Supplementary information

Genome-wide association study implicates *CHRNA2* in cannabis use disorder

Table of Contents

Supplementary Table 1. Association of previously identified cannabis risk variants

Association (P-value) and odds ratio (OR) in the present study of index variants passing the threshold for genome-wide significance in previous GWAS studies of cannabis related phenotypes.

Variant	Phenotype	study	OR (current study)	P-value (current study)
rs143244591	Cannabis criterion count	Sherva et al.[1]	NA (MAF < 0.01 in Europeans)	NA
rs77378271	Cannabis criterion count	Sherva et al.[1]	0.96	0.46
rs146091982	Cannabis criterion count	Sherva et al.[1]	NA (monomorphic in Europeans)	NA
rs1409568	Cannabis dependence vs cannabis exposed controls	Agrawal et al.[2]	0.97	0.72

NA = information not available

Supplementary Table 2. Distribution of psychiatric diagnoses in the sample

Number of individuals among CUD cases and the control group having a diagnosis of one or more of five major psychiatric disorders studied in iPSYCH.

Disorder	CUD cases	% of cases	controls	% controls
Schizophrenia	569	23.84	1,789	3.65
Bipolar disorder	188	7.88	1,176	2.40
Autism spectrum disorder	153	6.41	8,698	17.76
Attention deficit hyperactivity disorder	926	38.79	9,701	19.80
Major depression disorder	1,134	47.51	16,393	33.47
Individuals without any of the above disorders	97	4.06	16,269	33.21

Note: numbers are after QC (genetic outliers and related individuals are excluded) numbers does not represent unique individuals, meaning that individuals could count more than once if diagnosed with more than one psychiatric disorder in combination with CUD.

Supplementary Table 3. Tissues and number of genes tested using PrediXcan

Overview of tissues and number of genes included in the analysis of imputed genetically regulated expression using PrediXcan[3].

Tissue	Number of genes tested	Direction of expression in cases compared to controls (Beta)	P-value for association of *CHRNA2* imputed expression with cannabis use disorder
Dorsolateral prefrontal cortex	10,929	-1.210	5.912×10^{-4} (SE = 0.35)
Anterior cingulate cortex	2,658		NA
Caudate basal ganglia	3,620		NA
Cerebellar Hemisphere	4,234	-0.121	5.304×10^{-3} (SE = 0.04)
Cerebellum	4,836	-0.210	2.713×10^{-6} (SE = 0.05)
Cortex	3,586		NA
Frontal cortex	3,212		NA
Hippocampus	2,537		NA
Hypothalamus	2,459		NA
Nucleus accumbens basal ganglia	3,133		NA
Putamen basal ganglia	2,810		NA

NA = results not available, due to no valid model for gene expression in the tissue.

Supplementary Table 4. Association of SNPs related to smoking behavior

List of SNPs previously found strongly associated with smoking behavior and/or nicotine dependence. The studies reporting the associations are listed together with the P-value of association of the SNP in the studies. The P-value from association with CUD in this study is also reported.

Chr	SNP	GWAS study	Gene	Phenotype	P-value in published GWAS	P-value in GWAS of CUD
15	rs1051730	Thorgeirsson et al.[4]	*CHRNA3*	Cigarettes per day	2×10^{-69}	0.306
		Tobacco and Genetics Consortium[5]		Cigarettes per day	3×10^{-73}	
19	rs8102683	Kumasaka et al.[6]	*CYP2A6*	Cigarettes per day	4×10^{-42}	0.383
15	rs6011779*	Hancock et al.[7]	*CHRNA4*	Nicotine dependence	8×10^{-9}	0.54
8	rs1451240	Rice et al.[8]	*CHRNB3*	Nicotine dependence	7×10^{-16}	0.950
19	rs4105144	Thorgeirsson et al.[4]	*CYP2A6*	Cigarettes per day	2×10^{-12}	0.718
10	rs1329650	Tobacco and Genetics Consortium[5]	*LOC100188947*	Cigarettes per day	6×10^{-10}	0.791

The top SNP (rs2273500) identified in the study by Hancock et al. is not present in our data, instead we report results for rs6011779, which is located within the same smoking risk locus.

