Supplementary Note for 'Association analysis in over 329,000 individuals identifies 116 independent variants influencing neuroticism'

Luciano et al.

Supplementary Note

UK Biobank Sample

The sample was drawn from an open resource, UK Biobank, established to study determinants of disease in middle-aged and older adults [2]. Recruitment to the study occurred between 2006 and 2010 targeting both urban and rural community-dwelling individuals with a broad range of socio-economic circumstances. This resulted in 502,655 participants assessed at baseline on a range of cognitive and other psychological measures, physical functioning, physical and mental health, lifestyle variables, and biological samples (including blood, urine, and saliva). Restricting the sample to only those with White British ancestry and high quality genotyping, the genome-wide association analysis was performed on 329,821 participants (152,710 male) aged 39-73 years (mean 56.9, SD 8.0 years). Ethical approval for UK Biobank was received from the Research Ethics Committee (REC reference 11/NW/0382).

Independent replication of the top association signals was sought from the 23andMe and the Genetics of Personality Consortium (GPC-2) meta-analysis of genome-wide association results for neuroticism [3]. The meta-analysis sample included 59,225 unrelated

individuals with European ancestry from the US consumer genomics company, 23andMe, and 63,661 individuals with European ancestry from 29 European, Australian and US cohorts forming the GPC-2. GWA results for 10,037,522 SNPs were available for look-up.

Neuroticism Measurement in UK Biobank

A 12-item scale (Eysenck Personality Questionnaire-Revised (EPQ-R) Short Form) was used to measure the personality trait of Neuroticism [1]. This required binary responses (1-yes; 0 –no) to the following items: 1) *Does your mood often go up and down?*; 2) *Do you ever feel 'just miserable' for no reason?*; 3) *Are you an irritable person?*; 4) *Are your feelings easily hurt?*; 5) *Do you often feel 'fed-up'?*; 6) *Would you call yourself a nervous person?*; 7) *Are you a worrier?*; 8) *Would you call yourself tense or 'highly strung'?*; 9) *Do you worry too long after an embarrassing experience?*; 10) *Do you suffer from 'nerves'?*; 11) *Do you often feel lonely?*; and 12) *Are you often troubled by feelings of guilt?* Scores from this neuroticism scale show high internal consistency and concurrent validity [1,4]; importantly, genetic correlations between the short and long EPQ forms exceed 0.90 and, between the EPQ and more widely-used scales, such as the NEO Personality Inventory, are upwards of 0.82 [5].

Imputation of Neuroticism Data in UK Biobank

There were 501,278 individuals with neuroticism data in the UK Biobank sample. Of these, 99,604 provided 'Do not Know' or 'Prefer not to Answer' for between 1 and 12 neuroticism items. 'Do not Know' responses were more frequent than 'Prefer not to Answer', and there was variation in the endorsement of these between items (see Table S1). No single item had more than 5% of these non-usable responses (3.9% in the genotyped sample). Individuals who had more than four 'Do not Know' or 'Prefer not to Answer' responses were set to missing (n = 4683; 0 in the genotyped sample), because their overall score was considered unreliable. The remaining missing items (affecting 94,921 cases; 59,762 cases in the genotyped sample) were imputed by a logistic regression multiple imputation procedure including sex and age as predictors using the mice package in R [6]. A total neuroticism score was then calculated by summing the responses to the 12 items, with a Yes response scoring 1, and a No response scoring 0. In comparison with the subsample of complete data, the median of the imputed neuroticism variable was the same (4), but the mean was slightly higher (4.12 ±3.27 vs 4.28 ±3.28). Age and sex (0 = female, 1 = male) correlations were .02 and -.15,

respectively. The positive skew of neuroticism scores in the UK Biobank was comparable to a Scottish study of 21,340 individuals (aged 18 to 99 years) measured on the same scale and reported a similar mean and variability: 3.9 (±3.2) [7]. Other scales of neuroticism (e.g., IPIP Big Five Factor markers, NEO Five Factor Inventory) show more normal distributions of scores in UK samples, yet they correlate strongly (upwards of .80) with the EPQ-R Short Form [4]. In the GWA sample, neuroticism was residualized for the effects of age, sex, assessment centre, genotype batch, array, and 40 genetic principal components. This residual score (mean 0, SD = 3.21) correlated 0.98 with the raw imputed neuroticism score; the distribution of scores is shown in Supplementary Figure 1.

Neuroticism Measurement in the Replication Cohorts

23andMe

Neuroticism was measured using a web-based version of the Big Five Inventory (BFI) [8,9]. This test includes 8 neuroticism questions (five scored positively and three scored negatively), each of which had five possible answers (Strongly disagree, disagree a little, neither agree nor disagree, agree a little, or strongly agree) and were scored from 0 to 4. A total of N = 59,206 took this test and had genetic data available for analysis.

The Genetics of Personality Consortium (GPC)

Neuroticism was measured in the GPC-2 [10] by harmonising the measures used across the 9 inventories (NEO Personality Inventory—PI-R, FFI, FFI-30—, Eysenck Personality Questionnaire—EPQ-R, JEPQ, EPI—, the International Personality Item Pool inventory, all item data for harm avoidance from the Cloninger's Tridimensional Personality Questionnaire, and all item data for negative emotionality from the Multidimensional Personality Questionnaire) and 29 cohorts (from Europe, the United States, and Australia) using Item Response Theory. This prevented neuroticism scores being affected by differences in each inventory, as well as the number of items used in each inventory. The final sample size for neuroticism was 63,661.

Genotyping and Imputation

UK Biobank

Full details of the UK Biobank genotyping procedure can be found elsewhere [11]. In short, two custom genotyping arrays were used to genotype 49,950 participants (UK BiLEVE Axiom Array) and 438,427 participants (UK Biobank Axiom Array) [11,12]. Genotype data (805,426 markers) were available for 488,377 individuals, with imputation to the HRC reference panel (39,131,578 autosomal SNPs) available for 487,442 individuals in this study [11]. Allele frequency checks [13] against the HRC [14] and 1000G [15] site lists were run and any variants with minor allele frequencies (MAF) that differed more than +/- 0.2 from the reference sets were removed.

