Supplementary note

Population descriptions for GWAS discovery cohorts in the Alcohol Genome-wide Association (AlcGen) consortium

Cohorte Lausannoise study (CoLaus)

The cohort is a random population sample of the city of Lausanne aged 35-75 years. Recruitment began in June 2003 and ended in May 2006. The CoLaus study was approved by the Institutional Ethics Committee of the University of Lausanne and informed consent was appropriately obtained by all participants. All participants attended the outpatient clinic of the University Hospital of Lausanne in the morning after an overnight fast. Data were collected by trained field interviewers in a single visit lasting about 60 min. Alcohol consumption was assessed by questionnaire and measured in units per week. In total 3,121 individuals were included in the analysis.

The Estonian Biobank Cohort (EGCUT)

The Estonian Biobank Cohort is a population-based cohort of 52000 Estonian residents (81% ethnic Estonians), recruited on volunteer-basis in 2002-2010 (www.biobank.ee), managed by the Estonian Genome Center, University of Tartu.

The European Prospective Investigation of Cancer - Norfolk study (EPIC-Norfolk)

The EPIC-Norfolk sample includes 2,566 participants randomly selected from the EPIC-Norfolk Study, a population-based cohort study of 25,663 men and women of European descent aged 39 - 79 years recruited in Norfolk, UK between 1993 and 1997.

The Erasmus Rucphen Family study (ERF)

The ERF[1] is a family based study that includes over 3,000 participants descending from 22 couples living in the Rucphen region in the 19th century. All living descendants of these couples and their spouses were invited to take part in the study. The medical ethics committee of Erasmus MC constituted according to the WMO (National Act Medical-scientific research in human beings) approved the Study (MEC 213.575/2002/114).

The genotyping for the ERF study was supported by EUROSPAN (European Special Populations Research Network) through 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 (NWO), Erasmus MC, the Centre for Medical Systems Biology (CMSB1 and CMSB2) and the Netherlands Genomics Initiative (NGI) and also received funding from the European Community's Seventh Framework Programme (FP7/2007-2013)/grant agreement HEALTH-F4-2007-201413 by the European Commission under the

programme “Quality of Life and Management of the Living Resources” of 5th Framework Programme (no. QLG2-CT-2002-01254). High-throughput analysis of the ERF data was supported by joint grant from Netherlands Organisation for Scientific Research and the Russian Foundation for Basic Research (NWO-RFBR 047.017.043). Exome sequencing analysis in ERF was supported by the ZonMw grant (project 91111025).

The Fenland study (Fenland)

The Fenland study is a population-based cohort study that uses objective measures of disease exposure, such as accurate methods of body composition and energy expenditure, to study the interactions between genetic and lifestyle factors that cause obesity and diabetes. The volunteers are recruited from general practice lists in and around Cambridgeshire (Cambridge, Ely, and Wisbech) in the United Kingdom from birth cohorts from 1950–1975.

The Younger Finnish Twin Cohort (FinnTwin12)

The FinnTwin12 cohort is composed of twins born in Finland during 1983-87. The study has a two-stage sampling design. The larger, first-stage study is an epidemiological investigation of five consecutive and complete birth cohorts of Finnish twin children, including questionnaire assessments of both twins and parents at baseline, starting with a family questionnaire (returned by 2724 families, 87% participation rate) that was mailed late in the year before the twins reach age 12, with follow-up of all twins at age 14, 17.5 years and ~22. Nested within this epidemiological, population-based study, is the second-stage of FinnTwin12, an intensive assessment of a sub-sample of twin families. Most of the sub-sample was selected at random, but this random sample (~72%) was then enriched with twins at elevated familial risk for alcoholism. Genome-wide genotyping was performed on the subjects of the intensive sub-sample.

The Older Finnish Twin Cohort (FinnTwinOld)

This sample originates from the Older Finnish Twin Cohort. The 1975 , 1981 and 1990 questionnaires for the same-sex twins and the 1996-97 questionnaires for opposite-sex twins requested identical information on the frequency and quantity of alcohol used during an average week (or month), the frequency of passouts experienced during the preceding year, and required a yes/no response to a question on drinking density. Frequency of alcohol use, measured as days' use per month on 5-point scales ('never' to "over 16 days a month') was assessed separately for beer, wine, and spirits. Similarly, quantity was measured on three 7-point scales, with the upper limits defined as consuming >48 bottles of beer (or 10- bottles of wine) per week, or >20 bottles of spirits per month. Wine use did contribute to the consumption measure. For each type of beverage, consumption was converted into grams of absolute alcohol and summed to yield an estimate of total consumption in grams per month using the class midpoints of the categories and the average alcohol content of each beverage type. The script for computing alcohol amount is available from Jaakko Kaprio on request.

The Helsinki Birth Cohort Study (HBCS)

The Helsinki Birth Cohort Study (HBCS) is composed of 8,760 individuals born between the years 1934-44 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.

The population-based Cooperative Health Research in the Region of Augsburg F3 Study (KORA F3)

The population-based Cooperative Health Research in the Region of Augsburg (KORA) F3 Study was carried out in 2004–2005 as a follow-up of the MONICA/KORA S3 baseline study (1994–1995). In S3, 4,856 participants were recruited out of a randomised two-stage cluster sample of 6,640 subjects, with equal-sized sex- and age-strata, from the target population of all German residents in the region of Augsburg aged 25–74 years. The F3 Study included 3,007 participants aged 35–84 years. 1,644 randomly drawn participants aged 35–79 with Affymetrix genotype data and data on alcohol intake were included in the investigations reported.

The population-based Cooperative Health Research in the Region of Augsburg F4 Study (KORA F4)

The population-based Cooperative Health Research in the Region of Augsburg (KORA) F4 Study was carried out in 2006–2008 as a follow-up of the KORA S4 baseline study (1999–2001). In S4, 4,261 participants were recruited out of a randomised two-stage cluster sample of 6,640 subjects, with equal-sized sex- and age-strata, from the target population of all German residents in the region of Augsburg aged 25–74 years. The F4 Study included 3,080 participants aged 32–81 years. 1814 randomly drawn participants aged 32–81 with Affymetrix genotype data and data on alcohol intake were included in the investigations reported.

LifeLines Cohort Study & Biobank (Lifelines)

LifeLines is a multi-disciplinary prospective population-based cohort study examining in a unique three-generation design the health and health-related behaviours of 165,000 persons living in the North East region of The Netherlands. It employs a broad range of investigative procedures in assessing the biomedical, socio-demographic, behavioural, physical and psychological factors which contribute to the health and disease of the general population, with a special focus on multimorbidity and complex genetics.

The London Life Sciences Prospective Population Study (LOLIPOP)

LOLIPOP is a population based prospective study of 17,606 Indian Asian and 7,766 European men and women aged 35-75 years, recruited from the lists of 58 General Practitioners in West London, United Kingdom between 2003 and 2008[2,3]. Europeans were of self-reported white

ancestry. Assessments of participants were carried out by trained research nurses with an interviewer-administered questionnaire. Anthropometric measurements and blood samples were taken on site. Alcohol consumption was measured in units per week. One unit is equivalent to: 1 small glass of wine, 1 single pub measure of spirits, or half a pint of beer / lager. Aliquots of whole blood were stored at -80C and DNA was extracted and genotyping was carried out thereafter. The LOLIPOP study is approved by the local Research Ethics Committees [3]. All participants provided written consent for the study.

The Netherlands Study of Depression and Anxiety (NESDA)

The Netherlands Study of Depression and Anxiety (NESDA)[4], an ongoing cohort study into the long-term course and consequences of depressive and anxiety disorders. Briefly, in 2004-2007 participants aged 18 to 65 years were recruited from the community (19%), general practice (54%) and secondary mental health care (27%), reflecting therefore various settings and developmental stages of psychopathology in order to obtain a full and generalizable picture of the course of psychiatric disorders. A total of 2,981 participants were included, consisting of persons with a current or past depressive and/or anxiety disorder and healthy controls. The research protocol was approved by the ethical committee of participating universities, and all respondents provided written informed consent.

The Northern Finland Birth Cohort 1966 (NFBC1996)

The North Finland Birth Cohort of 1966 (NFBC1966, n=12,058 live born) was designed to study factors affecting preterm birth, low birth weight, and subsequent morbidity and mortality (http://kelo.oulu.fi/NFBC/). The longitudinal data collection includes clinical examination and blood sampling at age 31 years, from which data in the current study are drawn. The attendees in the follow-up (71% response rate) were adequately representative of the original cohort as is the final study sample in the present analyses. A total of 4,763 genotyped samples were available from the NFBC1966.

