Title: Genetic Associations with Subjective Well-Being
Also Implicate Depression and Neuroticism

Authors: All authors and their affiliations appear at the end of the paper*

*Correspondence to: Daniel Benjamin (djbenjam@usc.edu), Meike Bartels (m.bartels@vu.nl), or Philipp Koellinger (p.d.koellinger@vu.nl)

Summary: We conducted a genome-wide association study of subjective well-being (SWB) in 298,420 individuals. We also performed auxiliary analyses of depressive symptoms ("DS"; N = 161,460) and neuroticism (N = 170,910), both of which have a substantial genetic correlation with SWB ($\hat{\rho} \approx -0.8$). We identify three SNPs associated with SWB at genome-wide significance. Two of them are significantly associated with DS in an independent sample. In our auxiliary analyses, we identify 13 additional genome-wide-significant associations: two with DS and eleven with neuroticism, including two inversion polymorphisms. Across our phenotypes, loci regulating expression in central nervous system and adrenal/pancreas tissues are enriched. The discovery of genetic loci associated with the three phenotypes we study has proven elusive; our findings illustrate the payoffs from studying them jointly.

One Sentence Summary: Using both genome-wide association studies and proxy-phenotype studies, we identify genetic variants associated with subjective well-being, depressive symptoms, and neuroticism.

Main Text: Subjective well-being (SWB)—as measured by survey questions on life satisfaction, positive affect, or happiness—is a major topic of research within psychology, economics, and epidemiology and is the focus of policy initiatives instigated by many governments and international bodies (*1–3*). Twin studies suggest that roughly 35% of the variation in SWB across individuals can be explained by genetic factors (*4*). The discovery of genetic variants associated with SWB could pave the way for investigations into how environmental conditions moderate genetic effects. Moreover, it may provide insights not only into the molecular mechanisms underlying normal-range variation in SWB, but also those underlying genetically correlated phenotypes, such as depression and neuroticism (*5–7*), for which the discovery of genetic associations has proven elusive (*8, 9*).

Here, we report a series of analyses that exploit the relationships between SWB, depressive symptoms (DS), and neuroticism. Our primary analysis is a genome-wide association study (GWAS) of SWB based on data from 59 cohorts (N = 298,420). We supplement this primary analysis with auxiliary GWAS meta-analyses of DS and neuroticism. The auxiliary analyses are performed by combining publicly available summary statistics from published studies with new genome-wide analyses of additional data. All analyses in this paper are restricted to European-ancestry individuals.

Following a pre-specified analysis plan, our primary analysis combines survey measures of life satisfaction (LS) and positive affect (PA). Although LS and PA are distinguishable (*10, 11*), these two facets of SWB are phenotypically correlated and load on a common genetic factor (*12*). To reduce measurement error, cohorts with repeated measures averaged responses, and cohorts with survey questions measuring both LS and PA combined the two measures into a single variable. For details on SWB phenotype construction, see tables S1.1 and S2.1.

All SWB analyses were performed at the cohort level and subsequently meta-analyzed. Our analysis plan specified controls for age, age^2, sex, and (to reduce confounding from population stratification) four principal components of the genome-wide data (*13*). A uniform set of quality-control (QC) procedures, described in tables S2.3-2.4, were applied to the cohort-level summary statistics, including but not limited to the EasyQC protocol recently recommended (*14*).

We conducted a sample-size-weighted meta-analysis of ~2.3M HapMap2 SNPs and adjusted the standard errors using the estimated intercept from an LD Score regression (*15*). The estimated intercept implies that a modest ~5% of the observed inflation in the unadjusted mean χ^2 is accounted for by bias rather than polygenic signal (fig. S3.1). The slope estimate implies that our phenotype has a SNP-based heritability estimate of 4.0% (SE = 0.2%; cf. (*16*)).

In our primary analysis of SWB, we found three approximately independent genome-wide significant SNPs (hereafter "lead SNPs"). Fig. 1a shows the Manhattan plot from the main analysis. All three lead SNPs have estimated effects in the range 0.015 to 0.018 standard deviations (SDs) per allele ($R^2 \approx 0.01\%$). For comparison, we calculated, using publicly available survey data from the U.S. Behavioral Risk Factor Surveillance System, that the average difference in SWB between "no" and "yes" responders to "ever diagnosed with a depressive disorder" is 0.66 SDs (table S1.2). We also conducted separate meta-analyses of LS ($N = 166,205$) and PA ($N = 180,281$), using procedures identical to those described above. Consistent with our theoretical calculations (fig. S4.1), these analyses yielded fewer signals across a range of *p*-value thresholds (table S2.6). We found two lead SNPs in the LS analyses (both distinct from the SWB lead-SNPs) and none in the PA analyses.

In our auxiliary analysis of DS we combined publicly available results from a study performed by the Psychiatric Genomics Consortium (PGC) (*17*) with new results from analyses of the initial release of the U.K. Biobank data (UKB) and the Resource for Genetic Epidemiology Research on Aging (GERA) Cohort. In UKB ($N = 105,739$), our measure is constructed from responses to two questions about the frequency with which the respondent experienced feelings of hopelessness or depression in the past two weeks. The other cohorts have case-control data on major depressive disorder (GERA: $N_{cases} = 7,231$, $N_{controls} = 49,316$; PGC: $N_{cases} = 9,240$, $N_{controls} = 9,519$). In the DS meta-analysis, we weight UKB by sample size and the two case-control studies by effective sample size (*18*).

In our auxiliary analysis of neuroticism ($N = 170,910$), we pooled summary statistics from a study by the Genetics of Personality Consortium (GPC) (*8*) with results from a new analysis of UKB data. In UKB, our measure is the respondent's score on a 12-item version of the Eysenck Personality Inventory Neuroticism scale (*19*). The GPC harmonized different neuroticism batteries (table S5.2).

Figs. 1b and 1c show the Manhattan plots. For DS, we identified two genome-wide significant SNPs. For neuroticism, our meta-analysis procedure yielded 16 seemingly-independent genome-wide significant SNPs. However, 6 of these SNPs reside within a well-

known inversion polymorphism on chromosome 8 (*20*). We established that all genome-wide significant signals in the inversion region are attributable to the inversion (fig. S6.3) and confirmed that the inversion is associated with neuroticism in both the GPC and UKB data. Another SNP, rs193236081, is located on a well-known inversion polymorphism on chromosome 17, and we also established that this association is attributable to the inversion polymorphism (fig S6.6). However, because this inversion yields only one significant SNP and is genetically complex (*21*), we hereafter simply use rs193236081 as its proxy. Our neuroticism GWAS therefore identified 9 lead SNPs and 2 inversion polymorphisms. A concurrent, unpublished neuroticism GWAS using overlapping data reports similar findings (*22*).

The estimated effects of the genome-wide significant associations on DS and neuroticism are similar, with coefficient estimates in the range 0.020 to 0.031 SDs per allele ($R^2 \approx 0.02\%$). In the UKB cohort we estimated the effect of an additional allele of the chromosome 8 inversion polymorphism on neuroticism to be 0.035 SDs. For the 16 polymorphisms identified across our main and auxiliary analyses—three for SWB, two for DS, and eleven for neuroticism— Bayesian analyses show that for a wide range of priors, the evidence for association is highly credible (fig. S7.1).

