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

The Coffee and Caffeine Genetics Consortium

Marilyn C Cornelis[1,2], Enda M Byrne[3,117], Tõnu Esko[4-7,117], Michael A Nalls[8,117], Andrea Ganna[9], Nina Paynter[10], Keri L Monda[11], Najaf Amin[12], Krista Fischer[4], Frida Renstrom[13], Julius S Ngwa[14], Ville Huikari[15], Alana Cavadino[16], Ilja M Nolte[17], Alexander Teumer[18], Kai Yu[19], Pedro Marques-Vidal[20], Rajesh Rawal[21], Ani Manichaikul[22], Mary K Wojczynski[23], Jacqueline M Vink[24], Jing Hua Zhao[25], George Burlutsky[26], Jari Lahti[27,28], Vera Mikkilä[29,30], Rozenn N Lemaitre[31], Joel Eriksson[32], Solomon K Musani[33], Toshiko Tanaka[34], Frank Geller[35], Jian'an Luan[25], Jennie Hui[36-39], Reedik Mägi[4], Maria Dimitriou[40], Melissa E Garcia[41], Weang-Kee Ho[42], Margaret J Wright[43], Lynda M Rose[10], Patrik KE Magnusson[9], Nancy L Pedersen[9], David Couper[44], Ben A Oostra[45], Albert Hofman[12], Mohammad Arfan Ikram[12,46,47], Henning W Tiemeier[12,48], Andre G Uitterlinden[12,49], Frank JA van Rooij[12], Inês Barroso[50,51], Ingegerd Johansson[52], Luting Xue[14], Marika Kaakinen[15,53,54], Lili Milani[4], Chris Power[16], Harold Snieder[17], Ronald P Stolk[17], Sebastian E Baumeister[55], Reiner Biffar[56], Fangyi Gu[19], François Bastardot[57], Zoltán Kutalik[58-60], David R Jacobs Jr [61], Nita G Forouhi[25], Evelin Mihailov[4], Lars Lind[62], Cecilia Lindgren[63], Karl Michaëlsson[64], Andrew Morris[63], Majken Jensen[2], Kay-Tee Khaw[42], Robert N Luben[42], Jie Jin Wang[26], Satu Männistö[65], Mia-Maria Perälä[65], Mika Kähönen[66], Terho Lehtimäki[67], Jorma Viikari[68], Dariush Mozaffarian[1,2,69], Kenneth Mukamal[70], Bruce M Psaty[31,71-73], Angela Döring[74], Andrew C Heath[75], Grant W Montgomery[43], Norbert Dahmen[76], Teresa Carithers[77], Katherine L Tucker[78], Luigi Ferrucci[34], Heather A Boyd[35], Mads Melbye[35], Jorien L Treur[24], Dan Mellström[32], Jouke Jan Hottenga[24], Inga Prokopenko[63,79], Anke Tönjes[80,81], Panos Deloukas[50,82,83], Stavroula Kanoni[82], Mattias Lorentzon[32], Denise K Houston[84], Yongmei Liu[84], John Danesh[42], Asif Rasheed[85], Marc A Mason[86], Alan B Zonderman[87], Lude Franke[88], Bruce S Kristal[89,90], International Parkinson's Disease Genomics Consortium (IPDGC)[91], North American Brain Expression Consortium (NABEC)[91], UK Brain Expression Consortium (UKBEC)[91], Juha Karjalainen[88], Danielle R Reed[92], Harm-Jan Westra[88], Michele K Evans[86], Danish Saleheen[42,85], Tamara B Harris[41], George Dedoussis[40], Gary Curhan[1], Michael Stumvoll[80,81], John Beilby[36-38], Louis R Pasquale[1,93], Bjarke Feenstra[35], Stefania Bandinelli[94], Jose M Ordovas[95], Andrew T Chan[1,96], Ulrike Peters[97], Claes Ohlsson[32], Christian Gieger[21], Nicholas G Martin[43], Melanie Waldenberger[98], David S Siscovick[31,71], Olli Raitakari [30,99], Johan G Eriksson[28,100,101], Paul Mitchell[26], David J Hunter[1,102], Peter Kraft[102], Eric B Rimm[1,2,69], Dorret I Boomsma[24], Ingrid B Borecki[23], Ruth JF Loos[25,103,104], Nicholas J Wareham[25], Peter Vollenweider[57], Neil Caporaso[19], Hans Jörgen Grabe[105], Marian L Neuhouser[106], Bruce HR Wolffenbuttel[107], Frank B Hu[1,2,69], Elina Hyppönen[16,108,109], Marjo-Riitta Järvelin[15,53,54,110,111], L Adrienne Cupples[14,112], Paul W Franks[2,13,113], Paul M Ridker[10], Cornelia M van Duijn[12,114], Gerardo Heiss[11], Andres Metspalu[4], Kari E North[11], Erik Ingelsson[62,63], Jennifer A Nettleton[115], Rob M van Dam[116], Daniel I Chasman[10]