Downstream quality control steps for the present study involved excluding (1) those with non-British ancestry based on self-report and a principal components analysis, (2)

extreme scores based on heterozygosity and missingness, (3) individuals with neither XX nor XY sex chromosomes, (4) individuals whose reported sex was inconsistent with genetically inferred sex, and (5) individuals with >10 putative third degree relatives from the kinship table. This left 408,095 individuals. Related individuals were removed based on a genetic relationship threshold of 0.025 ascertained using GCTA-GREML on 131,790 reportedly-related participants [16]. After implementing these steps, the sample size was 332,050, with neuroticism data available for 329,821 individuals. The following GWA quality control thresholds were applied: minor allele frequency > 0.0005 (i.e., a minimum allele count of 164, comparable with a previous GWA of neuroticism in a subsample of UK Biobank [17]), imputation quality score > 0.1, and inclusion of bi-allelic SNPs only; this resulted in 18,485,882 analysed autosomal SNPs. For the gene-based analysis, only common (MAF > 1%) SNPs were used.

23andMe

Summary GWA results for neuroticism were provided by 23andMe for data imputed to the 1000 Genomes Project phase 1 version 3 reference panel [18]. Details of the genotyping procedure, quality control parameters, genotype imputation and association analysis can be found in Lo and colleagues [3]. SNPs were filtered by Hardy–Weinberg equilibrium ($P < 10^{-20}$), call rate < 95%, and allele frequencies different to those found in the European 1000 Genomes Project reference panel. Following the removal of SNPs on chromosomes X and Y, and on mitochondria, a total of 13,341,935 SNPs were retained. Additional quality control was applied to the current study where a MAF of 0.5% was applied and only SNPs with an

average imputation r^2 of > .5 across all batches of results were included. This resulted in a final sample size of 9,763,840 autosomal SNPs.

Genetics of Personality Consortium (GPC-2)

Genotyping was performed on the cohorts of the GPC-2 [10,19] using an Illumina or an Affymetrix platform. Genotype data underwent quality control within each cohort separately. Quality control consisted of checks for deviation from European ancestry, inconsistencies regarding sex, Mendelian errors, genome-wide homozygosity, checks on relatedness, as well as minor allele frequencies, SNP call rates, sample call rate, and Hardy–Weinberg equilibrium. The genotype data in the GPC-2 were imputed using the reference panel from 1000 Genomes Project phase 1 version 3 [10,19,20]. Following imputation, additional quality control was performed with poorly-imputed SNPs being removed ($r^2 < 0.4$) along with those with a low MAF (< sqrt(5/N)).

Data Sources for LDSC

GWA summary statistics for performing LDSC with UK Biobank neuroticism GWA results are shown in Supplementary Table 11 (extracted from a table originally published in *Translational Psychiatry,* [21] and updated here with further psychiatric, mental and physical health traits).

Replicating SNPs previously identified for Neuroticism

The two SNPs—rs6981523 and rs9611519—previously identified for neuroticism in 23andMe [3] were similarly significant in our larger sample, with respective p-values of 4.7×10^{-22} and 1.17×10^{-10} and consistent direction of allelic effect. SNP rs35855737, significant in the GPC-2 GWA of neuroticism [22], was not significant, P = .069. Previous association of SNPs within the 8p23.1 inversion region was even stronger in our study, with the lead SNPs, rs2572431 [23] and rs12682352 [17], showing respective p-values of 1.33×10^{-18} and 1.11×10^{-24}. The 8p23.1 locus was previously cited as important in developmental neuropsychiatric disorders [24] which may be relevant for personality traits, which emerge early in life, and could therefore be influenced by genes that are expressed during development.

Replicating SNPs previously identified for MDD

Neuroticism may be a more tractable trait for the genetic study of major depressive disorder, which is affected by noise in diagnosis and other factors (e.g., lifestyle, socio-demographic factors and co-morbidity) that influence clinical status. Of the 44 recently discovered SNP variants for major depressive disorder [25], 36 were available in our GWA: 5

were significant at a genome-wide level, 4 at a suggestive level, and a further 18 at a nominal .05 alpha level (See Supplementary Table 4). Further, the genes in which three of our top replicated SNPs resided were highlighted among the MDD results: *CACNA1E* from their gene-based tests, *LINC00461* from their SNP association results, and *CELF4* from their pathway analysis [25]. Given our estimate of a .69 genetic correlation between neuroticism and MDD (consistent with [25]), then this partial overlap between our results is expected. Overlapping genetic findings from both neuroticism and MDD studies are likely to be the most reliable candidates to take forward in MDD research.

Relation of Replicated Genetic Loci to Inversions on Chromosomes 8 and 17

The 8p23.1 locus has been linked to an inversion polymorphism (based partly on a subsample of UK Biobank), so we used the lead tagging SNP (rs13270267) of this inversion already established in UK Biobank [23] to evaluate whether our 5 replicated loci were independent of this inversion. We found that only rs10097870 was in high LD (r^2 = .98) with the tag SNP (and this was one of the 20 SNPs showing the strongest correlation with the inversion principal component [23]); the other four replicated loci in the extended inversion

region showed r^2 ranging between .15 and .43 (see Supplementary Fig. 5) suggesting that the inversion—previously associated with neuroticism—may not fully explain these signals. However, given the complex nature of LD within inversion regions we will cautiously treat this as a single locus in line with previous findings [23]. Our replicated chromosome 17 locus is not in the region of the previously associated inversion with neuroticism [23]; although its tag SNP, rs79959255, was in high LD (r^2 = .89) with our second strongest independent SNP, rs77804065, which was nominally significant in 23andMe.