In independent samples, a polygenic score constructed from our primary SWB GWAS results explains ~0.9% of the variance in SWB, ~0.5% in DS, and ~0.7% in neuroticism (table S8.1). Applying bivariate LD Score regression (*23*) to our GWAS results, we found substantial pairwise genetic correlations between SWB and our two auxiliary phenotypes: -0.81 (SE = 0.046) for DS (cf., (*24*)) and -0.75 (SE = 0.034) for neuroticism (table S9.1). As a placebo test, we also examined height, which we expected to have much less genetic overlap with SWB (*25*). Indeed, we find that the polygenic score explains ~0.2% of the variance in height, and the genetic correlation with SWB is only 0.07 (SE = 0.028).

Motivated by this evidence of genetic overlap with DS and neuroticism, which is in line with large genetic correlations estimated from twin-family studies (*5, 6*), we conducted a series of "quasi-replication" exercises in which we examined whether (i) the three genome-wide significant SNPs in the SWB analyses are associated with DS and neuroticism, and (ii) the genome-wide significant hits found in the auxiliary GWAS analyses of DS and neuroticism are associated with SWB. To avoid sample overlap, the analyses of the "second-stage" phenotype were always restricted to cohorts that did not contribute to the "first-stage" analysis.

Fig. 2 illustrates the results from these analyses. For interpretational ease, we choose reference alleles so that each SNP's estimated effect on SWB is negative; that way, its predicted effect on DS and neuroticism is positive. Panel A shows that two out of the three SWB lead-SNPs are significantly associated with DS in the predicted direction (*p* = 0.004 and *p* = 0.001), and the third has the opposite sign and is not significantly associated with DS. For neuroticism, the SWB-increasing allele has the predicted sign for all three SNPs, but none reach significance (Panel B). Panel C shows that the two DS lead-SNPs have the predicted sign for SWB, and one is nominally significant (*p* = 0.04). Finally, of the eleven polymorphisms associated with neuroticism, four have the predicted sign and are significantly associated with SWB, five have the predicted sign and are not significantly associated with SWB, and two have the opposite sign and are not significantly associated with SWB (Panel D). One of the four significant polymorphisms is the SNP tagging the inversion on chromosome 8 (*20*). That SNP's association with neuroticism (and likely with SWB) is driven by its correlation with the inversion (fig. S6.3).

Because of the strong genetic correlation between SWB and DS, the two successful quasi-replications of the SWB lead-SNPs for DS lend credibility to their truly being associated with *both* SWB and DS. Indeed, our formal Bayesian calculations (table S11.1) show that for each SNP's effect on each phenotype, we can exclude a zero effect size with greater than 95% probability. Similarly, across the 19 quasi-replication attempts depicted in fig. 2, with at least 95% probability we can exclude a zero effect size for 18 of the first-stage effects and 9 of the second-stage effects. This Bayesian analysis also yields effect-size estimates that are corrected for the winner's curse: relative to the initial estimates, on average these estimates are smaller by a factor of four for the first-stage phenotype and two for the second-stage phenotype.

We also conducted two similar "proxy-phenotype" analyses (*26*). We examine whether the independent SNPs associated with SWB at $p < 10^{-4}$ and available in the second-stage sample—163 SNPs for DS and 170 for neuroticism—are enriched for association with DS and neuroticism. For both second-stage phenotypes, the observed level of enrichment is stronger than the expected level of enrichment for a randomly drawn set of SNPs matched on allele frequency and second-stage sample size ($p < 3\times10^{-5}$ and $p < 0.04$, respectively). For DS, 116 out of 163 SNPs (71%) have the predicted direction of effect. Twenty of the 163 SNPs are significantly associated with DS at the 5% level (19 with the predicted direction), and 2 remain significant after Bonferroni correction (fig. S10.1). For neuroticism, 129 out of 170 SNPs (76%) have the predicted direction of effect, all 28 significant SNPs have the predicted sign, and 4 remain significant after Bonferroni correction (fig. S10.2). As a placebo test, we also test SWB-associated SNPs for enriched association with height. In stark contrast to what we find for DS and neuroticism, we find no evidence for enrichment and cannot reject the null hypothesis that the effects have the predicted sign only 50% of the time (*25*).

Table 1 gives a summary overview of the SNPs identified across our analyses. In the upper panel, we list all lead SNPs from the meta-analyses of SWB, DS, or neuroticism, and we indicate which have significant associations in the quasi-replication analyses. In the lower panel, we list the SNPs identified in the proxy-phenotype analyses. For each SNP, we report its estimated effect. Finally, we also report if each SNP (or a variant in strong linkage disequilibrium with that SNP) is (i) nonsynonymous (i.e., alters the protein product) or is (ii) an eQTL (i.e., affects the extent to which certain genes are expressed), and we report (iii) the nearest protein-coding gene; for details, see tables S2.6, S10.2-10.3, S12.5-12.6, and S12.8.

We further investigated the biological mechanisms underlying the GWAS results by applying stratified LD Score regression (*27*) to our meta-analysis results. The method takes as given some categorization of SNPs, and for each category, it estimates the expected increase in the explanatory power of a SNP due to its being in that category. In our first analysis, we report estimates for all 53 functional categories included in the "baseline model" (tables S12.1-12.3); the results for SWB, DS, and neuroticism are broadly similar and are in line with what has been found for other phenotypes (*27*). In our second analysis, the categories are groupings of SNPs likely to regulate gene expression in cells of a specific tissue. The estimates for SWB, DS, and neuroticism are shown in fig. 3a. As a benchmark, we also provide results from the GIANT consortium's most recent study of height (*28*). The height results are clearly distinguished by the much higher estimates for the categories CONNECTIVE/BONE, CARDIOVASCULAR, and SKELETAL MUSCLE. In contrast, the estimates for the category CENTRAL NERVOUS SYSTEM are significantly positive for all three of our phenotypes, and the estimates for ADRENAL/PANCREAS for two of the three, SWB and DS. This latter result is in line with previous research on the role of hypothalamic-pituitary-adrenal (HPA)-axis in depressed patients (*29*).

We also investigated potential molecular mechanisms underlying the effects of the inversion on chromosome 8. The chromosome 8 inversion, and the 7 genes for which we find the inversion is a significant *cis*-eQTL at FDR < 0.05, are depicted in fig. 3b. The fact that these genes are all positioned in close proximity to the inversion breakpoints suggests that the molecular mechanism could involve the relocation of regulatory sequences. Three of the genes (*FDFT1*, *MSRA*, *MTMR9*) are known to be highly expressed in tissues and cell types that belong to the central nervous system, and two (*BLK*, *MFHAS1*) in the immune system.

To summarize, we report a number of credible genetic associations with our primary phenotype of SWB as well as our auxiliary phenotypes of DS and neuroticism. The small effect sizes help shed light on why earlier studies of these phenotypes have had limited success. For example, given our winner's-curse-adjusted estimates of the effect sizes, we calculate that the power for detecting the genome-wide significant SNPs from our analysis in the largest previous GWAS of depression and neuroticism was only ~0.012% and ~0.00005%, respectively (table S11.2).

