogenet Genomics* 2013; **23**(5)**:** 286-292.

359. Hung WT, Lambert GH, Huang PW, Patterson DG, Jr., Guo YL. Genetic susceptibility to dioxin-like chemicals' induction of cytochrome P4501A2 in the human adult linked to specific AhRR polymorphism. *Chemosphere* 2013; **90**(9)**:** 2358-2364.

360. Legendre A, Baudoin R, Alberto G, Paullier P, Naudot M, Bricks T *et al.* Metabolic characterization of primary rat hepatocytes cultivated in parallel microfluidic biochips. *Journal of pharmaceutical sciences* 2013; **102**(9)**:** 3264-3276.

361. Vaynshteyn D, Jeong H. Caffeine induces CYP1A2 expression in rat hepatocytes but not in human hepatocytes. *Drug metabolism letters* 2012; **6**(2)**:** 116-119.

362. Kalthoff S, Ehmer U, Freiberg N, Manns MP, Strassburg CP. Coffee induces expression of glucuronosyltransferases via the aryl hydrocarbon receptor and Nrf2 in liver and stomach. *Gastroenterology* 2010; **139**(5)**:** 1699-1710.

363. de Waard PW, Peijnenburg AA, Baykus H, Aarts JM, Hoogenboom RL, van Schooten FJ *et al.* A human intervention study with foods containing natural Ah-receptor agonists does not significantly show AhR-mediated effects as measured in blood cells and urine. *Chemico-biological interactions* 2008; **176**(1)**:** 19-29.

364. Long JR, Egan KM, Dunning L, Shu XO, Cai Q, Cai H *et al.* Population-based case-control study of AhR (aryl hydrocarbon receptor) and CYP1A2 polymorphisms and breast cancer risk. *Pharmacogenet Genomics* 2006; **16**(4)**:** 237-243.

365. Fatima A, Andrabi S, Wolf G, Engelmann M, Spina MG. Urocortin 1 administered into the hypothalamic supraoptic nucleus inhibits food intake in freely fed and food-deprived rats. *Amino acids* 2013; **44**(3)**:** 879-885.

366. Fekete EM, Inoue K, Zhao Y, Rivier JE, Vale WW, Szucs A *et al.* Delayed satiety-like actions and altered feeding microstructure by a selective type 2 corticotropin-releasing factor agonist in rats: intra-hypothalamic urocortin 3 administration reduces food intake by prolonging the post-meal interval. *Neuropsychopharmacology* 2007; **32**(5)**:** 1052-1068.

367. Fekete EM, Zhao Y, Szucs A, Sabino V, Cottone P, Rivier J *et al.* Systemic urocortin 2, but not urocortin 1 or stressin 1-A, suppresses feeding via CRF2 receptors without malaise and stress. *Br J Pharmacol* 2011; **164**(8)**:** 1959-1975.

368. Wang C, Kotz CM. Urocortin in the lateral septal area modulates feeding induced by orexin A in the lateral hypothalamus. *American journal of physiology* 2002; **283**(2)**:** R358-367.

369. Wang C, Mullet MA, Glass MJ, Billington CJ, Levine AS, Kotz CM. Feeding inhibition by urocortin in the rat hypothalamic paraventricular nucleus. *American journal of physiology* 2001; **280**(2)**:** R473-480.

370. Benoit SC, Thiele TE, Heinrichs SC, Rushing PA, Blake KA, Steeley RJ. Comparison of central administration of corticotropin-releasing hormone and urocortin on food intake, conditioned taste aversion, and c-Fos expression. *Peptides* 2000; **21**(3)**:** 345-351.

371. Giardino WJ, Cote DM, Li J, Ryabinin AE. Characterization of Genetic Differences within the Centrally Projecting Edinger-Westphal Nucleus of C57BL/6J and DBA/2J Mice by Expression Profiling. *Frontiers in neuroanatomy* 2012; **6:** 5.

372. Fonareva I, Spangler E, Cannella N, Sabino V, Cottone P, Ciccocioppo R *et al.* Increased perioculomotor urocortin 1 immunoreactivity in genetically selected alcohol preferring rats. *Alcohol Clin Exp Res* 2009; **33**(11)**:** 1956-1965.