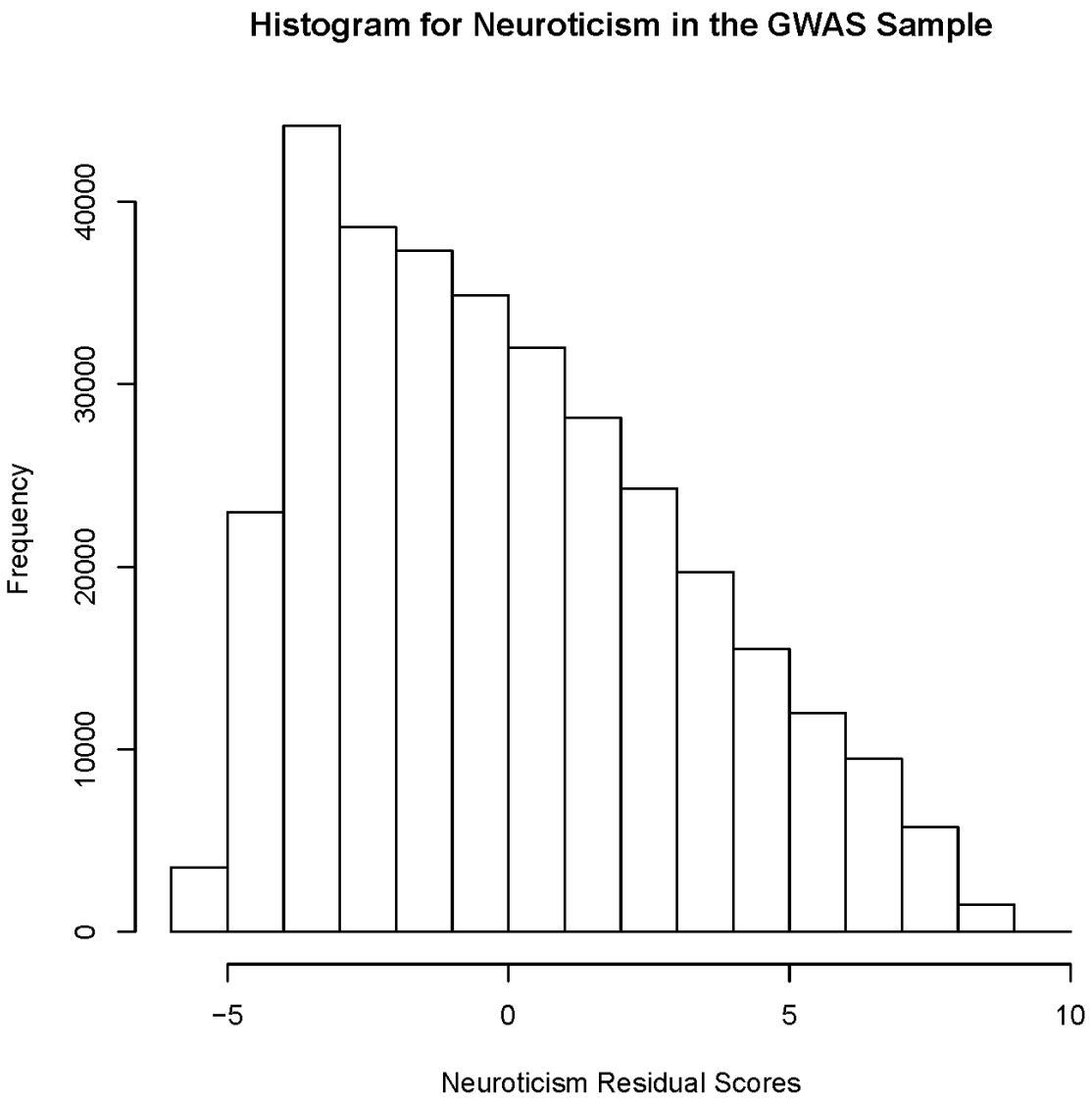
References

1. Eysenck, S.B.G., Eysenck, H.J. & Barrett, P. A revised version of the psychoticism scale. *Personality and Individual Differences* **6**, 21-29 (1985).
2. Sudlow, C. *et al.* UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age. *PLOS Medicine* **12**, e1001779 (2015).
3. Lo, M.-T. *et al.* Genome-wide analyses for personality traits identify six genomic loci and show correlations with psychiatric disorders. *Nat Genet* **49**, 152-156 (2017).
4. Gow, A.J., Whiteman, M.C., Pattie, A. & Deary, I.J. Goldberg's 'IPIP' Big-Five factor markers: Internal consistency and concurrent validation in Scotland. *Personality and Individual Differences* **39**, 317-329 (2005).
5. Wray, N.R., Birley, A.J., Sullivan, P.F., Visscher, P.M. & Martin, N.G. Genetic and Phenotypic Stability of Measures of Neuroticism Over 22 Years. *Twin Research and Human Genetics* **10**, 695-702 (2012).
6. van Buuren, S. & Groothuis-Oudshoorn, K. mice: Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software; Vol 1, Issue 3 (2011)* (2011).
7. Smith, B.H. *et al.* Cohort Profile: Generation Scotland: Scottish Family Health Study (GS:SFHS). The study, its participants and their potential for genetic research on health and illness. *International Journal of Epidemiology* **42**, 689-700 (2013).
8. John, O.P., Donahue, E.M. & Kentle, R.L. The big five inventory: Versions 4a and 54, institute of personality and social research. *University of California, Berkeley, CA* (1991).
9. Soto, C.J. & John, O.P. Ten facet scales for the Big Five Inventory: Convergence with NEO PI-R facets, self-peer agreement, and discriminant validity. *Journal of Research in Personality* **43**, 84-90 (2009).
10. van Den Berg, S.M. *et al.* Meta-analysis of Genome-Wide Association Studies for Extraversion: Findings from the Genetics of Personality Consortium. *Behavior genetics* **46**, 170-182 (2016).
11. Bycroft, C. *et al.* Genome-wide genetic data on ~500,000 UK Biobank participants. *bioRxiv* (2017).