Netherlands Twin Register cohort (NTR)

Netherlands Twin Register (NTR)[5,6] participants are ascertained because of the presence of twins or triplets in the family and consist of multiples, their parents, siblings and spouses. Twins are born in all strata of society and NTR represents a general sample from the Dutch population.

The Australian twin-family study of alcohol use disorder (OZALC)

This twin/family cohort was based on two groups of twins, born before 1964 and born 1964-71, enrolled in a voluntary Australia-wide twin registry. Twins, their spouses, and first-degree relatives were recruited for a study on alcohol dependence and related phenotypes [7]. Alcohol intake in the week preceding blood collection was self-reported, and history of alcohol use and dependence was obtained through structured telephone interviews.

The Prevention of REnal and Vascular ENd-stage Disease study (PREVEND)

The PREVEND study is an ongoing prospective study investigating the natural course of increased levels of urinary albumin excretion and its relation to renal and cardiovascular disease. Inhabitants 28 to 75 years of age (n=85,421) in the city of Groningen, The Netherlands, were asked to complete a short questionnaire, 47% responded, and individuals were then selected with a urinary albumin concentration of at least 10 mg/L (n = 7,768) and a randomly selected control group with a urinary albumin concentration less than 10 mg/L (n = 3,395).

The Study of Health in Pomerania (SHIP)

The Study of Health in Pomerania (SHIP) is a population-based project in West Pomerania, the north-east area of Germany[8,9]. A sample from the population aged 20 to 79 years was drawn from population registries. First, the three cities of the region (with 17,076 to 65,977 inhabitants) and the 12 towns (with 1,516 to 3,044 inhabitants) were selected, and then 17 out of 97 smaller towns (with less than 1,500 inhabitants), were drawn at random. Second, from each of the selected communities, subjects were drawn at random, proportional to the population size of each community and stratified by age and gender. Only individuals with German citizenship and main residency in the study area were included. Finally, 7,008 subjects were sampled, with 292 persons of each gender in each of the twelve five-year age strata. In order to minimize drop-outs by migration or death, subjects were selected in two waves. The net sample (without migrated or deceased persons) comprised 6,267 eligible subjects. Selected persons received a maximum of three written invitations. In case of non-response, letters were followed by a phone call or by home visits if contact by phone was not possible. The SHIP population finally comprised 4,308 participants (corresponding to a final response of 68.8%). Alcohol intake was assessed by questionnaire: drink-specific quantity-frequency 30d[10].

TwinsUK

TwinsUK is based on a sample of 5,654 individuals from the UK. Among these, 3,471 have been genotyped and have data on alcohol intake assessed by self-reported questionnaire, and 1,204 represent one co-twin per family which have been genotyped and have data on alcohol intake assessed by self-reported questionnaire.

The Cardiovascular Risk in Young Finns Study (YFS)

The YFS is a population-based follow up-study started in 1980. The main aim of the YFS is to determine the contribution made by childhood lifestyle, biological and psychological measures to the risk of cardiovascular diseases in adulthood. In 1980, over 3,500 children and adolescents all around Finland participated in the baseline study. The follow-up studies have been conducted mainly with 3-year intervals. The 27-year follow-up study was conducted in 2007 (ages 30-45 years) with 2,204 participants. The study was approved by the local ethics committees (University Hospitals of Helsinki, Turku, Tampere, Kuopio and Oulu) and was conducted

following the guidelines of the Declaration of Helsinki. All participants gave their written informed consent.

Population descriptions for GWAS discovery cohorts in the Heart and Aging Research in Genomic Epidemiology Plus (CHARGE+) Consortium

Age, Gene/Environment Susceptibility–Reykjavik (AGES-Reykjavik)

The AGES-Reykjavik Study[11] is a single center prospective cohort study based on the Reykjavik Study. The Reykjavik Study was initiated in 1967 by the Icelandic Heart Association to study cardiovascular disease and risk factors. The cohort included men and women born between 1907 and 1935 who lived in Reykjavik at the 1967 baseline examination. Re-examination of surviving members of the cohort was initiated in 2002 as part of the AGES-Reykjavik Study.

The Atherosclerosis Risk in Communities Study (**ARIC**)

The ARIC study[12] consists of a prospective cohort designed to identify the causes and outcomes of cardiovascular disease in 15,792 individuals from 4 communities (Forsyth County, NC; Jackson, MS; suburbs of Minneapolis, MN; and Washington County, MD). ARIC study participants underwent interviews, fasting venipuncture, and measurement of anthropometrics at the baseline and follow-up examinations. Trained interviewers ascertained basic demographic data, medical history, and information about personal diet habits. A full description of study design is available on the ARIC website (http://www2.cscc.unc.edu/aric/). In total, 4,106 individuals had both genotyping and alcohol phenotype.

Alcohol consumption was ascertained by means of an interviewer-administered dietary questionnaire. Frequency of alcohol consumption was determined as usual weekly intake, with the amount of alcohol consumed in grams per week calculated assuming different serving sizes and alcohol content for beer, wine, and hard liquor. Serving sizes and alcohol content were defined as follows: ‘one beer' (12 oz. bottles or cans of beer, 13.2 g), 'one glass of wine' (4 oz. glass, 10.8 g), or ‘one shot of liquor or one mixed drink' (1.5 oz. shot of hard liquor, 15.1 g). The total amount of absolute alcohol ingested weekly for past alcohol consumption was determined by multiplying the number of servings by the amount of alcohol in one serving of the type of alcohol ordinarily drunk. If more than one type was ordinarily drunk, the calculation was made assuming an equal number of drinks of each type. The total amount of absolute alcohol ingested weekly for present alcohol consumption resulted from the addition of absolute alcohol consumed for wine, beer, and hard liquor. The total amount of absolute alcohol drunk during the 24 hours prior to the clinic interview was determined by multiplying the number of drinks by the amount of absolute alcohol in the type of drink consumed. For a drinker who reported less than one drink per week, the alcohol consumption was recorded as zero grams per week. All questions were closed-ended and designed for direct data entry by a trained interviewer. In order to ensure

standardization, exact wording and order of questions were followed. Questions were skipped only if specified in the questionnaire instructions.

The Cardiovascular Health Study (CHS)

The CHS is a population-based cohort study of risk factors for coronary heart disease and stroke in adults ≥65 years conducted across four field centers[13]. The original predominantly European ancestry cohort of 5,201 persons was recruited in 1989-1990 from random samples of the Medicare eligibility lists; subsequently, an additional predominantly African-American cohort of 687 persons was enrolled for a total sample of 5,888. The CHS GWAS, which had the primary aim of studying incident cardiovascular events, focused on 3,980 participants who were free of clinical cardiovascular disease at study baseline, consented to genetic testing, and had DNA available for genotyping. A total of 1,908 persons were excluded from the GWAS study sample due to the presence at study baseline of coronary heart disease, congestive heart failure, peripheral vascular disease, valvular heart disease, stroke, or transient ischemic attack. Because the other cohorts were predominantly of European descent, the African American participants were excluded from this analysis. In total, 3009 participants with both genotype and alcohol phenotype were included in the analyses.

At the baseline visit and annually, participants separately reported their usual frequency of consumption of beer, wine, and liquor, and the usual number of 12-ounce cans or bottles of beer, 6-ounce glasses of wine, and shots of liquor that they drank on each occasion. The full text of the CHS nutritional questionnaire is publicly available (http://www.chs-nhlbi.org/forms/r25p3.htm). At baseline, participants also reported whether they changed their pattern of consumption during the past 5 years and whether they ever regularly consumed 5 or more drinks daily.

The Framingham Heart Study (FHS)

The FHS sample includes the Framingham Heart Study Offspring[14] and the third generation[15] cohorts. In 1971, children and spouses of children of the original FHS cohort participants were recruited into the Framingham offspring cohort, which consists of 5,124 men and women. The FHS offspring participants have been examined every four to eight years unless specified otherwise, common clinical phenotypes from all examinations were available for this investigation. From 2002 to 2005, a third generation cohort of 4,095 individuals was recruited to the FHS. The third generation cohort (n=4,095) includes children and spouses of children of the Offspring cohort. In total, 8,955 individuals had both genotyping and alcohol phenotype.