373. Sharpe AL, Phillips TJ. Central urocortin 3 administration decreases limited-access ethanol intake in nondependent mice. *Behav Pharmacol* 2009; **20**(4)**:** 346-351.

374. Tao J, Hildebrand ME, Liao P, Liang MC, Tan G, Li S *et al.* Activation of corticotropin-releasing factor receptor 1 selectively inhibits CaV3.2 T-type calcium channels. *Mol Pharmacol* 2008; **73**(6)**:** 1596-1609.

375. Zoumakis E, Rice KC, Gold PW, Chrousos GP. Potential uses of corticotropin-releasing hormone antagonists. *Ann N Y Acad Sci* 2006; **1083:** 239-251.

376. Dobrinas M, Cornuz J, Pedrido L, Eap CB. Influence of cytochrome P450 oxidoreductase genetic polymorphisms on CYP1A2 activity and inducibility by smoking. *Pharmacogenet Genomics* 2012; **22**(2)**:** 143-151.

377. Tomalik-Scharte D, Maiter D, Kirchheiner J, Ivison HE, Fuhr U, Arlt W. Impaired hepatic drug and steroid metabolism in congenital adrenal hyperplasia due to P450 oxidoreductase deficiency. *Eur J Endocrinol* 2010; **163**(6)**:** 919-924.

378. Das T, Yoo YS, Rhim H, Song EJ. Potential role of Hsp25 in calcium-modulated cardiomyocytes. *Proteomics* 2012; **12**(3)**:** 411-420.

379. Xu JW, Ikeda K, Kobayakawa A, Ikami T, Kayano Y, Mitani T *et al.* Downregulation of Rac1 activation by caffeic acid in aortic smooth muscle cells. *Life Sci* 2005; **76**(24)**:** 2861-2872.

380. Carr VM, Menco BP, Yankova MP, Morimoto RI, Farbman AI. Odorants as cell-type specific activators of a heat shock response in the rat olfactory mucosa. *The Journal of comparative neurology* 2001; **432**(4)**:** 425-439.

381. Mjahed H, Girodon F, Fontenay M, Garrido C. Heat shock proteins in hematopoietic malignancies. *Experimental cell research* 2012; **318**(15)**:** 1946-1958.

382. Moy GA, McNay EC. Caffeine prevents weight gain and cognitive impairment caused by a high-fat diet while elevating hippocampal BDNF. *Physiol Behav* 2012; **109:** 69-74.

383. Reyes-Izquierdo T, Nemzer B, Shu C, Huynh L, Argumedo R, Keller R *et al.* Modulatory effect of coffee fruit extract on plasma levels of brain-derived neurotrophic factor in healthy subjects. *Br J Nutr* 2013**:** 1-6.

384. Alzoubi KH, Srivareerat M, Aleisa AM, Alkadhi KA. Chronic caffeine treatment prevents stress-induced LTP impairment: the critical role of phosphorylated CaMKII and BDNF. *J Mol Neurosci* 2013; **49**(1)**:** 11-20.

385. Sallaberry C, Nunes F, Costa MS, Fioreze GT, Ardais AP, Botton PH *et al.* Chronic caffeine prevents changes in inhibitory avoidance memory and hippocampal BDNF immunocontent in middle-aged rats. *Neuropharmacology* 2013; **64:** 153-159.

386. Capiotti KM, Menezes FP, Nazario LR, Pohlmann JB, de Oliveira GM, Fazenda L *et al.* Early exposure to caffeine affects gene expression of adenosine receptors, DARPP-32 and BDNF without affecting sensibility and morphology of developing zebrafish (Danio rerio). *Neurotoxicology and teratology* 2011; **33**(6)**:** 680-685.

387. Alhaider IA, Aleisa AM, Tran TT, Alkadhi KA. Sleep deprivation prevents stimulation-induced increases of levels of P-CREB and BDNF: protection by caffeine. *Molecular and cellular neurosciences* 2011; **46**(4)**:** 742-751.

388. Connolly S, Kingsbury TJ. Caffeine modulates CREB-dependent gene expression in developing cortical neurons. *Biochem Biophys Res Commun* 2010; **397**(2)**:** 152-156.