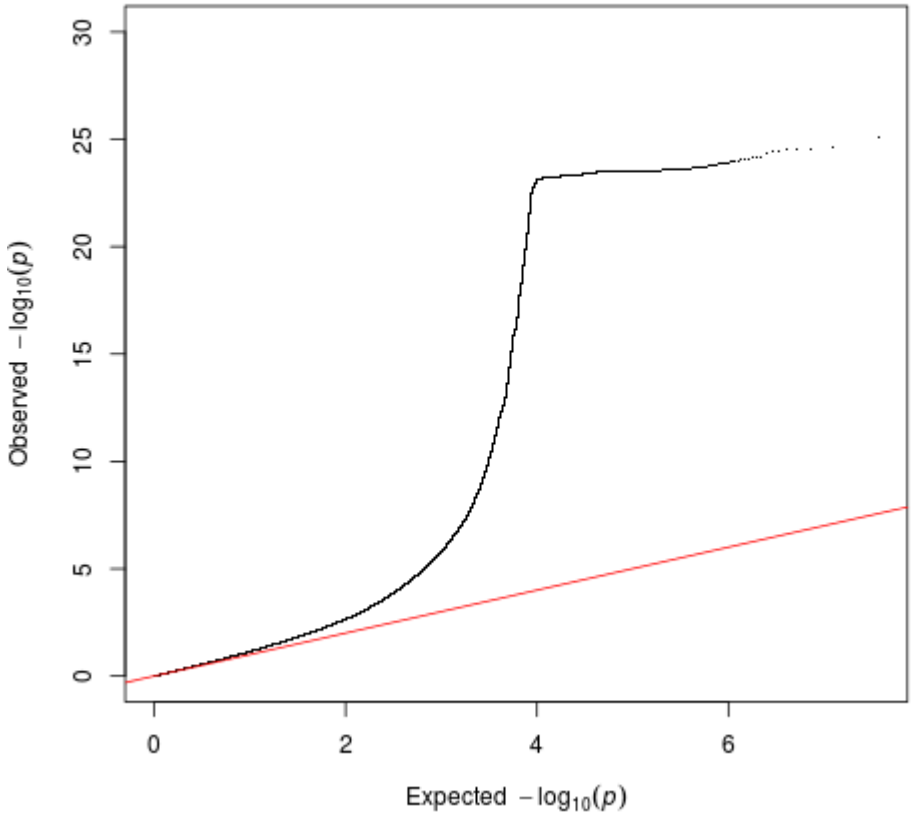
12. Wain, L.V. *et al.* Genome-wide association analyses for lung function and chronic obstructive pulmonary disease identify new loci and potential druggable targets. *Nat Genet* **49**, 416-425 (2017).
13. Winkler, T.W. *et al.* Quality control and conduct of genome-wide association meta-analyses. *Nature protocols* **9**, 1192 (2014).
14. Haplotype Reference, C. A reference panel of 64,976 haplotypes for genotype imputation. *Nature genetics* **48**, 1279-1283 (2016).
15. Genomes Project, C. An integrated map of genetic variation from 1,092 human genomes. *Nature* **491**, 56 (2012).
16. Yang, J., Lee, S.H., Goddard, M.E. & Visscher, P.M. GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet* **88**, 76-82 (2011).
17. Smith, D.J. *et al.* Genome-wide analysis of over 106 000 individuals identifies 9 neuroticism-associated loci. *Mol Psychiatry* **21**, 749-57 (2016).
18. 1000 Genomes Project Consortium. A map of human genome variation from population-scale sequencing. *Nature* **467**, 1061-1073 (2010).
19. De Moor, M.H. *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).
20. 1000 Genomes Project Consortium. A map of human genome variation from population-scale sequencing. *Nature* **467**, 1061-1073 (2012).
21. Gale, C.R. *et al.* Pleiotropy between neuroticism and physical and mental health: findings from 108038 men and women in UK Biobank. *Translational Psychiatry* **6**(2016).
22. Genetics of Personality, C. *et al.* Meta-analysis of Genome-wide Association Studies for Neuroticism, and the Polygenic Association With Major Depressive Disorder. *JAMA Psychiatry* **72**, 642-50 (2015).
23. Okbay, A. *et al.* Genetic variants associated with subjective well-being, depressive symptoms, and neuroticism identified through genome-wide analyses. *Nat Genet* **48**, 624-33 (2016).
24. Tabares-Seisdedos, R. & Rubenstein, J.L.R. Chromosome 8p as a potential hub for developmental neuropsychiatric disorders: implications for schizophrenia, autism and cancer. *Mol Psychiatry* **14**, 563-589 (2009).

25. Wray, N.R. & Sullivan, P.F. Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *bioRxiv* (2017).
26. Sekar, A. *et al.* Schizophrenia risk from complex variation of complement component 4. *Nature* **530**, 177-183 (2016).
27. Cross-Disorder Group of the Psychiatric Genomics, C. Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. *Lancet* **381**, 1371-1379 (2013).

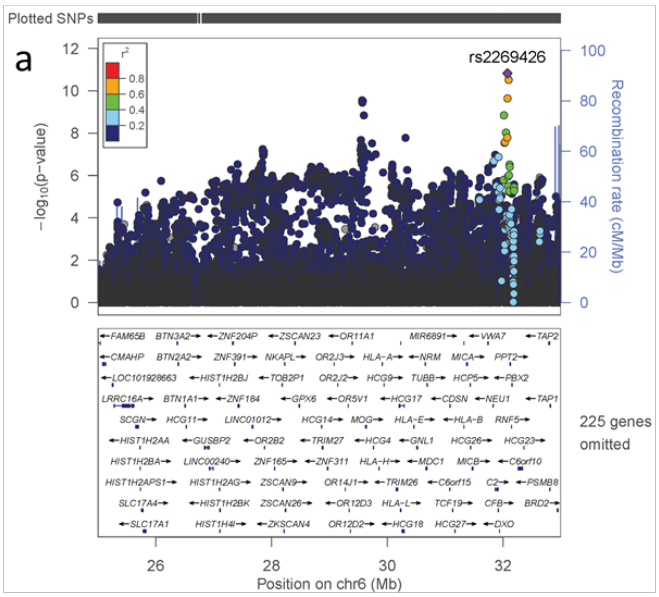
Supplementary Figure 1. Distribution of age-, sex-, assessment centre-, genotype batch-, array-, and 40 genetic principal components-residualized neuroticism scores in the GWA sample of 329,821 individuals.