Our study found many more credible associations than prior work due to two strategies. First, our GWAS had much larger sample sizes. Our findings support the view that GWAS can succeed even for highly polygenic phenotypes once sample sizes are large enough (*9*, *30*). Second, our proxy-phenotype strategy exploited the strong genetic overlap between SWB, depression, and neuroticism. We anticipate that future efforts to identify genetic variants associated with these phenotypes could similarly benefit from studying them in concert.

References and Notes:

1.   OECD, *OECD Guidelines on Measuring Subjective Well-being* (OECD Publishing, Paris, 2013; http://dx.doi.org/10.1787/9789264191655-en).

2.   C. Deeming, Addressing the social determinants of subjective wellbeing: The latest challenge for social policy. *J. Soc. Policy*. **42**, 541–565 (2013).

3.   J. F. Helliwell, R. Layard, J. Sachs, *World Happiness Report 2015* (Sustainable Development Solutions Network, New York, 2015; http://worldhappiness.report/wp-content/uploads/sites/2/2015/04/WHR15_Sep15.pdf).

4.   M. Bartels, Genetics of wellbeing and its components satisfaction with life, happiness, and quality of life: a review and meta-analysis of heritability studies. *Behav. Genet.* **45**, 137–56 (2015).

5.   K. S. Kendler, J. Myers, The genetic and environmental relationship between major depression and the five-factor model of personality. *Psychol. Med.* **40**, 801 (2009).

6.   A. Weiss, T. C. Bates, M. Luciano, Happiness is a personal(ity) thing: The genetics of personality and well-being in a representative sample. *Psychol. Sci.* **19**, 205–10 (2008).

7.   M. Bartels, J. T. Cacioppo, T. C. E. M. van Beijsterveldt, D. I. Boomsma, Exploring the association between well-being and psychopathology in adolescents. *Behav. Genet.* **43**, 177–190 (2013).

8.   M. H. M. de Moor *et al.*, Meta-analysis of genome-wide association studies for neuroticism, and the polygenic association with Major Depressive Disorder. *JAMA Psychiatry*. **72**, 642–650 (2015).

9.   S. Hyman, Mental health: Depression needs large human-genetics studies. *Nature*. **515**, 189–191 (2014).

10.   D. Kahneman, A. Deaton, High income improves evaluation of life but not emotional well-being. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 16489–16493 (2010).

11.   D. Kahneman, J. Riis, in *The Science of Well-Being*, F. A. Huppert, N. Baylis, B. Keverne, Eds. (Oxford University Press, New York, First., 2005), pp. 285–301.

12.   M. Bartels, D. I. Boomsma, Born to be happy? The etiology of subjective well-being. *Behav. Genet.* **39**, 605–615 (2009).

13.   A. L. Price *et al.*, Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.* **38**, 904–909 (2006).

14.   T. W. Winkler *et al.*, Quality control and conduct of genome-wide association meta-analyses. *Nat. Protoc.* **9**, 1192–212 (2014).

15.   B. K. Bulik-Sullivan *et al.*, LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.* **47**, 291–295 (2015).

16.   C. A. Rietveld *et al.*, Molecular genetics and subjective well-being. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 9692–7 (2013).

17.   S. Ripke *et al.*, A mega-analysis of genome-wide association studies for major depressive disorder. *Mol. Psychiatry*. **18**, 497–511 (2013).

18.   C. J. Willer, Y. Li, G. R. Abecasis, METAL: Fast and efficient meta-analysis of genomewide association scans. *Bioinformatics*. **26**, 2190–2191 (2010).

19.   H. J. Eysenck, S. B. G. Eysenck, *Manual of the Eysenck Personality Questionnaire* (Hodder and Stroughton, London, 1975).

20. C. Tian *et al.*, Analysis and application of European genetic substructure using 300 K SNP information. *PLoS Genet.* **4**, e4 (2008).

21. K. M. Steinberg *et al.*, Structural diversity and African origin of the 17q21.31 inversion polymorphism. *Nat. Genet.* **44**, 872–80 (2012).

22. D. J. Smith *et al.*, Genome-wide analysis of over 106,000 individuals identifies 9 neuroticism-associated loci. *bioRxiv*, 032417 (2015).

23. B. Bulik-Sullivan *et al.*, An atlas of genetic correlations across human diseases and traits. *Nat. Genet.* **47**, 1236–1241 (2015).

24. C. Gale *et al.*, Pleiotropy between neuroticism and physical and mental health: findings from 108 038 menand women in UK Biobank. *bioRxiv*. **031138** (2015), doi:10.1101/031138.

25. Please see the supplementary materials.

26. C. A. Rietveld *et al.*, Common genetic variants associated with cognitive performance identified using the proxy-phenotype method. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 13790–13794 (2014).

27. H. K. Finucane *et al.*, Partitioning heritability by functional category using GWAS summary statistics. *Nat. Genet.* **47**, 1228–1235 (2015).

28. A. R. Wood *et al.*, Defining the role of common variation in the genomic and biological architecture of adult human height. *Nat. Genet.* **46**, 1173–1186 (2014).

29. C. Stetler, G. E. Miller, Depression and hypothalamic-pituitary-adrenal activation: a quantitative summary of four decades of research. *Psychosom. Med.* **73**, 114–26 (2011).

30. P. F. Sullivan, Don't give up on GWAS. *Mol. Psychiatry*. **17**, 2–3 (2012).

Authors:

Aysu Okbay[1,2,3,*], Bart M.L. Baselmans[4,5,*], Jan-Emmanuel De Neve[6,*], Patrick Turley[7,*], Michel G. Nivard[4,*], Mark A. Fontana[8,*], S. Fleur W. Meddens[9,3,10*], Richard Karlsson Linnér[9,3,10*], Cornelius A. Rietveld[1,2,3,*], Jaime Derringer[11], Jacob Gratten[127], James J. Lee[12], Jimmy Z. Liu[133], Ronald de Vlaming[1,2,3], Tarunveer S. Ahluwalia[13,14,15], Jadwiga Buchwald[16], Alana Cavadino[17,97], Alexis C. Frazier-Wood[18], Nicholas A. Furlotte[19], Victoria Garfield[120], Marie Henrike Geisel[44], Juan R. Gonzalez[20,21,22], Saskia Haitjema[111], Robert Karlsson[23], Sander W. van der Laan[111], Karl-Heinz Ladwig[24], Jari Lahti[25,26,77], Sven J. van der Lee[2], Penelope A. Lind[27], Tian Liu[28,29], Lindsay Matteson[12], Evelin Mihailov[30], Michael B. Miller[12], Camelia C. Minica[4], Ilja M. Nolte[34], Dennis Mook-Kanamori[31,32,33], Peter J. van der Most[34], Christopher Oldmeadow[35,36], Yong Qian[37], Olli Raitakari[38,39], Rajesh Rawal[40], Anu Realo[41], Rico Rueedi[42,43], Börge Schmidt[44], Albert V. Smith[45,46], Evie Stergiakouli[47], Toshiko Tanaka[60], Kent Taylor[48], Juho Wedenoja[16], Juergen Wellmann[49], Harm-Jan Westra[50], Sara M. Willems[2], Wei Zhao[51], Behrooz Z. Alizadeh[34,52], Najaf Amin[2], Andrew Bakshi[127], Patricia A. Boyle[53], Simon Cox[54,55], Gail Davies[54,55], Oliver S.P. Davis[47], Jun Ding[37], Nese Direk[2], Peter Eibich[56,57], Rebecca T. Emeny[24], Ghazaleh Fatemifar[58], Jessica D. Faul[59], Luigi Ferrucci[60], Andreas J. Forstner[61,62], Christian Gieger[40], Tamara B. Harris[63], Juliette M. Harris[78], Elizabeth G. Holliday[35,36], Jouke-Jan Hottenga[4,5], Philip L. De Jager[64,65,66], Marika A. Kaakinen[67,70], Eero Kajantie[68,69], Ville Karhunen[70], Ivana Kolcic[114], Meena Kumari[71], Lenore J. Launer[72], LifeLines Cohort Study[73], Ruifang Li-Gao[31], Marisa Loitfelder[117], Anu Loukola[16], Pedro Marques-Vidal[74], Grant W. Montgomery[75], Lavinia Paternoster[47], Alison Pattie[55], Katja E. Petrovic[76], Laura Pulkki-Råback[25,77], Lydia Quaye[78], Katri Räikkönen[25], Igor Rudan[79], Rodney J. Scott[80,36], Jennifer A. Smith[51], Angelina R. Sutin[121,60], Maciej Trzaskowski[113,127], Anna E. Vinkhuyzen[127], Lei Yu[81], Delilah Zabaneh[113], John R. Attia[35,36], David A. Bennett[81], Klaus Berger[49], Lars Bertram[82,83], Dorret I. Boomsma[4,5,136], Ute Bultmann[84], Shun-Chiao Chang[85], Francesco Cucca[86], Ian J. Deary[54,55], Cornelia M. van Duijn[2], Johan G. Eriksson[87,88,89], Lude Franke[90], Eco J.C. de Geus[4,5,136], Patrick J.F. Groenen[3,91], Vilmundur Gudnason[45,46], Torben Hansen[14], Catharine A. Hartman[110], Claire M.A. Haworth[47], Caroline Hayward[92,93], Andrew C. Heath[94], David A. Hinds[19], Elina Hyppönen[95,96,97], William G. Iacono[12], Marjo-Riitta Järvelin[98,99,70,100], Karl-Heinz Jöckel[44], Jaakko Kaprio[16,101,102], Sharon L.R. Kardia[51], Liisa Keltikangas-Järvinen[25], Peter Kraft[103], Laura Kubzansky[104], Terho Lehtimäki[105,106], Patrik K.E. Magnusson[23], Nicholas G. Martin[107], Matt McGue[12], Andres Metspalu[30,108], Melinda Mills[109], Renée de Mutsert[31], Albertine J. Oldehinkel[110], Gerard Pasterkamp[111,112], Nancy L. Pedersen[23], Robert Plomin[113], Ozren Polasek[114], Christine Power[97], Stephen S. Rich[115], Frits R. Rosendaal[31], Hester M. den Ruijter[111], David Schlessinger[37], Helena Schmidt[116,117], Rauli Svento[118], Reinhold Schmidt[117], Harold Snieder[34], Thorkild I.A. Sørensen[14,47,119], Tim D. Spector[78], Andrew Steptoe[120], Antonio Terracciano[121,60], A. Roy Thurik[1,3,122,123], Nicholas J. Timpson[47], Henning Tiemeier[2,124,125], André G. Uitterlinden[2,3,126], Peter Vollenweider[74], Gert Wagner[56], David R. Weir[59], Jian Yang[127,128], Dalton C. Conley[129,130], George Davey Smith[47], Albert Hofman[2], Magnus Johannesson[131], David I. Laibson[7], Sarah E. Medland[27], Michelle N. Meyer[132], Joseph K. Pickrell[133,134], Tõnu Esko[30], Robert F. Krueger[12,#], Jonathan P. Beauchamp[7,#], Philipp D. Koellinger[9,3,10,#], Daniel J. Benjamin[135,#], Meike Bartels[4,5,136,#], David Cesarini[137,138,#]

9

1. Department of Applied Economics, Erasmus School of Economics, Erasmus University Rotterdam, 3062 PA, Rotterdam, The Netherlands

2. Department of Epidemiology, Erasmus Medical Center, Rotterdam, 3015 GE, The Netherlands

3. Erasmus University Rotterdam Institute for Behavior and Biology, Rotterdam, 3062 PA, The Netherlands

4. Department of Biological Psychology, Vrije Universiteit, Amsterdam, 1081 BT, The Netherlands

5. EMGO+ Institute for Health and Care Research, Amsterdam, 1081 BT, The Netherlands

6. Saïd Business School, University of Oxford, Oxford, OX1 1HP, UK

7. Department of Economics, Harvard University, Cambridge, MA 02138, USA

8. Department of Economics, University of Michigan, Ann Arbor, MI 48109-1220, USA

9. Department of Complex Trait Genetics, VU University, Center for Neurogenomics and Cognitive Research, Amsterdam, 1081 HV, The Netherlands

10. Amsterdam Business School, University of Amsterdam, Amsterdam, 1018 TV, The Netherlands

11. Psychology, University of Illinois, IL 61820, Champaign, USA

12. Department of Psychology, University of Minnesota Twin Cities, Minneapolis, MN 55455, USA

13. COPSAC, Copenhagen Prospective Studies on Asthma in Childhood, Herlev and Gentofte Hospital, University of Copenhagen, Copenhagen, 2820, Denmark

14. The Novo Nordisk Foundation Center for Basic Metabolic Research, Section of Metabolic Genetics, University of Copenhagen, Faculty of Health and Medical Sciences, Copenhagen, 2100, Denmark

15. Steno Diabetes Center, Gentofte, 2820, Denmark

16. Department of Public Health, University of Helsinki, Helsinki, FI-00014, Finland

17. Centre for Environmental and Preventive Medicine, Wolfson Institute of Preventive Medicine, Queen Mary University of London, London EC1M 6BQ, UK

18. USDA/ARS Children's Nutrition Research Center, Baylor College of Medicine, Houston, TX 77030, USA

19. 23andMe, Inc., Mountain View, CA 94041, USA

20. Centre for Research in Environmental Epidemiology, Institute for Global Health, Barcelona, Spain

21. Universitat Pompeu Fabra, Barcelona, Spain

22. CIBER Epidemiología y Salud Pública, Barcelona, Spain

23. Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, 171 77, Sweden

24. Institute of Epidemiology II, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, 85764, Germany

25. IBS, Unit of Personality, Work and Health, Institute of Behavioural Sciences, P.O. Box 9, 00014 University of Helsinki, Finland

26. Folkhälsan Research Centre, Helsingfors, FI-00014, Finland

27. Quantitative Genetics, QIMR Berghofer Medical Research Institute, Brisbane, QLD 4029, Australia

28. Lifespan Psychology, Max Planck Institute for Human Development, Berlin, 14195, Germany

29. Max Planck Institute for Molecular Genetics, Department of Vertebrate Genomics, Berlin, 14195, Germany

30. Estonian Genome Center, University of Tartu, Tartu, 51010, Estonia

31. Clinical Epidemiology, Leiden University Medical Center, Leiden, 2300 RC, The Netherlands

32. Public Health and Primary Care, Leiden University Medical Center, Leiden, 2300 RC, The Netherlands

33. BESC, King Faisal Specialist Hospital and Research Centre, Riyadh 11211, Saudi Arabia

34. Department of Epidemiology, University of Groningen, University Medical Center Groningen, Groningen, 9700 RB, The Netherlands