389. Alhaider IA, Aleisa AM, Tran TT, Alkadhi KA. Caffeine prevents sleep loss-induced deficits in long-term potentiation and related signaling molecules in the dentate gyrus. *The European journal of neuroscience* 2010; **31**(8)**:** 1368-1376.

390. Bairam A, Kinkead R, Lajeunesse Y, Joseph V. Neonatal caffeine treatment does not induce long-term consequences on TrkB receptors or BDNF expression in chemosensory organs of adult rats. *Neurosci Lett* 2010; **468**(3)**:** 292-296.

391. Prakash YS, Iyanoye A, Ay B, Mantilla CB, Pabelick CM. Neurotrophin effects on intracellular Ca2+ and force in airway smooth muscle. *Am J Physiol Lung Cell Mol Physiol* 2006; **291**(3)**:** L447-456.

392. Wang Y, Zhang TY, Xin J, Li T, Yu H, Li N *et al.* Differential involvement of brain-derived neurotrophic factor in reconsolidation and consolidation of conditioned taste aversion memory. *PLoS One* 2012; **7**(11)**:** e49942.

393. Runge EM, Hoshino N, Biehl MJ, Ton S, Rochlin MW. Neurotrophin-4 is more potent than brain-derived neurotrophic factor in promoting, attracting and suppressing geniculate ganglion neurite outgrowth. *Developmental neuroscience* 2012; **34**(5)**:** 389-401.

394. Adaikkan C, Rosenblum K. The role of protein phosphorylation in the gustatory cortex and amygdala during taste learning. *Experimental neurobiology* 2012; **21**(2)**:** 37-51.

395. de Souza FT, Amaral TM, dos Santos TP, Abdo EN, Aguiar MC, Teixeira AL *et al.* Burning mouth syndrome: a therapeutic approach involving mechanical salivary stimulation. *Headache* 2012; **52**(6)**:** 1026-1034.

396. Nosrat CA, Blomlof J, ElShamy WM, Ernfors P, Olson L. Lingual deficits in BDNF and NT3 mutant mice leading to gustatory and somatosensory disturbances, respectively. *Development* 1997; **124**(7)**:** 1333-1342.

397. Nosrat IV, Margolskee RF, Nosrat CA. Targeted taste cell-specific overexpression of brain-derived neurotrophic factor in adult taste buds elevates phosphorylated TrkB protein levels in taste cells, increases taste bud size, and promotes gustatory innervation. *J Biol Chem* 2012; **287**(20)**:** 16791-16800.

398. Patel AV, Krimm RF. Neurotrophin-4 regulates the survival of gustatory neurons earlier in development using a different mechanism than brain-derived neurotrophic factor. *Dev Biol* 2012; **365**(1)**:** 50-60.

399. Martinez-Moreno A, Rodriguez-Duran LF, Escobar ML. Late Protein Synthesis-Dependent Phases in CTA Long-Term Memory: BDNF Requirement. *Frontiers in behavioral neuroscience* 2011; **5:** 61.

400. Zimmerberg B, Foote HE, Van Kempen TA. Olfactory association learning and brain-derived neurotrophic factor in an animal model of early deprivation. *Developmental psychobiology* 2009; **51**(4)**:** 333-344.

401. Cao L, Dhilla A, Mukai J, Blazeski R, Lodovichi C, Mason CA *et al.* Genetic modulation of BDNF signaling affects the outcome of axonal competition in vivo. *Curr Biol* 2007; **17**(11)**:** 911-921.

402. Kranjac D, McLinden KA, Deodati LE, Papini MR, Chumley MJ, Boehm GW. Peripheral bacterial endotoxin administration triggers both memory consolidation and reconsolidation deficits in mice. *Brain, behavior, and immunity* 2012; **26**(1)**:** 109-121.

403. DePoy LM, Noble B, Allen AG, Gourley SL. Developmentally divergent effects of Rho-kinase inhibition on cocaine- and BDNF-induced behavioral plasticity. *Behav Brain Res* 2013; **243:** 171-175.