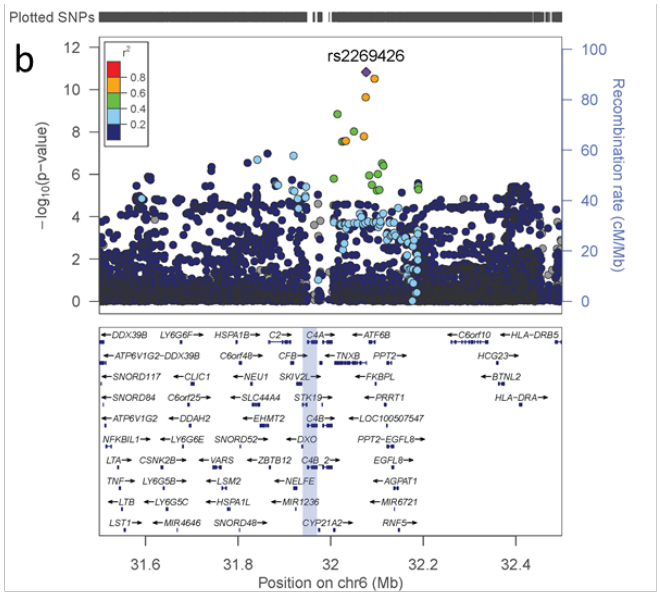


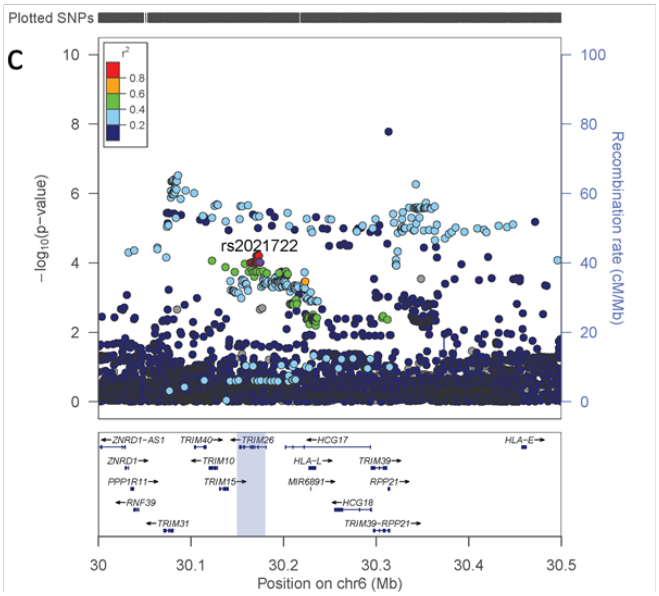
Supplementary Figure 2. Quantile-quantile plot for the GWA of neuroticism in UK Biobank. Lambda = 1.15.

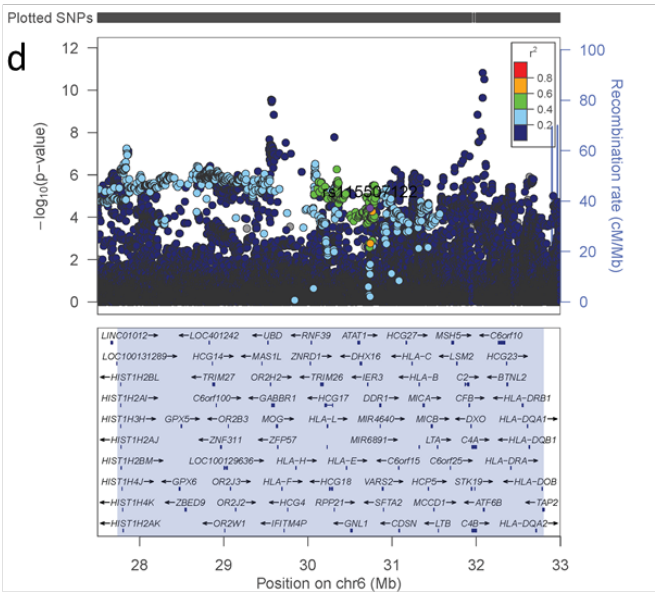


Supplementary Figure 3. Regional association plots of UK Biobank results for a) the entire MHC region on chromosome 6, b) the *C4* gene region previously associated with schizophrenia [26], c) region surrounding rs2021722, the primary marker associated with schizophrenia [27], and d) region surrounding rs115507122, the primary marker associated with MDD [25].