35. Public Health Stream, Hunter Medical Research Institute, New Lambton, NSW 2305, Australia

36. Faculty of Health and Medicine, University of Newcastle, Newcastle, NSW 2300, Australia

37. Laboratory of Genetics, National Institute on Aging, Baltimore, MD 21224, USA

38. Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku, Turku, 20521, Finland

39. Department of Clinical Physiology, Turku University Hospital, Turku 20520, Finland

40. Research Unit of Molecular Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, 85764, Germany

41. Department of Psychology, University of Tartu, Tartu 50409, Estonia

42. Department of Medical Genetics, University of Lausanne, Lausanne, 1005, Switzerland

43. Swiss Institute of Bioinformatics, Lausanne, 1015, Switzerland

44. Institute for Medical Informatics, Biometry and Epidemiology, University Hospital of Essen, Essen, 45147, Germany

45. Icelandic Heart Association, Kopavogur, 201, Iceland

46. Faculty of Medicine, University of Iceland, Reykjavik, 101, Iceland

47. MRC Integrative Epidemiology Unit, University of Bristol, Bristol, BS8 2BN, UK

48. Los Angeles Biomedical Research Institute and Department of Pediatrics, Harbor-UCLA, Torrence, 90505 CA, USA

49. Institute of Epidemiology and Social Medicine, University of Muenster, Muenster, 48149, Germany

50. Divisions of Genetics and Rheumatology, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115, USA

51. Department of Epidemiology, University of Michigan, Ann Arbor, MI 48104, USA

52. Department of Gastroenterology and Hepatology, University of Groningen, University Medical Center Groningen, Groningen, 9713 GZ, The Netherlands

53. Department of Behavioral Sciences, Rush University Medical Center, Chicago, IL 60612, USA

54. Centre for Cognitive Ageing and Cognitive Epidemiology, University of Edinburgh, Edinburgh, EH8 9JZ, UK

55. Department of Psychology, University of Edinburgh, Edinburgh, EH8 9JZ, UK

56. German Socio-Economic Panel Study, DIW Berlin, Berlin, 10117, Germany

57. Health Economics Research Centre, Nuffield Department of Population Health, University of Oxford, Oxford, OX3 7LF, UK

58. The Farr Institute of Health Informatics, University College London, London NW1 2DA, UK

59. Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor, MI 48104, USA

60. National Institute on Aging, National Institutes of Health, Baltimore, MD 21224, USA

61. Institute of Human Genetics, University of Bonn, Bonn, 53127, Germany

62. Department of Genomics, Life and Brain Center, University of Bonn, Bonn, 53127, Germany

63. Laboratory of Epidemiology, Demography, National Institute on Aging, National Institutes of Health, Bethesda, MD 20892-9205, United States

64. Program in Translational NeuroPsychiatric Genomics, Departments of Neurology & Psychiatry, Brigham and Women's Hospital, Boston, MA 02115, USA

65. Harvard Medical School, Boston, MA 02115, USA

66. Program in Medical and Population Genetics, Broad Institute of Harvard and Massachusetts Institute of Technology, Cambridge, MA, 02142, USA

67. Department of Genomics of Common Disease, Imperial College London, London, W12 0NN, UK

68. Department of Pediatrics, University of Helsinki, Helsinki, Finland

69. National Institute for Health and Welfare, Helsinki, Finland

70. Center for Life Course Health Research, University of Oulu; Oulu University Hospital, Oulu, Finland

71. Institute for Social & Economic Research, University of Essex, Wivenhoe Park CO4 3SQ, UK

72. Neuroepidemiology Section, National Institute on Aging, National Institutes of Health, Bethesda, MD 20892-9205, USA

73.   LifeLines Cohort Study, University of Groningen, University Medical Center Groningen, Groningen, 9713 BZ, The Netherlands

74.   Department of Internal Medicine, Internal Medicine, Lausanne University Hospital (CHUV), Lausanne, 1011, Switzerland

75.   Molecular Epidemiology, QIMR Berghofer Medical Research Institute, Brisbane, QLD 4029, Australia

76.   Centre for Genomic and Experimental Medicine, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, EH4 2XU, UK

77.   Helsinki Collegium for Advanced Studies, University of Helsinki, Helsinki 00014, Finland

78.   Department of Twin Research and Genetic Epidemiology, King's College London, London, SE1 7EH, UK

79.   Centre for Global Health Research, The Usher Institute for Population Health Sciences and Informatics, University of Edinburgh, Edinburgh, EH8 9AG, UK

80.   Information Based Medicine Stream, Hunter Medical Research Institute, New Lambton, NSW 2305, Australia

81.   Department of Neurological Sciences, Rush University Medical Center, Chicago, IL 60612, USA

82.   Lübeck Interdisciplinary platform for Genome Analytics (LIGA), Institutes of Neurogenetics and Integrative & Experimental Genomics, University of Lübeck, Lübeck, 23562, Germany

83.   Neuroepidemiology and Ageing Research Unit, School of Public Health, Faculty of Medicine, Imperial College, London SW7 2AZ, UK

84.   Department of Health Sciences, Community & Occupational Medicine, University of Groningen, University Medical Center Groningen, Groningen, 9713 AV, The Netherlands

85.   Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital, Boston, MA 02115, USA

86.   Istituto di Ricerca Genetica e Biomedica (IRGB), Consiglio Nazionale delle Ricerche, c/o Cittadella Universitaria di Monserrato, Monserrato, Cagliari, 9042, Italy

87.   Department of General Practice and Primary Health Care, University of Helsinki, Helsinki, 00014, Finland

88.   Department of Chronic Disease Prevention, National Institute for Health and Welfare, Helsinki, Finland

89.   Unit of General Practice, University Central Hospital, Helsinki, Finland

90.   Department of Genetics, University Medical Center Groningen, University of Groningen, Groningen. 9700 RB, The Netherlands

91.   Econometric Institute, Erasmus School of Economics, Erasmus University Rotterdam, Rotterdam, 3062 PA, The Netherlands

92.   MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, EH4 2XU, UK

93. Generation Scotland, Centre for Genomics and Experimental Medicine, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, EH4 2XU, UK

94. Department of Psychiatry, Washington University School of Medicine, St. Louis, MO 63110, USA

95. Centre for Population Health Research, School of Health Sciences and Sansom Institute, University of South Australia, SA5000, Adelaide, Australia

96. South Australian Health and Medical Research Institute, Adelaide, SA5000, Australia

97. Population, Policy and Practice, UCL Institute of Child Health, London, WC1N 1EH, UK

98. Department of Epidemiology and Biostatistics, MRC Health Protection Agency (HPE) Centre for Environment and Health, School of Public Health, Imperial College London, UK

