

Meta-analysis of Genome-Wide Association Studies for Extraversion: Findings from the Genetics of Personality Consortium

Stéphanie M. van den Berg¹ · Marleen H. M. de Moor^{2,3,4} · Karin J. H. Verweij^{5,6} · Robert F. Krueger⁷ · Michelle Luciano^{8,9} · Alejandro Arias Vasquez^{10,11,12,13} · Lindsay K. Matteson⁷ · Jaime Derringer¹⁴ · Tõnu Esko¹⁵ · Najaf Amin¹⁶ · Scott D. Gordon⁵ · Narelle K. Hansell⁵ · Amy B. Hart¹⁷ · Ilkka Seppälä¹⁸ · Jennifer E. Huffman¹⁹ · Bettina Konte²⁰ · Jari Lahti^{21,22} · Minyoung Lee²³ · Mike Miller⁷ · Teresa Nutile²⁴ · Toshiko Tanaka²⁵ · Alexander Teumer²⁶ · Alexander Viktorin²⁷ · Juho Wedenoja²⁸ · Abdel Abdellaoui² · Goncalo R. Abecasis²⁹ · Daniel E. Adkins³⁰ · Arpana Agrawal³¹ · Jüri Allik^{32,33} · Katja Appel³⁴ · Timothy B. Bigdeli²³ · Fabio Busonero³⁵ · Harry Campbell³⁶ · Paul T. Costa³⁷ · George Davey Smith³⁸ · Gail Davies^{8,9} · Harriet de Wit³⁹ · Jun Ding⁶⁷ · Barbara E. Engelhardt⁴⁰ · Johan G. Eriksson^{22,41,42,43,44} · Iryna O. Fedko² · Luigi Ferrucci²⁵ · Barbara Franke^{10,11,12} · Ina Giegling²⁰ · Richard Grucza³¹ · Annette M. Hartmann²⁰ · Andrew C. Heath³¹ · Kati Heinonen²¹ · Anjali K. Henders⁵ · Georg Homuth⁴⁵ · Jouke-Jan Hottenga² · William G. Iacono⁷ · Joost Janzing¹¹ · Markus Jokela²¹ · Robert Karlsson²⁷ · John P. Kemp^{38,46} · Matthew G. Kirkpatrick³⁹ · Antti Latvala^{41,28} · Terho Lehtimäki¹⁸ · David C. Liewald^{8,9} · Pamela A. F. Madden³¹ · Chiara Magri⁴⁷ · Patrik K. E. Magnusson²⁷ · Jonathan Marten¹⁹ · Andrea Maschio³⁵ · Hamdi Mbarek² · Sarah E. Medland⁵ · Evelin Mihailov^{15,48} · Yuri Milaneschi⁴⁹ ·

Edited by Stacey Cherny.

Stéphanie M. van den Berg and Marleen H. M. de Moor shared first authorship.

Electronic supplementary material The online version of this article (doi:[10.1007/s10519-015-9735-5](https://doi.org/10.1007/s10519-015-9735-5)) contains supplementary material, which is available to authorized users.

✉ Stéphanie M. van den Berg
stephanie.vandenberg@utwente.nl

¹ Department of Research Methodology, Measurement and Data-Analysis (OMD), Faculty of Behavioural, Management, and Social Sciences, University of Twente, PO Box 217, 7500 AE Enschede, The Netherlands

² Department of Biological Psychology, VU University Amsterdam, Amsterdam, The Netherlands

³ Department of Clinical Child and Family Studies, VU University Amsterdam, Amsterdam, The Netherlands

⁴ Department of Methods, VU University Amsterdam, Amsterdam, The Netherlands

⁵ QIMR Berghofer Medical Research Institute, Brisbane, Australia

⁶ Department of Developmental Psychology and EMGO Institute for Health and Care Research, VU University Amsterdam, Amsterdam, The Netherlands

⁷ Department of Psychology, University of Minnesota, Minneapolis, USA

⁸ Department of Psychology, University of Edinburgh, Edinburgh, UK

⁹ Centre for Cognitive Ageing and Cognitive Epidemiology, University of Edinburgh, Edinburgh, UK

¹⁰ Donders Institute for Cognitive Neuroscience, Radboud University Nijmegen, Nijmegen, The Netherlands

¹¹ Department of Psychiatry, Radboud University Nijmegen Medical Center, Nijmegen, The Netherlands

¹² Department of Human Genetics, Radboud University Nijmegen Medical Center, Nijmegen, The Netherlands

¹³ Department of Cognitive Neuroscience, Radboud University Nijmegen Medical Center, Nijmegen, The Netherlands

¹⁴ Department of Psychology, University of Illinois at Urbana-Champaign, Champaign, IL, USA

¹⁵ Estonian Genome Center, University of Tartu, Tartu, Estonia

Grant W. Montgomery⁵ · Matthias Nauck⁵⁰ · Michel G. Nivard² · Klaasjan G. Ouwens² · Aarno Palotie^{51,52} · Erik Pettersson²⁷ · Ozren Polasek⁵³ · Yong Qian⁶⁷ · Laura Pulkki-Råback²¹ · Olli T. Raitakari^{54,55} · Anu Realo³² · Richard J. Rose⁵⁶ · Daniela Ruggiero²⁴ · Carsten O. Schmidt²⁶ · Wendy S. Slutske⁵⁷ · Rossella Sorice²⁴ · John M. Starr^{9,58} · Beate St Pourcain^{38,59,60} · Angelina R. Sutin^{25,61} · Nicholas J. Timpson³⁸ · Holly Trochet¹⁹ · Sita Vermeulen^{12,62} · Eero Vuoksimaa²⁸ · Elisabeth Widen⁵² · Jasper Wouda^{1,2} · Margaret J. Wright⁵ · Lina Zgaga^{36,63} · Generation Scotland⁶⁴ · David Porteous⁶⁵ · Alessandra Minelli⁴⁷ · Abraham A. Palmer^{17,39} · Dan Rujescu²⁰ · Marina Ciullo²⁴ · Caroline Hayward¹⁹ · Igor Rudan³⁶ · Andres Metspalu^{15,33} · Jaakko Kaprio^{41,28,52} · Ian J. Deary^{8,9} · Katri Räikkönen²¹ · James F. Wilson^{19,36} · Liisa Keltikangas-Järvinen²¹ · Laura J. Bierut³¹ · John M. Hettema²³ · Hans J. Grabe^{34,66} · Brenda W. J. H. Penninx⁴⁹ · Cornelia M. van Duijn¹⁶ · David M. Evans³⁸ · David Schlessinger⁶⁷ · Nancy L. Pedersen²⁷ · Antonio Terracciano^{22,25} · Matt McGue^{7,68} · Nicholas G. Martin⁵ · Dorret I. Boomsma²

Received: 30 October 2014 / Accepted: 10 August 2015

© The Author(s) 2015. This article is published with open access at Springerlink.com

Abstract Extraversion is a relatively stable and heritable personality trait associated with numerous psychosocial, lifestyle and health outcomes. Despite its substantial heritability, no genetic variants have been detected in previous genome-wide association (GWA) studies, which may be due to relatively small sample sizes of those studies. Here, we report on a large meta-analysis of GWA studies for extraversion in 63,030 subjects in 29 cohorts. Extraversion

item data from multiple personality inventories were harmonized across inventories and cohorts. No genome-wide significant associations were found at the single nucleotide polymorphism (SNP) level but there was one significant hit at the gene level for a long non-coding RNA site (LOC101928162). Genome-wide complex trait analysis in two large cohorts showed that the additive variance explained by common SNPs was not significantly different

¹⁶ Department of Epidemiology, Erasmus University Medical Center, Rotterdam, The Netherlands

¹⁷ Department of Human Genetics, University of Chicago, Chicago, IL, USA

¹⁸ Department of Clinical Chemistry, Fimlab Laboratories and School of Medicine, University of Tampere, Tampere, Finland

¹⁹ MRC Human Genetics Unit, MRC IGMM, Western General Hospital, University of Edinburgh, Edinburgh, UK

²⁰ Department of Psychiatry, University of Halle, Halle, Germany

²¹ Institute of Behavioural Sciences, University of Helsinki, Helsinki, Finland

²² Folkhälsan Research Center, Helsinki, Finland

²³ Department of Psychiatry, Virginia Institute for Psychiatric and Behavioral Genetics, Virginia Commonwealth University, Richmond, VA, USA

²⁴ Institute of Genetics and Biophysics “A. Buzzati-Traverso” – CNR, Naples, Italy

²⁵ National Institute on Aging, NIH, Baltimore, MD, USA

²⁶ Institute for Community Medicine, University Medicine Greifswald, Greifswald, Germany

²⁷ Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden

²⁸ Department of Public Health, University of Helsinki, Helsinki, Finland

²⁹ Department of Biostatistics, Center for Statistical Genetics, University of Michigan School of Public Health, Ann Arbor, MI, USA

³⁰ Pharmacotherapy & Outcomes Science, Virginia Commonwealth University, Richmond, VA, USA

³¹ Department of Psychiatry, Washington University School of Medicine, St. Louis, MO, USA

³² Department of Psychology, University of Tartu, Tartu, Estonia

³³ Estonian Academy of Sciences, Tallinn, Estonia

³⁴ Department of Psychiatry and Psychotherapy, University Medicine Greifswald, Greifswald, Germany

³⁵ Istituto di Ricerca Genetica e Biomedica (IRGB), CNR, Monserrato, Italy

³⁶ Usher Institute for Population Health Sciences and Informatics, University of Edinburgh, Edinburgh, UK

³⁷ Behavioral Medicine Research Center, Duke University School of Medicine, Durham, NC, USA

³⁸ Medical Research Council Integrative Epidemiology Unit, School of Social and Community Medicine, University of Bristol, Bristol, UK

from zero, but polygenic risk scores, weighted using linkage information, significantly predicted extraversion scores in an independent cohort. These results show that extraversion is a highly polygenic personality trait, with an architecture possibly different from other complex human traits, including other personality traits. Future studies are required to further determine which genetic variants, by what modes of gene action, constitute the heritable nature of extraversion.

