Supplement

Material and Methods

Genotyping and Quality control (QC)

Cases and controls were genotyped using Illumina's HumanOmniExpress-12v1_H (n=730525 marker) and HumanOmni2.5-4v1_B BeadChips (n=2450000 marker) respectively. For individuals with call rates of < 97%, and in the case of duplicates or cryptic related samples (average identity by state across autosomal markers > 1.65), the sample with the lower call rate was removed. Conformity between reported sex and genotypic sex was required. Outlier status was determined using multi-dimensional scaling, first separately for cases and controls, and then in the combined data. Multidimensional scaling was based on pruned Single Nucleotide Polymorphisms (SNPs) with a Hardy–Weinberg equilibrium (HWE) $p > 0.2$, a minor allele frequency (MAF) > 0.2, a call rate of 100%, and a pairwise LD-pruning of $r^2 < 0.1$. A sliding window of 200 SNPs was considered with shifting of 50 SNPs. After visual inspection, we decided to take five principal components (PCs) for outlier detection. Cases exceeding more than six standard deviations on any of the first five PCs were excluded.

In both subsamples, only autosomal SNPs were taken into account, only SNPs with a call rate ≥ 98% were included, and SNPs with a MAF < 0.01 were removed. Non-random-missingness was accounted for by excluding SNPs with differences in the call rate between cases and control s significant with P-values of $< 1x\ 10^{-5}$. A haplotype-based test was performed for non-random missing genotype data $p < 1x\ 10^{-10}$, and conformity with HWE was considered by only selecting SNPs with $p_{HWE} \geq 1x\ 10^{-4}$ in controls and $p_{HWE} \geq 1x\ 10^{-6}$ in cases.

Statistical analysis

Data preparation and statistical analysis were conducted using PLINK (http://pngu.mgh.harvard.edu/~purcell/plink/) and R version 3.1 (http://www.r-project.org/).

For the logistic regression, correction for population stratification was performed using consistently the first five PCs resulting from a principal component analysis across independent autosomal markers. These markers were created by pruning SNPs remaining after quality control. A pair-wise $r^2 < 0.1$ was applied within a sliding window of 200 SNPs, shifting 50 SNPs, all of which had a $p_{HWE} > 0.2$ and had a MAF ≥ 0.2. These five PCs were included in the logistic regression model as covariates. In a second approach, we included these five PCs together with age and sex as covariates. Both approaches were used for all of the following analyses.

First analysis: correction for PC 1 to 5 only

Second analysis: correction for PC 1 to 5, age, and sex.

Cluster plots of top hits < 5 x 10^{-5} were visually inspected, and markers with poor cluster quality were removed.

Gene-based test

We used VEGAS2, downloaded from https://vegas2.qimrberghofer.edu.au/zVEGAS2offline.tgz, version 16:09:002, using the 1000Genomes data to model SNP correlations [1], updating the genome build form hg18 to hg19. With VEGAS2, tests for association are performed for the combined effect of SNPs grouped together per gene. We used the CEU population as reference and all SNPs belonging to a gene, defining gene boundaries as +/- 50 kb of 5' and 3' UTRs according the Vegas programme [2].

Polygenic risk scores

Strand-ambiguous SNPs, as well as SNPs showing no overlap between samples were removed. SNPs with MAF < 0.1 (training and test sample) and those inside the extended MHC-region (chr6:25-34 Mb according to UCSC hg19/NCBI Build 37) were also removed. In determining the score, only SNPs remaining after clumping were considered. LD-pruning was performed using only those SNPs present in a "clumped" version of the file containing independent SNPs (pairwise r^2 < 0.1 within a 500 kb window). Markers with P-values of < 0.01, <0.05, <0.1, <0.2, <0.3, <0.4, or <0.5 were included in the polygenic risk score analysis in alternative approaches. In the training samples, all markers below the respective threshold were used for calculating a weighted value. Marker weights of alcohol dependence (AD) were calculated as the natural logarithm of odds ratios provided by the association results from a GWAS of AD [3] with n=3501 individuals (1333 cases). To determine the risk score for disordered gambling (DG, n=1312), beta values provided by Lind et al. [4] were used directly. A set of 39930 (AD) or 44324 (DG) independent markers for p<0.5 was considered in the score analysis.

For every individual in our dataset, a weighted sum of these associated alleles was constructed. A logistic regression approach was then applied to test for association between PG cases and controls using the polygenic score and the first five PC components, or the PC components, sex and age for the prediction of the phenotype.

Pathway/gene-set download

Gene Information for gene-sets was obtained from the following databases: Kyoto Encyclopaedia of Genes and Genomes (dbKEGG, [5] http://www.genome.jp/kegg/), downloaded with R package *KEGGREST*, version 1.2.2 ; Reactome (dbRC [6] http://www.reactome.org/), downloaded with R package *reactome.db*, version 1.46.1; and Gene Ontology (dbGO [7] www.geneontology.org), downloaded with R package *org.Hs.eg.db* , version 2.14.0 [8]).

KEGG cancer pathways were removed and only gene-sets with 5-200 genes were taken into account according [9], reducing the gene-set number by 32.

Pruning

The Global Test was performed on a reduced SNP-set in order to adjust to the assumption of independence between variables. Therefore, a pruned set SNPs of the GWAS data was used by applying a variance inflation factor of 10 (VIF=10), and using a window size of 50, shifted by 5 SNPs per step, as implemented in PLINK (version 1.07). Of the previous 595861 SNPs, 298286 SNPs remained.

Mapping SNPs to genes

Mapping of markers to genes was performed according to [9]. SNPs were annotated using information from dbSNP build 131 (Assembly GRCh37). The start and end of a gene are defined as the start position of its first exon and the end position of its last exon. RefSeq FTP release 61, distributed in September 2013 [10].

Assignment of genes

To account for important regulatory regions, markers were assigned to a gene if they were located within the genomic sequence or within a frame of 20kb of the 5' and 3'ends of the first and last exon. SNPs occurring within regions shared by multiple genes were assigned to all of the respective genes.

Accounting for possible bias

The number of SNPs that were mapped to the pathways differs. This factor could introduce bias into the pathway association. In small pathways, even single SNPs could influence the results, while in larger pathways, chance association may be observed. To control for this type of bias, the gene-constraint was applied to KEGG and Reactome. It was not applied to the GO gene-sets in view of its nested structure.

Additionally, a SNP label permutation without replacement was performed to correct for bias due to different pathway lengths. This comprised different numbers of genes and different genes with different numbers of mapped SNPs. A significant P-value of the SNP-label permutation test indicates a low probability of obtaining test statistics with even more extreme values if the test is performed with randomly selected markers other than that observed.

After running the permutation test with all KEGG and Reactome pathways for 100 permutations, those pathways with a P-value < 0.05 were selected. For these pathways, the SNP-shuffling test was run again 900 times. Table S2 lists the P-values for the 1000 permutation tests.

To test for bias due to a random variation at the individual level, a subject-sampling test was performed according to Efron and Tibshirani [11,9]. Here, case-control status was randomized 10,000 times, as in the main global test.

For the GO test, neither SNP shuffling nor the permutation test were used due to the hierarchical structure of GO. None of the gene-sets in GO had a Benjamini-Hochberg corrected P-value of <0.05 in the global test due to the large number of gene-sets (8474).