99. Biocenter Oulu, University of Oulu, Oulu, Finland

100. Unit of Primary Care, Oulu University Hospital, Oulu, Finland

101. Institute for Molecular Medicine Finland (FIMM), University of Helsinki, Helsinki, 00014, Finland

102. Department for Health, THL-National Institute for Health and Welfare, Helsinki, FI-00271, Finland

103. Department of Epidemiology and Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA 02115, USA

104. Department of Social and Behavioral Sciences, Harvard T.H. Chan School of Public Health, Boston, MA 02115, USA

105. Fimlab Laboratories, Tampere, 33520, Finland

106. Department of Clinical Chemistry, University of Tampere, School of Medicine, Tampere, 33014, Finland

107. Genetic Epidemiology, QIMR Berghofer Medical Research Institute, Brisbane, QLD 4029, Australia

108. Institute of Molecular and Cell Biology, University of Tartu, Tartu, 51010, Estonia

109. Department of Sociology, University of Oxford, Oxford OX1 3UQ, UK

110. Department of Psychiatry, University of Groningen, University Medical Center Groningen, 9700 RB, Groningen, The Netherlands

111. Laboratory of Experimental Cardiology, University Medical Center Utrecht, Utrecht, 3584 CX, The Netherlands

112. Laboratory of Clinical Chemistry and Hematology, Division Laboratories and Pharmacy, University Medical Center Utrecht, Utrecht, 3584 CX, The Netherlands

113. Social, Genetic and Developmental Psychiatry Centre, King's College London, De Crespigny Park SE5 8AF, UK

114. Department of Public Health, Faculty of Medicine, University of Split, Croatia, Split 21000, Croatia

115. Department of Public Health Sciences, University of Virginia, Charlottesville, VA 22904, USA

116. Research Unit for Genetic Epidemiology, Institute of Molecular Biology and Biochemistry, Center of Molecular Medicine, General Hospital and Medical University, Graz, Graz, 8010, Austria

117. Department of Neurology, General Hospital and Medical University Graz, Graz, 8036, Austria

118. Department of Economics, Oulu Business School, Oulu, Finland

119. Institute of Preventive Medicine, Bispebjerg and Frederiksberg Hospitals, The Capital Region, Frederiksberg, 2000, Denmark

120. Department of Epidemiology & Public Health, University College London, London WC1E 7HB, UK

121. Department of Geriatrics, Florida State University College of Medicine, Tallahassee, FL 32306, USA

122. Montpellier Business School, Montpellier, 34080, France

123. Panteia, Zoetermeer, 2715 CA, The Netherlands

124. Department of Psychiatry, Erasmus Medical Center, Rotterdam, 3015 GE, The Netherlands

125. Department of Child and Adolescent Psychiatry, Erasmus Medical Center, Rotterdam, 3015 GE, The Netherlands

126. Department of Internal Medicine, Erasmus Medical Center, Rotterdam, 3015 GE, The Netherlands

127. Queensland Brain Institute, The University of Queensland, Brisbane, QLD 4072, Australia

128. The University of Queensland Diamantina Institute, The Translation Research Institute, Brisbane, QLD 4102, Australia