Keywords Personality · Phenotype harmonization · Common genetic variants · Imputation · Polygenic risk

Introduction

Extraversion is a personality trait characterized by the tendency to experience positive emotions, to be active and feel energetic, to be talkative and to enjoy social interactions. Extraversion is associated with numerous psychosocial, lifestyle and health outcomes, such as academic and job performance, well-being, obesity, substance use, physical activity, bipolar disorder, borderline personality disorder, Alzheimer's disease, and longevity (De Moor et al. 2006, 2011; Distel et al. 2009a; Furnham et al. 2013; Judge et al. 2013; Middeldorp et al. 2011; Rhodes and

Smith 2006; Sutin et al. 2011; Terracciano et al. 2008; Terracciano et al. 2014; Weiss et al. 2008).

Extraversion can be measured with multiple inventories that have been developed as part of different personality theories. For example, extraversion is one of the five personality domains as assessed with the Neuroticism–Extraversion–Openness to Experience (NEO) personality inventories (Costa and McCrae 1992). Extraversion is also included in Eysenck's three-dimensional theory of personality (Eysenck and Eysenck 1964, 1975; Eysenck et al. 1985). In Cloninger's theory on temperaments and characters (Cloninger 1987; Cloninger et al. 1993), Harm Avoidance, Novelty Seeking and Reward Dependence are related to extraversion (De Fruyt et al. 2000). Tellegen's personality theory posits the higher order domain of Positive Emotionality (Patrick et al. 2002), which resembles and is highly correlated with extraversion (Church 1994).

We showed recently, by performing an Item Response Theory (IRT) analysis using test linking (Kolen and Brennan 2004), that item data on Extraversion, Reward dependence and Positive Emotionality can be harmonized to broadly assess the same underlying extraversion construct (van den Berg et al. 2014). This harmonization was performed in over 160,000 individuals from 23 cohorts participating in the Genetics of Personality Consortium (GPC). Briefly, harmonization was carried out in each cohort separately by first fitting an IRT model to data from

³⁹ Department of Psychiatry and Behavioral Neuroscience, University of Chicago, Chicago, USA

⁴⁰ Department of Biostatistics and Bioinformatics, Duke University, Durham, NC, USA

⁴¹ National Institute for Health and Welfare (THL), Helsinki, Finland

⁴² Department of General Practice and Primary Health Care, University of Helsinki, Helsinki, Finland

⁴³ Unit of General Practice and Primary Health Care, University of Helsinki, Helsinki, Finland

⁴⁴ Vasa Central Hospital, Vaasa, Finland

⁴⁵ Interfaculty Institute for Genetics and Functional Genomics, University of Greifswald, Greifswald, Germany

⁴⁶ Translational Research Institute, University of Queensland Diamantina Institute, Brisbane, Australia

⁴⁷ Department of Molecular and Translational Medicine, University of Brescia, Brescia, Italy

⁴⁸ Department of Biotechnology, University of Tartu, Tartu, Estonia

⁴⁹ Department of Psychiatry, EMGO+ Institute, Neuroscience Campus Amsterdam, VU University Medical Center, Amsterdam, The Netherlands

⁵⁰ Institute of Clinical Chemistry and Laboratory Medicine, University Medicine Greifswald, Greifswald, Germany

⁵¹ Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, UK

⁵² Institute for Molecular Medicine Finland (FIMM), University of Helsinki, Helsinki, Finland

⁵³ Department of Public Health, Faculty of Medicine, University of Split, Split, Croatia

⁵⁴ Department of Clinical Physiology and Nuclear Medicine, Turku University Hospital, Turku, Finland

⁵⁵ Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku, Turku, Finland

⁵⁶ Department of Psychological & Brain Sciences, Indiana University, Bloomington, IN, USA

⁵⁷ Department of Psychological Sciences and Missouri Alcoholism Research Center, University of Missouri, Columbia, MO, USA

⁵⁸ Alzheimer Scotland Dementia Research Centre, University of Edinburgh, Edinburgh, UK

⁵⁹ School of Oral and Dental Sciences, University of Bristol, Bristol, UK

⁶⁰ School of Experimental Psychology, University of Bristol, Bristol, UK

⁶¹ College of Medicine, Florida State University, Tallahassee, FL, USA

individuals who had completed at least two different personality questionnaires. Next, based on calibrated item parameters, personality scores were estimated based on all available data for each individual, irrespective of what personality questionnaire was used. The harmonized extraversion phenotype was heritable. A broad-sense heritability of 49 % was estimated, based on a meta-analysis in six twin cohorts that are included in the GPC (29,501 twin pairs), of which 24 % was due to additive genetic variance and 25 % due to non-additive genetic variance. The broad-sense heritability estimate is similar to heritability estimates obtained for extraversion as assessed with single measurement instruments (Bouchard and Loehlin 2001; Distel et al. 2009b; Finkel and McGue 1997; Keller et al. 2005; Rettew et al. 2008; Yamagata et al. 2006). Some evidence for qualitative sex differences in the genetic influences on extraversion was suggested by a genetic correlation in opposite-sex twin pairs of 0.38 (van den Berg et al. 2014). Extraversion becomes more genetically stable during adolescence until it is almost perfectly genetically stable in adulthood (Briley and Tucker-Drob 2014; Kandler 2012), that is, the same genes are responsible for extraversion measured at different ages.

A handful of genome-wide association (GWA) studies for extraversion have been published, aimed at detecting specific single nucleotide polymorphisms (SNPs) that explain part of the heritability. The first GWA study for personality, which focused on the five NEO personality traits, was conducted in 3972 adults (Terracciano et al. 2010). No genome-wide significant SNP associations were found for extraversion, although some interesting associations with P -values $<10^{-5}$ were seen with SNPs in two cadherin genes and the brain-derived neurotrophic factor (*BDNF*) gene. A subsequent meta-analysis of GWA results for the NEO personality traits, conducted in 17,375 subjects,

also did not yield any genome-wide significant associations for extraversion (De Moor et al. 2012). Two other GWA studies reported a similar lack of genome-wide significance for Cloninger's temperament scales (Service et al. 2012; Verweij et al. 2010). Interestingly, a study that performed a genetic complex trait analysis (GCTA; Yang et al. 2010) for neuroticism and extraversion in around 12,000 unrelated individuals reported that 12 % (SE = 3 %) of the variance in extraversion was explained by common SNPs of additive effect (Vinkhuyzen et al. 2012). Taken together, the results from twin and genome-wide studies suggest that common SNPs of additive effect are important, that genetic non-additivity may play a role, and that large sample sizes are likely to be required to identify specific variants.

In this paper, we report the results of the largest meta-analysis of GWA results for extraversion so far, carried out in 29 cohorts that participate in the GPC. A total of 63,030 subjects with harmonized extraversion and genome-wide genotype data were included in the meta-analysis. A 30th cohort was used for replication. In this consortium we reported earlier on a genome-wide significant hit for neuroticism (De Moor et al. 2015), indicating that we may begin to analyze data from sufficiently large samples, to obtain the first significant findings from GWA studies for personality. In addition to meta-analysis of GWA results, we computed weighted polygenic scores in an independent cohort and associated them with extraversion, and estimated variance explained by SNPs in two large cohorts.

Materials and methods

Cohorts

The full meta-analysis was performed on 63,030 subjects from 29 discovery cohorts. All samples were of European origin. Twenty-one cohorts were from Europe, six from the United States and two from Australia. Sample sizes of the individual cohorts ranged from 177 to 7210 subjects. Please note that some cohorts were also part of previously published GWA studies on extraversion. The Generation Scotland: Scottish Family Health Study (GS:SFHS) cohort was included as a replication sample (9,783 subjects). A brief overview of all cohorts is provided in Table 1. A description of each individual cohort is found in the Supplementary materials and methods (see also De Moor et al. 2015).