Results

Marker-wide associations

Supplementary table S1a: All top SNPs with a P-value of < 10^{-4}, first analysis including PC 1 to 5

CHR	SNP	Position (hg19)	Minor allele	OR	L95	U95	P-value
20	rs6065904	44534651	A	0.53	0.409	0.686	1.48×10^{-6}
20	rs4810479	44545048	C	0.565	0.443	0.722	4.67×10^{-6}
16	rs3943418	17337724	A	1.711	1.355	2.162	6.61×10^{-6}
4	rs11723785	178136407	T	1.706	1.343	2.167	1.19×10^{-5}
4	rs4690502	178141976	A	1.698	1.339	2.154	1.25×10^{-5}
10	rs10995114	64074412	T	3.046	1.844	5.031	1.30×10^{-5}
4	rs10031235	5324465	C	1.859	1.405	2.459	1.44×10^{-5}
4	rs6853653	77725242	C	1.696	1.335	2.155	1.56×10^{-5}
17	rs8078855	78225055	T	1.563	1.275	1.916	1.69×10^{-5}
4	rs6853980	5324579	A	1.761	1.361	2.28	1.71×10^{-5}
6	rs9396970	169966644	C	2.36	1.595	3.494	1.76×10^{-5}
3	rs1868488	29964567	C	1.848	1.391	2.455	2.29×10^{-5}
6	rs2745599	1613686	G	1.54	1.261	1.881	2.30×10^{-5}
15	rs3803497	63053858	C	2.401	1.592	3.62	2.92×10^{-5}
6	rs2860492	169930402	T	2.196	1.514	3.186	3.41×10^{-5}
6	rs2745596	1606031	T	0.659	0.54	0.805	4.23×10^{-5}
1	rs764656	48694259	C	0.571	0.44	0.75	5.06×10^{-5}
9	rs10815757	8097977	C	1.517	1.24	1.856	5.19×10^{-5}
11	rs11035648	5116955	G	2.025	1.439	2.85	5.21×10^{-5}
20	rs215543	2780151	G	0.616	0.486	0.779	5.45×10^{-5}
9	rs987073	116546918	G	1.539	1.248	1.899	5.59×10^{-5}
6	rs2997887	169912559	A	2.232	1.508	3.302	5.90×10^{-5}
2	rs10166009	133471789	T	0.63	0.502	0.789	6.03×10^{-5}
10	rs11257470	6277556	T	0.664	0.544	0.811	6.04×10^{-5}
9	rs768703	18070475	A	1.507	1.233	1.843	6.26×10^{-5}
3	rs10049438	133294203	C	1.623	1.28	2.057	6.39×10^{-5}
5	rs1541077	26157740	A	1.5	1.229	1.832	6.70×10^{-5}
11	rs7947494	10997718	A	0.662	0.541	0.811	6.77×10^{-5}

14	rs761530	97568613	T	0.664	0.542	0.812	6.91×10^{-5}
8	rs10086260	9158475	A	0.58	0.443	0.758	6.95×10^{-5}
2	rs10497460	177691497	G	0.654	0.531	0.807	7.17×10^{-5}
20	rs17447545	44547068	G	0.572	0.434	0.754	7.22×10^{-5}
7	rs4559136	16946486	A	0.388	0.243	0.619	7.26×10^{-5}
10	rs17143250	8177255	T	2.036	1.432	2.894	7.44×10^{-5}
10	rs12773241	130232922	G	0.614	0.483	0.782	7.45×10^{-5}
3	rs9858736	29982654	T	1.842	1.361	2.492	7.56×10^{-5}
9	rs2385188	138955201	C	1.529	1.238	1.889	8.07×10^{-5}
14	rs10136662	64597186	G	1.744	1.322	2.3	8.14×10^{-5}
2	rs6737220	235375794	A	0.433	0.286	0.657	8.35×10^{-5}
22	rs7289240	29854959	C	1.483	1.218	1.81	8.83×10^{-5}
2	rs13021421	216811307	T	1.834	1.353	2.486	9.26×10^{-5}
7	rs7780145	19198132	G	1.511	1.228	1.859	9.49×10^{-5}
3	rs1121119	8302786	G	0.681	0.561	0.826	9.76×10^{-5}
7	rs579864	154539863	A	0.661	0.536	0.814	9.77×10^{-5}
7	rs17351688	19193072	G	1.522	1.232	1.88	9.80×10^{-5}
9	rs4534200	25652508	C	1.468	1.21	1.781	9.80×10^{-5}
8	rs6989065	12609188	T	1.612	1.268	2.051	9.85×10^{-5}

Supplementary table S1b: All top SNPs with a P-value of $< 10^{-4}$ second analysis including PC 1 to 5, age and sex

CHR	SNP	Position (hg19)	Minor allele	OR	L95	U95	P-value
2	rs7591351	46063406	T	1.673	1.339	2.09	5.88×10^{-6}
2	rs6738409	46062550	C	0.5986	0.4783	0.7492	7.39×10^{-6}
12	rs6582294	76034992	A	1.691	1.339	2.137	1.07×10^{-5}
2	rs13021421	216811307	T	2.191	1.543	3.111	1.16×10^{-5}
15	rs17255585	54107802	C	0.2609	0.1429	0.4766	1.24×10^{-5}
9	rs10815757	8097977	C	1.696	1.338	2.151	1.28×10^{-5}
12	rs3898937	75947644	G	1.671	1.326	2.106	1.35×10^{-5}
1	rs2359854	198561100	A	0.5954	0.4678	0.7577	2.49×10^{-5}
4	rs6853653	77725242	C	1.808	1.373	2.381	2.52×10^{-5}
9	rs10815753	8093954	G	1.657	1.31	2.096	2.53×10^{-5}
15	rs8036417	78419476	G	1.714	1.331	2.207	2.99×10^{-5}
10	rs10825357	56323564	T	0.5108	0.3722	0.7009	3.17×10^{-5}
18	rs190166	24499317	A	1.621	1.291	2.036	3.25×10^{-5}
7	rs579864	154539863	A	0.5967	0.4674	0.7618	3.43×10^{-5}
10	rs1411823	20077441	A	0.6154	0.489	0.7744	3.49×10^{-5}
3	rs6550215	33277828	G	1.713	1.327	2.211	3.57×10^{-5}
12	rs7965173	127897824	T	1.613	1.285	2.025	3.76×10^{-5}
2	rs10497460	177691497	G	0.6043	0.4751	0.7688	4.12×10^{-5}

9	rs768703	18070475	A	1.624	1.288	2.048	4.23×10^{-5}
15	rs4776181	54122875	C	0.216	0.1037	0.4499	4.25×10^{-5}
7	rs17351688	19193072	G	1.66	1.302	2.117	4.32×10^{-5}
15	rs2289524	78390414	C	1.59	1.272	1.987	4.50×10^{-5}
16	rs3943418	17337724	A	1.753	1.338	2.298	4.65×10^{-5}
13	rs1465661	75906289	C	2.129	1.479	3.065	4.76×10^{-5}
17	rs7208143	1811983	T	1.655	1.297	2.112	5.13×10^{-5}
6	rs9444074	84169154	G	4.874	2.263	10.5	5.23×10^{-5}
11	rs7947494	10997718	A	0.6168	0.4879	0.7796	5.30×10^{-5}
2	rs828867	74334462	G	0.62	0.4916	0.782	5.41×10^{-5}
16	rs182928	26603412	T	0.4806	0.3366	0.6862	5.54×10^{-5}
11	rs12280713	5116109	C	1.92	1.396	2.641	6.06×10^{-5}
7	rs4534036	16959685	C	0.4108	0.2659	0.6347	6.12×10^{-5}
20	rs4810479	44545048	C	0.5671	0.4292	0.7494	6.66×10^{-5}
3	rs9858736	29982654	T	2.057	1.443	2.932	6.70×10^{-5}
10	rs10741187	131055789	A	1.589	1.265	1.995	6.80×10^{-5}
6	rs1885634	169075136	A	2.657	1.642	4.299	6.86×10^{-5}
7	rs4559136	16946486	A	0.3432	0.2027	0.5811	6.88×10^{-5}
20	rs6065904	44534651	A	0.5453	0.4045	0.7351	6.93×10^{-5}
10	rs10995114	64074412	T	3.398	1.859	6.214	7.10×10^{-5}
11	rs17129771	96852063	A	0.6037	0.4707	0.7744	7.11×10^{-5}
6	rs6928575	106950833	C	1.694	1.306	2.198	7.14×10^{-5}
22	rs7289240	29854959	C	1.581	1.261	1.983	7.27×10^{-5}
2	rs6761327	46066261	G	1.566	1.255	1.954	7.31×10^{-5}
10	rs6482515	18827828	T	1.663	1.292	2.14	7.68×10^{-5}
5	rs1559090	62874332	C	1.652	1.287	2.119	8.00×10^{-5}
5	rs10036059	62942257	A	1.646	1.285	2.11	8.13×10^{-5}
22	rs5992629	17602839	G	1.933	1.393	2.684	8.17×10^{-5}
9	rs4391483	9711904	G	1.584	1.26	1.992	8.31×10^{-5}
2	rs6732900	46066236	T	1.565	1.252	1.957	8.40×10^{-5}
2	rs355895	165635869	T	1.568	1.253	1.962	8.44×10^{-5}
19	rs12978300	32573668	C	2.064	1.437	2.963	8.63×10^{-5}
9	rs4838118	126935255	A	3.997	1.999	7.992	8.86×10^{-5}
10	rs12242391	71201504	T	0.4779	0.3301	0.6918	9.15×10^{-5}
19	rs7247279	17769508	C	1.607	1.267	2.039	9.19×10^{-5}
7	rs7780145	19198132	G	1.609	1.267	2.044	9.62×10^{-5}
2	rs4952781	46073997	T	1.571	1.252	1.971	9.73×10^{-5}
15	rs17820305	57746815	G	1.689	1.297	2.198	9.81×10^{-5}
4	rs6853980	5324579	A	1.82	1.346	2.461	9.96×10^{-5}