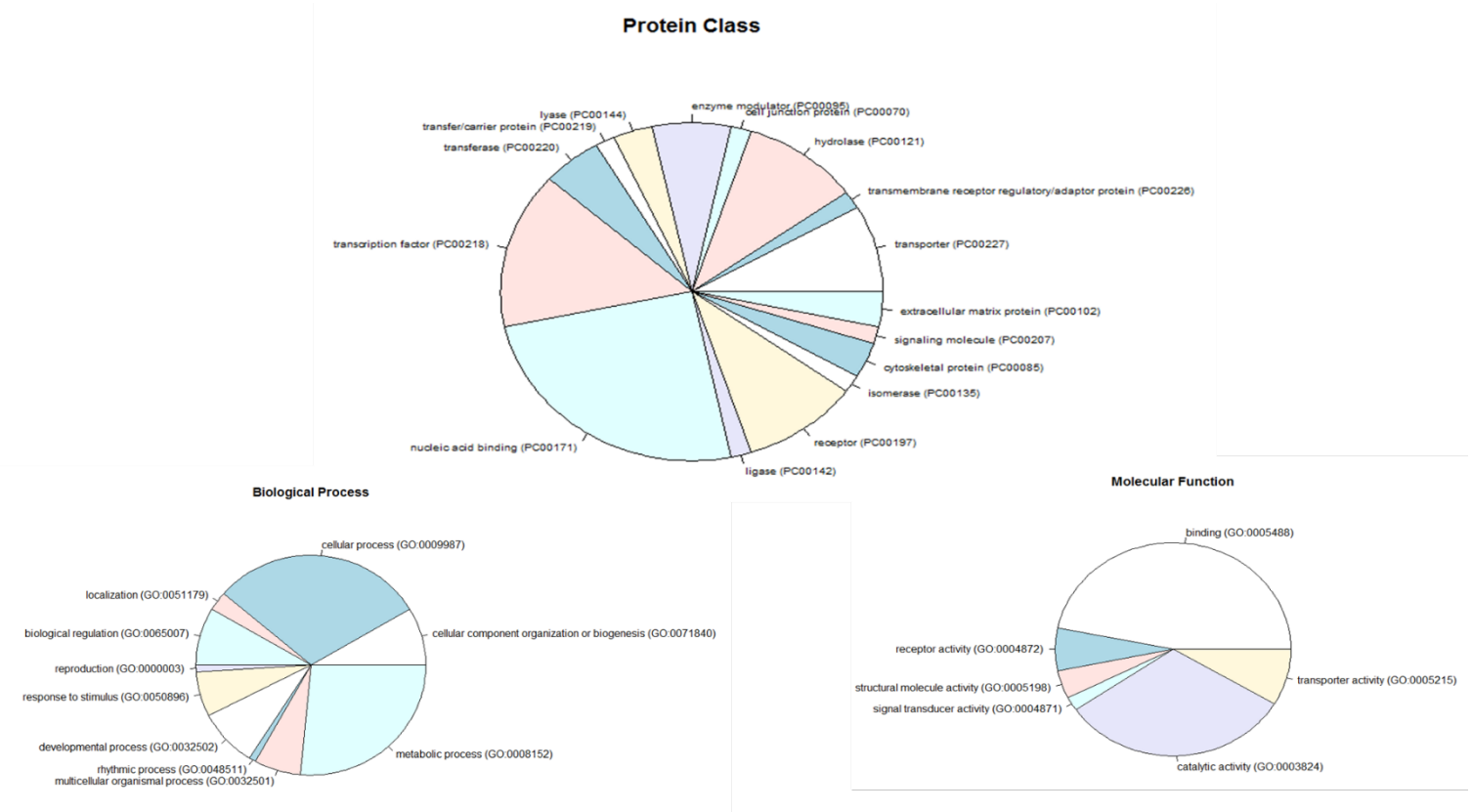

Plotted SNPs
a
rs2269426
225 genes omitted
Position on chr6 (Mb)
Recombination rate (cM/Mb)


Plotted SNPs
b
rs2269426
Position on chr6 (Mb)
Recombination rate (cM/Mb)

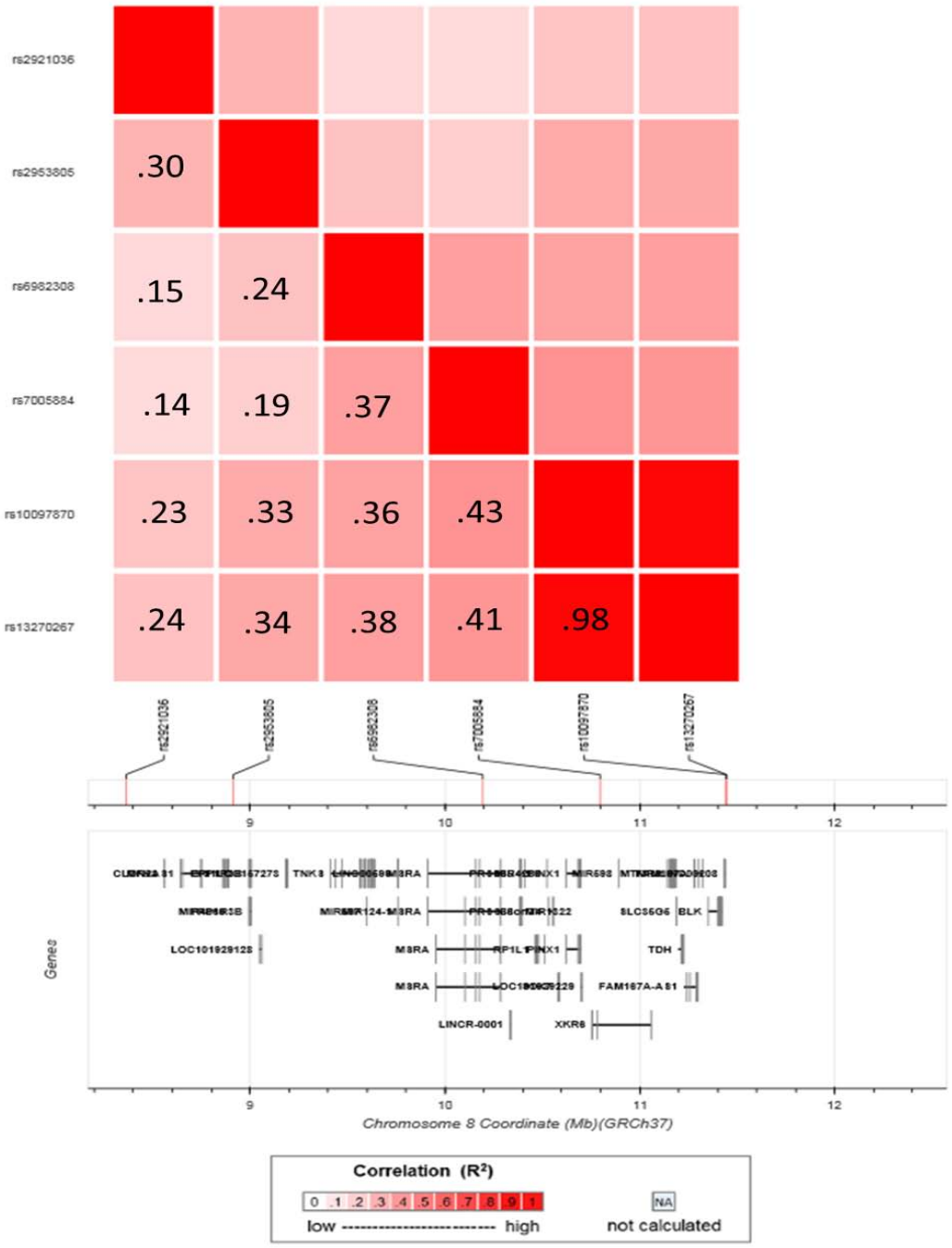

Plotted SNPs
c
rs2021722
Position on chr6 (Mb)
Recombination rate (cM/Mb)


Plotted SNPs
d
Position on chr6 (Mb)
Recombination rate (cM/Mb)

Supplementary Figure 4. Pie charts of the diverse protein classes (60 class hits), biological processes (106 process hits) and molecular functions (47 function hits) of the 69 genes located in the region of the 15 replicated loci associated with neuroticism. Classification is based on PANTHER Version 12.0 (http://pantherdb.org/) and is presented for descriptive purpose.