404. Legastelois R, Botia B, Naassila M. Blockade of Ethanol-Induced Behavioral Sensitization by Sodium Butyrate: Descriptive Analysis of Gene Regulations in the Striatum. *Alcohol Clin Exp Res* 2013; **37**(7)**:** 1143-1153.

405. Schmidt HD, McGinty JF, West AE, Sadri-Vakili G. Epigenetics and psychostimulant addiction. *Cold Spring Harbor perspectives in medicine* 2013; **3**(3)**:** a012047.

406. Barish PA, Xu Y, Li J, Sun J, Jarajapu YP, Ogle WO. Design and functional evaluation of an optically active mu-opioid receptor. *European journal of pharmacology* 2013; **705**(1-3)**:** 42-48.

407. Kaczmarczyk MM, Machaj AS, Chiu GS, Lawson MA, Gainey SJ, York JM *et al.* Methylphenidate prevents high-fat diet (HFD)-induced learning/memory impairment in juvenile mice. *Psychoneuroendocrinology* 2013; **38**(9)**:** 15553-15564.

408. Geisel O, Banas R, Schneider M, Hellweg R, Muller CA. Serum levels of brain-derived neurotrophic factor in patients with internet use disorder. *Psychiatry research* 2013.

409. Li X, DeJoseph MR, Urban JH, Bahi A, Dreyer JL, Meredith GE *et al.* Different roles of BDNF in nucleus accumbens core versus shell during the incubation of cue-induced cocaine craving and its long-term maintenance. *J Neurosci* 2013; **33**(3)**:** 1130-1142.

410. Harte-Hargrove LC, Maclusky NJ, Scharfman HE. Brain-derived neurotrophic factor-estrogen interactions in the hippocampal mossy fiber pathway: Implications for normal brain function and disease. *Neuroscience* 2012; **239:** 46-66.

411. Zhang XY, Chen da C, Xiu MH, Luo X, Zuo L, Haile CN *et al.* BDNF Val66Met variant and smoking in a Chinese population. *PLoS One* 2012; **7**(12)**:** e53295.

412. McCarthy DM, Brown AN, Bhide PG. Regulation of BDNF expression by cocaine. *The Yale journal of biology and medicine* 2012; **85**(4)**:** 437-446.

413. Sim MS, Mohamed Z, Hatim A, Rajagopal VL, Habil MH. Association of brain-derived neurotrophic factor (Val66Met) genetic polymorphism with methamphetamine dependence in a Malaysian population. *Brain research* 2012; **1357:** 91-96.

414. Montag C, Reuter M, Newport B, Elger C, Weber B. The BDNF Val66Met polymorphism affects amygdala activity in response to emotional stimuli: evidence from a genetic imaging study. *NeuroImage* 2008; **42**(4)**:** 1554-1559.

415. Tsai SJ. Increased central brain-derived neurotrophic factor activity could be a risk factor for substance abuse: Implications for treatment. *Med Hypotheses* 2007; **68**(2)**:** 410-414.

416. Kim DJ, Noh JH, Cho NH, Lee BW, Choi YH, Jung JH *et al.* Serum gamma-glutamyltransferase within its normal concentration range is related to the presence of diabetes and cardiovascular risk factors. *Diabet Med* 2005; **22**(9)**:** 1134-1140.

417. Hou H, Qing Z, Jia S, Zhang X, Hu S, Hu J. Influence of brain-derived neurotrophic factor (val66met) genetic polymorphism on the ages of onset for heroin abuse in males. *Brain research* 2010; **1353:** 245-248.

418. Xu D, Zhang B, Liang G, Ping J, Kou H, Li X *et al.* Caffeine-induced activated glucocorticoid metabolism in the hippocampus causes hypothalamic-pituitary-adrenal axis inhibition in fetal rats. *PLoS One* 2012; **7**(9)**:** e44497.

419. Toyoshima K, Seta Y, Toyono T, Kataoka S. Immunohistochemical identification of cells expressing steroidogenic enzymes cytochrome P450scc and P450 aromatase in taste buds of rat circumvallate papillae. *Archives of histology and cytology* 2007; **70**(4)**:** 215-224.

420. Kravchenko LV, Trusov NV, Aksenov IV, Avren'eva LI, Guseva GV, Lashneva NV *et al.* [Effects of green tea extract and its components on antioxidant status and activities of xenobiotic metabolizing enzymes of rats]. *Voprosy pitaniia* 2011; **80**(2)**:** 9-15.