Alcohol consumption was assessed via questionnaire at the study examination closest to the timepoint of DNA collection.

The Health, Aging, and Body Composition (HABC)

The Health ABC study[16] is a prospective cohort study investigating the associations between body composition, weight-related health conditions, and incident functional limitation in older adults. Health ABC 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. The current study sample consists of 1559 white participants who attended the second exam in 1998-1999 with available genotyping data.
Alcohol consumption at baseline was assessed by asking the participant how many alcoholic drinks he/she consumed in a typical week, during the past 12 months. Furthermore, it was asked whether a person ever drank more than what he/she typically drank in the past 12 months.

The Multi-Ethnic Study of Atherosclerosis (MESA)

The MESA is a study of the characteristics of subclinical cardiovascular disease (disease detected non-invasively before it has produced clinical signs and symptoms) and the risk factors that predict progression to clinically overt cardiovascular disease or progression of the subclinical disease[17]. MESA researchers study a diverse, population-based sample of 6,814 asymptomatic men and women aged 45-84. Thirty-eight percent of the recruited participants are White, 28 percent African-American, 22 percent Hispanic, and 12 percent Asian, predominantly of Chinese descent. Participants were recruited from six field centers across the United States: Wake Forest University, Columbia University, Johns Hopkins University, University of Minnesota, Northwestern University and University of California - Los Angeles. The current analysis was limited to n=1596 White participants with data available on alcohol consumption through the Food-frequency questionnaire.
Data on alcoholic beverage consumption (drinks/day) were obtained on 2,382 Caucasian individuals with genotypes available through MESA SHARe. As a part of a 120-item food frequency questionnaire, participants were asked the frequency they consumed each beer, wine, and liquor or mixed drinks (9 frequency options ranging from rarely/never to six or more drinks/day)[18,19]. Responses to these three line items were summed to estimate total alcoholic drinks consumed each day.

The Rotterdam Study (RS)

The RS[20] is a prospective, population-based study from the well-defined district of Ommoord within the city of Rotterdam, designed to investigate the occurrence and determinants of diseases in the elderly. The cohort was initially defined in 1990 among 7 983 persons who underwent a home interview and extensive physical examination at baseline and during follow-up examinations occurring every 3-4 years (RS-I). The cohort was further extended in 2000 (RS-II) and 2005 (RS-III), establishing a total of 14926 participants.

The Women's Genome Health Study (WGHS)

The WGHS[21] is a prospective cohort of initially healthy, female North American health care professionals at least 45 years old at baseline in 1992-1994, representing participants in the Women's Health Study (WHS) who provided a blood sample at baseline and consent for blood-based analyses. These WHS was 2x2 randomized, placebo controled trial of aspirin and vitamin E in prevention of cardiovascular disease and cancer over 10 years. Since the end of the trial, follow-up in the WHS/WGHS has continued in observational mode.

Population descriptions for replication cohorts

Airwave Health Monitoring Study (Airwave)

The Airwave Health Monitoring Study[22] was established to evaluate possible health risks associated with use of TETRA, a digital communication system used by police forces and other emergency services in Great Britain since 2001. The study has been broadened to investigate more generally the health of the work force. From 2004, participants from each force who agreed to participate were enrolled either with an enrolment questionnaire or a comprehensive health screening performed locally. This includes questionnaire, 7-day food diaries, anthropometry, measurements of cardiovascular and cognitive function, blood chemistry, coagulation and haematology. By March 2015, the study had recruited 53,606 participants, of whom 45,433 had attended the health screening. 12,930 participants with genotype data were included in this analysis.

The Austrian Stroke Prevention Study (ASPS)

The ASPS study is a single center prospective follow-up study on the effects of vascular risk factors on brain structure and function in the normal elderly population of the city of Graz, Austria. The procedure of recruitment and diagnostic work-up of study participants has been described previously[23,24]. A total of 2,007 participants were randomly selected from the official community register stratified by gender and 5 year age groups. Individuals were excluded from the study if they had a history of neuropsychiatric disease, including previous stroke, transient ischemic attacks, and dementia, or an abnormal neurologic examination determined on the basis of a structured clinical interview and a physical and neurologic examination. During 2 study periods between September 1991 and March 1994 and between January 1999 and December 2003 an extended diagnostic work-up including neuropsychological testing was done in 1,076 individuals aged 45 to 85 years randomly selected from the entire cohort: 509 from the first period and 567 from the second. In 1992, blood was drawn from all study participants for DNA extraction. They were all European Caucasians. Genotyping was performed in 996 participants, and those 829 who passed genotyping quality control and have data on alcohol intake were available for these analyses.

The British 1958 birth cohort (B58C)

The British 1958 birth cohort[25] is a follow-up study of persons born throughout England, Scotland and Wales one week in March 1958. Alcohol consumption was self-reported at a biomedical examination at age 44-45 years, at which blood sampling was performed with consent for DNA extraction and creation of immortalised cell lines. Genotyping of three non-overlapping subsets of the cohort was performed by the Wellcome Trust Case-Control Consortium, the Type 1 Diabetes Genetics Consortium and the GABRIEL Asthma Genetics Consortium. The three subsets were combined for imputation using the 1000-genomes phase 1 reference panel, and for subsequent statistical analysis.

Data from an Epidemiological Study on the Insulin Resistance syndrome (DESIR)
General population from ten French Social Security Health Examination Centres.

The Finnish Twin Cohort replication sample (FinnTwin_replication)
Sample used for replication consists of subjects from the Older Finnish Twin Cohort and the Younger Finnish Twin Cohorts (non-overlapping with the discovery sample). Please see cohort descriptions of the discovery sample.

Genetic Regulation of Arterial Pressure of Humans in the Community Study (GRAPHIC)
The GRAPHIC Study comprises 2024 individuals from 520 nuclear families recruited from the general population in Leicestershire, UK between 2003-2005 for the purpose of investigating the genetic determinants of blood pressure and related cardiovascular traits. Families were included if both parents aged 40-60 years and two offspring $\geqslant$18 years wished to participate.
A detailed medical and lifestyle history including alcohol intake was obtained from study subjects by standardized questionnaires and clinical examination was performed by research nurses following standard procedures.

Generation Scotland: Scottish Family Health Study (GS:SFHS)
GS:SFHS[26] consists of 23,960 individuals recruited at random from general medical practices across Scotland, 21,516 of these attended the research clinic. Eligibility criteria specified that participants were over 18 years of age and had one first-degree relative also willing to participate. Genome-wide SNP data were ascertained for 10,000 individuals, and after quality control, genotype data were available for 9,863 participants, which are the participants used in this study. 7,281 of these individuals self-reported as currently consuming alcohol. Alcohol consumption was assessed using a pre-clinical questionnaire. Participants were identified as current drinkers, former drinkers or never drinkers. Consumption was measured in self-reported units of alcohol consumed in the previous week. The cohort has been described in further detail elsewhere[26].

The INGI - Carlantino study (INGI_CARL)
This cohort comprises the samples coming from a small village from the southern region of Italy

Puglia. For all samples a wide range of information are available including alcohol intake and anthropometric measurements. Moreover for all samples a DNA sample was acquired and was used for genotyping with high density SNP arrays.

The INGI - Friuli Venezia Giulia study (INGI_FVG)

This cohort comprises the samples coming from a 6 small villages from the northern region of Italy Friuli Venezia Giulia. For all samples a wide range of information are available including alcohol intake and anthropometric measurements. Moreover for all samples a DNA sample was acquired and was used for genotyping with high density SNP arrays.

The INGI - Val Borbera study (INGI_VB)

The INGI-Val Borbera population is a collection of 1,785 genotyped samples collected in the Val Borbera Valley, a geographically isolated valley located within the Appenine Mountains in Northwest Italy.

The Lothian Birth Cohort 1921 (LBC1921)

LBC1921 consists of 550 (234 male) relatively healthy individuals, assessed on cognitive and medical traits at a mean age of 79.1 years (SD = 0.6). They were born in 1921, most took part in the Scottish Mental Survey of 1932, and almost all lived independently in the Lothian region (Edinburgh City and surrounding area) of Scotland. Data on alcohol intake is available.

The Lothian Birth Cohort 1936 (LBC1936)

LBC1936 consists of 1091 (548 male) relatively healthy individuals who underwent cognitive and medical testing at a mean age of 69.6 years (SD = 0.8). They were born in 1936, most took part in the Scottish Mental Survey of 1947, and almost all lived independently in the Lothian region of Scotland. Data on alcohol intake is available.