Phenotyping

A harmonized latent extraversion score was estimated for all participants in all 29 cohorts that were included in the GWA meta-analysis. This score was based on all available extraversion item data for each individual (for a detailed

⁶² Department for Health Evidence, Radboud University Medical Center, Nijmegen, The Netherlands

⁶³ Department of Public Health and Primary Care, Trinity College Dublin, Dublin, Ireland

⁶⁴ Scottish Family Health Study, A Collaboration Between the University Medical Schools and NHS, Aberdeen, Dundee, Edinburgh and Glasgow, UK

⁶⁵ Medical Genetics Section, Centre for Genomics and Experimental Medicine, Institute of Genetics and Molecular Medicine, Western General Hospital, The University of Edinburgh, Edinburgh, UK

⁶⁶ Department of Psychiatry and Psychotherapy, HELIOS Hospital Stralsund, Stralsund, Germany

⁶⁷ Laboratory of Genetics, National Institute on Aging, National Institutes of Health, Baltimore, MD, USA

⁶⁸ Institute of Public Health, University of Southern Denmark, Odense, Denmark

Table 1 Overview of 29 discovery cohorts and 1 replication cohort of the Genetics of Personality Consortium

	Cohort	# Subjects ^a	# SNPs ^b
1	ALSPAC	4705	6,454,153
2	BLSA	820	4,989,411
3	BRESCIA	177	3,549,919
4	CHICAGO	311	3,755,416
5	CILENTO	627	1,123,089
6	COGA	647	5,127,101
7	COGEND	1279	5,932,838
8	EGCUT	1184	5,574,695
9	ERF	2300	5,142,865
10	FTC EPI	567	4,870,096
11	FTC NEO	813	5,092,018
12	HBCS	1456	5,612,790
13	CROATIA-Korcula	808	5,094,034
14	LBC1921	437	4,363,611
15	LBC1936	952	5,168,754
16	MCTFR	7099	6,569,999
17	MGS	2101	5,900,898
18	NBS	1832	5,603,447
19	NESDA	2227	4,707,569
20	NTR	6416	5,339,798
21	ORCADES	1650	4,265,590
22	PAGES	476	4,547,293
23	QIMR adolescents	2842	5,957,064
24	QIMR adults	7210	6,343,920
25	SardiNIA	4332	6,291,135
26	SHIP	2213	5,913,428
27	STR	4903	6,519,094
28	CROATIA-Vis	909	5,327,671
29	YFS	1737	5,914,679
	Total	63,030	7,460,147
30	GS:SFHS	9783	74

NA Not Applicable for replication cohort because only top hits were sought to replicate

^a Number of subjects with valid latent score for Extraversion and SNP data (after imputation and cleaning)

^b Number of SNPs (after imputation and cleaning) with valid association results that entered the meta-analysis

description see van den Berg et al. 2014). Extraversion item data came from the extraversion scales of the NEO Personality Inventory, the NEO Five Factor Inventory, the 50-item Big-Five version of the International Personality Item Pool inventory, the Eysenck Personality Questionnaire and the Eysenck Personality Inventory, from the Reward Dependence scale of the Cloninger's Tridimensional Personality Questionnaire, and from the Positive Emotionality scale of the Multidimensional Personality

Questionnaire (see van den Berg et al. 2014 and Supplementary materials and methods). In the GS:SFHS cohort that was included for replication of top signals, extraversion was based on the summed score of the extraversion scale of the EPQ Revised Short Form.

Genotyping and imputation

Genotyping in all cohorts was carried out on Illumina or Affymetrix platforms, after which quality control (QC) was performed, followed by imputation of genotypes. QC of genotype data was performed in each cohort separately, with comparable but cohort specific criteria. Standard QC checks included tests of European ancestry, sex inconsistencies, Mendelian errors, and high genome-wide homozygosity. Checks for relatedness were conducted in those cohorts that aimed to include unrelated individuals only. Other checks of genotype data were based on minor allele frequencies (MAF), SNP call rate (% of subjects with missing genotypes per SNP), sample call rate (% of missing SNPs per subject) and Hardy–Weinberg Equilibrium (HWE). Genotype data were imputed using the 1000Genomes phase 1 version 3 (build37, hg19) reference panel with standard software packages such as IMPUTE, MACH, or Minimac, see Supplementary Table 1.

Statistical analyses

GWA analysis per cohort

GWA analyses were conducted independently in each cohort. Since the cohorts used different research designs (case–control, population twin studies, extended pedigrees, etc.), GWA methods were optimized for each cohort. Extraversion scores were regressed on each SNP under an additive model, with sex and age included as covariates. Covariates such as ancestry Principal Components (PCs) were added if deemed necessary for a particular cohort. In all analyses, the uncertainty of the imputed genotypes was taken into account, either using dosage scores or mixtures of distributions. In those cohorts that included related individuals, the dependency among participants was accounted for using cohort-specific methods. Standard software packages for GWA analyses were used (see Supplementary Table 1).

Meta-analysis of GWA results across cohorts

A meta-analysis of the GWA results was conducted with the weighted inverse variance method in METAL (<http://www.sph.umich.edu/csg/abecasis/metal/index.html>). Excluded from meta-analysis were poorly imputed SNPs ($r^2 < 0.30$ or $\text{proper_info} < 0.40$) and SNPs with low

MAF ($MAF < \sqrt{(5/N)}$, which corresponds to less than 5 estimated individuals in the least frequent genotype group, under the assumption of HWE). This resulted in a total number of 7,460,147 unique SNPs in the final meta-analysis (with 1.1–6.6 M SNPs across cohorts). For 2182 SNPs, SNP locations could not be matched with rs names. For an additional 516,362 SNPs, results were based on one cohort only and therefore left out of the analysis, so that the results are based on 6,941,603 SNPs. Genomic control inflation factors (λ), Manhattan plots and quantile–quantile plots per cohort are provided in Supplementary Table 2 and Supplementary Figs. 1, 2. A P value of 5×10^{-8} was used as the threshold for genome-wide significance.

The meta-analysis results (P -values per SNP) were used as the input to compute P -values at the gene level. We performed these analyses in KGG (Li et al. 2012). A P -value of 2.87×10^{-6} was used as the threshold for genome-wide significance in these gene-wide analyses, based on controlling for the false-discovery rate (Benjamini and Hochberg 1995).

All GWAS SNP top hits with a P -value smaller than 1×10^{-5} were selected for replication in the GS:SFHS cohort.

Polygenic risk score analysis

Additional analyses were conducted to test whether extraversion could be predicted in an independent target cohort based on the GWA meta-analysis results. The target cohort was the Netherlands Twin Register (NTR) cohort (8648 subjects). Polygenic risk scores for this cohort were estimated using LDpred (Vilhjalmsson et al. 2015) that takes into account linkage disequilibrium among the SNPs. The estimation was based on a GWA meta-analysis in which the NTR and NESDA cohorts were excluded (further referred to as the discovery set). With the LD-corrected polygenic risk scores, generalized estimating equation (GEE) modeling was applied to test whether the polygenic risk scores predicted extraversion in the target cohort. The covariates age, sex and ten PCs were included as fixed effects in the model. The model also included a random intercept with family number as the cluster variable, to account for dependency among family members. Outliers on the PCs, including ethnic outliers, were excluded from the analysis.

Variance explained by SNPs

In the NTR cohort and the QIMR Berghofer Medical Research Institute (QIMR) adult cohort (see also Supplementary materials and methods), GCTA software (Visscher et al. 2010; Yang et al. 2010) was used to estimate the proportion of variance in extraversion that can be explained

by common SNPs of additive effect. In the NTR, this analysis was carried out in a set of 3597 unrelated individuals and in the QIMR adult cohort this was done in 3369 unrelated individuals (in each cohort one member per family was selected with harmonized extraversion and genome-wide SNP data). GCTA analysis was based on best guess genotypes obtained in PLINK using a threshold of a maximum genotype probability >0.70 , and additionally filtering on r -squared >0.80 . Next, in estimating the GRM matrix in the GCTA software, SNPs with $MAF < 0.05$ were excluded. The additive genetic relationship matrices (GRM) estimated based on SNPs for all individuals formed the basis to estimate the proportion of phenotypic variance explained by SNPs in the NTR and QIMR cohorts. In other words, it was determined to what extent phenotypic similarity between individuals corresponds to genetic similarity (at the SNP level). For both NTR and QIMR, sex, age and a set of population-specific PCs were included as covariates.

Results

Meta-analysis of GWA results

Meta-analysis of GWA results across the 29 discovery cohorts did not yield genome-wide significant SNPs associated with extraversion. The lowest P -value observed was 2.9×10^{-7} for a SNP located on chromosome 2. There were 74 SNPs with P -values $< 1 \times 10^{-5}$. The Manhattan and quantile–quantile plots are provided in Figs. 1 and 2. A list with the top five SNPs is given in Table 2. A list with all SNPs that reached the level of suggestive genome-wide significance ($P < 1 \times 10^{-5}$) is found in Supplementary Table 3. The results of all SNPs can be downloaded from www.tweelingenregister.org/GPC. A gene-based test showed one significant hit for *LOC101928162*, a long non-coding RNA site, $P = 2.87 \times 10^{-6}$. A list with the top five genes from the gene-based analysis is provided in Table 3. Supplementary Table 4 provides the top 30 genes. Among the top 30 genes was Brain-Derived Neurotrophic Factor (*BDNF*, $P = 0.0003$), a gene also implicated, though not genome-wide significant, in Terracciano et al. (2010), as was the *BDNF* anti-sense RNA gene ($P = 0.0001$).