421. Uno Y, Uehara S, Murayama N, Yamazaki H. CYP1D1, pseudogenized in human, is expressed and encodes a functional drug-metabolizing enzyme in cynomolgus monkey. *Biochemical pharmacology* 2011; **81**(3)**:** 442-450.

422. Mills BM, Zaya MJ, Walters RR, Feenstra KL, White JA, Gagne J *et al.* Current cytochrome P450 phenotyping methods applied to metabolic drug-drug interaction prediction in dogs. *Drug Metab Dispos* 2010; **38**(3)**:** 396-404.

423. Eugster HP, Probst M, Wurgler FE, Sengstag C. Caffeine, estradiol, and progesterone interact with human CYP1A1 and CYP1A2. Evidence from cDNA-directed expression in Saccharomyces cerevisiae. *Drug Metab Dispos* 1993; **21**(1)**:** 43-49.

424. Wiercinska P, Lou Y, Squires EJ. The roles of different porcine cytochrome P450 enzymes and cytochrome b5A in skatole metabolism. *Animal* 2012; **6**(5)**:** 834-845.

425. Smolowitz RM, Schultz ME, Stegeman JJ. Cytochrome P4501A induction in tissues, including olfactory epithelium, of topminnows (Poeciliopsis spp.) by waterborne benzo[a]pyrene. *Carcinogenesis* 1992; **13**(12)**:** 2395-2402.

426. Bromek E, Haduch A, Daniel WA. The ability of cytochrome P450 2D isoforms to synthesize dopamine in the brain: An in vitro study. *European journal of pharmacology* 2010; **626**(2-3)**:** 171-178.

427. Lin W, Zhang J, Ling X, Yu N, Li J, Yang H *et al.* Evaluation of the effect of TM208 on the activity of five cytochrome P450 enzymes using on-line solid-phase extraction

HPLC-DAD: A cocktail approach. *J Chromatogr B Analyt Technol Biomed Life Sci* 2013; **923-924:** 29-36.

428. Ke AB, Nallani SC, Zhao P, Rostami-Hodjegan A, Isoherranen N, Unadkat JD. A Physiologically Based Pharmacokinetic Model to Predict Disposition of CYP2D6 and CYP1A2 Metabolized Drugs in Pregnant Women. *Drug Metab Dispos* 2013; **41**(4)**:** 801-813.

429. Lowcock EC, Cotterchio M, Anderson LN, Boucher BA, El-Sohemy A. High coffee intake, but not caffeine, is associated with reduced estrogen receptor negative and postmenopausal breast cancer risk with no effect modification by CYP1A2 genotype. *Nutrition and cancer* 2013; **65**(3)**:** 398-409.

430. Doroshyenko O, Rokitta D, Zadoyan G, Klement S, Schlafke S, Dienel A *et al.* Drug Cocktail Interaction Study on the Effect of the Orally Administered Lavender Oil Preparation Silexan on Cytochrome P-450 Enzymes in Healthy Volunteers. *Drug Metab Dispos* 2013; **41**(5)**:** 987-993.

431. Perera V, Gross AS, McLachlan AJ. Diurnal variation in CYP1A2 enzyme activity in South Asians and Europeans. *The Journal of pharmacy and pharmacology* 2013; **65**(2)**:** 264-270.

432. Li Q, Liu Y, Jiao J, Zhang C, Lou J. Assessment of effects of IR and IPC on activities of cytochrome P450 isozymes in rats by a five-drug cocktail approach. *Drug development and industrial pharmacy* 2013.

433. Dumond JB, Vourvahis M, Rezk NL, Patterson KB, Tien HC, White N *et al.* A phenotype-genotype approach to predicting CYP450 and P-glycoprotein drug interactions with the mixed inhibitor/inducer tipranavir/ritonavir. *Clin Pharmacol Ther* 2010; **87**(6)**:** 735-742.

434. Blake MJ, Abdel-Rahman SM, Pearce RE, Leeder JS, Kearns GL. Effect of diet on the development of drug metabolism by cytochrome P-450 enzymes in healthy infants. *Pediatr Res* 2006; **60**(6)**:** 717-723.