[1]Channing Division of Network Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts, USA. [2]Department of Nutrition, Harvard School of Public Health, Boston, Massachusetts, USA. [3]The University of Queensland, Queensland Brain Institute, Queensland, Australia. [4]Estonian Genome Center,

University of Tartu, Tartu, Estonia. [5]Division of Endocrinology, Children's Hospital Boston, Boston, Massachusetts, USA. [6]Department of Genetics, Harvard Medical School, Boston, Massachusetts, USA. [7]Program in Medical and Population Genetics, Broad Institute, Cambridge, Massachusetts, USA. [8]Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health (NIH), Bethesda, Maryland, 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, Massachusetts, USA. [11]Department of Epidemiology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, 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, Massachusetts, 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, Maryland, 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, Virginia, USA. [23]Washington University School of Medicine, Department of Genetics, Division of Statistical Genomics, St Louis, Missouri, 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, 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, Washington, USA. [32]Centre for Bone and Arthritis Research, Institute of Medicine, Sahlgrenska Academy, University of Gothenburg, Sweden. [33]University of Mississippi Medical Center, Jackson, Mississippi, USA. [34]Translational Gerontology Branch, National Institute on Aging, NIH, Baltimore, Maryland, USA. [35]Statens Serum Institut, Department of Epidemiology Research, Copenhagen, Denmark. [36]Busselton Population Medical Research Foundation Inc., Busselton, Australia [37]PathWest Laboratory Medicine WA, Nedlands, Western Australia. [38]School of Pathology & Laboratory Medicine, The University of Western Australia, Nedlands, Western Australia. [39]School of Population Health, The University of Western Australia, Nedlands, Western 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 North Carolina at Chapel Hill, Chapel Hill, North Carolina, 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, Minnesota, 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, Massachusetts, USA. [70]Beth Israel Deaconess Medical Center, Boston, Massachusetts, USA. [71]Cardiovascular Health Research Unit, Department of Epidemiology, University of Washington, Seattle, Washington, USA. [72]Department of Health Services, University of Washington, Seattle, Washington, USA. [73]Group Health Research Institute, Group Health Cooperative, Seattle, Washington, USA. [74]Institute of Epidemiology, Helmholtz Zentrum-München, Munich-Neuherberg, Germany. [75]Department of Psychiatry, Washington University, St.Louis, Missouri, USA. [76]Department for Psychiatry, Johannes-Gutenberg-University, Mainz, Germany. [77]School of Applied Sciences, University of Mississippi, Oxford, Mississippi, USA. [78]Clinical Laboratory & Nutritional Sciences, University of Massachusetts Lowell, Lowell, Massachusetts, USA. [79]Department of Genomics of Common Diseases, Imperial College London, London, UK. [80]Medical Department, University of Leipzig, Germany. [81]IFB Adiposity Diseases, University of Leipzig, Leipzig, Germany. [82]William Harvey Research Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, UK. [83]King Abdulaziz University, Jeddah, Saudi Arabia. [84]Wake Forest School of Medicine, Winston-Salem, North Carolina, USA. [85]Center for Non-Communicable Diseases, Pakistan. [86]Health Disparities Research Section, Clinical Research Branch, National Institute on Aging, NIH, Baltimore, Maryland, USA. [87]Laboratory of Personality and Cognition, National Institute on Aging, NIH, Baltimore, Maryland, USA. [88]Department of Genetics, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands. [89]Department of Neurosurgery, Brigham and Women's Hospital, Boston, Massachusetts, USA. [90]Department of Surgery, Harvard Medical School, Boston, Massachusetts, USA. [91]see Appendix. [92]Monell Chemical Senses Center, Philadelphia, Pennsylvania, USA. [93]Mass Eye and Ear Infirmary, Boston, Massachusetts, USA. [94]Geriatric Unit, Azienda Sanitaria Firenze, Florence, Italy [95]Jean Mayer USDA Human Nutrition Research Center on Aging, Tufts University, Boston, Massachusetts, USA. [96]Division of Gastroenterology, Massachusetts General Hospital, Boston, Massachusetts, USA. [97]Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, Washington, USA. [98]Research Unit of Molecular Epidemiology, Helmholtz Zentrum-München, Munich-Neuherberg, Germany. [99]Department of Clinical Physiology and Nuclear Medicine, Turku University Hospital, Turku, Finland. [100]Department of General Practice and Primary health Care, University of Helsinki, Helsinki, Finland. [101]Helsinki University Central Hospital, Unit of General Practice, Helsinki, Finland. [102]Program in Genetic Epidemiology and Statistical Genetics, Harvard School of Public Health, Boston, Massachusetts, USA. [103]The Genetics of Obesity and Related Metabolic Traits Program, The Icahn School of Medicine at Mount Sinai, New York, New York, USA. [104]The Charles Bronfman Institute for Personalized Medicine, The Icahn School of Medicine at Mount Sinai, New York, New York, USA. [105]Department of Psychiatry and Psychotherapy, University Medicine Greifswald, HELIOS Hospital Stralsund, Germany. [106]Fred Hutchinson Cancer Research Center, Seattle, Washington, USA. [107]Department of Endocrinology, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands. [108]School of Population Health, University of South Australia, Adelaide, Australia. [109]South Australian Health and Medical Research Institute, Adelaide, 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, Massachusetts, USA. [113]Department of Public Health & 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, Texas, USA. [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, Singapore.
[117]These authors contributed equally to this work.
Correspondence: Dr Marilyn C Cornelis at the Department of Nutrition, Harvard School of Public Health, 401 Park Drive, Boston, MA 02215 USA
Email: mcorneli@hsph.harvard.edu
Dr. Daniel I Chasman at Division of Preventive Medicine, Brigham and Women's Hospital, Harvard Medical School, 900 Commonwealth Ave, Boston, MA 02215 USA
Email: dchasman@research.bwh.harvard.edu.