129. Department of Sociology, New York University, New York, NY 10012, USA

130. School of Medicine, New York University, NY 10016, New York, USA

131. Department of Economics, Stockholm School of Economics, Stockholm, 113 83, Sweden

132. Bioethics Program, Union Graduate College - Icahn School of Medicine at Mount Sinai, Schenectady, NY 12308, USA

133. New York Genome Center, New York, NY 10013, USA

134. Department of Biological Sciences, Columbia University, 600 Fairchild Center, New York, NY 10027, USA

135. Center for Economic and Social Research, University of Southern California, Los Angeles, CA 90089-3332, USA

136. Neuroscience Campus Amsterdam, Amsterdam, The Netherlands

137. Department of Economics, New York University, New York, NY 10012, USA

138. Research Institute for Industrial Economics, Stockholm, 10215, Sweden

Fig. 1. Manhattan plots. (A) Subjective well-being (N = 298,420), **(B)** Depressive symptoms (N_{eff} = 149,707), **(C)** Neuroticism (N = 170,911). The x-axis is chromosomal position, and the y-axis is the significance on a $-\log_{10}$ scale. The upper dashed line marks the threshold for genome-wide significance ($p = 5\times10^{-8}$); the lower line marks the threshold for nominal significance ($p = 10^{-5}$). Each independent genome-wide significant association ("lead SNP") is marked by ×. We find three associations for SWB, two for DS, and eleven for neuroticism.
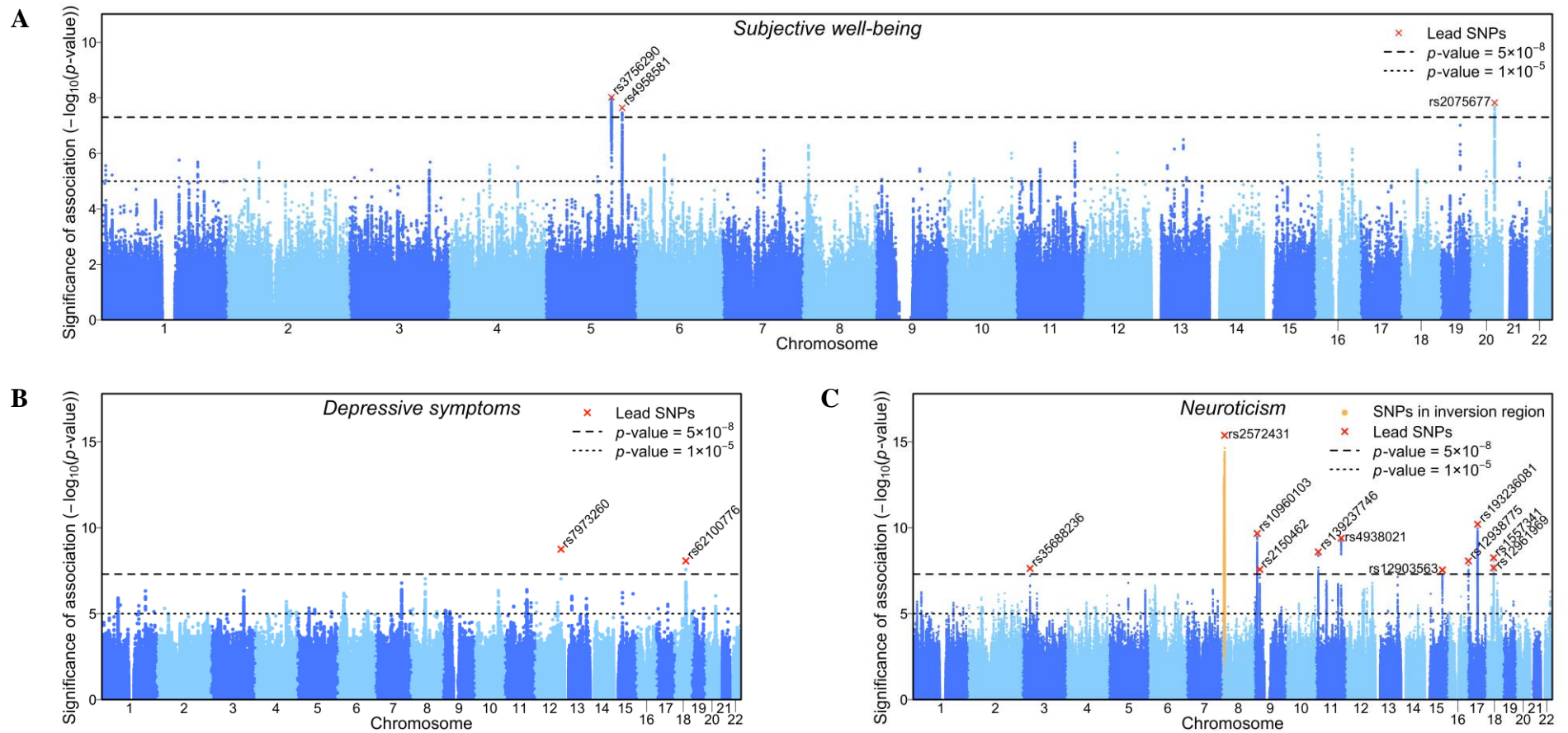
Fig. 2. Quasi-replication of lead SNPs. We examined whether (i) lead SNPs identified in the SWB meta-analyses are associated with DS and neuroticism, and (ii) lead SNPs identified in the analyses of DS and neuroticism are associated with SWB. The quasi-replication sample is always restricted to non-overlapping cohorts. The bars represent 95% CIs (not adjusted for multiple testing). The sample sizes (maximum across SNPs) are: **(A)** SWB (N = 294,043) → DS (N_{eff} = 123,506), **(B)** SWB (N = 294,043) → Neuroticism (N = 68,201), **(C)** DS (N_{eff} = 124,498) → SWB (N = 238,254), **(D)** Neuroticism (N = 170,908) → SWB (N = 198,360). For interpretational ease, we choose the reference allele so that positive coefficients imply that the estimated effect is in the predicted direction.
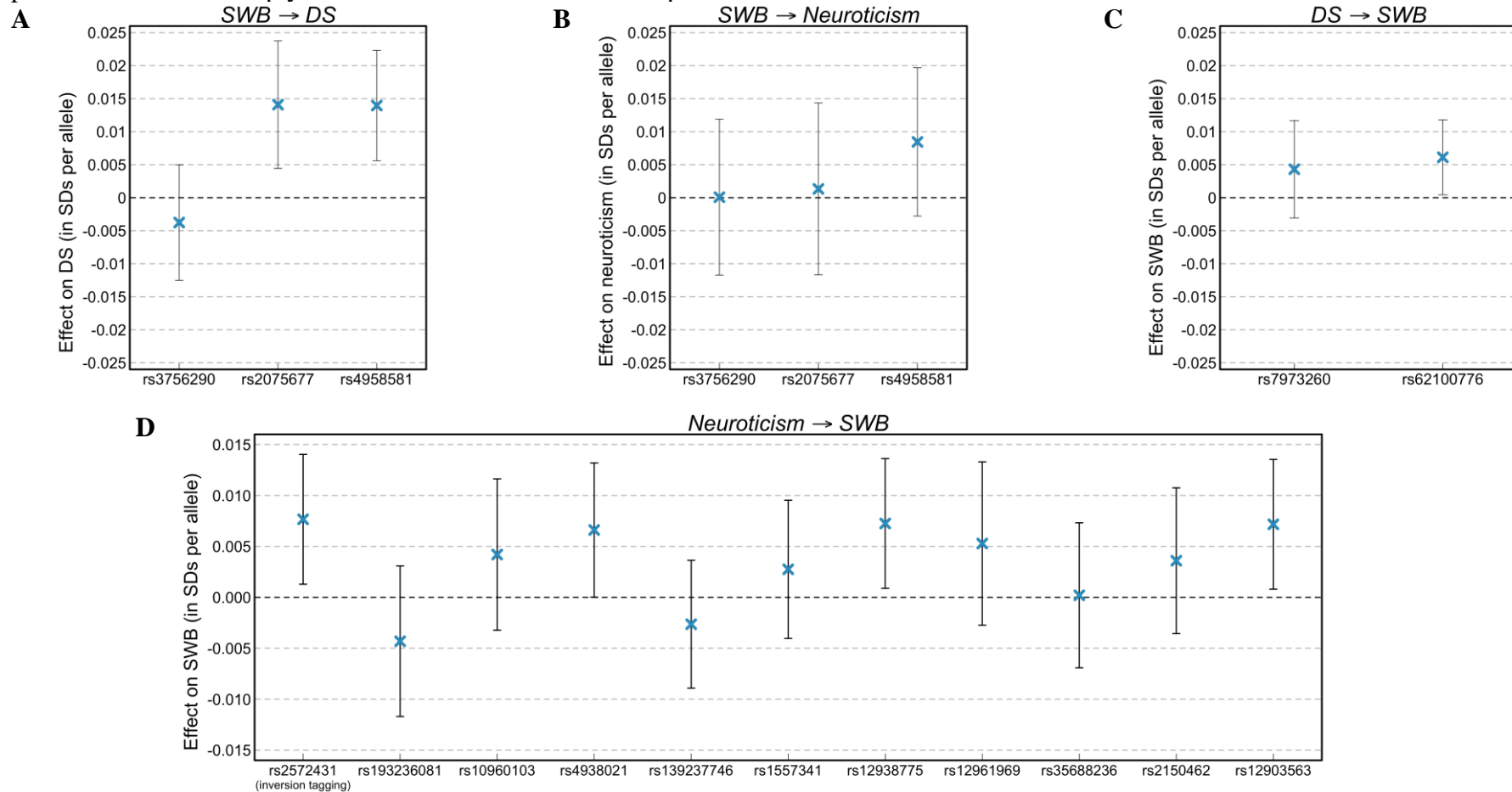
Fig. 3. Results from selected biological analyses. (**A**) Estimates of the expected increase in the phenotypic variance accounted for by a SNP due to the SNP's being in a given category (τ_c), divided by the LD Score heritability of the phenotype (h^2). Each estimate of τ_c comes from a separate stratified LD Score regression, controlling for the 52 functional annotation categories in the "baseline model." The bars represent 95% CIs (not adjusted for multiple testing). To benchmark the estimates, we compare them to those obtained from a recent study of height (*28*). (**B**) Inversion polymorphism on chromosome 8 and the 7 genes for which the inversion is a significant *cis*-eQTL at FDR < 0.05. The upper half of the figure shows the Manhattan plot for neuroticism for the inversion and surrounding regions. The bottom half shows the squared correlation between the SNPs and the principal component that captures the inversion. The inlay plots the relationship, for each SNP in the inversion region, between the SNP's significance and its squared correlation with the principal component that captures the inversion.