Supplementary Figure 5. Linkage disequilibrium (r^2) among 5 replicated loci on chromosome 8 and with inversion tagging SNP, rs13270267, estimated in the Great British subpopulation from 1000G (http://analysistools.nci.nih.gov/LDlink/.**).**



Supplementary Table 1. Eysenck Personality Questionnaire-Revised Short Form (EPQ-R-S) neuroticism scale [1] items, their associated UK Biobank data-field, and number of imputed 'Do not Know' and 'Prefer not to Answer' responses.

		UK Biobank data-field	Do not Know	Prefer not to Answer
1	Does your mood often go up and down?	1920	12448 (2.5%)	965 (0.19%)
2	Do you ever feel 'just miserable' for no reason?	1930	8459 (1.69%)	1033 (0.21%)
3	Are you an irritable person?	1940	22362 (4.46%)	1267 (0.25%)
4	Are your feelings easily hurt?	1950	14365 (2.87%)	1064 (0.21%)
5	Do you often feel 'fed-up'?	1960	10288 (2.05%)	1378 (0.27%)
6	Would you call yourself a nervous person?	1970	13233 (2.64%)	794 (0.16%)
7	Are you a worrier?	1980	12955 (2.58%)	1027 (0.20%)
8	Would you call yourself tense or 'highly strung'?	1990	17951 (3.58%)	980 (0.19%)
9	Do you worry too long after an embarrassing	2000	20447	993

	experience?		(4.08%)	(0.20%)
10	Do you suffer from 'nerves'?	2010	18687 (3.73%)	1043 (0.21%)
11	Do you often feel lonely?	2020	7534 (1.50%)	1435 (0.29%)
12	Are you often troubled by feelings of guilt?	2030	12860 (2.56%)	1636 (0.33%)

Supplementary Table 2. Genome-wide significant ($P < 5 \times 10^{-8}$) and suggestive ($P < 1 \times 10^{-5}$) SNP association results for the GWA of neuroticism in UK Biobank. [see Online Excel file]

Supplementary Table 3. Genome-wide significant ($p < 5 \times 10^{-8}$) 1000G SNP association results for the GWA of neuroticism in UK Biobank with gene annotation. [see Online Excel file]

Supplementary Table 4. Look-up of 44 major depressive disorder significant SNPs from Wray et al (2017) in UK Biobank neuroticism GWA. [see Online Excel file]

Supplementary Table 5. One hundred and sixteen independent genome-wide significant ($p < 5 \times 10^{-8}$) SNPs associated with neuroticism in UK Biobank and the extent of their

LD across genes. Replication p-value in 23andMe and GPC is also shown (those significant at P<.00045 are in bold). [see Online Excel file]

Supplementary Table 6. Genome-wide significant (p < 5 x 10-8) SNP association results for the GWA of neuroticism in UK Biobank and their effect size and significance in the meta-analysis of 23andMe and GPC cohorts, and in the meta-analysis of UK Biobank and the replication cohorts. [see Online Excel file]

Supplementary Table 7. PsyGeNET (v2.0) lookups for the 69 genes located in the region of the 15 replicated loci associated with neuroticism. Four genes were found in this genetic association database which covers eight classes of psychiatric disorder. [see Online Excel file]

Supplementary Table 8. GTEx results (indicating the gene regulated, associated p-value, tissue type) for the the 15 replicated SNPs associated with neuroticism. Brain expressed genes are shown in bold. [see Online Excel file]

Supplementary Table 9. The most affected genes regulated by each of the 15 replicated SNPs in each brain region as identified by searching the brain eQTL database BRAINEAC (http://www.braineac.org/). [see Online Excel file]

Supplementary Table 10. Significant gene-based results ($P < 2.77 \times 10^{-6}$) for neuroticism in UK Biobank. [see Online Excel file]

Supplementary Table 11. Sources of GWA results from consortia investigating psychiatric disease, mental and physical health.

Health Trait	Consortium/Cohort	URL	Reference	No. of individuals in GWAS
ADHD	Psychiatric Genetics Consortium (PGC)	https://www.med.unc.edu/pgc/downloads	Demontis, D et al. (2017). Discovery Of The First Genome-Wide Significant Risk Loci For ADHD. *bioRxiv*.	19,099 cases 34,194 controls
Alzheimer's disease	International Genomics of Alzheimer's Project (IGAP)	http://www.pasteur-lille.fr/en/recherche/u744/igap/igap_download.php	Lambert et al. (2013) Nat Genet; 45: 1452-1458.	17,008 cases, 37,154 controls
Anorexia nervosa	Genetic Consortium for Anorexia Nervosa (GCAN)	http://www.med.unc.edu/pgc/downloads	Boraska et al. (2014) Molecular psychiatry 19(10): 1085-1094.	2,907 cases 14,860 controls
Bipolar disorder	Psychiatric Genetics Consortium (PGC)	https://www.med.unc.edu/pgc/downloads	Psychiatric GWAS Consortium Bipolar Disorder Working Group. (2011) Nat Genet; 43: 977-983.	7,481 cases, 9,250 controls
BMI	GIANT	http://www.broadinstitute.org/collaboration/giant/index.php/GIANT_consortium_data_files	Locke et al. (2015) Nature; 518: 197-206.	339,224
Coronary artery disease	CARDIoGRAM	http://www.cardiogramplusc4d.org/downloads/	Schunkert et al. (2011) Nat Genet; 43: 333-338.	22,233 cases, 64762 controls
Depressive symptoms	Social Science Genetic Association Consortium	https://www.thessgac.org/data	Okbay et al. (2016) Nat Genet 48, 624–633	161,460
Educational attainment	Social Science Genetic Association Consortium	https://www.thessgac.org/data	Okbay et al. (2016) Nature 533, 539–542	217,568
HbA1c	MAGIC consortium	https://www.magicinvestigators.org/downloads/	Soranzo et al. Diabetes 2010; 59: 3229-3239	46,368
HDL/LDL cholesterol/triglycerides	Global Lipids Consortium	http://csg.sph.umich.edu//abecasis/public/lipids2013/	Willer et al. Nat Genet 2013; 45: 1274-1283	188,577
Major depressive disorder (ICD-10)	UK Biobank	NA	Howard et al (2017) Biorxiv doi: 10.1101/168732	8,276 cases, 209,308 controls
Neuroticism	Genetics of Personality Consortium	http://www.tweelingenregister.org/GPC/	De Moor et al. (2015) JAMA Psychiatry, 72(7): 642-650.	63,661
Neuroticism (Smith 2016)	UK Biobank	NA	Smith et al (2016) Molecular Psychiatry, 21(6): 749–757.	91,370
Rheumatoid arthritis	NA	ftp://ftp.broadinstitute.org/pub/rheumatoid_arthritis/Stahl_etal_2010NG	Stahl et al (2010) Nat Genet 42, 508–514	5,539 cases, 20,169 controls