The Northern Swedish Population Health Study (NSPHS)

The NSPHS was initiated in 2006 to provide a health survey of the population in the parish of Karesuando, county of Norrbotten, Sweden, and to study the medical consequences of lifestyle and genetics. This parish has about 1,500 inhabitants who meet the eligibility criteria in terms of age (≥15 years), of which 1066 individuals participated in the study.

The Orkney Complex Disease Study (ORCADES)

The Orkney Complex Disease Study (ORCADES) is a family-based study of 2078 individuals with ancestry from the isolated Scottish archipelago of Orkney. Fasting blood samples were collected and over 300 health-related phenotypes and environmental exposures were measured in each individual. All participants gave informed consent and the study was approved by Research Ethics Committees in Orkney and Aberdeen.

the PROspective Study of Pravastatin in the Elderly at Risk (PROSPER)

All data come from the PROspective Study of Pravastatin in the Elderly at Risk (PROSPER). A detailed description of the study has been published elsewhere. PROSPER was a prospective multicenter randomized placebo-controlled trial to assess whether treatment with pravastatin diminishes the risk of major vascular events in elderly. Between December 1997 and May 1999, we screened and enrolled subjects in Scotland (Glasgow), Ireland (Cork), and the Netherlands (Leiden). Men and women aged 70-82 years were recruited if they had pre-existing vascular disease or increased risk of such disease because of smoking, hypertension, or diabetes. A total number of 5,804 subjects were randomly assigned to pravastatin or placebo. A large number of prospective tests were performed including Biobank tests and cognitive function measurements.

The IMAGEN cohort

The IMAGEN consortium *(www.imagen-europe.com)*, "Reinforcement behaviour in normal brain function and psychopathology" has recruited n=2090 participants from four European countries assessed at baseline at age 14 years, with current follow-up assessments at 16 years, 19 years and one follow-up at 22 years planned in 2016. Its measures include functional neuroimaging, including tasks assessing reward processing (Monetary Incentive Delay, MID, impulsiveness (Stop Signal Reaction Time, SSRT, social-emotional processing (Emotional Faces Task, EFT and resting state, as well as structural neuroimaging, including diffusion tensor imaging. –omics characterisation comprise genome-wide genomic, gene expression and epigenetic data. Furthermore, there are detailed neuropsychological, behavioural and clinical characterisations.

Funding and Acknowledgements

Study	Funding	Acknowledgements
Colaus	The CoLaus study was and is supported by research grants from GlaxoSmithKline, the Faculty of Biology and Medicine of Lausanne, and the Swiss National Science Foundation (grants 33CSCO-122661, 33CS30-139468 and 33CS30-148401).	The authors also express their gratitude to the participants in the Lausanne CoLaus study and to the investigators who have contributed to the study in particular Gérard Waeber, Vincent Mooser and Dawn Waterworth, and the research nurses for data collection. ZK received financial support from Swiss National Science Foundation (grant no: 31003A-143914) and the Leenaards Foundation
EGCUT	NA	NA
EPIC-Norfolk	The EPIC Norfolk Study is funded by Cancer Research United Kingdom and the Medical Research Council.	We would like to thank all participants who contributed the study, and colleagues and collaborators who performed the genotyping and facilitated the analysis.
ERF	The genotyping for the ERF study was supported by EUROSPAN (European Special Populations Research Network) through 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 (NWO), Erasmus MC, the Centre for Medical Systems Biology (CMSB1 and CMSB2) and the Netherlands Genomics Initiative (NGI) and also received funding from the European Community's Seventh Framework Programme (FP7/2007-2013)/grant agreement HEALTH-F4-2007-201413 by the European Commission under the programme “Quality of Life and Management of the Living Resources” of 5th Framework Programme (no. QLG2-CT-2002-01254). High-throughput analysis of the	We are grateful to all study participants and their relatives, general practitioners and neurologists for their contributions and to P. Veraart for her help in genealogy, J. Vergeer for the supervision of the laboratory work and P. Snijders for his help in data collection.

	ERF data was supported by joint grant from Netherlands Organisation for Scientific Research and the Russian Foundation for Basic Research (NWO-RFBR 047.017.043). Exome sequencing analysis in ERF was supported by the ZonMw grant (project 91111025). Najaf Amin is supported by the Netherlands Brain Foundation (project number F2013(1)-28)Exome sequencing analysis in ERF was supported by the ZonMw grant (project 91111025).	
Fenland	Medical Research Council (MC_U106179471)	The Fenland Study is funded by the Wellcome Trust and the Medical Research Council (MC_U106179471). We are grateful to all the volunteers for their time and help, and to the General Practitioners and practice staff for assistance with recruitment. We thank the Fenland Study Investigators, Fenland Study Co-ordination team and the Epidemiology Field, Data and Laboratory teams.
FinnTwin12	Phenotyping and genotyping of the Finnish twin cohorts has been supported by the Academy of Finland Center of Excellence in Complex Disease Genetics (grants 213506, 129680), the Academy of Finland (grants 100499, 205585, 118555, 141054, 265240, 263278 and 264146 to J. Kaprio), National Institute of Alcohol Abuse and Alcoholism (grants AA-12502, AA-00145, and AA-09203 to R. J. Rose and AA15416 and K02AA018755 to D. M. Dick), and the Welcome Trust Sanger Institute, UK. Antti-Pekka Sarin and Samuli Ripatti are acknowledged for genotype data quality controls and imputation.	We warmly thank the participating twin pairs and their family members for their contribution. We would like to express our appreciation to the skilled study interviewers A-M Iivonen, K Karhu, H-M Kuha, U Kulmala-Gråhn, M Mantere, K Saanakorpi, M Saarinen, R Sipilä, L Viljanen and E Voipio. Anja Häppölä and Kauko Heikkilä are acknowledged for their valuable contribution in recruitment, data collection, and data management.
FinnTwinOld_1	Phenotyping and genotyping of the Finnish twin cohorts has been supported by the Academy of Finland Center of Excellence in Complex Disease Genetics (grants 213506,	We warmly thank the participating twin pairs and their family members for their contribution. We would like to express our appreciation to the

	129680), the Academy of Finland (grants 100499, 205585, 118555, 141054, 265240, 263278 and 264146 to J. Kaprio), Global Research Award for Nicotine Dependence, Pfizer Inc. (to J. Kaprio), and the Welcome Trust Sanger Institute, UK. Antti-Pekka Sarin and Samuli Ripatti are acknowledged for genotype data quality controls and imputation.	skilled study interviewers A-M Iivonen, K Karhu, H-M Kuha, U Kulmala-Gråhn, M Mantere, K Saanakorpi, M Saarinen, R Sipilä, L Viljanen and E Voipio. Anja Häppölä and Kauko Heikkilä are acknowledged for their valuable contribution in recruitment, data collection, and data management.
FinnTwinOld_2	The Finnish Twin Cohort study received financial support from the Academy of Finland Center of Excellence in Complex Disease Genetics, ENGAGE project and grant agreement HEALTH-F4-2007-201413 and the GenomeEUtwin project, which was supported by the European Commission under the program 'Quality of Life and Management of the Living Resources' of 5th Framework Program (no. QLG2-CT-2002-01254). The DNA extractions, sample quality controls, biobank up-keep and aliquotting was performed in the National Public Health Institute, Helsinki, Finland.	We warmly thank the participating twin pairs and their family members for their contribution. We would like to express our appreciation to the skilled study interviewers A-M Iivonen, K Karhu, H-M Kuha, U Kulmala-Gråhn, M Mantere, K Saanakorpi, M Saarinen, R Sipilä, L Viljanen and E Voipio. Anja Häppölä and Kauko Heikkilä are acknowledged for their valuable contribution in recruitment, data collection, and data management.
FinnTwinOld_3	Phenotyping and genotyping of the Finnish twin cohorts has been supported by the Academy of Finland Center of Excellence in Complex Disease Genetics (grants 213506, 129680), the Academy of Finland (grants 100499, 205585, 118555, 141054, 265240, 263278 and 264146 to J. Kaprio), and the Welcome Trust Sanger Institute, UK. Antti-Pekka Sarin and Samuli Ripatti are acknowledged for genotype data quality controls and imputation.	We warmly thank the participating twin pairs and their family members for their contribution. We would like to express our appreciation to the skilled study interviewers A-M Iivonen, K Karhu, H-M Kuha, U Kulmala-Gråhn, M Mantere, K Saanakorpi, M Saarinen, R Sipilä, L Viljanen and E Voipio. Anja Häppölä and Kauko Heikkilä are acknowledged for their valuable contribution in recruitment, data collection, and data management.
HBCS	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	We thank all study participants as well as everybody involved in the Helsinki Birth Cohort Study.