Results of the follow-up analysis of the top five SNPs in the GS:SFHS cohort can be found in Table 2. Of the top five SNPs, none showed a significant effect. For an overview of the replication results of all top SNPs with P -value $< 1 \times 10^{-5}$ see Supplementary Table 3. Of the 74 SNPs tested in the replication cohort, three SNPs showed nominal evidence of association ($P < 0.05$), which is less than the number expected based on chance alone ($0.05 \times 74 = 3.7$).

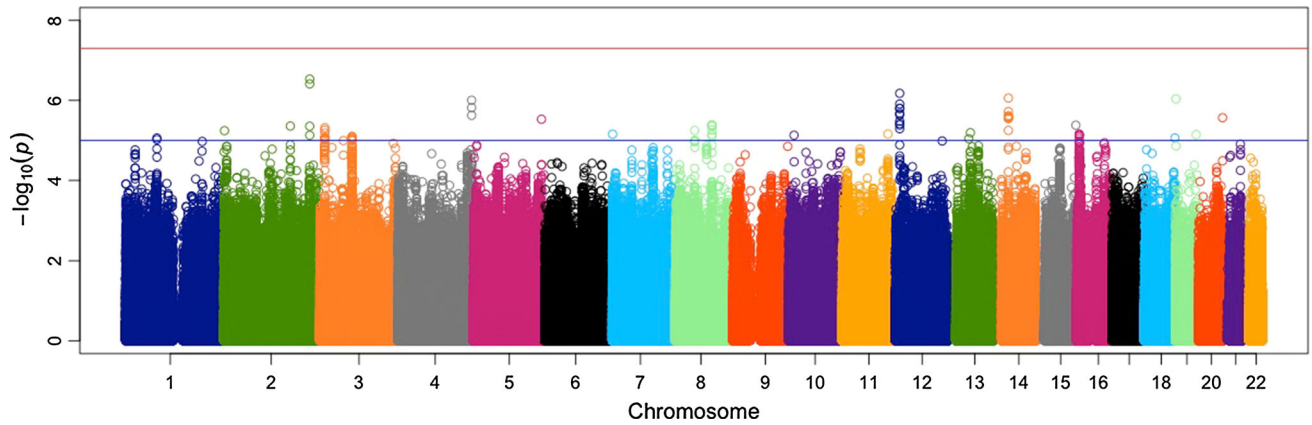


Fig. 1 Manhattan plot for meta-analysis results of 29 discovery cohorts for extraversion in the Genetics of Personality Consortium

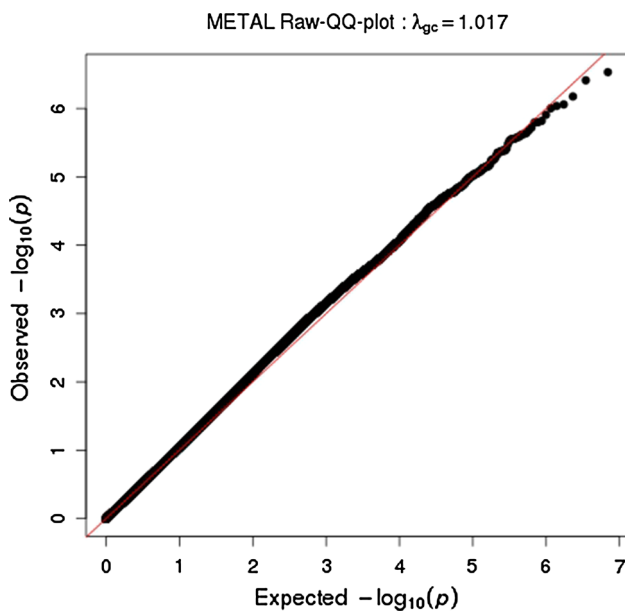


Fig. 2 Quantile-Quantile plots for meta-analysis results of 29 discovery cohorts for extraversion in the Genetics of Personality Consortium

Polygenic risk score analysis

There were 8201 persons individuals with polygenic scores for prediction of extraversion. The LDpred-based genetic

risk scores significantly predicted extraversion in the target cohort, $B = 0.059$, $X^2(1) = 27.30$, $P < 0.001$.

Variance explained by SNPs

In the NTR cohort, an estimated 5.0 % (SE = 7.2) of the variance in extraversion was explained by all SNPs, but this estimate was not significantly different from zero ($P = 0.24$). In the QIMR cohort, 0.0001 % (SE = 15) of the variance was explained by SNPs ($P = 0.46$).

Discussion

This study assessed the influence of common genetic variants on extraversion in 63,030 individuals from 29 cohorts in the GPC. First, a meta-analysis of GWA analyses across 29 discovery cohorts showed no genome-wide significant SNPs. Top SNPs detected in the meta-analysis of GWA results in the discovery phase were not replicated in the GS:SFHS cohort. The SNPs with lowest P -values have no previously reported relationship with personality, psychopathology or brain functioning. Polygenic risk scores based on the meta-analysis results predicted extraversion in an independent data set. SNP-based

Table 2 Top SNPs from the meta-analysis of GWA results in 29 discovery cohorts for extraversion, and their replication in the GS:SFHS cohort, in the Genetics of Personality Consortium

SNP	Chr_BP	Alleles	Closest gene	Discovery results			Replication results		
				Effect	SE	P -value	Effect	SE	P -value
rs2024488	2_217662968	A/G	<i>LOC101928250</i>	-0.0303	0.0059	2.939×10^{-7}	0.0285	0.0164	0.08244
rs2712162	2_217661788	T/C	<i>LOC101928250</i>	-0.0300	0.0059	3.872×10^{-7}	0.0278	0.0164	0.08947
rs797182	12_10900487	A/G	<NA>	-0.0277	0.0056	6.673×10^{-7}	-0.0135	0.0153	0.37721
rs8010306	14_37150160	A/G	<i>SLC25A21</i>	0.0629	0.0128	8.730×10^{-7}	0.0180	0.0314	0.56650
rs117292860	19_2227621	A/C	<i>DOTIL</i>	0.0553	0.0113	9.191×10^{-7}	-0.0350	0.0239	0.14368

Table 3 Top genes from the meta-analysis of GWA results in 29 discovery cohorts for Extraversion in the Genetics of Personality Consortium

Gene	Full gene name	Pathways	P-value
<i>LOC101928162</i>	[Long non-coding RNA]	Unknown	0.00000287
<i>LOC729506</i>	[Long non-coding RNA]	Unknown	0.00000893
<i>PLEKHJ1</i>	Pleckstrin Homology Domain Containing, Family J Member 1	Phospholipid binding, circadian clock functioning	0.0000132
<i>POU2F3</i>	POU Class 2 Homeobox 3	Influenza A	0.0000179
<i>CRTAC1</i>	Cartilage Acidic Protein 1	Unknown	0.0000297

heritabilities for extraversion were not significantly different from zero in two large cohorts of the GPC.

Although there were no genome-wide significant results for individual SNPs, in the gene-based analysis, there was a significant hit for one locus, *LOC101928162*. This is long noncoding RNA site whose function remains elusive. Interestingly, among the top 30 genes were genes previously implicated in extraversion or in psychiatric disorders associated with extraversion. The low *P*-value for *CRTAC1* ($P = 2.97 \times 10^{-5}$), harks back to an interesting extraversion SNP (rs7088779) in a previous GWAS on personality (Amin et al. 2013) that is located between *CRTAC1* and *C10orf28*. *RELN* ($P = 5.69 \times 10^{-5}$) has been reported to increase the risk for schizophrenia and bipolar disorder (Kuang et al. 2011; Ovadia and Shifman 2011), while *ADAM12* (7.65×10^{-5}) was previously found to be involved in schizophrenia (Farkas et al. 2010), and bipolar disorder treatment (Nadri et al. 2007). The *BDNF* gene was also implicated in a previous extraversion GWAS (Terracciano et al. 2010), though not genome-wide significant. Liu et al. (2005) reported a trend towards association of *BDNF* variants with substance abuse, Jiao et al. (2011) reported an association with obesity, and Lang et al. (2007) and Beuten et al. (2005) reported associations with smoking behavior. As extraversion is known to be associated with lifestyle, obesity and substance abuse, we deem *BDNF* to be an interesting candidate gene for extraversion in future studies, along with *CRTAC*, *ADAM12* and *RELN*.

With the current meta-analysis we more than tripled the sample size as compared to the largest previously published meta-analysis for extraversion (De Moor et al. 2012). In contrast to neuroticism, no genome-wide significant SNPs were found. Some have argued (Turkheimer et al. 2014) that the heritability of personality traits represents nonspecific genetic background, which is composed of so many genetic variants with extremely small effect sizes that individually these have no causal biological interpretation. It may be that extraversion differs in this respect from neuroticism. One other difference was indicated from the analyses of the IRT-based extraversion and neuroticism scores: whereas for neuroticism no evidence for genotype x sex interaction was

seen (van den Berg et al. 2014), for extraversion there was significant evidence for sex limitation. It also is interesting to note that despite the fact that for extraversion no genome-wide significant findings emerged for single SNPs, we were able to predict extraversion in an independent dataset, based on the polygenic risk cohorts from the discovery set. This indicates that some true signal is entailed in the meta-analysis results.