435. Ma JD, Nafziger AN, Villano SA, Gaedigk A, Bertino JS, Jr. Maribavir pharmacokinetics and the effects of multiple-dose maribavir on cytochrome P450 (CYP) 1A2, CYP 2C9, CYP 2C19, CYP 2D6, CYP 3A, N-acetyltransferase-2, and xanthine oxidase activities in healthy adults. *Antimicrobial agents and chemotherapy* 2006; **50**(4)**:** 1130-1135.

436. Kot M, Daniel WA. Caffeine as a marker substrate for testing cytochrome P450 activity in human and rat. *Pharmacol Rep* 2008; **60**(6)**:** 789-797.

437. Pardo Lozano R, Alvarez Garcia Y, Barral Tafalla D, Farre Albaladejo M. [Caffeine: a nutrient, a drug or a drug of abuse]. *Adicciones* 2007; **19**(3)**:** 225-238.

438. Keating E, Lemos C, Monteiro R, Azevedo I, Martel F. The effect of a series of organic cations upon the plasmalemmal serotonin transporter, SERT. *Life Sci* 2004; **76**(1)**:** 103-119.

439. Alaux-Cantin S, Warnault V, Legastelois R, Botia B, Pierrefiche O, Vilpoux C *et al.* Alcohol intoxications during adolescence increase motivation for alcohol in adult rats and induce neuroadaptations in the nucleus accumbens. *Neuropharmacology* 2013; **67:** 521-531.

440. Wheeler JM, Reed C, Burkhart-Kasch S, Li N, Cunningham CL, Janowsky A *et al.* Genetically correlated effects of selective breeding for high and low methamphetamine consumption. *Genes, brain, and behavior* 2009; **8**(8)**:** 758-771.

441. Gomez C, Brinon JG, Barbado MV, Weruaga E, Valero J, Alonso JR. Heterogeneous targeting of centrifugal inputs to the glomerular layer of the main olfactory bulb. *Journal of chemical neuroanatomy* 2005; **29**(4)**:** 238-254.

442. Ren Y, Shimada K, Shirai Y, Fujimiya M, Saito N. Immunocytochemical localization of serotonin and serotonin transporter (SET) in taste buds of rat. *Brain Res Mol Brain Res* 1999; **74**(1-2)**:** 221-224.

443. Elkins RL, Orr TE, Li JQ, Walters PA, Whitford JL, Carl GF *et al.* Serotonin reuptake is less efficient in taste aversion resistant than in taste aversion-prone rats. *Pharmacol Biochem Behav* 2000; **66**(3)**:** 609-614.

444. Hansson SR, Mezey E, Hoffman BJ. Serotonin transporter messenger RNA expression in neural crest-derived structures and sensory pathways of the developing rat embryo. *Neuroscience* 1999; **89**(1)**:** 243-265.

445. Sorensen L, Andersen J, Thomsen M, Hansen SM, Zhao X, Sandelin A *et al.* Interaction of antidepressants with the serotonin and norepinephrine transporters: mutational studies of the S1 substrate binding pocket. *J Biol Chem* 2012; **287**(52)**:** 43694-43707.

446. Naumenko VS, Tsybko AS, Bazovkina DV, Popova NK. [Implication of 5-HT2A receptors in the genetic mechanisms of the brain 5-HT system autoregulation]. *Molekuliarnaia biologiia* 2012; **46**(3)**:** 416-422.

447. Mulligan MK, Ponomarev I, Hitzemann RJ, Belknap JK, Tabakoff B, Harris RA *et al.* Toward understanding the genetics of alcohol drinking through transcriptome meta-analysis. *Proc Natl Acad Sci U S A* 2006; **103**(16)**:** 6368-6373.

448. Inouye M, Ripatti S, Kettunen J, Lyytikainen LP, Oksala N, Laurila PP *et al.* Novel Loci for metabolic networks and multi-tissue expression studies reveal genes for atherosclerosis. *PLoS Genet* 2012; **8**(8)**:** e1002907.