S1a: Results of the single-marker analysis including the first five PCs.

In addition to those markers described in the main article, the top SNPs included (i) rs10031235 (P-value = 1.44 x 10^{-5}, OR=1.86; CI = [1.41, 2.46], ranked 7th; age and sex corrected P-value: 2.26 x 10^{-4}). This is located inside the intron of the *STK32B*, which encodes a serine/threonine kinase associated with AD [12], and (ii) rs8078855 (P-value = 1.69 x 10^{-5}, OR = 1.56, CI= [1.28, 1.92], ranked 9th; age and sex corrected P-value 6.75 x 10^{-4}), an intronic SNP in *SLC26A11*. The protein of *SLC26A11* acts as voltage-gated CI(-) channel, activated upon neuronal depolarisation [13].

S1b: Results of the single-marker analysis including the first five PCs, age and sex.

The top hits after age and sex correction included rs13021421 (P-value 1.16 x 10^{-5}, OR 2.19, CI= [1.54, 3.11], uncorrected 9.26 x 10^{-5}). This is located inside the gene encoding melanoregulin. Melanoregulin may play a role in membrane fusion and the regulation of the biogenesis of disk membranes of photoreceptor rod cells [14]. This gene has shown significant association with AD at a genome-wide level [14].

***C*omparison with top hits of the Australian GWAS**

The results were compared with the six top SNPs from the Australian GWAS of DG, which was performed in the community-based Australian twin study cohort [4]. In the present PC 1 to 5 correction analysis, the top Australian GWAS hit, rs8064100, obtained a one-sided P-value of 0.045 with the same allele (OR = 1.18; CI = [0.974, 1.43]. This result is not corrected for multiple testing for the number of SNPs. In the analysis including age and sex correction, rs8064100 had a one-sided P-value of 0.077. In the Australian GWAS, this SNP achieved a P-value of 2.57 x 10^{-6} (after correction using genomics controls) [4]. The SNP is located downstream of *MT1X* encoding metallothionein 1X, which is involved in metal ion binding. Metallothioneins are metal- and cysteine-rich proteins with zinc binding- and antioxidant properties. They also have antioxidant and anti-inflammatory properties, and are involved in diverse physiological mechanisms, including tissue regeneration and cell survival [15]. Metallothionein 1 proteins have been implicated in neuroprotection and neuroregeneration [15], and *MT1* shows differential expression in alcohol related phenotypes [16].

The only other top SNP from the 6 top hits of the Australian GWAS that was available in our dataset, was rs9383153. This achieved a P-value of 0.87 and 0.55 in the first and second approach, respectively.

Gene-based associations

Description of the top hits of the first analysis (PC 1 to 5), Table 2a in the main article:

PCIF1. The protein of *PCIF1* binds to the phosphorylated C-terminal domain of the largest subunit of RNA polymerase II. Although its functional consequences remain unclear, previous authors have suggested that it negatively regulates gene expression of the polymerase II via the modulation of the phosphorylation status of the C-terminal domain [17]. PCIF1 is also thought to play a role in either transcription elongation or in coupling transcription to pre-mRNA processing through its association with the phosphorylated C-terminal domain (CTD) of RNAPII largest subunit.

PLTP is a phospholipid transfer protein found in human plasma. It plays an important role in PLTP-mediated HDL conversion. It regulates the size and composition of HDL in the circulation [18,19], and controls levels of plasma HDL.

CTSA encodes the protective protein/cathepsin A. Mutations in this gene lead to a secondary deficiency of ß-galactosidase and neuraminidase 1 [17].

NEURL2 encodes a protein being involved in the regulation of myofibril organization. Research suggests that it represents the adaptor component of the E3 ubiquitin ligase complex in striated muscle and regulates the ubiquitin-mediated degradation of beta-catenin during myogenesis.

C20orf165, also known as *SPATA25*, spermatogenesis associated 25 [20].

***MIR3926-2*:** microRNAs (miRNAs) are short non-coding RNAs that are involved in post-transcriptional regulation of gene expression by affecting the stability as well as the translation of mRNA.

ZSWIM1 encodes a protein in leukocytes with no exactly known function [21]. It contains a zinc finger SWIM motif. Research suggests that it may play a novel role in the development or function of T helper cells. It is located near NEURL2, CTSA, SPATA25, and PLTP.

***MIR3926-1*:** see MIR3926-2.

ZNF335, zinc finger protein 335. The protein encoded by this gene enhances transcriptional activation via ligand-bound nuclear hormone receptors.

ZSWIM3, zinc finger, SWIM containing 3. An important paralog of this gene is ZSWIM1.

LONRF1, *LON Peptidase N-Terminal Domain and Ring Finger 1* is thought to participate in proteolysis. Proteins are expected to have ATP-dependent peptidase activity, metal ion binding, protein binding, and zinc ion binding, and to be located in cytoplasm.

GAPVD1 is a G protein regulator which acts as both a GTPase-activating protein (GAP) and a guanine nucleotide exchange factor (GEF). GAPVD1 has GEF activity for Rab5 and GAP activity for Ras. It is involved in processes such as endocytosis, insulin receptor internalisation, and GLUT4 trafficking.

ACOT8, *acyl-CoA thioesterase 8* encodes an Acyl-CoA thioesterase protein that catalyses the hydrolysis of acyl-CoA to the free fatty acid.

DNAI2, axonemal dynein intermediate chain 2 is part of the dynein complex of respiratory cilia and sperm flagella (disease: Primary ciliary dyskinesia [22]). DNAI2 (rs7219585) was reported in a GWAS of information processing speed [23]

DNAH7, dynein heavy chain 7 (axonemal) is a component of the inner arm of human cilia. It is a force generating protein of respiratory cilia; dynein has ATPase activity. It is detected in brain, testis, and trachea, (in protein level) detected in bronchial cells.

FERD3L, *Fer3-like bHLH* transcription factor is a transcription factor that binds to the E-box and functions as inhibitor of transcription. DNA binding requires dimerization with an E protein (Uniprot).

HSPA5 the glucose regulated heat shock 70kD protein 5. It is involved in the folding and assembly of proteins in the endoplasmic reticulum. It has been associated with alcohol preference in mice (Kerns et al., 2005) and alcohol consumption and preference in rats [24]. A possible association with bipolar disorder has been reported [25]

TWIST1, class A basic helix-loop-helix protein 38, is a HLH transcription factor. Loss-of-function mutations of the TWIST 1 gene cause the Saethre-Chotzen craniosynostosis syndrome (SCS) [26]. This gene was reported in a GWAS of obesity-related traits with a P-value of 4.18×10^{-7} (Urinary free epinephrine).