Supplementary Table 5. Association of cannabis risk locus with smoking behaviour

Association of the risk locus for cannabis in GWASs of smoking behaviour performed by the Tobacco and Genetics Consortium[5]. The index SNP (rs56372821) identified in this present GWAS of CUD was not present in the Tobacco and Genetics Consortium GWAS results, therefore P-values for rs11783093, which is in complete LD ($r^2 = 0.992$) with rs56372821, is reported in the table.

Phenotype	Sample size of smoking GWAS	P-value of rs11783093
Cigarettes per dag	68,028	0.007
Ever smoked	74,035	0.003
Former vs current smoker	41,969	0.589
Log age of onset	24,114	0.009

Supplementary Table 6. Polygenic risk score analysis

Results from PRS analysis in CUD for phenotypes related to cognition, personality, psychiatric disorders, reproduction and smoking behavior. Association P-value of PRS with CUD (P-value). The P-value threshold in the discovery GWAS (P-value threshold), the Z-score from logistic regression (Z-score) and Nagelkerke's R^2 (R^2) as estimate of the percentage of the variance in the phenotype explained by PRS. Values in bold represents significance after Bonferroni correction.

Phenotype/GWAS study	Trait type	Sample size	Z score	P-value	R^2	Threshold
Attainment of college/university degree (UK Biobank)[9]	Cognition/education	111,114	-2.81	4.94×10^{-03}	4.6×10^{-04}	0.5
Memory (UK Biobank)[9]	Cognition/education	112,067	1.23	2.18×10^{-01}	8.9×10^{-05}	1×10^{-03}
Reaction time (UK Biobank)[9]	Cognition/education	111,483	1.96	5.00×10^{-02}	2.2×10^{-04}	0.1
verbal-numerical reasoning (UK Biobank)[9]	Cognition/education	36,035	1.43	1.53×10^{-01}	1.2×10^{-04}	0.2
Human Intelligence[10]	Cognition/education	78,308	-3.51	**4.33×10^{-04}**	7.2×10^{-04}	0.5
Number of educational years (SSGAC)[11]	Cognition/education	328,917	-7.78	**7.44×10^{-15}**	3.5×10^{-03}	5×10^{-02}
College completion (SSGAC)[12]	Cognition/education	126,559	-5.10	**3.33×10^{-07}**	1.5×10^{-03}	1×10^{-02}
Openness (GPC-1)[13]	Personality	17,375	1.47	1.43×10^{-01}	1.3×10^{-04}	1×10^{-04}
Neuroticism (GPC-2)[14]	Personality	63,661	-1.95	5.02×10^{-02}	2.2×10^{-04}	1×10^{-03}
Extraversion (GPC-2)[15]	Personality	63,030	-0.52	6.03×10^{-01}	1.6×10^{-05}	1×10^{-04}
Agreeableness (GPC-1)[13]	Personality	17,375	-1.73	8.23×10^{-02}	1.8×10^{-04}	1×10^{-02}
Conscientiousness (GPC-1)[13]	Personality	17,375	-2.08	3.74×10^{-02}	2.5×10^{-04}	1×10^{-04}
ADHD (iPSYCH+PGC)[16]	Psychiatric*	53,293	5.10	**3.45×10^{-07}**	2.3×10^{-03}	0.5
Schizophrenia (PGC)[17]	Psychiatric*	150,064	5.47	**4.45×10^{-08}**	2.2×10^{-03}	0.5
Major depressive disorder (PGC)[18]	Psychiatric*	18,759	-0.55	5.81×10^{-01}	3.1×10^{-05}	1×10^{-03}
Depressive symptoms (SSGAC)[19]	Psychiatric*	161,460	4.34	**6.58×10^{-06}**	2.05×10^{-03}	0.1
Age of first birth (SSGAC)[20]	Reproduction	251,151	-7.41	**1.26×10^{-13}**	3.22×10^{-03}	0.2
Number of children ever born (SSGAC)[20]	Reproduction	343,072	2.81	5.01×10^{-03}	4.61×10^{-04}	1.0
Smoking age of onset (TAG)[5]	Smoking behaviour	24,114	1.17	2.43×10^{-01}	7.9×10^{-05}	1×10^{-03}
life-time smoking (TAG)[5]	Smoking behaviour	74,035	4.01	**6.08×10^{-05}**	9.4×10^{-04}	0.1
Current vs former smoker (TAG)[5]	Smoking behaviour	41,969	-1.19	2.33×10^{-01}	8.3×10^{-05}	1
Cigarettes per day (TAG)[5]	Smoking behaviour	68,028	1.75	8.01×10^{-02}	1.8×10^{-04}	1×10^{-03}