449. Osman W, Okada Y, Kamatani Y, Kubo M, Matsuda K, Nakamura Y. Association of common variants in TNFRSF13B, TNFSF13, and ANXA3 with serum levels of non-albumin protein and immunoglobulin isotypes in Japanese. *PLoS One* 2012; **7**(4)**:** e32683.

450. Kettunen J, Tukiainen T, Sarin AP, Ortega-Alonso A, Tikkanen E, Lyytikainen LP *et al.* Genome-wide association study identifies multiple loci influencing human serum metabolite levels. *Nat Genet* 2012; **44**(3)**:** 269-276.

451. Gieger C, Radhakrishnan A, Cvejic A, Tang W, Porcu E, Pistis G *et al.* New gene functions in megakaryopoiesis and platelet formation. *Nature* 2011; **480**(7376)**:** 201-208.

452. Chambers JC, Zhang W, Sehmi J, Li X, Wass MN, Van der Harst P *et al.* Genome-wide association study identifies loci influencing concentrations of liver enzymes in plasma. *Nat Genet* 2011; **43**(11)**:** 1131-1138.

453. Middelberg RP, Ferreira MA, Henders AK, Heath AC, Madden PA, Montgomery GW *et al.* Genetic variants in LPL, OASL and TOMM40/APOE-C1-C2-C4 genes are associated with multiple cardiovascular-related traits. *BMC Med Genet* 2011; **12:** 123.

454. Kathiresan S, Willer CJ, Peloso GM, Demissie S, Musunuru K, Schadt EE *et al.* Common variants at 30 loci contribute to polygenic dyslipidemia. *Nat Genet* 2009; **41**(1)**:** 56-65.

455. Chambers JC, Elliott P, Zabaneh D, Zhang W, Li Y, Froguel P *et al.* Common genetic variation near MC4R is associated with waist circumference and insulin resistance. *Nat Genet* 2008; **40**(6)**:** 716-718.

456. Dehghan A, Dupuis J, Barbalic M, Bis JC, Eiriksdottir G, Lu C *et al.* Meta-analysis of genome-wide association studies in >80 000 subjects identifies multiple loci for C-reactive protein levels. *Circulation* 2011; **123**(7)**:** 731-738.

457. Johansen CT, Wang J, Lanktree MB, Cao H, McIntyre AD, Ban MR *et al.* Excess of rare variants in genes identified by genome-wide association study of hypertriglyceridemia. *Nat Genet* 2010; **42**(8)**:** 684-687.

458. Kottgen A, Pattaro C, Boger CA, Fuchsberger C, Olden M, Glazer NL *et al.* New loci associated with kidney function and chronic kidney disease. *Nat Genet* 2010; **42**(5)**:** 376-384.

459. Waterworth DM, Ricketts SL, Song K, Chen L, Zhao JH, Ripatti S *et al.* Genetic variants influencing circulating lipid levels and risk of coronary artery disease. *Arterioscler Thromb Vasc Biol* 2010; **30**(11)**:** 2264-2276.

460. Coviello AD, Haring R, Wellons M, Vaidya D, Lehtimaki T, Keildson S *et al.* A genome-wide association meta-analysis of circulating sex hormone-binding globulin

reveals multiple Loci implicated in sex steroid hormone regulation. *PLoS Genet* 2012; **8**(7)**:** e1002805.

461. Franke A, McGovern DP, Barrett JC, Wang K, Radford-Smith GL, Ahmad T *et al.* Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci. *Nat Genet* 2010; **42**(12)**:** 1118-1125.

462. Kristiansson K, Perola M, Tikkanen E, Kettunen J, Surakka I, Havulinna AS *et al.* Genome-wide screen for metabolic syndrome susceptibility Loci reveals strong lipid gene contribution but no evidence for common genetic basis for clustering of metabolic syndrome traits. *Circ Cardiovasc Genet* 2012; **5**(2)**:** 242-249.

463. Suhre K, Shin SY, Petersen AK, Mohney RP, Meredith D, Wagele B *et al.* Human metabolic individuality in biomedical and pharmaceutical research. *Nature* 2011; **477**(7362)**:** 54-60.