Running title: GWAS of coffee consumption

Coffee, a major dietary source of caffeine, is amongst 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/day) among up to 91,462 coffee consumers of European ancestry with top single-nucleotide polymorphisms (SNPs) followed-up in ~30,062 and 7,964, 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 ($_{log10}$Bayes-factor>5.64) with per allele effect sizes of 0.03-0.14 cups/day. Six are located in or near genes potentially involved in pharmacokinetics (*ABCG2*, *AHR*, *POR*, *CYP1A2*) and pharmacodynamics (*BDNF*, *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 ($P<5\times10^{-8}$).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.

INTRODUCTION

Coffee is amongst 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 43 and 58%[6]. Genome-wide association studies (GWAS) of habitual caffeine and coffee intake have identified variants near *CYP1A2* and *AHR*[7-9]. Cytochrome P450 (CYP)1A2 is responsible for ~95% of caffeine metabolism in humans and aryl hydrocarbon receptor (AHR) plays 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 GW meta-analysis of coffee consumption including over 120,000 coffee consumers sourced from population-based studies of European and African American ancestry.

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 (e.g. 2-3 cups/day) was converted to cups/day by taking the median value of each category (e.g. 2.5 cups/day). A secondary analysis was performed comparing high to 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/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 (MAF) <0.02 or with low imputation quality scores were removed prior to meta-analysis (Supplementary Table S5). The GWAtoolbox (see Supplementary Information for URLs) was used for initial quality control. Plots of study (1/median standard error of effect size) *vs.* (square root of sample size) and MAFs were also viewed for outliers and these were addressed prior to 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 (GC) 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 subject to pathway analysis using MAGENTA[15] (Supplementary Information).