The results of the polygenic risk score analysis are in contrast with the results from the GCTA analysis, in which no significant proportion of variance explained by SNPs was detected in two large cohorts of the GPC. Our study on neuroticism reported a SNP-based heritability of 15 % (De Moor et al. 2015). The current extraversion GCTA findings are also somewhat at odds with two previous GCTA studies for personality traits. One study focused on neuroticism and extraversion as measured with different instruments in four cohorts, and found on average 12 % explained variance for extraversion, although across cohorts these estimates varied widely (0–27 %) (Vinkhuyzen et al. 2012). Estimates for neuroticism also varied, but were generally lower than for extraversion in this study, with an average of 6 % explained variance. In another study, between 4.2 and 9.9 % of explained variances were found for the four Cloninger temperaments in a combined sample of four cohorts (Verweij et al. 2012). The proportions of variances for Harm Avoidance, Novelty Seeking and Persistence were significant at $P < 0.05$, whereas interestingly the proportion of variance for Reward Dependence was not. It should be noted that both these studies included the QIMR cohort in their analyses, so there is some overlap in subjects across studies. The difference is that in the earlier studies extraversion and reward dependence were based on single personality inventories, while in our study extraversion scores harmonized among different personality inventories were analyzed. What our results and the results in the previous studies have in common though, is that the estimates are considerably smaller than the heritability estimates based on twin studies. Given that about half of the heritability of extraversion consists of non-additive genetic variance (van den Berg et al. 2014), it is not unlikely that this discrepancy is caused by the influence of

common variants that interact within loci (dominance) or across loci (epistasis). In addition, the influence of rare variants may be implicated. The relatively limited influence of common additive genetic variation, as well as a previously reported finding that higher levels of inbreeding are associated with less socially desirable personality trait levels, has led to the idea that the genetic variation in personality traits may have been maintained by mutation–selection balance (Verweij et al. 2012), and our results are consistent with this idea.

This study comes with some limitations. Genotyping, QC, and imputation were carried out separately in each cohort. Any difference in procedures may have caused some loss of statistical power to detect SNPs in the meta-analysis. Similarly, extraversion item data were harmonized as much as possible (van den Berg et al. 2014), but the Reward Dependence item data from the TCI were least successfully linked to the extraversion data from the other inventories. This may also have caused some loss in power. Importantly however, it should be noted that by combining genotype and phenotype data across cohorts as performed in this study, a substantial increase in sample size was obtained. It is nontrivial that the gain in power associated with this increase in sample size largely outweighs any potential loss in power due to any remaining genotyping or phenotyping differences across cohorts.

In conclusion, extraversion is a heritable, highly polygenic personality trait with a genetic background that may be qualitatively different from that of other complex behavioral traits. Future studies are required to increase our knowledge of which types of genetic variants, by which modes of gene action, constitute the heritable nature of extraversion. Ultimately, this knowledge can be used to increase our understanding of how extraversion is related to various important psychosocial and health outcomes.

Acknowledgments We would like to thank all participating subjects. Analyses were carried out on the Genetic Cluster Computer (<http://www.geneticcluster.org>), which is financially supported by the Netherlands Organization for Scientific Research (NWO 480-05-003).

ALSPAC We are extremely grateful to all the families who took part in this study, the midwives for their help in recruiting them and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses. The UK Medical Research Council (Grant 74882), the Wellcome Trust (Grant 076467) and the University of Bristol provide core support for ALSPAC. We thank 23andMe for funding the genotyping of the ALSPAC children's sample. This publication is the work of the authors, and they will serve as guarantors for the contents of this paper.

BLSA We acknowledge support from the Intramural Research Program of the NIH, National Institute on Aging. We thank Robert McCrae.

BRESCIA We acknowledge support from the Italian Ministry of Health (RC and RF2007 Conv. 42) and Regione Lombardia (ID: 17387 SAL-13). We thank Ilaria Gandin for imputation analysis support.

CHICAGO This work was supported by NIH Grants, DA007255 (ABH), HG006265 (to BEE), DA02812 (to HdW), and DA021336 and DA024845 (to AAP). BEE was also funded through the Bioinformatics Research Development Fund, supported by Kathryn and George Gould. We wish to thank Andrew D. Skol for providing advice about genotype calling.

CILENTO We acknowledge Dr Maria Enza Amendola for the test administration and thank the personnel working in the organization of the study in the villages. MC received funding support from the Italian Ministry of Universities (FIRB - RBNE08NKH7, INTEROMICS Flagship Project), the Assessorato Ricerca Regione Campania, the Fondazione con il SUD (2011-PDR-13), and the Fondazione Banco di Napoli.

SAGE – COGA/CONGEND Funding support for the Study of Addiction: Genetics and Environment (SAGE) was provided through the NIH Genes, Environment and Health Initiative [GEI] (U01 HG004422). SAGE is one of the genome-wide association studies funded as part of the Gene Environment Association Studies (GENEVA) under GEI. Assistance with phenotype harmonization and genotype cleaning, as well as with general study coordination, was provided by the GENEVA Coordinating Center (U01 HG004446). Assistance with data cleaning was provided by the National Center for Biotechnology Information. Support for collection of datasets and samples was provided by the Collaborative Study on the Genetics of Alcoholism (COGA; U10 AA008401) and the Collaborative Genetic Study of Nicotine Dependence (COGEND; P01 CA089392). Funding support for genotyping, which was performed at the Johns Hopkins University Center for Inherited Disease Research, was provided by the NIH GEI (U01HG004438), the National Institute on Alcohol Abuse and Alcoholism, the National Institute on Drug Abuse, and the NIH contract “High throughput genotyping for studying the genetic contributions to human disease” (HHSN268200782096C). The Collaborative Study on the Genetics of Alcoholism (COGA), Principal Investigators B. Porjesz, V. Hesselbrock, H. Edenberg, L. Bierut, includes ten different centers: University of Connecticut (V. Hesselbrock); Indiana University (H.J. Edenberg, J. Nurnberger Jr., T. Foroud); University of Iowa (S. Kuperman, J. Kramer); SUNY Downstate (B. Porjesz); Washington University in St. Louis (L. Bierut, A. Goate, J. Rice, K. Bucholz); University of California at San Diego (M. Schuckit); Rutgers University (J. Tischfield); Texas Biomedical Research Institute (L. Almasy), Howard University (R. Taylor) and Virginia Commonwealth University (D. Dick). Other COGA collaborators include: L. Bauer (University of Connecticut); D. Koller, S. O'Connor, L. Wetherill, X. Xuei (Indiana University); Grace Chan (University of Iowa); S. Kang, N. Manz, M. Rangaswamy (SUNY Downstate); J. Rohrbaugh, J-C Wang (Washington University in St. Louis); A. Brooks (Rutgers University); and F. Aliev (Virginia Commonwealth University). A. Parsian and M. Reilly are the NIAAA Staff Collaborators. This national collaborative study is supported by NIH Grant U10AA008401 from the National Institute on Alcohol Abuse and Alcoholism (NIAAA) and the National Institute on Drug Abuse (NIDA). The Collaborative Genetic Study of Nicotine Dependence (COGEND) project is a collaborative research group and part of the NIDA Genetics Consortium. Subject collection was supported by NIH Grant P01 CA089392 (L.J. Bierut) from the National Cancer Institute. Phenotypic and genotypic data are stored in the NIDA Center for Genetic Studies (NCGS) at <http://zork.wustl.edu/> under NIDA Contract HHSN271200477451C (J. Tischfield and J. Rice). Jaime Derringer was supported by NIH T32 MH016880.

EGCUT AM and TE received support from FP7 Grants (201413 ENGAGE, 212111 BBMRI, ECOGENE (No. 205419, EBC)) and OpenGENE. AM and TE also received targeted financing from Estonian Government SF0180142s08 and by EU via the European Regional Development Fund, in the frame of Centre of Excellence in Genomics. The genotyping of the Estonian Genome Project samples were performed in Estonian Biocentre Genotyping Core Facility, AM

and TE thank Mari Nelis and Viljo Soo for their contributions. AR and JA were supported by a grant from the Estonian Ministry of Science and Education (SF0180029s08).

ERF The ERF study as a part of EUROSPAN (European Special Populations Research Network) was supported by European Commission FP6 STRP Grant Number 018947 (LSHG-CT-2006-01947) 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). The ERF study was further supported by ENGAGE consortium and CMSB. 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). ERF was further supported by the ZonMw Grant (Project 91111025). 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.

Finnish Twin Cohort (FTC) We acknowledge support from the Academy of Finland Center of Excellence in Complex Disease Genetics (Grant Numbers: 213506, 129680), the Academy of Finland (Grants 100499, 205585, 118555 and 141054 to JK, Grant 257075 to EV), Global Research Awards for Nicotine Dependence (GRAND), ENGAGE (European Network for Genetic and Genomic Epidemiology, FP7-HEALTH-F4-2007, Grant Agreement Number 201413), DA12854 to P A F Madden, and AA-12502, AA-00145, and AA-09203 to RJRose, AA15416 and K02AA018755 to DM Dick.