464. Lemaitre RN, Tanaka T, Tang W, Manichaikul A, Foy M, Kabagambe EK *et al.* Genetic loci associated with plasma phospholipid n-3 fatty acids: a meta-analysis of genome-wide association studies from the CHARGE Consortium. *PLoS Genet* 2011; **7**(7)**:** e1002193.

465. Kolz M, Johnson T, Sanna S, Teumer A, Vitart V, Perola M *et al.* Meta-analysis of 28,141 individuals identifies common variants within five new loci that influence uric acid concentrations. *PLoS Genet* 2009; **5**(6)**:** e1000504.

466. Aulchenko YS, Ripatti S, Lindqvist I, Boomsma D, Heid IM, Pramstaller PP *et al.* Loci influencing lipid levels and coronary heart disease risk in 16 European population cohorts. *Nat Genet* 2009; **41**(1)**:** 47-55.

467. Ridker PM, Pare G, Parker A, Zee RY, Danik JS, Buring JE *et al.* Loci related to metabolic-syndrome pathways including LEPR,HNF1A, IL6R, and GCKR associate with plasma C-reactive protein: the Women's Genome Health Study. *Am J Hum Genet* 2008; **82**(5)**:** 1185-1192.

468. Kathiresan S, Melander O, Guiducci C, Surti A, Burtt NP, Rieder MJ *et al.* Six new loci associated with blood low-density lipoprotein cholesterol, high-density lipoprotein cholesterol or triglycerides in humans. *Nat Genet* 2008; **40**(2)**:** 189-197.

469. Willer CJ, Sanna S, Jackson AU, Scuteri A, Bonnycastle LL, Clarke R *et al.* Newly identified loci that influence lipid concentrations and risk of coronary artery disease. *Nat Genet* 2008; **40**(2)**:** 161-169.

470. Chasman DI, Giulianini F, MacFadyen J, Barratt BJ, Nyberg F, Ridker PM. Genetic determinants of statin-induced low-density lipoprotein cholesterol reduction: the Justification for the Use of Statins in Prevention: an Intervention Trial Evaluating Rosuvastatin (JUPITER) trial. *Circ Cardiovasc Genet* 2012; **5**(2)**:** 257-264.

471. Kottgen A, Albrecht E, Teumer A, Vitart V, Krumsiek J, Hundertmark C *et al.* Genome-wide association analyses identify 18 new loci associated with serum urate concentrations. *Nat Genet* 2012; **45**(2)**:** 145-154.

472. Yang Q, Kottgen A, Dehghan A, Smith AV, Glazer NL, Chen MH *et al.* Multiple genetic loci influence serum urate levels and their relationship with gout and cardiovascular disease risk factors. *Circ Cardiovasc Genet* 2010; **3**(6)**:** 523-530.

473. Dehghan A, Kottgen A, Yang Q, Hwang SJ, Kao WL, Rivadeneira F *et al.* Association of three genetic loci with uric acid concentration and risk of gout: a genome-wide association study. *Lancet* 2008; **372**(9654)**:** 1953-1961.

474. Kim YJ, Go MJ, Hu C, Hong CB, Kim YK, Lee JY *et al.* Large-scale genome-wide association studies in East Asians identify new genetic loci influencing metabolic traits. *Nat Genet* 2011; **43**(10)**:** 990-995.

475. Thorleifsson G, Walters GB, Gudbjartsson DF, Steinthorsdottir V, Sulem P, Helgadottir A *et al.* Genome-wide association yields new sequence variants at seven loci that associate with measures of obesity. *Nat Genet* 2009; **41**(1)**:** 18-24.

476. Consortium TaG. Genome-wide meta-analyses identify multiple loci associated with smoking behavior. *Nat Genet* 2010; **42**(5)**:** 441-447.

477. Newton-Cheh C, Johnson T, Gateva V, Tobin MD, Bochud M, Coin L *et al.* Genome-wide association study identifies eight loci associated with blood pressure. *Nat Genet* 2009; **41**(6)**:** 666-676.

478. Chasman DI, Pare G, Mora S, Hopewell JC, Peloso G, Clarke R *et al.* Forty-three loci associated with plasma lipoprotein size, concentration, and cholesterol content in genome-wide analysis. *PLoS Genet* 2009; **5**(11)**:** e1000730.