For regions achieving association *P* values $<5\times10^{-8}$ (7p21, 7q23.11, 11p13, 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 selecting 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 GC (Supplementary Information).

Studies from all stages were included in an overall meta-analysis using Meta-ANalysis of TRans-ethnic Association studies (MANTRA)[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 GC 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 potential candidates. Finally, we searched the National Human Genome Research Institute GWAS catalogue[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 Figs S2-S5. Little evidence for genomic inflation (λ<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, approximately 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 $<5\times10^{-8}$ (7p21, 7q23.11, 11p13, 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 33 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 Figs S7 and S8). Confirmed loci have effect sizes of 0.03-0.14 cups/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 nonsynonymous 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 *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).

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 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 CYP450s 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, 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 189102 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, Val66Met) in *BDNF* (Supplementary Table S26). Brain-derived neurotrophin factor (BDNF) modulates the activity of serotonin, dopamine and glutamate, neurotransmitters involved in mood-related circuits and plays 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* 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 Women's Genome Health Study (WGHS, n~17,000) and plasma glucose in TwinGene (n~8,800) 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 nonsynonymous change in the encoded, glucokinase regulatory protein (GKRP) leading to increased hepatic glucokinase activity[34]. GKRP 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 (Table 2, Figure 1, Supplementary Tables S23, S26, S30, S31). 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/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 ($P<5\times10^{-8}$, 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, $P=4.85\times10^{-3}$), birth weight (rs7800944, $P=2.10\times10^{-3}$), plasma high-density lipoprotein (HDL, rs7800944, $P=2.24\times10^{-3}$), risk of Parkinson's disease (rs1481012, $P=7.11\times10^{-3}$), reduced blood pressure (rs6265, $P=6.58\times10^{-4}$; rs2472297, $P<6.80\times10^{-5}$ and rs9902453, $P=6.05\times10^{-3}$), HDL (rs6968554, $P=1.18\times10^{-3}$), risk of major depressive disorder (rs17685, $P=6.98\times10^{-3}$) and bipolar disorder (rs1260326, $P=2.31\times10^{-3}$). 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 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 current study is the first to link *GCKR* and *MLXIPL* variation to a behavioral trait. The nonsynonymous 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. mQTL and binding motif analysis suggests *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 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/day (e.g. 2-3 cups/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 FFQs 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 subject to similar limitations in GWAS[40, 41].

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.

CONFLICTS 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.

Supplementary information is available at *Molecular Psychiatry*'s website

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, 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 C A 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, Maryland 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, Maryland 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, Guy¹s 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).

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**(1)**:** 83-133.

2. Cornelis M. Gene-coffee interactions and health. *Curr Nutr Rep* 2014; **(in press)**.

3. Spiller MA. The chemical components of coffee. In: Spiller GA (ed). *Caffeine*. Boca Raton: CRC, 1998, pp 97-161.

4. American Psychiatric Association. *Diagnostic and statistical manual of mental disorders*. 5 edn. American Psychiatric Publishing: Arlington, VA, 2013.

5. Cornelis MC. Coffee intake. *Progress in molecular biology and translational science* 2012; **108:** 293-322.

6. Yang A, Palmer AA, de Wit H. Genetics of caffeine consumption and responses to caffeine. *Psychopharmacology* 2010; **211**(3)**:** 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**(4)**:** 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**(10)**:** 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. *Molecular psychiatry* 2011; **17**(11)**:** 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. *Biochemical pharmacology* 2008; **76**(4)**:** 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**(6)**:** 1775-1781.

12. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* 2010; **26**(17)**:** 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**(1)**:** 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**(8)**:** 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**(4)**:** 369-375, S361-363.

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**(4)**:** 744-753.

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

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

20. 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.

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**(6)**:** 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**(5)**:** 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**(1)**:** 27-40.