* In PRS analyses of psychiatric disorders, individuals having a diagnosis of the respective psychiatric disorder being analysed (ADHD, schizophrenia and major depressive disorder) were excluded in the CUD target sample.

Supplementary Figure 1. Null distribution for rs56372821 odds ratio for schizophrenia
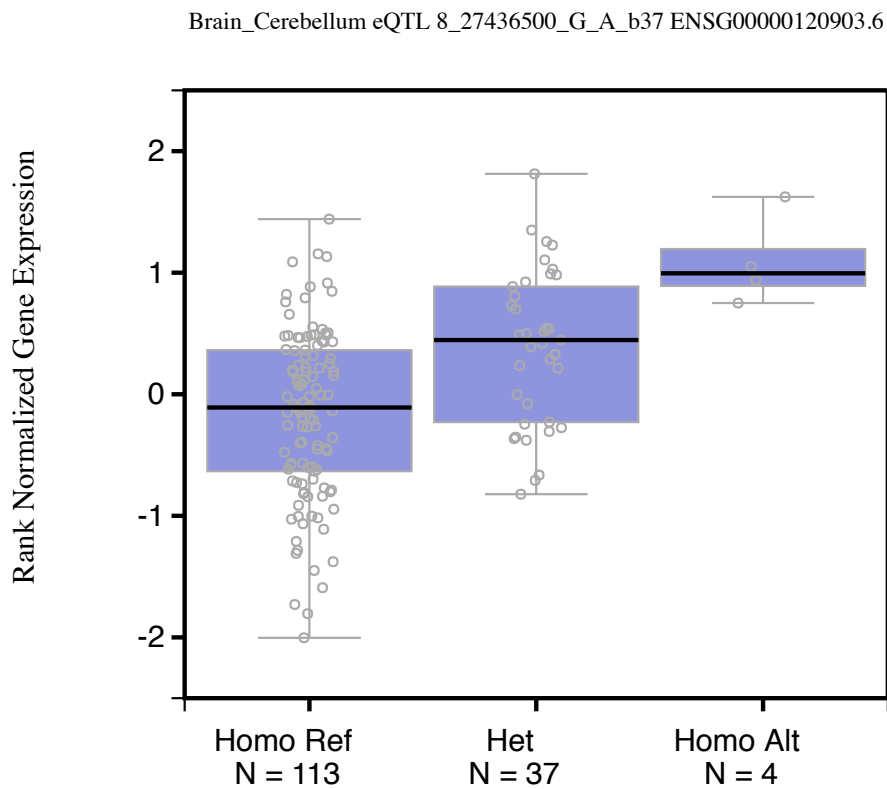
Null distribution of the odds ratio for rs56372821 when applying 10,000 permutations removing randomly 554 schizophrenia cases and 101 individuals. The Red lines indicate the observed $\ln(OR_{rs56372821})$ when excluding individuals with CUD (554 SZ cases and 101 controls with CUD) mirrored around the mean $\ln(OR_{rs56372821})$ of the null distribution obtained by random removal of the same number of SZ cases and control.
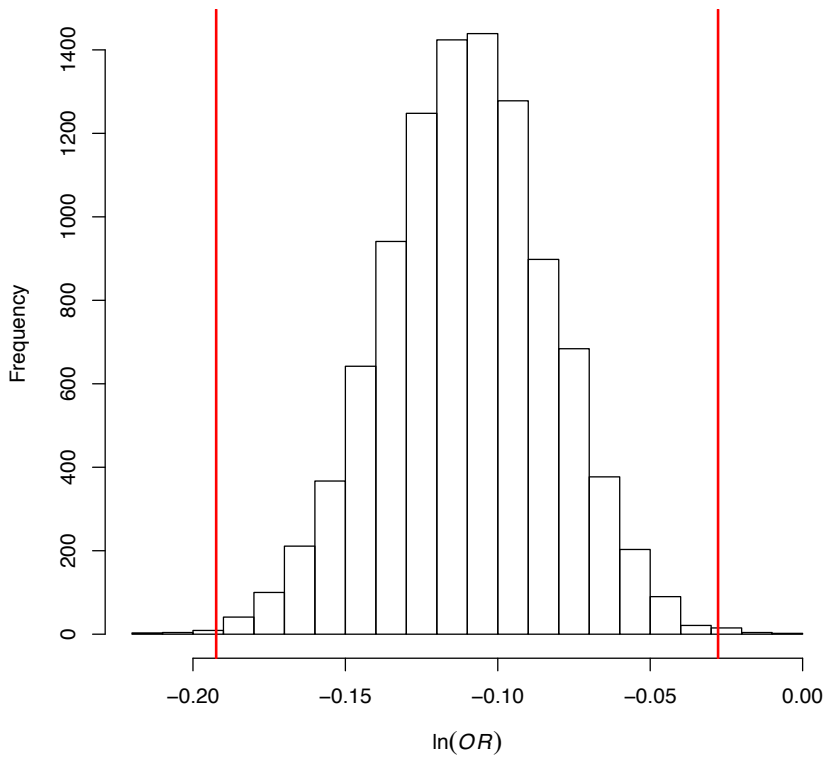
Supplementary Figure 2. *CHRNA2* **expression in cerebellum among rs56372821 genotypes**