KIF19, *kinesin family member 19* encodes a motor protein that regulates the length of motile cilia [27].

It may be of interest that the top genes with P-values $<10^{-3}$ including only PC components 1 to 5 included *HSPA5*, remaining significant in the second approach with age and sex included with a P-value of 0.0031. The encoded heat shock protein A5 belongs to the family of heat shock proteins, which are involved in important cellular processes such as glucose metabolism and protein folding. Expression studies of alcohol exposure in animal models have also implicated *Hspa5* in addiction

phenotypes [28,24]. To exclude the possibility that the result was due to the 40% of patients with comorbid AD, the association was also tested in PG patients without AD, and remained nominally significant. Thus, this association - if genuine - would be explained by genes common to both PG and AD.

Description of the top hits of the second analysis (PC 1 to 5, age and sex corrected), Table 2b in the main article:

RBM33, RNA binding motif protein 33, is located closely to *En2* (P-value: 0.79); sonic hedgehog (*SHH*, P-value: 0.0013); Insulin induced gene1 (*INSIG1*, P-value: 0.44); Canopy1 homolog (*CNPY1*, P-value: 0.62), Serotonin receptor 5A (*HTR5A*, P-value: 0.469). All five genes are co-expressed during brain development and have similar biological functions [29].

MIR3926-1, Micro RNAs are non-coding RNAs involved in post-transcriptional regulation of gene expression in multicellular organisms by affecting both the stability and translation of mRNAs. This micro RNA ranked 8th in the analysis without age and sex correction.

LONRF1, see rank 11 in the approach without age and sex correction.

MIR3926-2, see rank 6 in the analysis without age and sex correction.

PPY encodes a protein belonging to the neuropeptide Y (NPY) family of peptides. The small preproprotein is synthesised in the pancreatic islets of Langerhans. Two peptide products are generated by proteolytically processing creating the active pancreatic hormone and an icosapeptide of unknown function. The active hormone regulates pancreatic and gastrointestinal functions, and may be important in the regulation of food intake [30]. It has been implicated in brain-mediated effects on skeletal metabolism and as a regulator of energy homeostatic processes [31,32], and may also inhibit sexual behaviour in response to low-energy conditions [33]. NPY in noradrenergic neurons within the dorsomedial hypothalamus modulates the release and effects of catecholamines in a prolonged stress response [34], and its overexpression induces obesity in rodents [32,34].

MIR5003 also belongs to the group of non-coding RNAs involved in post-transcriptional regulation of gene expression in multicellular organisms by affecting both the stability and translation of mRNAs.

SH2D7 is the SH2 domain containing protein 7. Src homology 2 domains are involved in signal transduction [35,36].

FAM215A, Family with Sequence Similarity 215, Member A, is a non-protein coding gene which is also called APR-2. It is an RNA Gene, affiliated with the non-coding RNA class.

CNST, encodes the Consortin, Connexin Sorting Protein, alias C1orf71. This is an integral membrane protein, which acts as a binding partner of connexins, the building block of gap junctions. CNST is located in the trans-Golgi network, the plasma membrane, and tubulovesicular transport organelles. The receptor is involved in connexin targeting to the plasma membrane and recycling from the cell surface [37].

CTSA, Cathepsin A, is a glycoprotein that associates with the lysosomal enzymes beta-galactosidase and neuraminidase forming a complex of high molecular weight multimers. The protein can act as a protease, but also as a protective protein. Deficiencies in this gene are related to multiple forms of galactosialidosis [38].

PLTP, see above, in the descriptions of genes with PC 1 to 5 corrections, rank 2.

FERD3L, Fer3-Like BHLH Transcription Factor. This transcription factor inhibits transcription.

MAFB, V-Maf Avian Musculoaponeurotic Fibrosarcoma Oncogene Homolog B. The protein encoded by this gene is a basic leucine zipper transcription factor, with an important role in the regulation of lineage-specific hematopoiesis.

TFB2M, Transcription Factor B2, Mitochondrial. This gene encodes an S-adenosyl-L-methionine-dependent methyltransferase which specifically dimethylates mitochondrial 12S rRNA at the conserved stem loop. The protein is required for transcription of mitochondrial DNA and stimulates transcription independently of methyltransferase activity

ZSWIM1, ranked 7th in the first analysis.

SPATA25, ranked 5th in the first analysis.

NEURL2, ranked 4th in the first analysis.

ACTG1, Actins are highly conserved proteins that are involved in various types of cell motility, as well cytoskeleton maintenance. In vertebrates, the three main groups of known actin isoforms are alpha, beta, and gamma. The alpha actins are found in muscle, and are a major constituent of the contractile apparatus. This protein is a cytoplasmic actin found in non-muscle cells.

CIB2, calcium and integrin binding family member 2. The encoded protein is a calcium-binding regulatory protein that interacts with DNA-dependent protein kinase catalytic subunits (DNA-PKcs). It is involved in photoreceptor cell maintenance.

TWIST1 ranked 18th in the first analysis.

Table S2a: Comparison with top genes and candidate genes of the Australian GWAS [4]