Schizophrenia	Psychiatric Genetics Consortium (PGC)	https://www.med.unc.edu/pgc/downloads	Schizophrenia Working Group of the Psychiatric Genomics Consortium. Nature 2014; 511: 421-427.	36,989 cases, 113,075 controls
Self-rated health	UK Biobank	http://www.psy.ed.ac.uk/ccace/downloads/Harris2016_IJE_self_rated_health.zip	Harris et al (2016) International Journal of Epidemiology doi:10.1093/ije/dyw219	111,749
Smoking status	TAG	https://www.med.unc.edu/pgc/files/resultfiles/tag.evrsmk.tbl.gz	Furberg et al Nature Genetics 2010; 42: 441-447	74,053
Subjective wellbeing	Social Science Genetic Association Consortium	https://www.thessgac.org/data	Okbay et al. (2016) Nat Genet 48, 624–633	298,420
Tiredness	UK Biobank	http://www.psy.ed.ac.uk/ccace/downloads/Deary2017_Mol_Psych_tiredness.zip	Deary et al (2017) Molecular Psychiatry doi:10.1038/mp.2017.5	108,976
Type 2 diabetes	DIAGRAM	http://diagram-consortium.org/downloads.html	Morris et al. Nat Genet 2012; 44: 981-990. PMID: 22885922	12,171 cases, 56,862 controls
Waist-hip ratio	GIANT	http://portals.broadinstitute.org/collaboration/giant/index.php/GIANT_consortium_data_files#GWAS_Anthropometric_2015_Waist	Shungin D et al. (2015). Nature 518: 187-196.	224,459

Supplementary Table 12. Genetic correlations between neuroticism and 22 health outcomes. The heritability Z-score and the mean χ2 indicate the level of power to detect association where a heritability Z-score of >4 and a mean χ2 >1.02 being considered well powered. Significant genetic correlations (FDR p-value threshold =< 0.0037) are highlighted in bold.

Trait	Genetic correlation	Standard error	P-value	Heritability Z-score	Mean χ2
ADHD	0.22	0.03	**1.67×10^{-10}**	15.15	1.3
Alzheimer's disease	0.12	0.06	0.053	2.11	1.11
Alzheimer's disease (500kb)	0.10	0.05	0.0386	5.39	1.11
Anorexia nervosa	0.18	0.03	**6.02×10^{-11}**	18.55	1.05
Bipolar disorder	0.11	0.04	**0.0037**	10.45	1.19
BMI	-0.01	0.02	0.6342	18.43	1.26
Coronary artery disease	0.08	0.04	0.0524	7.47	1.15
Depressive symptoms	0.82	0.03	**4.73×10^{-223}**	12.2	1.16
Educational attainment	-0.20	0.02	**5.99×10^{-21}**	28.63	1.65
HbA1c	0.04	0.05	0.4485	5.36	1.06
HDL cholesterol	-0.01	0.02	0.7921	6.53	1.21
LDL cholesterol	0.03	0.03	0.2393	4.17	1.19
Major depressive disorder	0.69	0.07	**1.13×10^{-21}**	5.53	1.08
Neuroticism (GPC-2)	1.02	0.09	**6.43×10^{-27}**	4.86	1.06
Neuroticism (Smith	1.02	0.02	**0**	9.17	1.24

2016)

Rheumatoid arthritis	0.02	0.04	0.6152	3.65	1.06
Schizophrenia	0.21	0.03	**5.99×10^{-11}**	22.53	1.81
Self-rated health	0.41	0.05	**2.95×10^{-19}**	16.27	1.26
Smoking status	0.11	0.04	**0.0029**	11.04	1.1
Subjective wellbeing	-0.68	0.03	**2.89×10^{-85}**	12.2	1.16
Tiredness	0.62	0.03	**1.28×10^{-97}**	11.2	1.14
Triglycerides	0.04	0.03	0.0937	4.74	1.22
Type 2 diabetes	-0.05	0.05	0.2805	8.98	1.13
Waist-hip ratio	0.05	0.02	0.0427	16.40	1.1

Note: Self-rated health is scored in the direction of higher scores indicating poorer health.

Supplementary Table 13. Associations between polygenic profiles based on neuroticism in UK Biobank. Significant associations (FDR correct p-value =5.75×10^{-6}) are shown in bold.

	Phenotypes in Generation Scotland							
	Neuroticism				Depression status			
Threshold	Beta	95% CI	R^2 %	p	OR	95% CI	R^2 %*	P
0.01	0.187	0.155 - 0.219	1.79	**7.22×10^{-31}**	1.249	1.134 – 1.375	0.52	**6.31×10^{-6}**
0.05	0.199	0.171 - 0.227	2.60	**3.28×10^{-44}**	1.258	1.155 – 1.370	0.71	**1.53×10^{-7}**
0.1	0.189	0.162 - 0.216	2.55	**2.64×10^{-43}**	1.246	1.147 – 1.353	0.71	**1.40×10^{-7}**
0.5	0.188	0.163 - 0.213	2.79	**2.65×10^{-47}**	1.252	1.158 – 1.353	0.82	**1.53×10^{-8}**
1	0.186	0.161 - 0.211	2.75	**1.26×10^{-46}**	1.245	1.152 – 1.346	0.79	**2.80×10^{-8}**

***** Nagelkerke R^2