The distribution of *CHRNA2* expression over rs56372821 genotypes (Hom Ref= GG; He =AG; Homo

Alt =AA), for individuals in the GTEx database[21]. The major allele (G) is the risk allele for CUD.

(picture downloaded from: https://www.gtexportal.org/home/eqtls/bySnp?snpId=rs56372821&tissueName=All)

Brain_Cerebellum eQTL 8_27436500_G_A_b37 ENSG00000120903.6

Supplementary Figure 3. Gene expression of *CHRNA2* and *CNR1* in Allan Brain Atlas

Gene expression of *CHRNA2* and *CNR1* in Allan Brain Atlas (http://www.brain-map.org/) in 26 brain regions from 6 donors. Doners are marked in six different colors in the first row, and the 26 different regions is marked by different colors in the second row. Cerebellar cortex and cerebellar nuclei are marked with arrows. The results are shown in 'Coarse' mode, where the brain is divided into large neuroanatomical divisions or regions and the samples within each of the regions are averaged together to provide a summary value for the partition. The colors of the heat map are normalized expression values. The color green should be interpreted as relatively low expression and red as relatively high expression within the scope of each probe. There are two probes tagging *CHRNA2* expression (both shown) and 89 probes tagging *CNR1*. The six probes tagging *CNR1* expression demonstrating strongest negative correlation with *CHRNA2* expression is shown (r_{max}= -0.498). 67% of the probes tagging *CNR1* correlated with *CHRNA2* expression with r > -0.4.

(Picture downloaded from: http://human.brain-map.org/microarray/search/show?domain=&donors=9861,10021,12876,14380,15496,15697&search_type=correlation&search_term=&seed=1058207)
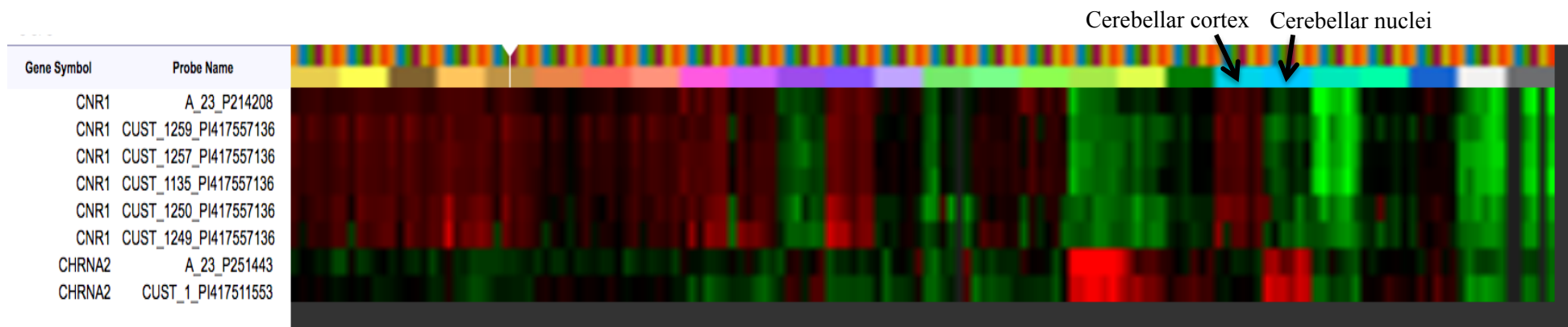
Supplementary Figure 4. A-C. Quintile plots of Odds Ratio for CUD by PRS for cognitive performance/educational attainment