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

CROATIA-Korcula The CROATIA-Korcula study was funded by grants from the Medical Research Council (UK), European Commission Framework 6 project EUROSPAN (Contract No. LSHG-CT-2006-018947) and Republic of Croatia Ministry of Science, Education and Sports research Grants to I.R. (108-1080315-0302). We would like to acknowledge the invaluable contributions of the recruitment team in Korcula, the administrative teams in Croatia and Edinburgh and the people of Korcula. The SNP genotyping for the CROATIA-Korcula cohort was performed in Helmholtz Zentrum München, Neuherberg, Germany.

LBC1921 & LBC1936 For the Lothian Birth Cohorts, we thank Paul Redmond for database management; Alan Gow, Michelle Taylor, Janie Corley, Caroline Brett and Caroline Cameron for data collection and data entry; nurses and staff at the Wellcome Trust Clinical Research Facility, where blood extraction and genotyping was performed; staff at the Lothian Health Board, and the staff at the SCRE Centre, University of Glasgow. The research was supported by a program grant from Research Into Ageing. The research continues with program grants from Age UK (The Disconnected Mind). 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 Biotechnology and Biological Sciences Research Council (BBSRC) and Medical Research Council (MRC) is gratefully acknowledged. IJD, DJP and colleagues receive support from Wellcome Trust Strategic Award 104036/Z/14/Z.

MCTFR We would like to thank Rob Kirkpatrick for his help running analyses.

Research reported in this publication was supported by the National Institutes of Health under award numbers R37DA005147, R01AA009367, R01AA011886, R01DA013240, R01MH066140, and U01DA024417.

MGS Samples were collected under the following grants: NIMH Schizophrenia Genetics Initiative U01s: MH46276, MH46289, and MH46318; and Molecular Genetics of Schizophrenia Part 1 (MGS1) and Part 2 (MGS2) R01s: MH67257, MH59588, MH59571, MH59565, MH59587, MH60870, MH60879, MH59566, MH59586, and MH61675. Genotyping and analyses were funded under the MGS U01s: MH79469 and MH79470.

NBS Principal investigators of the Nijmegen Biomedical Study are L.A.L.M. Kiemeny, M. den Heijer, A.L.M. Verbeek, D.W. Swinkels and B. Franke.

NESDA The Netherlands Study of Depression and Anxiety (NESDA) were funded by the Netherlands Organization for Scientific Research (Geestkracht program Grant 10-000-1002); the Center for Medical Systems Biology (CMSB, NWO Genomics), Biobanking and Biomolecular Resources Research Infrastructure (BBMRI-NL), VU University's EMGO Institute for Health and Care Research and Neuroscience Campus Amsterdam. Genotyping was funded by the US National Institute of Mental Health (RC2MH089951) as part of the American Recovery and Reinvestment Act of 2009. BP is financially supported by NWO-VIDI Grant No. 91811602.

NTR We acknowledge financial support from the Netherlands Organization for Scientific Research (NWO): Grants 575-25-006, 480-04-004, 904-61-090; 904-61-193, 400-05-717 and Spinozapremie SPI 56-464-14192 and the European Research Council (ERC-230374). MHMdeM is supported by NWO VENI Grant No. 016-115-035. Genotyping was funded by the Genetic Association Information Network (GAIN) of the Foundation for the US National Institutes of Health, and analysis was supported by grants from Genetic Association Information Network and the NIMH (MH081802). Genotype data were obtained from dbGaP (<http://www.ncbi.nlm.nih.gov/dbgap>, accession number phs000020.v1.p1).

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 research nurses in Orkney, the administrative team in Edinburgh and the people of Orkney.

PAGES none.

QIMR Berghofer adolescents/adults We thank Marlene Grace and Ann Eldridge for sample collection; Megan Campbell, Lisa Bowdler, Steven Crooks and staff of the Molecular Epidemiology Laboratory for sample processing and preparation; Harry Beeby, David Smyth and Daniel Park for IT support. We acknowledge support from the Australian Research Council Grants A79600334, A79906588, A79801419, DP0212016, DP0343921, DP0664638, and DP1093900 (to NGM and MJW), Beyond Blue and the Borderline Personality Disorder Research Foundation (to NGM), NIH Grants DA12854 (to PAFM), AA07728, AA07580, AA11998, AA13320, AA13321 (to ACH) and MH66206 (to WSS); and grants from the Australian National Health and Medical Research Council; MLP is supported by DA019951. Genotyping was partly funded by the National Health and Medical Research Council (Medical Bioinformatics Genomics Proteomics Program, 389891) and the 5th Framework Programme (FP-5) GenomEUtwin Project (QLG2-CT-2002-01254). Further genotyping at the Center for Inherited Disease Research was supported by a grant to the late Richard Todd, M.D., Ph.D., former Principal Investigator of Grant AA13320. SEM and GWM are supported by the National Health and Medical Research Council Fellowship Scheme. Further, we gratefully acknowledge Dr Dale R Nyholt for his substantial

involvement in the QC and preparation of the QIMR GWA data sets. Dr Nyholt also contributed 8 % of the GWAS for the QIMR adult cohort (NHMRC IDs 339462, 442981, 389938, 496739).

SardiNIA We acknowledge support from the Intramural Research Program of the NIH, National Institute on Aging. Funding was provided by the National Institute on Aging, NIH Contract No. NO1-AG-1-2109 to the SardiNIA ('ProgeNIA') team.

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 and the Social Ministry of the Federal State of Mecklenburg-West Pomerania. Genome-wide data have been supported by the Federal Ministry of Education and Research (Grant No. 03ZIK012) and a joint grant from Siemens Healthcare, Erlangen, Germany and the Federal State of Mecklenburg-West Pomerania. The University of Greifswald is a member of the 'Center of Knowledge Interchange' program of the Siemens AG. This work was also funded by the German Research Foundation (DFG: GR 1912/5-1).

STR The STR was supported by grants from the Ministry for Higher Education, the Swedish Research Council (M-2005-1112 and 2009-2298), GenomEUtwin (EU/QLRT-2001-01254; QLG2-CT-2002-01254), NIH Grant DK U01-066134, The Swedish Foundation for Strategic Research (SSF; ICA08-0047), the Swedish Heart-Lung Foundation, the Royal Swedish Academy of Science, and ENGAGE (within the European Union Seventh Framework Programme, HEALTH-F4-2007-201413).

CROATIA-Vis The CROATIA-Vis study was funded by grants from the Medical Research Council (UK) and Republic of Croatia Ministry of Science, Education and Sports research Grants to I.R. (108-1080315-0302). We would like to acknowledge the staff of several institutions in Croatia that supported the field work, including but not limited to The University of Split and Zagreb Medical Schools, the Institute for Anthropological Research in Zagreb and Croatian Institute for Public Health. The SNP genotyping for the CROATIA-Vis cohort was performed in the core genotyping laboratory of the Wellcome Trust Clinical Research Facility at the Western General Hospital, Edinburgh, Scotland.

YFS The Young Finns Study has been financially supported by the Academy of Finland (Grants 126925, 121584, 124282, 129378 (Salve), 117787 (Gendi), 41071 (Skidi), and 265869 (Mind)), the Social Insurance Institution of Finland, Kuopio, Tampere and Turku University Hospital Medical Funds (Grant 9N035 for Dr. Lehtimäki), Juho Vainio Foundation, Paavo Nurmi Foundation, Finnish Foundation of Cardiovascular Research and Finnish Cultural Foundation, Tampere Tuberculosis Foundation and Emil Aaltonen Foundation (for Dr. Lehtimäki). The expert technical assistance in statistical analysis by Irina Lisinen, Mika Helminen, and Ville Aalto is gratefully acknowledged.

GS:SHFHS GS:SHFHS is funded by the Scottish Executive Health Department, Chief Scientist Office, Grant Number CZD/16/6. Exome array genotyping for GS:SHFHS was funded by the Medical Research Council UK and performed at the Wellcome Trust Clinical Research Facility Genetics Core at Western General Hospital, Edinburgh, UK. We would like to acknowledge the invaluable contributions of the families who took part in the GS:SHFHS, the general practitioners and Scottish School of Primary Care for their help in recruiting them, and the whole GS:SHFHS team, which includes academic researchers, IT staff, laboratory technicians, statisticians and research managers.