	Helsinki, Ministry of Education, Ahokas Foundation, Emil Aaltonen Foundation.	
KORA F3 and F4	The KORA study was initiated and financed by the Helmholtz Zentrum München – German Research Center for Environmental Health, which is funded by the German Federal Ministry of Education and Research (BMBF) and by the State of Bavaria. Furthermore, KORA research was supported within the Munich Center of Health Sciences (MC-Health), Ludwig-Maximilians-Universität, as part of LMUinnovativ.	NA
Lifelines	NA	NA
LOLIPOP	The LOLIPOP study is funded by the British Heart Foundation (SP/04/002), the Medical Research Council (G0601966, G0700931), the Wellcome Trust (084723/Z/08/Z), the NIHR (RP-PG-0407-10371), European Union FP7 (EpiMigrant, 279143) and Action on Hearing Loss (G51).	The LOLIPOP study is supported by the National Institute for Health Research (NIHR) Comprehensive Biomedical Research Centre Imperial College Healthcare NHS Trust. The work was carried out in part at the NIHR/Wellcome Trust Imperial Clinical Research Facility. We thank the participants and research staff who made the study possible.
NESDA	Funding was obtained from the Netherlands Organization for Scientific Research (Geestkracht program grant 10-000-1002); the Center for Medical Systems Biology (CSMB, NWO Genomics), Biobanking and Biomolecular Resources Research Infrastructure (BBMRI-NL), VU University's Institutes for Health and Care Research (EMGO+) and Neuroscience Campus Amsterdam, University Medical Center Groningen, Leiden University Medical Center, National Institutes of Health (NIH, R01D0042157-01A, MH081802, Grand Opportunity grants 1RC2 MH089951 and 1RC2 MH089995). Part of the genotyping and analyses were funded by the Genetic Association Information Network (GAIN) of the Foundation for the National Institutes of Health.	NA

	Computing was supported by BiG Grid, the Dutch e-Science Grid, which is financially supported by NWO	
NFBC1966	NFBC1966 received financial support from the Academy of Finland (project grants 104781, 120315, 129269, 1114194, 24300796, Center of Excellence in Complex Disease Genetics and SALVE), University Hospital Oulu, Biocenter, University of Oulu, Finland (75617), NHLBI grant 5R01HL087679-02 through the STAMPEED program (1RL1MH083268-01), NIH/NIMH (5R01MH63706:02), ENGAGE project and grant agreement HEALTH-F4-2007-201413, EU FP7 EurHEALTHAgeing -277849, the Medical Research Council, UK (G0500539, G0600705, G1002319, PrevMetSyn/SALVE) and the MRC, Centenary Early Career Award. The program is currently being funded by the H2020-633595 DynaHEALTH action and academy of Finland EGEA-project (285547). The DNA extractions, sample quality controls, biobank up-keeping and aliquotting was performed in the National Public Health Institute, Biomedicum Helsinki, Finland and supported financially by the Academy of Finland and Biocentrum Helsinki.	We thank the late Professor Paula Rantakallio (launch of NFBCs), and Ms Outi Tornwall and Ms Minttu Jussila (DNA biobanking). The authors would like to acknowledge the contribution of the late Academian of Science Leena Peltonen.
NTR	Netherlands Scientific Organization (NWO); Grant number: 480-05-003; Netherlands Organization for Scientific Research; Grant numbers: ZonMW (Addiction program) 31160008, ZonMW 940-37-024, NWO/SPI 56-464-14192, NWO-400-05-717, NWO-MW 904-61-19, NWOMagW 480-04-004, European Research Council ERC-230374 and ERC-284167; Centre for Medical Systems Biology (NWO Genomics); Netherlands Bioinformatics Center/BioAssist/RK/2008.024; National Institute of Mental Health (NIH); MH081802, 1RC2 MH089951, 1RC2 MH089995; US National Institute of	NA

	Mental Health; Grant number: RC2 MH089951	
OZALC	Supported by National Institutes of Health grants AA07535,AA07728, AA13320, AA13321, AA14041, AA11998, AA17688,DA012854, and DA019951; by grants from the Australian National Health and Medical Research Council (241944, 339462, 389927,389875, 389891, 389892, 389938, 442915, 442981, 496739, 552485, and 552498) and the Australian Research Council(A7960034, A79906588, A79801419, DP0770096, DP0212016, and DP0343921); and by the 5th Framework Programme (FP-5) GenomEUtwin Project (QLG2-CT-2002-01254). Genome-wide genotyping at the Center for Inherited Disease Research was supported by a grant to the late Richard Todd, M.D., Ph.D. GWM is supported by the National Health and Medical Research Council Fellowship Scheme.	We acknowledge the contribution of the participants in this study, and the work of the interviewers and laboratory staff at QIMR Berghofer Medical Research Institute in obtaining the phenotype information and genotypes. We also acknowledge the contributions of Dr Grant Montgomery and Ms Anjali Henders to project management, sample acquisition and genotyping. Aspects of the data collection and genotyping were made possible by grants to Dr Pamela AF Madden and to the late Dr Richard Todd.
PREVEND	PREVEND genetics is supported by the Dutch Kidney Foundation (Grant E033), the EU project grant GENECURE (FP-6 LSHM CT 2006 037697), the National Institutes of Health (grant 2R01LM010098), The Netherlands organisation for health research and development (NWO-Groot grant 175.010.2007.006, NWO VENI grant 916.761.70, ZonMw grant 90.700.441), and the Dutch Inter University Cardiology Institute Netherlands (ICIN). N.Verweij is supported by ICIN-NHI and Marie Sklodowska-Curie GF (call: H2020-MSCA-IF-2014, Project ID: 661395)	NA

SHIP	SHIP is part of the Community Medicine Research net of the University of Greifswald, Germany, which is funded by the Federal Ministry of Education and Research (grants no. 01ZZ9603, 01ZZ0103, and 01ZZ0403), the Ministry of Cultural Affairs as well as the Social Ministry of the Federal State of Mecklenburg-West Pomerania, and the network 'Greifswald Approach to Individualized Medicine (GANI_MED)' funded by the Federal Ministry of Education and Research (grant 03IS2061A). 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 Caché Campus program of the InterSystems GmbH.	NA
TwinsUK	The study was funded by the Wellcome Trust; European Community's Seventh Framework Programme (FP7/2007-2013). The study also receives support from the National Institute for Health Research (NIHR)-funded BioResource, Clinical Research Facility and Biomedical Research Centre based at Guy's and St Thomas' NHS Foundation Trust in partnership with King's College London.	The authors are extremely grateful to all the twins who took part in this study, the midwives for recruiting them and the whole TwinsUK team, which includes interviewers, com-puter and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses. SNP Genotyping was performed by The Wellcome Trust Sanger Institute and National Eye Institute via NIH/CIDR
YFS	The Young Finns Study has been financially supported by the Academy of Finland: grants 286284, 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; Juho Vainio	The expert technical assistance in the statistical analyses by Ville Aalto and Irina Lisinen is gratefully acknowledged.