18
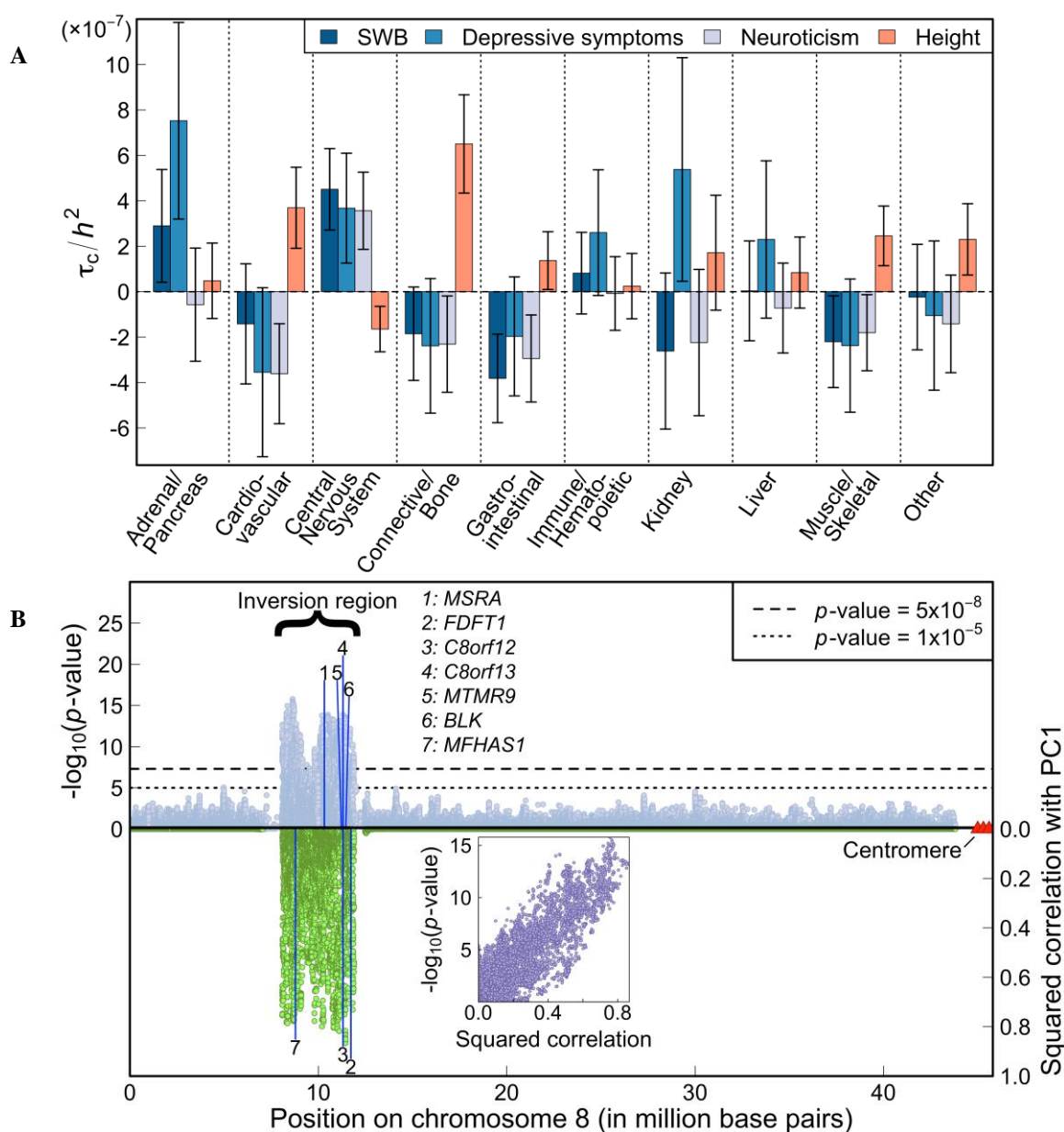
Table 1. Summary Overview of Polymorphisms Identified Across Analyses. All effect sizes are reported in units of SDs per allele. "Q-rep": Whether SWB lead-SNP quasi-replicated in DS or neuroticism, and whether DS/neuroticism lead-SNP quasi-replicated in SWB, all at $p < 0.05$. "Location": type of genomic region according to dbsnp and Ensembl. "Non-syn": Whether the SNP or any LD partner ($R^2 \geq 0.6$ and within 250kb) is non-synonymous according to HaploReg (version 4). "eQTL": Whether the SNP is a *cis*-eQTL in whole blood (at FDR < 0.05). "Nearest protein-coding gene": The nearest-protein coding gene (5' start site) of the lead SNP according to GENCODE coordinates archived at http://www.broadinstitute.org/mpg/snpsnap/. Proxy SNP rs4346787 ($R^2 = 0.98$) was used to tag rs4481363 in the DS data. All details are in SOM Section 12.2.

[*] Note: rs10960103 tags the inversion polymorphism on chromosome 8.

Panel A. Genome-Wide Significant Associations

SWB ($N = 298,420$)

SNP	Beta	SE	p-value	Q-rep	Location	Non-syn	eQTL	Nearest protein-coding gene
rs3756290	-0.0177	0.0031	9.55×10^{-9}		intronic		Y	*RAPGEF6*
rs2075677	0.0175	0.0031	1.49×10^{-8}	DS	intronic	Y	Y	*CSE1L*
rs4958581	0.0153	0.0028	2.29×10^{-8}	DS	intronic			*NMUR2*

DS ($N = 161,460$)

SNP	Beta	SE	p-value	Q-rep	Location	Non-syn	eQTL	Nearest protein-coding gene
rs7973260	0.0306	0.0051	1.78×10^{-9}		intronic		Y	*KSR2*
rs62100776	-0.0252	0.0044	8.45×10^{-9}	SWB	intronic			*DCC*

Neuroticism ($N = 107,245$)

SNP	Beta	SE	p-value	Q-rep	Location	Non-syn	eQTL	Nearest protein-coding gene
rs2572431	0.0283	0.0035	4.20×10^{-16}		intergenic		Y	*MTMR9*
rs193236081	-0.0284	0.0043	6.26×10^{-11}		intronic	Y	Y	*KANSL1*
rs10960103[*]	0.0264	0.0042	2.14×10^{-10}	SWB	intergenic			*TYRP1*
rs4938021	0.0233	0.0337	4.03×10^{-10}		intergenic			*DRD2*
rs139237746	-0.0204	0.0034	2.55×10^{-9}	SWB	intronic		Y	*SBF2*
rs1557341	0.0213	0.0036	5.58×10^{-9}		intronic			*CELF4*
rs12938775	-0.0202	0.0035	8.54×10^{-9}	SWB	intronic			*PAFAH1B1*
rs12961969	0.0253	0.0045	2.16×10^{-8}		intergenic			*CELF4*
rs35688236	0.0213	0.0038	2.35×10^{-8}		intronic			*PDCD6IP*
rs2150462	-0.0217	0.0039	2.66×10^{-8}	SWB	intergenic			*ELAVL2*
rs12903563	0.0198	0.0036	2.86×10^{-8}		intronic			*LINGO1*

Panel B. Additional DS and Neuroticism SNPs Identified Via Proxy-Phenotype Analysis

SWB ($N = 278,956$) → DS ($N_{eff} = 111,106$)

SNP	Beta_{DS}	SE_{DS}	p-value$_{DS}$	Location	Non-syn	eQTL	Nearest protein-coding gene
rs4481363	0.0136	0.0038	3.06×10^{-4}	intronic			*ZNF184*
rs6904596	-0.023	0.0059	9.79×10^{-5}	exonic		Y	*MAT2B*

SWB ($N = 239,982$) → Neuroticism ($N = 131,867$)

SNP	Beta_{Neur}	SE_{Neur}	p-value$_{Neur}$	Location	Non-syn	eQTL	Nearest protein-coding gene
rs4481363	0.0151	0.0040	1.86×10^{-4}	intronic			*ZNF184*
rs6904596	-0.0264	0.0072	2.49×10^{-4}	exonic		Y	*MAT2B*
rs10838738	0.0178	0.0039	5.03×10^{-6}	intronic		Y	*MTCH2*
rs10774909	-0.015	0.0039	1.20×10^{-4}	intronic			*NOS1*