Authors' contributions Writing group: SMvdB, MHMdeM, KJHV, RfK, ML, AAV, LKM, JD, TE, DIB. Analytic group: MHMdeM, SMvdB, KJHV, ML, AAV, LKM, JDe, TE, NA, SG, NKH, ABH, JH, BK, JL, ML, MM, TT, ATeu, AV, JW, IOF, NT, DME, TL, IS, EP, GRA, JM, HM, AA, MN. Study design and project management: LF, LPR, JGE, AAP, GWM, MJW, PAFM, DP, AMin,

AP, DR, MC, IG, CH, IR, AMet, JK, IJD, KR, JFW, LKJ, JMH, HJG, BWJHP, CMvd, DME, NLP, PTC, ATer, MMG, NGM, DIB, RfK, AAV, GDS, TL, OTR, PKEM, KH, JMS, DS, GRA, HC, WGI, JDi. Sample and phenotype data collection: BWJHP, AMet, AR, JA, PAFM, ACH, NGM, MJW, KA, MN, LJB, JGE, LF, PTC, IG, AMH, ATer, GDS, MGK, HdW, AAP, AKH, WSS, RS, DR, Amin, MJ, LPR, LKJ, OTR, PKEM, EV, KH, AM, FB, OP, LZ. Data preparation: SMvdB, MHMdeM, JW, KGO, JJH, SEM, NKH, YM, TE, AR, GD, ML, RG, AA, JD, EW, GD, BEE, COS, GH, KJHV, SDG, DEA, TBB, JPK, NJT, BP, BK, CM, MJ, AL, ARS, ATer, DCL, HT.

Conflict of Interest The authors declare that they have no conflict of interest.

Human and Animal Rights and Informed Consent All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Informed consent was obtained from all individual participants included in the study.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Amin N, Hottenga JJ, Hansell NK, Janssens ACJW, de Moor MHM, Madden PAF, Zorkoltseva IV, Penninx BW, Terracciano A, Uda M, Tanaka T, Esko T, Realo A, Ferrucci L, Luciano M, Davies G, Metspalu A, Abecasis GR, Deary IJ, Raikonen K, Bierut LJ, Costa PT, Saviouk V, Zhu G, Kirichenko AV, Isaacs A, Aulchenko YS, Willemsen G, Heath AC, Pergadia ML, Medland SE, Axenovich TI, de Geus E, Montgomery GW, Wright MJ, Oostra BA, Martin NG, Boomsma DI, van Duijn CM (2013) Refining genome-wide linkage intervals using a meta-analysis of genome-wide association studies identifies loci influencing personality dimensions. *Eur J Hum Genet* 21(8):876–882
- Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate—a practical and powerful approach to multiple testing. *J R Stat Soc Series B-Methodol* 57(1):289–300
- Beuten J, Ma JZ, Payne TJ, Dupont RT, Quezada P, Huang WH, Crews KA, Li MD (2005) Significant association of BDNF haplotypes in European-American male smokers but not in European-American female or African-American smokers. *Am J Med Genet B* 139b(1):73–80
- Bouchard TJ, Loehlin JC (2001) Genes, evolution, and personality. *Behav Genet* 31(3):243–273
- Briley DA, Tucker-Drob EM (2014) Genetic and environmental continuity in personality development: a meta-analysis. *Psychol Bull* 140(5):1303–1331
- Church AT (1994) Relating the tellegen and 5-factor models of personality structure. *J Pers Soc Psychol* 67(5):898–909
- Cloninger CR (1987) A systematic method for clinical description and classification of personality variants—a proposal. *Arch Gen Psychiatry* 44(6):573–588

- Cloninger CR, Svrakic DM, Przybeck TR (1993) A psychobiological model of temperament and character. *Arch Gen Psychiatry* 50(12):975–990
- Costa PT, McCrae RR (1992) Professional manual: revised NEO personality inventory (NEO-PI-R) and NEO five-factor-inventory (NEO-FFI). Psychol Assess Resour, Odessa
- De Fruyt F, Van De Wiele L, Van Heeringen C (2000) Cloninger's psychobiological model of temperament and character and the five-factor model of personality. *Pers Indiv Differ* 29(3):441–452
- De Moor MHM, Beem AL, Stubbe JH, Boomsma DI, de Geus EJC (2006) Regular exercise, anxiety, depression and personality: a population-based study. *Prev Med* 42(4):273–279
- De Moor MHM, Vink JM, van Beek JHDA, Geels LM, Bartels M, de Geus EJC, Willemsen AHM, Boomsma DI (2011) Heritability of problem drinking and the genetic overlap with personality in a general population sample. *Front Behav Psychiatr Genet* 2:76
- De Moor MH, Costa PT, Terracciano A, Krueger RF, de Geus EJ, Toshiko T, Penninx BW, Esko T, Madden PA, Derringer J, Amin N, Willemsen G, Hottenga JJ, Distel MA, Uda M, Sanna S, Spinoven P, Hartman CA, Sullivan P, Realo A, Allik J, Heath AC, Pergadia ML, Agrawal A, Lin P, Grucza R, Nutile T, Ciullo M, Rujescu D, Giegling I, Konte B, Widen E, Cousminer DL, Eriksson JG, Palotie A, Peltonen L, Luciano M, Tenesa A, Davies G, Lopez LM, Hansell NK, Medland SE, Ferrucci L, Schlessinger D, Montgomery GW, Wright MJ, Aulchenko YS, Janssens AC, Oostra BA, Metspalu A, Abecasis GR, Deary IJ, Raikonen K, Bierut LJ, Martin NG, van Duijn CM, Boomsma DI (2012) Meta-analysis of genome-wide association studies for personality. *Mol Psychiatry* 17(3):337–349
- De Moor MHM, van den Berg SM, Consortium GoP, Boomsma DI (2015) Meta-analysis of genome-wide association studies for neuroticism, and the polygenic association with major depressive disorder. *JAMA Psychiatry* 72:642–650
- Distel MA, De Moor MHM, Boomsma DI (2009a) Dutch translation of the Personality Assessment Inventory-Borderline features scale (PAI-BOR): norms, factor structure and reliability. *Psychol Gezondh* 37(1):38–46
- Distel MA, Trull TJ, Willemsen G, Vink JM, Derom CA, Lynskey MT, Martin NG, Boomsma DI (2009b) The Five Factor Model of personality and borderline personality disorder: a genetic analysis of comorbidity. *Biol Psychiatry* 66:1131–1138
- Eysenck HJ, Eysenck SBG (1964) Eysenck personality inventory. Educational and Industrial Testing Service, San Diego
- Eysenck HJ, Eysenck SBG (1975) Manual of the Eysenck personality questionnaire. Hodder & Stoughton, London
- Eysenck SBG, Eysenck HJ, Barrett P (1985) A revised version of the psychoticism scale. *Pers Indiv Differ* 6(1):21–29
- Farkas N, Lendeckel U, Dobrowolny H, Funke S, Steiner J, Keilhoff G, Schmitt A, Bogerts B, Bernstein HG (2010) Reduced density of ADAM 12-immunoreactive oligodendrocytes in the anterior cingulate white matter of patients with schizophrenia. *World J Biol Psychiatry* 11(3):556–566
- Finkel D, McGue M (1997) Sex differences and nonadditivity in heritability of the multidimensional personality questionnaire scales. *J Pers Soc Psychol* 72(4):929–938
- Furnham A, Nuygards S, Chamorro-Premuzic T (2013) Personality, assessment methods and academic performance. *Instr Sci* 41(5):975–987
- Jiao H, Arner P, Hoffstedt J, Brodin D, Dubern B, Czernichow S, van't Hooft F, Axelsson T, Pedersen O, Hansen T, Sorensen TIA, Hebebrand J, Kere J, Dahlman-Wright K, Hamsten A, Clement K, Dahlman I (2011) Genome wide association study identifies KCNMA1 contributing to human obesity. *BMC Med Genomics* 4:51
- Judge TA, Rodell JB, Klinger RL, Simon LS, Crawford ER (2013) Hierarchical representations of the five-factor model of personality in predicting job performance: integrating three organizing frameworks with two theoretical perspectives. *J Appl Psychol* 98(6):875–925
- Kandler C (2012) Nature and nurture in personality development: the case of neuroticism and extraversion. *Curr Dir Psychol Sci* 21(5):290–296
- Keller MC, Coventry WL, Heath AC, Martin NG (2005) Widespread evidence for non-additive genetic variation in Cloninger's and Eysenck's personality dimensions using a twin plus sibling design. *Behav Genet* 35(6):707–721
- Kolen MJ, Brennan RL (2004) Test equating, scaling, and linking: methods and practices. Springer, New York
- Kuang WJ, Sun RF, Zhu YS, Li SB (2011) A new single-nucleotide mutation (rs362719) of the reelin (RELN) gene associated with schizophrenia in female Chinese Han. *Genet Mol Res* 10(3):1650–1658
- Lang UE, Sander T, Lohoff FW, Hellweg R, Bajbouj M, Winterer G, Gallinat J (2007) Association of the met66 allele of brain-derived neurotrophic factor (BDNF) with smoking. *Psychopharmacology (Berl)* 190(4):433–439
- Li MX, Kwan JS, Sham PC (2012) HYST: a hybrid set-based test for genome-wide association studies, with application to protein-protein interaction-based association analysis. *Am J Hum Genet* 91(3):478–488
- Liu Q-R, Walther D, Drgon T, Poleskaya O, Lesnick TG, Strain KJ, de Andrade M, Bower JH, Maraganore DM, Uhl GR (2005) Human brain derived neurotrophic factor (BDNF) genes, splicing patterns, and assessments of associations with substance abuse and Parkinson's Disease. *Am J Med Genet B Neuropsychiatr Genet* 134B(1):93–103
- Middeldorp CM, de Moor MH, McGrath LM, Gordon SD, Blackwood DH, Costa PT, Terracciano A, Krueger RF, de Geus EJ, Nyholt, Tanaka T, Esko T, Madden PA, Derringer J, Amin N, Willemsen G, Hottenga JJ, Distel MA, Uda M, Sanna S, Spinoven P, Hartman CA, Ripke S, Sullivan PF, Realo A, Allik J, Heath AC, Pergadia ML, Agrawal A, Lin P, Grucza RA, Widen E, Cousminer DL, Eriksson JG, Palotie A, Barnett JH, Lee PH, Luciano M, Tenesa A, Davies G, Lopez LM, Hansell NK, Medland SE, Ferrucci L, Schlessinger D, Montgomery GW, Wright MJ, Aulchenko YS, Janssens AC, Oostra BA, Metspalu A, Abecasis GR, Deary IJ, Raikonen K, Bierut LJ, Martin NG, Wray NR, van Duijn CM, Smoller JW, Penninx BW, Boomsma DI (2011) The genetic association between personality and major depression or bipolar disorder. A polygenic score analysis using genome-wide association data. *Transl Psychiatry* 1:e50
- Nadri C, Bersudsky Y, Belmaker RH, Agam G (2007) Elevated urinary ADAM12 protein levels in lithium-treated bipolar patients. *J Neural Transm* 114(4):473–477
- Ovadia G, Shifman S (2011) The genetic variation of RELN expression in schizophrenia and bipolar disorder. *PLoS One* 6(5):e19955
- Patrick CJ, Curtin JJ, Tellegen A (2002) Development and validation of a brief form of the multidimensional personality questionnaire. *Psychol Assess* 14(2):150–163
- Rettew DC, Rebollo-Mesa I, Hudziak JJ, Willemsen G, Boomsma DI (2008) Non-additive and additive genetic effects on extraversion in 3314 Dutch adolescent twins and their parents. *Behav Genet* 38(3):223–233
- Rhodes RE, Smith NEI (2006) Personality correlates of physical activity: a review and meta-analysis. *Br J Sports Med* 40(12):958–965
- Service SK, Verweij KJ, Lahti J, Congdon E, Ekelund J, Hintsanen M, Raikonen K, Lehtimäki T, Kahonen M, Widen E, Taanila A, Veijola J, Heath AC, Madden PA, Montgomery GW, Sabatti C, Jarvelin MR, Palotie A, Raitakari O, Viikari J, Martin NG, Eriksson JG, Keltikangas-Jarvinen L, Wray NR, Freimer NB