	Foundation; Paavo Nurmi Foundation; Finnish Foundation of Cardiovascular Research; Finnish Cultural Foundation; Tampere Tuberculosis Foundation; Emil Aaltonen Foundation; and Yrjö Jahnsson Foundation.	
AGES	This study has been funded by NIH contracts N01-AG-1-2100 and 271201200022C, the NIA Intramural Research Program, Hjartavernd (the Icelandic Heart Association), and the Althingi (the Icelandic Parliament). The study is approved by the Icelandic National Bioethics Committee, VSN: 00-063.	The researchers are indebted to the participants for their willingness to participate in the study.
ARIC	The Atherosclerosis Risk in Communities 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.	NA

CHS	Infrastructure for the CHARGE Consortium is supported in part by the National Heart, Lung, and Blood Institute grant R01HL105756. This CHS research was supported by NHLBI contracts HHSN268201200036C, HHSN268200800007C, N01HC55222, N01HC85079, N01HC85080, N01HC85081, N01HC85082, N01HC85083, N01HC85086; and NHLBI grants U01HL080295, R01HL087652, R01HL105756, R01HL103612, and R01HL120393 with additional contribution from the National Institute of Neurological Disorders and Stroke (NINDS). Additional support was provided through R01AG023629 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. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.	NA
FHS	The Framingham Heart Study is funded by National Institutes of Health contract N01-HC-25195. The laboratory work for this investigation was funded by the Division of Intramural Research, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, MD. The analytical component of this project was funded by the Division of Intramural Research, National Heart, Lung, and Blood Institute, and the Center for Information Technology, National Institutes of	We thank Dr. Curtis Ellison at Boston University School of Medicine for defining phenotypes

	Health, Bethesda, MD.	
HABC	Health Aging and Body Composition Study (Health ABC): This study was funded by the National Institutes of Aging. This research was supported by NIA contracts N01AG62101, N01AG62103, and N01AG62106. 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.	NA
MESA	MESA and the MESA SHARe project are conducted and supported by the National Heart, Lung, and Blood Institute (NHLBI) in collaboration with MESA investigators. Support for MESA is provided by contracts N01-HC-95159, N01-HC-95160, N01-HC-95161, N01-HC-95162, N01-HC-95163, N01-HC-95164, N01-HC-95165, N01-HC-95166, N01-HC-95167, N01-HC-95168, N01-HC-95169, UL1-TR-001079, and UL1-TR-000040. Funding for SHARe genotyping was provided by NHLBI Contract N02-HL-64278. Genotyping was performed at Affymetrix (Santa Clara, California, USA) and the Broad Institute of Harvard and MIT (Boston, Massachusetts, USA) using the Affymetrix Genome-Wide Human SNP Array 6.0. 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.	NA

RS 1, 2 and 3	The Rotterdam Study is funded by Erasmus Medical Center and Erasmus University, Rotterdam, Netherlands Organization for the Health Research and Development (ZonMw), the Research Institute for Diseases in the Elderly (RIDE), the Ministry of Education, Culture and Science, the Ministry for Health, Welfare and Sports, the European Commission (DG XII), and the Municipality of Rotterdam. The generation and management of GWAS genotype data for the Rotterdam Study (RS I, RS II, RS III) was executed by the Human Genotyping Facility of the Genetic Laboratory of the Department of Internal Medicine, Erasmus MC, Rotterdam, The Netherlands. The GWAS datasets are supported by the Netherlands Organisation of Scientific Research NWO Investments (nr. 175.010.2005.011, 911-03-012), the Genetic Laboratory of the Department of Internal Medicine, Erasmus MC, the Research Institute for Diseases in the Elderly (014-93-015; RIDE2), the Netherlands Genomics Initiative (NGI)/Netherlands Organisation for Scientific Research (NWO) Netherlands Consortium for Healthy Aging (NCHA), project nr. 050-060-810 .	The authors are grateful to the study participants, the staff from the Rotterdam Study and the participating general practitioners and pharmacists. We thank Pascal Arp, Mila Jhamai, Marijn Verkerk, Lizbeth Herrera and Marjolein Peters, MSc, and Carolina Medina-Gomez, MSc, for their help in creating the GWAS database, and Karol Estrada, PhD, Yurii Aulchenko, PhD, and Carolina Medina-Gomez, MSc, for the creation of imputed data.
WGHS	The WGHS is supported by HL043851 and HL080467 from the National Heart, Lung, and Blood Institute and CA047988 from the National Cancer Institute, with collaborative scientific support and funding for genotyping provided by Amgen.	NA
Airwave	The study is funded by the Home Office (Grant number 780-TETRA) with additional support from the National Institute for Health Research (NIHR), Imperial College Healthcare NHS Trust (ICHNT) and Imperial College Biomedical Research Centre (BRC) (Grant number BRC-P38084). P.E. is supported by the ICHNT and Imperial College BRC, the MRC-PHE Centre for Environment	We thank all participants in the Airwave Health Monitoring Study. We also thank Louisa Cavaliero who assisted in data collection and management as well as Peter McFarlane and the Glasgow CARE, Patricia Munroe at Queen Mary University of London, Joanna Sarnecka and Ania Zawodniak at Northwick Park for their

	and Health, the NIHR Health Protection Research Unit on Health Impact of Environmental Hazards and is an NIHR Senior Investigator.	contributions to the study.
ASPS	The research reported in this article was funded by the Austrian Science Fond (FWF) grant number P20545-P05 and P13180. The Medical University of Graz supports the databank of the ASPS.	The authors thank the staff and the participants for their valuable contributions. We thank Birgit Reinhart for her long-term administrative commitment, Elfi Hofer for the technical assistance at creating the DNA bank, Ing. Johann Semmler and Anita Harb for DNA sequencing and DNA analyses by TaqMan assays and Irmgard Poelzl for supervising the quality management processes after ISO9001 at the biobanking and DNA analyses
B58C	We acknowledge use of phenotype and genotype data from the British 1958 Birth Cohort DNA collection, funded by the Medical Research Council grant G0000934 and the Wellcome Trust grant 068545/Z/02. Genotyping for the B58C-WTCCC subset was funded by the Wellcome Trust grant 076113/B/04/Z. The B58C-T1DGC genotyping utilized 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 (NIAID), National Human Genome Research Institute (NHGRI), National Institute of Child Health and Human Development (NICHD), and Juvenile Diabetes Research Foundation International (JDRF) and supported by U01 DK062418. B58C-T1DGC GWAS data were deposited by the Diabetes and Inflammation Laboratory, Cambridge Institute for Medical Research (CIMR), University of Cambridge, which is funded by Juvenile Diabetes Research Foundation International, the	NA

	Wellcome Trust and the National Institute for Health Research Cambridge Biomedical Research Centre; the CIMR is in receipt of a Wellcome Trust Strategic Award (079895). The B58C-GABRIEL genotyping was supported by a contract from the European Commission Framework Programme 6 (018996) and grants from the French Ministry of Research.	
DESIR	This work was supported by grants from the ‘‘Agence Nationale de la Recherche,’’ the ‘‘Conseil Regional Nord-Pas de Calais/Fonds europeen de dveloppement economique et r´egional,’’ Genome Quebec/Genome Canada and the Medical Research Council	We thank the subjects and families who participated in this study. We thank F. Allegaert, M. Deweirder for their technical support in DNA extraction and distribution.
FinnTwin_replication	Phenotyping and genotyping of the Finnish twin cohorts has been supported by the Academy of Finland Center of Excellence in Complex Disease Genetics (grants 213506, 129680), the Academy of Finland (grants 100499, 205585, 118555, 141054, 265240, 263278 and 264146 to J. Kaprio), National Institute of Alcohol Abuse and Alcoholism (grants AA-12502, AA-00145, and AA-09203 to R. J. Rose and AA15416 and K02AA018755 to D. M. Dick), Sigrid Juselius Foundation (to J. Kaprio), Global Research Award for Nicotine Dependence, Pfizer Inc. (to J. Kaprio), the Welcome Trust Sanger Institute, UK, and the Broad Institute, USA. Antti-Pekka Sarin and Samuli Ripatti are acknowledged for genotype data quality controls and imputation.	We warmly thank the participating twin pairs and their family members for their contribution. We would like to express our appreciation to the skilled study interviewers A-M Iivonen, K Karhu, H-M Kuha, U Kulmala-Gråhn, M Mantere, K Saanakorpi, M Saarinen, R Sipilä, L Viljanen and E Voipio. Anja Häppölä and Kauko Heikkilä are acknowledged for their valuable contribution in recruitment, data collection, and data management.
GRAPHIC	The GRAPHIC Study was funded by the British Heart Foundation (BHF). CPN and NJS are both funded by the BHF and NJS is a NIHR Senior Investigator	We would like to thank all the participants in the study for their contribution and support.
GS:SFHS	Generation Scotland has received core funding from the Chief Scientist Office of the Scottish Government Health Directorates CZD/16/6 and the Scottish Funding Council HR03006. Genotyping of the GS:SFHS samples was	We are grateful to all the families who took part, the general practitioners and the Scottish School of Primary Care for their help in recruiting them, and the whole Generation Scotland team, which