		Results Lind			Analysis including PC 1 to 5			Analysis including PC 1 to 5, age and sex		
Gene	**Purpose Lind analysis**	**P-value Gene**	**TopSNP**	**P-value SNP**	**P-value**	**TopSNP**	**TopSNP P-value**	**P-value**	**TopSNP**	**TopSNP P-value**
*CDK5RAP2**	top gene rank 3	**4.56×10^{-4}**	rs10984956	**2.80 E-05**	**3.28×10^{-2}**	rs4837771	**9.76×10^{-4}**	2.13×10^{-1}	rs10984917	**5.73×10^{-3}**
*INSM2**	top gene rank 49	**2.73×10^{-3}**	rs17103397	**5.74×10^{-4}**	**3.24×10^{-2}**	rs2296919	**9.58×10^{-3}**	9.69×10^{-2}	rs2296919	5.79×10^{-2}
ADORA2A	candidate gene	5.32×10^{-1}	rs8141793	5.38×10^{-2}	6.27×10^{-1}	rs5751862	1.85×10^{-1}	4.15×10^{-1}	rs2236624	1.27×10^{-1}
ADRA2C	candidate gene	**1.50×10^{-2}**	rs11725040	**3.66×10^{-2}**	7.60×10^{-1}	rs2748763	9.63×10^{-2}	1.42×10^{-1}	rs6822427	**1.98×10^{-2}**
COMT	candidate gene	3.05×10^{-1}	rs2531716	**1.83×10^{-2}**	8.17×10^{-1}	rs4633	1.06×10^{-1}	7.29×10^{-1}	rs4633	6.12×10^{-2}
CREB1	candidate gene	**2.20×10^{-2}**	rs12998817	**2.33×10^{-3}**	7.09×10^{-2}	rs2709373	**8.27×10^{-3}**	1.64×10^{-1}	rs2042484	**1.50×10^{-2}**
DDC	candidate gene	5.96×10^{-1}	rs10235371	**1.23×10^{-2}**	8.37×10^{-1}	rs12718729	1.26×10^{-1}	6.07×10^{-1}	rs6593011	6.19×10^{-2}
DRD1	candidate gene	3.09×10^{-1}	rs251937	**1.32×10^{-2}**	5.74×10^{-1}	rs1121582	**3.92×10^{-2}**	3.75×10^{-1}	rs265973	**3.53×10^{-2}**
DRD2	candidate gene	2.44×10^{-1}	rs17529477	**5.46×10^{-3}**	4.35×10^{-1}	rs4479021	**4.61×10^{-2}**	5.80×10^{-1}	rs12574471	**3.13×10^{-2}**
DRD3	candidate gene	3.87×10^{-1}	rs7620955	**3.21×10^{-2}**	5.01×10^{-2}	rs2630349	**9.01×10^{-4}**	1.11×10^{-1}	rs2630349	**7.32×10^{-4}**
DRD4	candidate gene	9.82×10^{-1}	rs6598007	9.25×10^{-2}	8.13×10^{-1}	rs3758653	1.86×10^{-1}	6.75×10^{-1}	rs3758653	2.86×10^{-2}
DRD5	candidate gene	5.16×10^{-1}	rs1519094	5.93×10^{-2}	5.72×10^{-1}	rs13106539	1.64×10^{-1}	2.91×10^{-1}	rs10001006	1.15×10^{-1}
FOS	candidate gene	5.21×10^{-1}	rs6574222	5.99×10^{-2}	9.77×10^{-1}	rs8021524	1.60×10^{-1}	9.68×10^{-1}	rs7146378	1.66×10^{-1}
GRIN1	candidate gene	5.19×10^{-1}	rs12238250	5.36×10^{-2}	3.92×10^{-1}	rs4880094	7.74×10^{-2}	3.19×10^{-1}	rs34499319	**3.19×10^{-2}**
GRIN2B	candidate gene	$3.12x\ 10^{-1}$	rs10772723	**1.69×10^{-3}**	6.76×10^{-1}	rs2110984	**4.65×10^{-3}**	8.09×10^{-1}	rs1805502	2.33×10^{-2}
HTR1A	candidate gene	$6.93x\ 10^{-1}$	rs13361335	1.44×10^{-1}	9.54×10^{-1}	rs16892399	3.82×10^{-1}	9.33×10^{-1}	rs7735151	3.50×10^{-1}
HTR2A	candidate gene	4.51×10^{-1}	rs2094591	**4.05×10^{-2}**	1.08×10^{-1}	rs7323079	**1.55×10^{-2}**	1.40×10^{-1}	rs2760345	**4.59×10^{-3}**
HTR2B	candidate gene	8.70×10^{-2}	rs13424110	**3.59×10^{-2}**	5.43×10^{-1}	rs10187149	2.12×10^{-1}	2.46×10^{-1}	rs16827801	8.21×10^{-2}
NCS1	candidate gene	4.81×10^{-1}	rs2240913	**1.55×10^{-2}**	4.69×10^{-1}	rs10819601	**1.96×10^{-2}**	5.00×10^{-1}	rs10819601	**2.04×10^{-2}**
PSEN1	candidate gene	4.77×10^{-1}	rs362353	**4.44×10^{-2}**	**3.16×10^{-2}**	rs362384	**1.20×10^{-3}**	**3.20×10^{-2}**	rs362384	**1.91×10^{-3}**
SLC18A1	candidate gene	1.73×10^{-1}	rs2410639	**1.30×10^{-2}**	2.79×10^{-1}	rs17411601	**1.97×10^{-2}**	6.08×10^{-1}	rs17411601	**3.63×10^{-2}**
SLC18A2	candidate gene	4.77×10^{-1}	rs363241	**1.84×10^{-2}**	4.76×10^{-1}	rs11197936	**1.54×10^{-2}**	5.34×10^{-1}	rs11197936	5.13×10^{-2}
SLC6A3	candidate gene	6.17×10^{-1}	rs7732456	**2.88×10^{-2}**	1.82×10^{-1}	rs12516758	**4.50×10^{-3}**	5.05×10^{-1}	rs12516758	**4.07×10^{-2}**
SLC6A4	candidate gene	4.64×10^{-1}	rs2020941	**2.86×10^{-2}**	5.96×10^{-1}	rs11544945	**6.07×10^{-2}**	4.11×10^{-1}	rs11653777	**4.24×10^{-2}**
TH	candidate gene	1.16×10^{-1}	rs2070762	**1.21×10^{-3}**	3.98×10^{-1}	rs6579002	**3.26 E-02**	**1.73 E-02**	rs10743182	**4.90×10^{-3}**
TPH2	candidate gene	5.30 E-02	rs11179002	**6.98×10^{-3}**	3.58×10^{-1}	rs1872824	**2.73 E-02**	6.90×10^{-1}	rs1872824	**3.84 E-02**

Results of gene-wide analyses in Lind et al.[4] compared to results. P-value -gene refers to the P-value of the gene shown in the first column. The top SNP is the best hit in the gene-based analysis, as shown with its P-value.

Table S2b: Results for previously examined SNPs of molecular genetic studies of gambling

Gene	SNP	Study	Trait	P-value in this study	
				PC 1 to 5	PC 1 to 5, age, and sex
DRD3	rs167771	Lobo et al. 2015 [40]	DG	4.61×10^{-1}	1.85×10^{-1}
DRD3	rs6280 n.s. (Ser9Gly)	da Silva Lobo et al. 2007 and Lim et al. 2012 [41,42]	PG	4.26×10^{-1}	2.32×10^{-1}
DRD3	rs2630349	Best SNP for DRD3 in this study	PG	9.01×10^{-4}	7.32×10^{-4}
SLC6A3	- n.s.	da Silva Lobo et al. 2007 [41]	PG	1.8×10^{-1}	2.07×10^{-1}
SLC6A3	rs12516758 (nearby gene)	Best SNP for SLC6A in this study	PG	4.5×10^{-3}	4.07×10^{-2}
5HTR2A	rs6313	Wilson et al. 2013 [43]	PG	1.33×10^{-1}	1.478×10^{-1}
HTR2A	rs7323079	Best hit for HTR2A in this study	PG	1.6×10^{-2}	5.5×10^{-3}
CAMK2D	rs3815072	Lobo et al.2015 [40]	DG	6.849×10^{-1}	2.60×10^{-1}
CAMK2D	rs7664824	Best hit for gene in this study	PG	1.67×10^{-2}	3.24×10^{-3}

Comparisons with results from other molecular genetic studies of gambling

Table S2a shows a comparison with: (i) genes, described by Lind et al [4] who tested a candidate gene set derived from a candidate gene study for pathological gambling by Comings et al. [44] and literature on dopamine agonist-induced DG (see *candidate genes*, second column) and (ii) genes referring to the top 50 gene list of Lind et al. [4] (see *Top gene rank*, second column). For the top hits of Lind et al., only those genes that were also significant in at least one of the present analyses are shown [4]. The dopamine receptors 1 to 5 genes are listed in table S2a. Results for further previously investigated SNPs of molecular genetic studies of gambling are provided in table 2b. Neither of these reported findings achieved nominal association in the present analysis. However, some SNPs belonging to these genes had small P-values.

Pathways

Supplementary table S3a: Results of the KEGG pathways analyses including PC1 to 5 with P-values < 0.01 including SNP- and case-control permutation tests

Pathway ID	Pathway	P-value	P-value*	P-value case-control test	P-value SNP shuffling test
hsa05016	**Huntington's disease**	**2.58×10^{-5}**	**6.63×10^{-3}**	**1.00×10^{-4}**	**9.99×10^{-4}**
hsa04152	**AMPK signalling pathway**	**7.45×10^{-5}**	**9.57×10^{-3}**	**1.00×10^{-4}**	**3.00×10^{-3}**
hsa04210	**Apoptosis**	**2.05×10^{-4}**	**1.75×10^{-2}**	**1.00×10^{-4}**	**9.99×10^{-4}**
hsa04920	Adipocytokine signalling pathway	1.13×10^{-3}	5.83×10^{-2}	1.60×10^{-4}	5.99×10^{-3}
hsa04668	TNF signalling pathway	1.13×10^{-3}	5.83×10^{-2}	1.00×10^{-4}	1.10×10^{-2}
hsa00051	Fructose and mannose metabolism	1.45×10^{-3}	6.23×10^{-2}	2.00×10^{-4}	2.00×10^{-3}
hsa04910	Insulin signalling pathway	2.9×10^{-3}	1.10×10^{-1}	1.00×10^{-4}	> 0.05
hsa00410	beta-Alanine metabolism	3.44×10^{-3}	1.10×10^{-1}	5.00×10^{-4}	4.00×10^{-3}
hsa04915	Estrogen signalling pathway	5.23×10^{-3}	1.49×10^{-1}	1.00×10^{-4}	> 0.05
hsa04350	TGF-beta signalling pathway	8.23×10^{-3}	1.76×10^{-1}	1.00×10^{-4}	> 0.05
hsa05010	Alzheimer's disease	7.56×10^{-3}	1.76×10^{-1}	1.00×10^{-4}	> 0.05
hsa04024	cAMP signalling pathway	7.40×10^{-3}	1.76×10^{-1}	1.00×10^{-4}	> 0.05
hsa05030	Cocaine addiction	9.39×10^{-3}	1.86×10^{-1}	1.00×10^{-4}	3.20×10^{-2}