- (2012) A genome-wide meta-analysis of association studies of Cloninger's Temperament Scales. *Transl Psychiatry* 2:e116
- Sutin AR, Ferrucci L, Zonderman AB, Terracciano A (2011) Personality and obesity across the adult life span. *J Pers Soc Psychol* 101(3):579–592
- Terracciano A, Lockenhoff CE, Zonderman AB, Ferrucci L, Costa PT (2008) Personality predictors of longevity: activity, emotional stability, and conscientiousness. *Psychosom Med* 70(6):621–627
- Terracciano A, Sanna S, Uda M, Deiana B, Usala G, Busonero F, Maschio A, Scally M, Patriciu N, Chen WM, Distel MA, Slagboom EP, Boomsma DI, Villafuerte S, Sliwerska E, Burmeister M, Amin N, Janssens ACJW, van Duijn CM, Schlessinger D, Abecasis GR, Costa PT (2010) Genome-wide association scan for five major dimensions of personality. *Mol Psychiatry* 15(6):647–656
- Terracciano A, Sutin AR, An Y, O'Brien RJ, Ferrucci L, Zonderman AB, Resnick SM (2014) Personality and risk of Alzheimer's disease: new data and meta-analysis. *Alzheimer's Dement J Alzheimer's Assoc* 10(2):179–186
- Turkheimer E, Pettersson E, Horn EE (2014) A phenotypic null hypothesis for the genetics of personality. *Annu Rev Psychol* 65:515–540
- van den Berg SM, de Moor MH, McGue M, Pettersson E, Terracciano A, Verweij KJ, Amin N, Derringer J, Esko T, van Grootheest G, Hansell NK, Huffman J, Konte B, Lahti J, Luciano M, Matteson LK, Viktorin A, Wouda J, Agrawal A, Allik J, Bierut L, Broms U, Campbell H, Smith GD, Eriksson JG, Ferrucci L, Franke B, Fox JP, de Geus EJ, Giegling I, Gow AJ, Gruca R, Hartmann AM, Heath AC, Heikkila K, Iacono WG, Janzing J, Jokela M, Kiemenev L, Lehtimaki T, Madden PA, Magnusson PK, Northstone K, Nutile T, Ouwens KG, Palotie A, Pattie A, Pesonen AK, Polasek O, Pulkkinen L, Pulkki-Raback L, Raitakari OT, Realo A, Rose RJ, Ruggiero D, Seppala I, Slutske WS, Smyth DC, Sorice R, Starr JM, Sutin AR, Tanaka T, Verhagen J, Vermeulen S, Vuoksimaa E, Widen E, Willemsen G, Wright MJ, Zgaga L, Rujescu D, Metspalu A, Wilson JF, Ciullo M, Hayward C, Rudan I, Deary IJ, Raikonen K, Arias Vasquez A, Costa PT, Keltikangas-Jarvinen L, van Duijn CM, Penninx BW, Krueger RF, Evans DM, Kaprio J, Pedersen NL, Martin NG, Boomsma DI (2014) Harmonization of Neuroticism and Extraversion phenotypes across inventories and cohorts in the Genetics of Personality Consortium: an application of Item Response Theory. *Behav Genet* 44(4):295–313
- Verweij KJ, Zietsch BP, Medland SE, Gordon SD, Benjamin B, Nyholt DR, McEvoy BP, Sullivan PF, Heath AC, Madden PA, Henders AK, Montgomery GW, Martin NG, Wray NR (2010) A genome-wide association study of Cloninger's temperament scales: implications for the evolutionary genetics of personality. *Biol Psychol* 85(2):306–317
- Verweij KJ, Yang J, Lahti J, Vejjola J, Hintsanen M, Pulkki-Raback L, Heinonen K, Pouta A, Pesonen AK, Widen E, Taanila A, Isohanni M, Miettunen J, Palotie A, Penke L, Service SK, Heath AC, Montgomery GW, Raitakari O, Kahonen M, Viikari J, Raikonen K, Eriksson JG, Keltikangas-Jarvinen L, Lehtimaki T, Martin NG, Jarvelin, Visscher PM, Keller MC, Zietsch BP (2012) Maintenance of genetic variation in human personality: testing evolutionary models by estimating heritability due to common causal variants and investigating the effect of distant inbreeding. *Evolution* 66(10):3238–3251
- Vilhjalmsson B, Yang J, Finucane HK, Gusev A, Lindstrom S, Ripke S, Genovese G, Loh P-R, Bhatia G, Do R, Hayeck T, Won H-H, Genomics Consortium SWGotP, Variants in Breast Cancer study tDB, Risk of I, Kathiresan S, Pato M, Pato C, Tamimi R, Stahl E, Zaitlen N, Pasaniuc B, Schierup M, De Jager P, Patsopoulos N, McCarroll SA, Daly M, Purcell S, Chasman D, Neale B, Goddard M, Visscher PM, Kraft P, Patterson NJ, Price AL (2015) Modeling Linkage Disequilibrium Increases Accuracy of Polygenic Risk Scores. *bioRxiv*. doi:10.1101/015859
- Vinkhuyzen AA, Pedersen NL, Yang J, Lee SH, Magnusson PK, Iacono WG, McGue M, Madden PA, Heath AC, Luciano M, Payton A, Horan M, Ollier W, Pendleton N, Deary IJ, Montgomery GW, Martin NG, Visscher PM, Wray NR (2012) Common SNPs explain some of the variation in the personality dimensions of neuroticism and extraversion. *Transl Psychiatry* 2:e102
- Visscher PM, Yang J, Goddard ME (2010) A commentary on 'Common SNPs explain a large proportion of the heritability for human height' by Yang et al. (2010). *Twin Res Human Genet* 13(6):517–524
- Weiss A, Bates TC, Luciano M (2008) Happiness is a personal(ity) thing—the genetics of personality and well-being in a representative sample. *Psychol Sci* 19(3):205–210
- Yamagata S, Suzuki A, Ando J, Ono Y, Kijima N, Yoshimura K, Ostendorf F, Angleitner A, Riemann R, Spinath FM, Livesley WJ, Jang KL (2006) Is the genetic structure of human personality universal? A cross-cultural twin study from North America, Europe, and Asia. *J Pers Soc Psychol* 90(6):987–998
- Yang J, Benjamin B, McEvoy BP, Gordon S, Henders AK, Nyholt DR, Madden PA, Heath AC, Martin NG, Montgomery GW, Goddard ME, Visscher PM (2010) Common SNPs explain a large proportion of the heritability for human height. *Nat Genet* 42(7):565–569