24. Swanson HI. DNA binding and protein interactions of the AHR/ARNT heterodimer that facilitate gene activation. *Chemico-biological interactions* 2002; **141**(1-2)**:** 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**(11)**:** 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**(9)**:** 1899-1912.

27. Pandey AV, Fluck CE. NADPH P450 oxidoreductase: structure, function, and pathology of diseases. *Pharmacol Ther* 2013; **138**(2)**:** 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 U S A* 2013; **110**(13)**:** 5223-5228.

29. Numakawa T, Suzuki S, Kumamaru E, Adachi N, Richards M, Kunugi H. BDNF function and intracellular signaling in neurons. *Histology and histopathology* 2010; **25**(2)**:** 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**(2)**:** 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**(20)**:** 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**(7345)**:** 43-49.

33. Canli T, Lesch KP. Long story short: the serotonin transporter in emotion regulation and social cognition. *Nature neuroscience* 2007; **10**(9)**:** 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**(21)**:** 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. *Journal of neurochemistry* 2002; **80**(1)**:** 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**(21)**:** 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**(1)**:** 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**(7)**:** 967-975.

40. Tobacco and Genetics Consortium. Genome-wide meta-analyses identify multiple loci associated with smoking behavior. *Nat Genet* 2010; **42**(5)**:** 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 U S A* 2011; **108**(17)**:** 7119-7124.

Figure Legends

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 (grey nodes). Edges indicate known interactions.

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] EUR n ≤ 30,062		Stage 2[b] AA n ≤ 7,964		Trans-Ethnic Meta-Analysis[c]			
					β (SE)	P	β (SE)	P	β (SE)	P	β (SE)	$Log_{10}BF$	Post Prob	n
2p24	**rs1260326**	*GCKR*	T/C	.41/.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	.89/.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	.37/.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	.39/.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	.72/.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	.29/.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	.19/.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	.31/.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	.24/.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	.54/.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: EA, effect allele; NEA, non-effect allele; EAF, effect allele frequency; EUR, European ancestry; AA, African American ancestry; SE, standard error; BF, Bayes-factor, Post Prob, posterior probability

[a]Genic SNPs are in bold-face.

[b]Effect coefficients (SE), representing cups/day per effect allele, and corresponding P-values from stage 1 fixed-effects meta-analysis (columns 6, 7) and stage 2 fixed-effects meta-analyses (columns 8-11).

[c]Effect coefficients (SE), representing cups/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 corresponding posterior probabilities (column 14) from trans-ethnic meta-analysis of all stage 1 and stage 2 studies. A posterior probability > 0.5 suggests heterogeneity in allelic effects.

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: CR, conserved region; eQTL, expression quantitative trait loci; mQTL, methylation quantitative trait loci

[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.

Check marks (✓) denote *presence* of

[d]non-synonymous SNPs in LD (CEU: $r^2 \geq 0.80$) with lead SNP (details provided for lead SNP only).

[e]a conserved region (spanning lead SNP and its correlated proxies, CEU: $r^2 \geq 0.8$).

[f]DNAse hypersensitivity sites at region spanning lead SNP and its correlated proxies, CEU: $r^2 \geq 0.8$.

[g]proteins bound at region spanning lead SNP and its correlated proxies, CEU: $r^2 \geq 0.8$.

[h]Enhancer (H3K4me1) or promoter (H3K4me3) histone marks (as defined by Ernst et al[32]) spanning lead SNP and its correlated proxies, CEU: $r^2 \geq 0.8$.

[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.

Table 3. Associations between coffee consumption loci and other traits

Lead SNP, allele ↑ coffee consumption[a] *closest gene*	Other Traits[b]	
	higher levels/risk[c]	lower levels/risk[c]
rs1260326, C *GCKR*	non-albumin protein fasting glucose HOMA-IR fasting insulin mannose	serum albumin 2-hr 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 ($P=2.33\times10^{-4}$) waist-to-hip-ratio ($P=3.40\times10^{-4}$) bipolar disorder ($P=2.31\times10^{-3}$)
rs1481012, A *ABCG2*		LDL response to statins ('responders') uric acid
	body mass index ($P=4.85\times10^{-3}$)	
rs6968554 , G *AHR*		caffeine
		HDL ($P=1.18\times10^{-3}$)
rs7800944, C *MLXIPL*		triglycerides
	HDL ($P=2.24\times10^{-3}$) birth weight ($P=2.10\times10^{-3}$)	
rs6265, C *BDNF*	smoking initiation body mass index	
		DBP ($P=6.58\times10^{-4}$)
rs2472297[d], T *CYP1A1_CYP1A2*		*caffeine*[e]
		SBP ($P=6.81\times10^{-5}$) DBP ($P=6.75\times10^{-6}$)