*Benjamini-Hochberg corrected P-values. P-values remaining significant after correction are shown in bold

The supplementary table S3a shows the P-values of the global test of all KEGG pathways resulting in a P-value <0.01; the corresponding P-values of the case-control test; and the SNP-shuffling permutation tests (1000 times). All listed pathways survived correction for multiple testing of the subject sampling method. Five of these pathways failed the SNP-label permutation test. Of 257 KEGG pathways, three pathways had a Benjamini-Hochberg corrected P-value of < 0.05 and a significant P-value < 0.01 in both the case-control permutation test and the SNP-shuffling test. These three pathways had Benjamini-Hochberg corrected P-value of < 0.05 and a significant P-value < 0.01 in the unpruned data set as well (data not shown).

Supplementary table S3b: Results of the KEGG pathways analyses including age and sex as covariates with P-values < 0.01, including SNP- and case-control permutation tests

Pathway ID	Pathway	P-value	P-value*	P-value case-control test	P-value SNP shuffling test
hsa04152	**AMPK signalling pathway**	5.36×10^{-4}	1.38×10^{-1}	4.92×10^{-4}	1.00×10^{-4}
hsa04340	Hedgehog signalling pathway	1.66×10^{-3}	1.64×10^{-1}	1.71×10^{-3}	6.00×10^{-4}
hsa05030	*Cocaine addiction*	1.94×10^{-3}	1.64×10^{-1}	4.92×10^{-4}	1.00×10^{-4}
hsa05410	Hypertrophic cardiomyopathy (HCM)	5.26×10^{-3}	1.64×10^{-1}	1.29×10^{-3}	4.00×10^{-4}
hsa04920	Adipocytokine signalling pathway	5.77×10^{-3}	1.64×10^{-1}	3.65×10^{-3}	1.60×10^{-4}
hsa05031	*Amphetamine addiction*	5.84×10^{-3}	1.64×10^{-1}	7.61×10^{-4}	2.00×10^{-4}
hsa05414	Dilated cardiomyopathy	6.13×10^{-3}	1.64×10^{-1}	7.61×10^{-4}	2.00×10^{-4}
hsa04910	Insulin signalling pathway	6.24×10^{-3}	1.64×10^{-1}	4.92×10^{-4}	1.00×10^{-4}
hsa05016	**Huntington disease (HD)**	6.64×10^{-3}	1.64×10^{-1}	4.92×10^{-4}	1.00×10^{-4}
hsa04932	Non-alcoholic fatty liver disease (NAFLD)	6.66×10^{-3}	1.64×10^{-1}	7.45×10^{-3}	3.80×10^{-4}
hsa04210	**Apoptosis**	7.01×10^{-3}	1.64×10^{-1}	4.92×10^{-4}	1.00×10^{-4}
hsa04921	Oxytocin signalling pathway	8.88×10^{-3}	1.90×10^{-1}	4.92×10^{-4}	1.00×10^{-4}

* Benjamini-Hochberg corrected P-values. Previously significant pathways (in the first analysis) shown in bold.

Supplementary table S3b shows the results of the KEGG analysis including age and sex in addition to PC 1 to 5, having P-values in the global test <0.01. Listed are also the corresponding P-values of the case-control test; and the SNP-shuffling permutation tests (1000 times). In the analysis including sex and age, no pathway had a Benjamini-Hochberg (BH) corrected P-value of < 0.05 (table 3b). Pathways that remained significant after BH correction in the previous analysis with PC components 1 to 5 (table 3a) are shown in bold.

Supplementary table S3c: Overlap of genes in the top KEGG pathways after including PC 1 to 5, with reference to results in table S3a

	hsa05016 (2008)	**hsa04152 (1743)**	**hsa04210 (823)**	**hsa04920 (1241)**	**hsa04668 (846)**	**hsa00051 (462)**	**hsa04910 (1549)**	**hsa00410 (417)**	**hsa04915 (1844)**	**hsa04350 (3580)**	**hsa05010 (1927)**	**hsa04024 (742)**	**hsa05030 (733)**
hsa05016 (2008)	1.00	0.10	0.16	0.04	0.10	0.02	0.05	0.06	0.16	0.07	0.62	0.10	0.27
hsa04152 (1743)	0.06	1.00	0.14	0.51	0.17	0.25	0.36	0.07	0.15	0.09	0.02	0.11	0.16
hsa04210 (823)	0.06	0.09	1.00	0.21	0.30	0.00	0.14	0.00	0.15	0.05	0.13	0.11	0.13
hsa04920 (1241)	0.01	0.32	0.21	1.00	0.21	0.00	0.19	0.00	0.06	0.05	0.04	0.06	0.04
hsa04668 (846)	0.05	0.14	0.39	0.27	1.00	0.01	0.13	0.00	0.30	0.08	0.07	0.16	0.27
hsa00051 (462)	0.00	0.07	0.00	0.00	0.00	1.00	0.03	0.03	0.00	0.01	0.01	0.00	0.00
hsa04910 (1549)	0.04	0.43	0.27	0.36	0.18	0.11	1.00	0.09	0.39	0.08	0.06	0.21	0.12
hsa00410 (417)	0.01	0.02	0.00	0.00	0.00	0.03	0.02	1.00	0.00	0.03	0.01	0.01	0.00
hsa04915 (1844)	0.08	0.13	0.20	0.08	0.30	0.01	0.27	0.00	1.00	0.04	0.10	0.27	0.41
hsa04350 (3580)	0.02	0.05	0.04	0.04	0.05	0.01	0.03	0.06	0.03	1.00	0.03	0.06	0.00
hsa05010 (1927)	0.54	0.03	0.30	0.08	0.12	0.03	0.08	0.06	0.18	0.08	1.00	0.12	0.24
hsa04024 (742)	0.09	0.17	0.26	0.15	0.30	0.01	0.27	0.07	0.49	0.17	0.13	1.00	0.60
hsa05030 (733)	0.07	0.08	0.09	0.03	0.15	0.00	0.05	0.00	0.22	0.00	0.07	0.18	1.00

Overlap between pathways shown in table S3a. Displaced values are the proportion of genes of the pathway in the column that are also part of the pathway listed in the row name. The number of genes of each pathway is shown in brackets.

Supplementary table S3d: Overlap of genes in the top KEGG pathways after including age and sex, with reference to results in table S3b

	hsa04152 (1745)	**hsa04340 (530)**	**hsa05030 (735)**	**hsa05410 (1950)**	**hsa04920 (848)**	**hsa05031 (1247)**	**hsa05414 (2092)**	**hsa04910 (1551)**	**hsa05016 (2010)**	**hsa04932 (1160)**	**hsa04210 (825)**	**hsa04921 (3604)**
hsa04152 (1745)	1.00	0.05	0.16	0.12	0.51	0.16	0.03	0.36	0.06	0.19	0.14	0.14
hsa04340 (530)	0.02	1.00	0.07	0.02	0.02	0.06	0.04	0.04	0.00	0.02	0.05	0.03
hsa05030 (735)	0.08	0.10	1.00	0.00	0.03	0.60	0.07	0.05	0.07	0.04	0.09	0.10
hsa05410 (1950)	0.08	0.04	0.01	1.00	0.14	0.06	0.83	0.05	0.00	0.08	0.03	0.17
hsa04920 (848)	0.32	0.04	0.04	0.13	1.00	0.00	0.04	0.19	0.01	0.20	0.21	0.07
hsa05031 (1247)	0.09	0.10	0.68	0.05	0.00	1.00	0.11	0.10	0.06	0.03	0.10	0.21
hsa05414 (2092)	0.02	0.10	0.11	0.90	0.05	0.15	1.00	0.03	0.00	0.04	0.08	0.22
hsa04910 (1551)	0.43	0.17	0.12	0.09	0.36	0.24	0.04	1.00	0.04	0.20	0.27	0.26
hsa05016 (2010)	0.10	0.02	0.27	0.01	0.04	0.20	0.01	0.05	1.00	0.55	0.16	0.06
hsa04932 (1160)	0.24	0.08	0.12	0.16	0.40	0.07	0.07	0.21	0.42	1.00	0.36	0.09
hsa04210 (825)	0.09	0.10	0.13	0.03	0.21	0.12	0.08	0.14	0.06	0.18	1.00	0.11
hsa04921 (3604)	0.17	0.13	0.26	0.32	0.13	0.48	0.38	0.26	0.05	0.09	0.21	1.00