	carried out by the Genetics Core Laboratory at the Wellcome Trust Clinical Research Facility, Edinburgh, Scotland and was funded by the UK's Medical Research Council and the Wellcome Trust. Ethics approval for the study was given by the NHS Tayside committee on research ethics (reference 05/S1401/89). financial support received for this work from the Dr Mortimer and Theresa Sackler Foundation and the MRC "QTL in Health and Disease" core programme.	includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists, healthcare assistants and nurses.
INGI_CARL	The INGI-FVG study was founded through the Italian Ministry of health	We would like to thank all the participants in the study for their contribution and support.
INGI_FVG	The INGI-CARL study was founded through the Italian Ministry of health	We would like to thank all the participants in the study for their contribution and support.
INGI_VB	The research was supported by funds from Compagnia di San Paolo, Torino, Italy; Fondazione Cariplo, Italy; Telethon Italy; Ministry of Health, Ricerca Finalizzata 2008 and 2011-2012 and Public Health Genomics Project 2010.	We thank all the participants to the project, the San Raffaele Hospital MDs who contributed to clinical data collection, prof. Clara Camaschella who coordinated the data collection, Corrado Masciullo and Massimiliano Cocca for help in the database informatics.
LBC1921	Phenotype collection in the LBC1921 was supported by the BBSRC, The Royal Society, and The Chief Scientist Office of the Scottish Government. Genotyping was funded by the BBSRC. The work was undertaken by The University of Edinburgh Centre for Cognitive Ageing and Cognitive Epidemiology, part of the cross council Lifelong Health and Wellbeing Initiative (MR/K026992/1). Funding from the BBSRC and MRC is gratefully acknowledged.	We thank the cohort participants and team members who contributed to these studies.
LBC1936	Phenotype collection in the LBC1936 was supported by Age UK (The Disconnected Mind project). Genotyping of the cohorts was funded by the BBSRC. The work was undertaken by The University of Edinburgh Centre for	We thank the cohort participants and team members who contributed to these studies.

	Cognitive Ageing and Cognitive Epidemiology, part of the cross council Lifelong Health and Wellbeing Initiative (MR/K026992/1). Funding from the BBSRC and MRC is gratefully acknowledged.	
NSPHS	The Northern Swedish Population Health Study (NSPHS) was funded by the Swedish Medical Research Council (Project Number K2007-66X-20270-01-3, 2011-5252, 2012-2884 and 2011-2354), the Foundation for Strategic Research (SSF). NSPHS as part of EUROSPAN (European Special Populations Research Network) was also supported by the European Commission FP6 STRP grant number 01947 (LSHG-CT-2006-01947). This work has also been supported by the Swedish Society for Medical Research (SSMF), the Kjell och Märta Beijers Foundation, the Marcus Borgström Foundation, the Åke Wiberg foundation and the Vleugels Foundation.	We are grateful for the contribution of district nurse Svea Hennix for data collection and Inger Jonasson for logistics and coordination of the health survey. We also thank all the participants from the community for their interest and willingness to contribute to this study. Illumina genotyping was performed by the SNP&SEQ Technology Platform in Uppsala, Sweden. The computations were performed on resources provided by SNIC through Uppsala Multidisciplinary Center for Advanced Computational Science (UPPMAX) under projects b2011203.
ORCADES	ORCADES was supported by the Chief Scientist Office of the Scottish Government, the Royal Society, the MRC Human Genetics Unit, Arthritis Research UK and the European Union framework program 6 EUROSPAN project (contract no. LSHG-CT-2006-018947).	DNA extractions were performed at the Wellcome Trust Clinical Research Facility in Edinburgh. We would like to acknowledge the invaluable contributions of Lorraine Anderson and the research nurses in Orkney, the administrative team in Edinburgh and the people of Orkney.
PROSPER	The PROSPER study was supported by an investigator initiated grant obtained from Bristol-Myers Squibb. Prof. Dr. J. W. Jukema is an Established Clinical Investigator of the Netherlands Heart Foundation (grant 2001 D 032). Support for genotyping was provided by the seventh framework program of the European commission (grant 223004) and by the Netherlands Genomics Initiative (Netherlands Consortium for Healthy Aging grant 050-060-810).	NA

IMAGEN	This work received support from the following sources: the European Union-funded FP6 Integrated Project IMAGEN (Reinforcement-related behaviour in normal brain function and psychopathology) (LSHM-CT- 2007-037286), the FP7 projects IMAGEMEND(602450; IMAging GEnetics for MENtal Disorders), AGGRESSOTYPE (602805) and MATRICS (603016), the Innovative Medicine Initiative Project EU-AIMS (115300-2), the Medical Research Council Grants ‘‘Developmental pathways into adolescent substance abuse’’ (93558) and Consortium on Vulnerability to Externalizing Disorders and Addictions [c-VEDA] (MR/N000390/1), the Swedish funding agencies VR, FORTE and FORMAS, the Medical Research Council and the Wellcome Trust (Behavioural and Clinical Neuroscience Institute, University of Cambridge), the National Institute for Health Research (NIHR) Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King’s College London, the Bundesministeriumfür Bildung und Forschung (BMBF grants 01GS08152; 01EV0711; eMED SysAlc01ZX1311A; Forschungsnetz AERIAL), the Deutsche Forschungsgemeinschaft (DFG grants SM 80/7-1, SM 80/7-2, SFB 940/1). Further support was provided by grants from: ANR (project AF12-NEUR0008-01 - WM2NA, and ANR-12-SAMA-0004), the Fondation de France, the Fondation pour la Recherche Médicale, the Mission Interministérielle de Lutte-contre-les-Drogues-et-les-Conduites-Addictives (MILDECA), the Assistance-Publique-Hôpitaux-de-Paris and INSERM (interface grant), Paris-Sud University Paris Saclay IDEX 2012; the National Institutes of Health, U.S.A. (Axon, Testosterone and Mental Health	NA

	during Adolescence; RO1 MH085772-01A1), and by NIH Consortium grant U54 EB020403, supported by a cross-NIH alliance that funds Big Data to Knowledge Centres of Excellence.	
Animal study (*KLB*)	This work was supported by National Institutes of Health grants R01DK067158 (S.A.K. and D.J.M.); the Robert A. Welch Foundation (grant I-1558 to S.A.K. and grant I-1275 to D.J.M.); and the Howard Hughes Medical Institute (D.J.M.).	We thank Yuan Zhang for assistance on animal experiments.