Odds Ratio (OR) for CUD by PRS within each quintile. The plots represent the five phenotypes related to cognitive/educational performance demonstrating significant association of PRS with CUD after correcting for multiple testing. Error bars indicate 95% confidence limits.
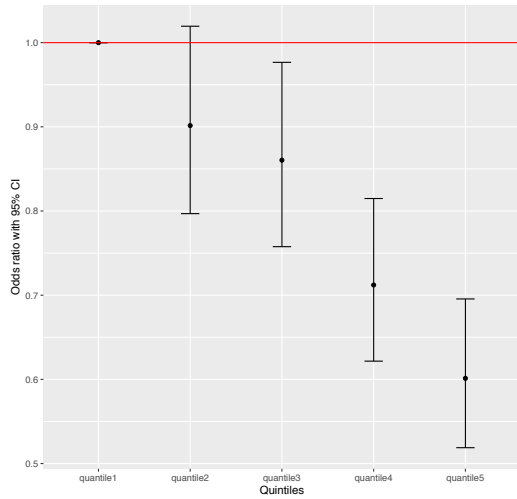
A. OR for CUD by PRS for educational years (SSGAC)
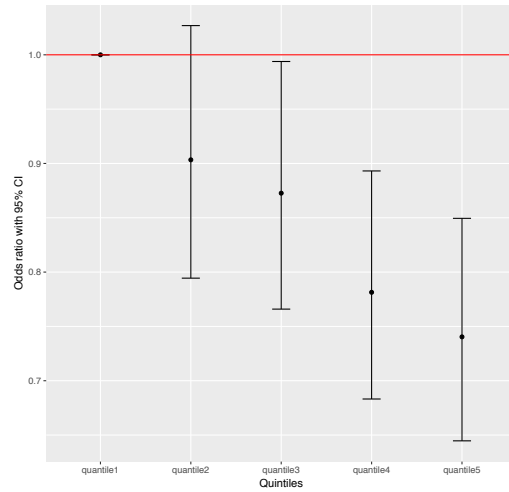
B. OR for CUD by PRS for college completion (SSGAC)
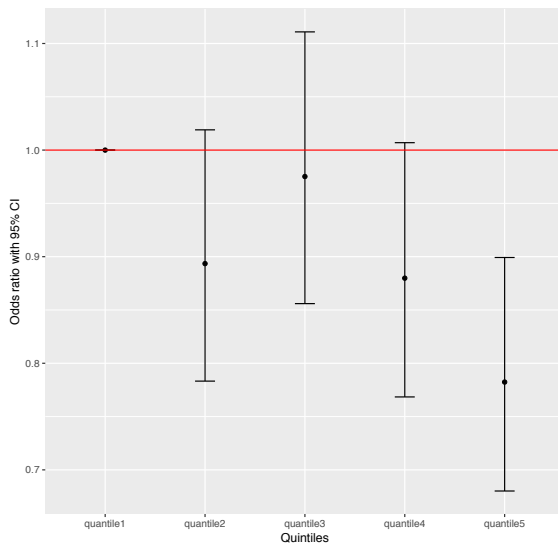
C. OR for CUD by PRS for human intelligence

References

1 Sherva, R. *et al.* Genome-wide Association Study of Cannabis Dependence Severity, Novel Risk Variants, and Shared Genetic Risks. *JAMA Psychiatry* **73**, 472-480, doi:10.1001/jamapsychiatry.2016.0036 (2016).