Overlap between pathways shown in table S3b. Displayed values are the proportion of genes of the pathway in the column that are also part of the pathway listed in the row name. The number of genes of each pathway is shown in brackets.

Supplementary table S3e: Additional interesting results of the KEGG pathways analyses with P-values < 0.05

PW ID	Pathway	P-value [1]	P-value [1*]	P-value [2]	P-value [2*]
hsa00760	Nicotinate and nicotinamide metabolism	**1.20×10^{-2}**	2.05×10^{-1}	**1.42×10^{-2}**	2.37×10^{-1}
hsa04261	Adrenergic signalling in cardiomyocytes	**2.51×10^{-2}**	3.05×10^{-1}	**1.49×10^{-2}**	2.37×10^{-1}
hsa04710	Circadian rhythm	**2.09×10^{-2}**	2.69×10^{-1}	**1.77×10^{-2}**	2.37×10^{-1}
hsa04024	cAMP signalling pathway	**Top list**	Top list	**1.91×10^{-2}**	2.37×10^{-1}
hsa04728	Dopaminergic synapse	7.35×10^{-2}	3.28×10^{-1}	**2.54×10^{-2}**	2.51×10^{-1}
hsa04340	Hedgehog signalling pathway	**2.75×10^{-2}**	3.05×10^{-1}	**Top list**	**Top list**
hsa04310	Wnt signalling pathway	**3.80×10^{-2}**	3.05×10^{-1}	3.93×10^{-1}	1.39×10^{-1}
hsa04725	Acetylcholine (ACh)	5.50×10^{-2}	3.28×10^{-1}	**4.76×10^{-2}**	2.91×10^{-1}
hsa04726	Serotonergic synapse	**4.61×10^{-2}**	3.28×10^{-1}	3.71×10^{-1}	1.13×10^{-1}
hsa05034	Alcoholism	2.22×10^{-1}	4.68×10^{-1}	**4.81×10^{-2}**	2.91×10^{-1}
hsa05012	Parkinson's disease	5.46×10^{-2}	3.28×10^{-1}	**4.93×10^{-2}**	2.91×10^{-1}

1. First analysis including PC 1 to 5 2: Second Analysis, including also age and sex. *BH-corrected P-values - P-values < 0.05 are shown in bold.

This table shows pathways ranking lower than those in the top list for at least one analysis, but which appeared interesting and whose P-values were still under the nominal threshold.

KEGG pathways

The three significant genome-wide significant pathways (table S3a) are described in the main paper. A possible link between these pathways is that diseases such as Huntington's, Alzheimer's, and Parkinson's are characterized by an over-activation of AMPK [42,45,46].

After correction for age and sex, no pathway remained significant (see table 3b). However, the previously 2nd ranked pathway, AMPK signalling, was ranked first, with a P-value of 5.36 x 10^{-4}.

KEGG analysis including corrections for age and sex generated two interesting new pathways (see table S3b). One of these was the Hedgehog signalling pathway, which ranked second. This pathway is involved in dopaminergic and serotonergic cell fate [47]. The gene *SHH*, sonic hedgehog, had a P-value of 0.0156 (first analysis) and 1.37 x 10^{-3} in the second analysis. Research has shown that Shh regulates granule cell precursors in the cerebellum. Treatment of these cells with Shh prevents differentiation, and induces a long lasting proliferative response. Blocking Shh function in vivo reduces granule cell proliferation [48]. Shh is expressed along the ventral neural tube. Together with FGF8, it creates induction sites for dopaminergic neurons in the mid- and forebrain. After induction by another signal, it defines an inductive centre for hindbrain 5-HT neurons [47].

The third ranked pathway in the second analysis was cocaine addiction, with a P-value of 1.94 x 10^{-3}. In the first analysis, this was a top finding with a P-value of 9.39 x 10^{-3}. In the disordered gambling GWAS [4], three pathways (synaptic long term potentiation, synaptic long term depression, gonadotrophin releasing hormone [GNRH] signalling) were under the most significant in their Ingenuity Pathway Analysis. Previous authors reported that they are enriched for substance addiction-related genes, with the synaptic long term depression and GNRH signalling pathways being common to cocaine-, alcohol-, opioid-, and nicotine addiction [4,16].

The pathway for amphetamine addiction is also novel (see table S3b). Cocaine and amphetamine regulated transcripts are widely expressed in the hypothalamus, involved in food intake control, and regulated by leptin. Leptin is suggested to have an effect on GnRH secretion [49]

Interestingly, pathways for the dopaminergic, serotonergic-, and cholinergic synapses all had P-values of < 0.05 in one of the analyses. The pathways for alcoholism and Parkinson's were nominally significant after correction for age and sex (see table S3e).

The cocaine addiction pathway was found in two, and the amphetamine addiction pathway was found in one of the two analyses as a KEGG top result with P-values < 0.01. This, and that the pathway for alcoholism was also nominally significant might be of interest given previous evidence that PG resembles substance-related addictions in many domains [50].

Comparison of KEGG pathways with Lind et al. [4]**:**

In the age and sex corrected analyses (see table S3b), the pathway hsa05410, Hypertrophic cardiomyopathy (HCM) was ranked 4th, and dilated cardiomyopathy, hsa05414 was ranked 7th. In the list of Enrichment of KEGG pathways in Lind et al. [4], these pathway were the 6thand 10th pathways respectively, with P-values of 1.45 x $10^{-7}$and 2.01 x 10^{-6}.

No other KEGG pathway in the top list of enriched KEGG pathways of Lind et al. [4] had P-values <0.05 in the present sample.

Of the three above mentioned (p.21) Ingenuity pathways reported in Lind et al. [4], synaptic long term potentiation, synaptic long term depression, and GnRH signalling pathway were not significant in the present analyses. Long term potentiation, hsa04720, had the lowest P-value with 0.0702 (BH-corrected: 0.328) and 0.106 (BH-corrected: 0.363) in the first and second analysis, respectively.