rs9902453, G *EFCAB5*		SBP ($P=6.05\times10^{-3}$)

Abbreviations: DBP, diastolic blood pressure; GGT, gamma-glutamyltransferase; HDL, high density lipoprotein; LDL, low density lipoprotein; SBP, systolic blood pressure; SHBG, sex hormone binding globulin

[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, 2014). Grey cells denote all GW-significant significant associations ($P<5.00\times10^{-8}$ [20] or $P<1.03\times10^{-10}$ (Shin et al, 2014) and white cells denote coffee-relevant trait associations ($P<6.25\times10^{-3}$). 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 ($P<1.46\times10^{-17}$) 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 ($P<5.00\times10^{-8}$).

[e]Borderline significant ($P<1.51\times10^{-10}$) according to Shin *et al* [21].

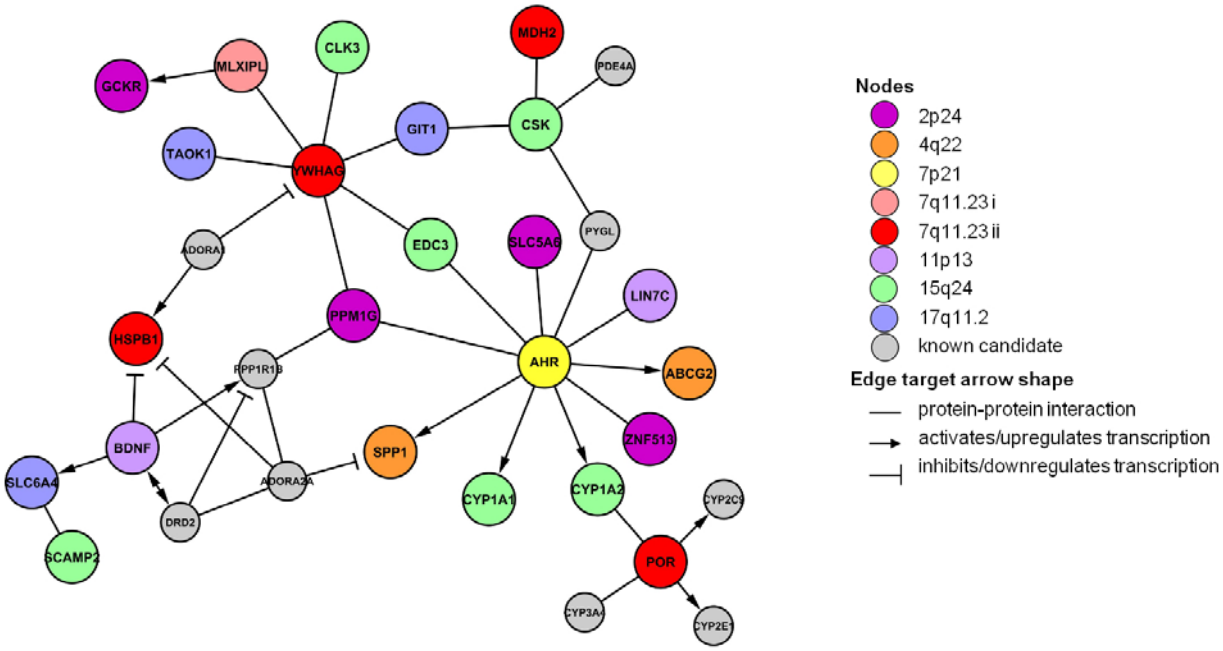


Figure 1. Network describing direct interactions between candidate genes of confirmed loci.