2 Agrawal, A. *et al.* Genome-wide association study identifies a novel locus for cannabis dependence. *Molecular psychiatry*, doi:10.1038/mp.2017.200 (2017).

3 Gamazon, E. R. *et al.* A gene-based association method for mapping traits using reference transcriptome data. *Nature genetics* **47**, 1091-1098, doi:10.1038/ng.3367 (2015).

4 Thorgeirsson, T. E. *et al.* Sequence variants at CHRNB3-CHRNA6 and CYP2A6 affect smoking behavior. *Nature genetics* **42**, 448-453, doi:10.1038/ng.573 (2010).

5 Tobacco & Genetics, C. Genome-wide meta-analyses identify multiple loci associated with smoking behavior. *Nature genetics* **42**, 441-447, doi:10.1038/ng.571 (2010).

6 Kumasaka, N. *et al.* Haplotypes with copy number and single nucleotide polymorphisms in CYP2A6 locus are associated with smoking quantity in a Japanese population. *PloS one* **7**, e44507, doi:10.1371/journal.pone.0044507 (2012).

7 Hancock, D. B. *et al.* Genome-wide meta-analysis reveals common splice site acceptor variant in CHRNA4 associated with nicotine dependence. *Transl Psychiatry* **5**, e651, doi:10.1038/tp.2015.149 (2015).

8 Rice, J. P. *et al.* CHRNB3 is more strongly associated with Fagerstrom test for cigarette dependence-based nicotine dependence than cigarettes per day: phenotype definition changes genome-wide association studies results. *Addiction* **107**, 2019-2028, doi:10.1111/j.1360-0443.2012.03922.x (2012).

9 Davies, G. *et al.* Genome-wide association study of cognitive functions and educational attainment in UK Biobank (N=112 151). *Molecular psychiatry* **21**, 758-767, doi:10.1038/mp.2016.45 (2016).

10 Sniekers, S. *et al.* Genome-wide association meta-analysis of 78,308 individuals identifies new loci and genes influencing human intelligence. *Nature genetics* **49**, 1107-1112, doi:10.1038/ng.3869 (2017).

11 Okbay, A. *et al.* Genome-wide association study identifies 74 loci associated with educational attainment. *Nature* **533**, 539-542, doi:10.1038/nature17671 (2016).

12 Rietveld, C. A. *et al.* GWAS of 126,559 individuals identifies genetic variants associated with educational attainment. *Science* **340**, 1467-1471, doi:10.1126/science.1235488 (2013).

13 de Moor, M. H. *et al.* Meta-analysis of genome-wide association studies for personality. *Molecular psychiatry* **17**, 337-349, doi:10.1038/mp.2010.128 (2012).

14 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**, 642-650, doi:10.1001/jamapsychiatry.2015.0554 (2015).

15 van den Berg, S. M. *et al.* Meta-analysis of Genome-Wide Association Studies for Extraversion: Findings from the Genetics of Personality Consortium. *Behav Genet* **46**, 170-182, doi:10.1007/s10519-015-9735-5 (2016).

16 Demontis, D. *et al.* Discovery Of The First Genome-Wide Significant Risk Loci For ADHD. *bioRxiv*, doi:10.1101/145581 (2017).

17 Schizophrenia Working Group of the Psychiatric Genomics, C. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* **511**, 421-427, doi:10.1038/nature13595 (2014).

18   Major Depressive Disorder Working Group of the Psychiatric, G. C. *et al.* A mega-analysis of genome-wide association studies for major depressive disorder. *Molecular psychiatry* **18**, 497-511, doi:10.1038/mp.2012.21 (2013).

19   Okbay, A. *et al.* Genetic variants associated with subjective well-being, depressive symptoms, and neuroticism identified through genome-wide analyses. *Nature genetics* **48**, 624-633, doi:10.1038/ng.3552 (2016).

20   Barban, N. *et al.* Genome-wide analysis identifies 12 loci influencing human reproductive behavior. *Nature genetics* **48**, 1462-1472, doi:10.1038/ng.3698 (2016).

21   Consortium, G. T. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science* **348**, 648-660, doi:10.1126/science.1262110 (2015).