Author contributions

Gunter Schumann, Paul Elliott, Dan Levy, Christian P. Müller, David Mangelsdorf and Steven Kliewer designed the study, acquired and analysed data, and wrote the manuscript; Chunyu Liu, Paul F. O'Reilly, He Gao and Envangelos Evangelou analysed GWAS data and contributed to writing the manuscript; Bing Xu, Parkyong Song, Barbara Ruggeri and Sylvane Desrivières carried out functional analyses, including neuroimaging, animal and epigenetic analyses, and contributed to writing the manuscript; Georgy Bakalkin and Yun Liu acquired and analysed epigenetic data; Najaf Amin, Tianye Jia, Sarah R. Preis, Marcelo P. Segura Lepe, and Kenneth Rice analysed GWAS data; The IMAGEN consortium contributed the IMAGEN data. The following authors contributed to the primary GWAS and replication by participating in **(i) study concept/design:** Dorret I. Boomsma, John C. Chambers, Daniel I. Chasman, Toniolo Daniela, Ian J. Deary, Eco J.C. de Geus, Johan G. Eriksson, Tõnu Esko, Oscar H. Franco, Philippe Froguel, Christian Gieger, Hans J. Grabe, Vilmundur Gudnason, Ulf Gyllensten, Tamara B. Harris, Anna-Liisa Hartikainen, Andrew C. Heath, Albert Hofman, Cornelia Huth, Marjo-Riitta Jarvelin, J. Wouter Jukema, Jaakko Kaprio, Jaspal S. Kooner, Jaana Laitinen, Claudia Langenberg, Terho Lehtimäki, Daniel Levy, Yongmei Liu, Pamela A.F. Madden, Nicholas G. Martin, Alanna C. Morrison, Jennifer A. Nettleton, Brenda W.J.H. Penninx, Nicola Pirastu, Bruce M. Psaty, Olli T. Raitakari, Paul M. Ridker, Richard J. Rose, Jerome I. Rotter, Nilesh J. Samani, Helena Schmidt, Reinhold Schmidt, Tim Spector, John M. Starr, David J. Stott, David P. Strachan, Stella Trompet, Ioanna Tzoulaki, Pim van der Harst, Cornelia M. van Duijn, Peter Vollenweider, Nick Wareham, James F. Wilson, Bruce H.R. Wolffenbuttel; **(ii) data acquisition:** Sebastian E. Baumeister, Dorret I. Boomsma, John C. Chambers, Daniel I. Chasman, Ian J. Deary, Eco J.C. de Geus, Ulf Gyllensten, Tõnu Esko, Oscar H. Franco, Philippe Froguel, Vilmundur Gudnason, Sarah E. Harris, Tamara B. Harris, Anna-Liisa Hartikainen, Andrew C. Heath, Lynne J. Hocking, Albert Hofman, Cornelia Huth, Marjo-Riitta Jarvelin, J. Wouter Jukema, Niina Kaartinen, Jaakko Kaprio, Jaspal S. Kooner, Jaana Laitinen, Claudia Langenberg, Terho Lehtimäki, Daniel Levy, Yongmei Liu, Pamela A.F. Madden, Nicholas G. Martin, Alanna C. Morrison, Kenneth Mukamal, Jennifer A. Nettleton, Brenda W.J.H. Penninx, Bruce M. Psaty, Olli T. Raitakari, Paul M. Ridker, Richard J. Rose, Jerome I. Rotter, Cinzia Sala, Nilesh J. Samani, Helena Schmidt, Reinhold Schmidt, John M. Starr, David J. Stott, David P. Strachan, Ioanna Tzoulaki, Pim van der Harst, Cornelia M. van Duijn, André G. Uitterlinden, Cristina Venturini, Veronique Vitart, Peter Vollenweider, Nick Wareham, John B. Whitfield, James F. Wilson, Bruce H.R. Wolffenbuttel; **(iii) data analysis:** Najaf Amin, Caterina Barbieri, Sebastian E. Baumeister, Stephane Cauchi, Toni-Kim Clarke, Stefan Enroth, Krista Fischer, Christian Gieger, Jenni Hällfors, Sarah E. Harris, Andrew C. Heath, Edith Hofer, Jouke-Jan Hottenga, Åsa Johansson, Peter K. Joshi, Zoltan Kutalik, Jari Lahti, Rozenn N. Lemaitre, Chunyu Liu, Yongmei Liu, Anu Loukola, Jian'an Luan, Leo-Pekka Lyytikäinen, Massimo Mangino, Ani Manichaikul, Nicholas G. Martin, Hamdi Mbarek, Yuri Milaneschi, Alireza Moayyeri, Christopher P. Nelson, Jennifer A. Nettleton, Eemil Partinen, Nicola Pirastu, Sara R. Preis, Rajesh Rawal, Antonietta Robino, Lynda M. Rose, Katharina E. Schraut, Robert Scott, Albert Vernon Smith, David P. Strachan, Alexander Teumer, Stella Trompet, Anne-Claire Vergnaud, Niek Verweij,

Pedro Marques Vidal, Veronique Vitart, Dragana Vuckovic, Juho Wedenoja, John B. Whitfield, Loic Yengo, Bing Yu, Weihua Zhang, Jinghua Zhao.

References for Supplementary Note

1 Aulchenko, Y. S. *et al.* Linkage disequilibrium in young genetically isolated Dutch population. *Eur J Hum Genet* **12**, 527-534, doi:10.1038/sj.ejhg.5201188 (2004).

2 Chambers, J. C. *et al.* Common genetic variation near melatonin receptor MTNR1B contributes to raised plasma glucose and increased risk of type 2 diabetes among Indian Asians and European Caucasians. *Diabetes* **58**, 2703-2708, doi:10.2337/db08-1805 (2009).

3 Chambers, J. C. *et al.* Common genetic variation near MC4R is associated with waist circumference and insulin resistance. *Nat Genet* **40**, 716-718, doi:10.1038/ng.156 (2008).

4 Penninx, B. W. *et al.* The Netherlands Study of Depression and Anxiety (NESDA): rationale, objectives and methods. *Int J Methods Psychiatr Res* **17**, 121-140, doi:10.1002/mpr.256 (2008).

5 Boomsma, D. I. *et al.* Genome-wide association of major depression: description of samples for the GAIN Major Depressive Disorder Study: NTR and NESDA biobank projects. *Eur J Hum Genet* **16**, 335-342, doi:10.1038/sj.ejhg.5201979 (2008).

6 Willemsen, G. *et al.* The Netherlands Twin Register biobank: a resource for genetic epidemiological studies. *Twin Res Hum Genet* **13**, 231-245, doi:10.1375/twin.13.3.231 (2010).

7 Heath, A. C. *et al.* A quantitative-trait genome-wide association study of alcoholism risk in the community: findings and implications. *Biol Psychiatry* **70**, 513-518, doi:10.1016/j.biopsych.2011.02.028 (2011).

8 John, U. *et al.* Study of Health In Pomerania (SHIP): a health examination survey in an east German region: objectives and design. *Soz Praventivmed* **46**, 186-194 (2001).

9 Volzke, H. *et al.* Cohort profile: the study of health in Pomerania. *Int J Epidemiol* **40**, 294-307, doi:10.1093/ije/dyp394 (2011).

10 Baumeister, S. E. *et al.* Alcohol consumption and out-patient services utilization by abstainers and drinkers. *Addiction* **101**, 1285-1291, doi:10.1111/j.1360-0443.2006.01538.x (2006).

11 Harris, T. B. *et al.* Age, Gene/Environment Susceptibility-Reykjavik Study: multidisciplinary applied phenomics. *Am J Epidemiol* **165**, 1076-1087, doi:10.1093/aje/kwk115 (2007).

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

13 Fried, L. P. *et al.* The Cardiovascular Health Study: design and rationale. *Ann Epidemiol* **1**, 263-276 (1991).

14 Kannel, W. B., Feinleib, M., McNamara, P. M., Garrison, R. J. & Castelli, W. P. An investigation of coronary heart disease in families. The Framingham offspring study. *Am J Epidemiol* **110**, 281-290 (1979).

15 Splansky, G. L. *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* **165**, 1328-1335, doi:10.1093/aje/kwm021 (2007).

16 Cesari, M. *et al.* Inflammatory markers and cardiovascular disease (The Health, Aging and Body Composition [Health ABC] Study). *Am J Cardiol* **92**, 522-528 (2003).

17 Bild, D. E. *et al.* Multi-ethnic study of atherosclerosis: objectives and design. *Am J Epidemiol* **156**, 871-881 (2002).

18 Mayer-Davis, E. J. *et al.* Validity and reproducibility of a food frequency interview in a Multi-Cultural Epidemiology Study. *Ann Epidemiol* **9**, 314-324 (1999).
19 Nettleton, J. A., Rock, C. L., Wang, Y., Jenny, N. S. & Jacobs, D. R. 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* **102**, 1220-1227, doi:10.1017/S0007114509382161 (2009).
20 Hofman, A. *et al.* The Rotterdam Study: 2014 objectives and design update. *Eur J Epidemiol* **28**, 889-926, doi:10.1007/s10654-013-9866-z (2013).
21 Ridker, P. M. *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* **54**, 249-255, doi:10.1373/clinchem.2007.099366 (2008).
22 Elliott, P. *et al.* The Airwave Health Monitoring Study of police officers and staff in Great Britain: rationale, design and methods. *Environ Res* **134**, 280-285, doi:10.1016/j.envres.2014.07.025 (2014).
23 Schmidt, R., Fazekas, F., Kapeller, P., Schmidt, H. & Hartung, H. P. MRI white matter hyperintensities: three-year follow-up of the Austrian Stroke Prevention Study. *Neurology* **53**, 132-139 (1999).
24 Schmidt, R. *et al.* Assessment of cerebrovascular risk profiles in healthy persons: definition of research goals and the Austrian Stroke Prevention Study (ASPS). *Neuroepidemiology* **13**, 308-313 (1994).
25 Strachan, D. P. *et al.* Lifecourse influences on health among British adults: effects of region of residence in childhood and adulthood. *Int J Epidemiol* **36**, 522-531, doi:10.1093/ije/dyl309 (2007).
26 Smith, B. H. *et al.* Cohort Profile: Generation Scotland: Scottish Family Health Study (GS:SFHS). The study, its participants and their potential for genetic research on health and illness. *Int J Epidemiol* **42**, 689-700, doi:10.1093/ije/dys084 (2013).