Reactome: Supplementary table S4a: Reactome global test results, including PC 1 to 5, with P-values of <0.01and SNP- and case-control permutation test results for pathways

Pathway	P-value	P-value*	P-value case-control	P-value SNP shuffling
Homo sapiens: Integration of energy metabolism	2.51×10^{-4}	2.96×10^{-1}	1.00×10^{-4}	9.99×10^{-4}
Homo sapiens: Translocation of GLUT4 to the Plasma Membrane	1.02×10^{-3}	3.48×10^{-1}	1.00×10^{-4}	9.99×10^{-4}
Homo sapiens: Regulation of Insulin Secretion	1.26×10^{-3}	3.48×10^{-1}	1.00×10^{-4}	9.99×10^{-4}
Homo sapiens: PKA-mediated phosphorylation of key metabolic factors	1.56×10^{-3}	3.48×10^{-1}	2.00×10^{-4}	9.99×10^{-4}
Homo sapiens: Glucose metabolism	3.11×10^{-3}	3.48×10^{-1}	1.00×10^{-4}	9.99×10^{-4}
Homo sapiens: ERK activation	3.13×10^{-3}	3.48×10^{-1}	2.03×10^{-2}	9.99×10^{-4}
Homo sapiens: DAP12 interactions	3.50×10^{-3}	3.48×10^{-1}	1.00×10^{-4}	9.99×10^{-4}
Homo sapiens: IRS-mediated signalling	3.86×10^{-3}	3.48×10^{-1}	1.00×10^{-4}	9.99×10^{-4}
Homo sapiens: Regulation of Rheb GTPase activity by AMPK	4.48×10^{-3}	3.48×10^{-1}	1.50×10^{-3}	9.99×10^{-4}
Homo sapiens: DAP12 signalling	5.03×10^{-3}	3.48×10^{-1}	1.00×10^{-4}	9.99×10^{-4}
Homo sapiens: G-protein mediated events	6.77×10^{-3}	3.48×10^{-1}	1.00×10^{-4}	4.00×10^{-3}
Homo sapiens: Meiosis	6.87×10^{-3}	3.48×10^{-1}	5.00×10^{-4}	9.99×10^{-4}
Homo sapiens: PLC beta mediated events	7.37×10^{-3}	3.48×10^{-1}	1.00×10^{-4}	4.00×10^{-3}
Homo sapiens: IRS-related events	8.37×10^{-3}	3.48×10^{-1}	2.00×10^{-4}	4.00×10^{-3}
Homo sapiens: Caspase-mediated cleavage of cytoskeletal proteins	8.55×10^{-3}	3.48×10^{-1}	7.80×10^{-3}	9.99×10^{-4}
Homo sapiens: N-Glycan antennae elongation	8.61×10^{-3}	3.48×10^{-1}	4.00×10^{-4}	9.99×10^{-4}
Homo sapiens: Binding and Uptake of Ligands by Scavenger Receptors	8.78×10^{-3}	3.48×10^{-1}	1.00×10^{-4}	3.00×10^{-3}
Homo sapiens: SMAC-mediated dissociation of IAP: caspase complexes	8.87×10^{-3}	3.48×10^{-1}	1.12×10^{-2}	2.00×10^{-3}
Homo sapiens: SMAC-mediated apoptotic response	8.87×10^{-3}	3.48×10^{-1}	1.20×10^{-2}	2.00×10^{-3}
Homo sapiens: SMAC binds to IAPs	8.87×10^{-3}	3.48×10^{-1}	1.19×10^{-2}	2.00×10^{-3}
Homo sapiens: Scavenging by Class F Receptors	9.15×10^{-3}	3.48×10^{-1}	2.31 x 10-2	2.00×10^{-3}
Homo sapiens: Insulin receptor signalling cascade	9.35×10^{-3}	3.48×10^{-1}	3.0 x 10-4	4.00×10^{-3}
Homo sapiens: Apoptotic cleavage of cellular proteins	9.89×10^{-3}	3.48×10^{-1}	3.7 x 10-3	3.00×10^{-3}

***Benjamini-Hochberg corrected**

Reactome:

Supplementary table S4b: Reactome global test results, including PC 1 to 5, age and sex, with P-values of <0.01, and SNP- and case-control permutation test results for pathways.

Pathway	P-value	P-value*	P-value case-control	P-value SNP shuffling
Homo sapiens: Translocation of GLUT4 to the Plasma Membrane	8.91×10^{-4}	5.89×10^{-1}	1.00×10^{-4}	9.99×10^{-4}
Homo sapiens: Scavenging by Class A Receptors	2.92×10^{-3}	5.89×10^{-1}	4.00×10^{-4}	9.99×10^{-4}
Homo sapiens: Regulation of Rheb GTPase activity by AMPK	4.60×10^{-3}	5.89×10^{-1}	1.50×10^{-3}	9.99×10^{-4}
Homo sapiens: Integration of energy metabolism	4.83×10^{-3}	5.89×10^{-1}	1.00×10^{-4}	9.99×10^{-4}
Homo sapiens: Signalling by Type 1 Insulin-like Growth Factor 1 Receptor (IGF1R)	4.99×10^{-3}	5.89×10^{-1}	1.00×10^{-4}	9.99×10^{-4}
Homo sapiens: IGF1R signalling cascade	4.99×10^{-3}	5.89×10^{-1}	1.00×10^{-4}	2.00×10^{-3}
Homo sapiens: PKA-mediated phosphorylation of key metabolic factors	5.02×10^{-3}	5.89×10^{-1}	2.00×10^{-4}	2.00×10^{-3}
Homo sapiens: Binding and Uptake of Ligands by Scavenger Receptors	5.88×10^{-3}	5.89×10^{-1}	1.00×10^{-4}	9.99×10^{-4}
Homo sapiens: IRS-related events triggered by IGF1R	6.38×10^{-3}	5.89×10^{-1}	2.00×10^{-4}	9.99×10^{-4}
Homo sapiens: Glucose metabolism	6.46×10^{-3}	5.89×10^{-1}	1.00×10^{-4}	2.00×10^{-3}
Homo sapiens: IRS-related events	6.48×10^{-3}	5.89×10^{-1}	2.00×10^{-4}	9.99×10^{-4}
Homo sapiens: Meiosis	6.97×10^{-3}	5.89×10^{-1}	5.00×10^{-4}	2.00×10^{-3}
Homo sapiens: Displacement of DNA glycosylase by APE1	7.09×10^{-3}	5.89×10^{-1}	2.06×10^{-1}	9.99×10^{-4}
Homo sapiens: Base-free sugar-phosphate removal via the single-nucleotide replacement pathway	7.09×10^{-3}	5.89×10^{-1}	1.98×10^{-1}	3.00×10^{-3}
Homo sapiens: Formation of tubulin folding intermediates by CCT TriC	7.72×10^{-3}	5.89×10^{-1}	3.24×10^{-2}	2.00×10^{-3}
Homo sapiens: Folding of actin by CCT TriC	8.19×10^{-3}	5.89×10^{-1}	2.76×10^{-2}	2.00×10^{-3}
Homo sapiens: IRS-mediated signalling	8.49×10^{-3}	5.89×10^{-1}	1.00×10^{-4}	9.99×10^{-4}
Homo sapiens: Insulin receptor signalling cascade	9.15×10^{-3}	5.93×10^{-1}	3.00×10^{-4}	5.00×10^{-3}

***Benjamini-Hochberg corrected**

Reactome pathways

In table S4a and b, results of analyses for the Reactome pathways are shown. For the first analysis, including PC 1 to PC5, the best pathway was *Integration of energy metabolism*. This pathway was ranked 4 in the age and sex corrected analysis.

In the age and sex corrected approach, translocation of GLUT4 to the Plasma Membrane was the best pathway, having being ranked 2 in the first analysis. When carbohydrates are ingested, insulin stimulated glucose transport into skeletal muscle is the major cellular mechanisms in terms of diminishing blood glucose. Glucose is stored there as glycogen and is oxidised to produce energy. The principal glucose transporter protein mediating this uptake is GLUT4, which therefore plays a key role in regulating glucose homeostasis [51].

Decreased expression of glucose transporter protein GLUT4, encoded by the solute carrier 2A4 gene, is involved in obesity-induced insulin resistance. Local tissue inflammation, via a nuclear factor-κB (NFκB)-mediated pathway, has been related to *Slc2a4* repression; a mechanism that could be modulated by statins [52].

Animal models have implicated both, energy metabolism, and GLUT 4 in Huntington's [53,54].

Supplementary table S5a: GO global test results of P-values < 10^{-3} for the first analysis

GO ID	GO Name	P-value	P-value*
GO:0003278	apoptotic process involved in heart morphogenesis	1.78×10^{-5}	2.16×10^{-1}
GO:0008037	cell recognition	4.05×10^{-5}	1.04×10^{-1}
GO:0005868	cytoplasmic dynein complex	5.10×10^{-5}	1.21×10^{-1}
GO:0009566	Fertilization	6.73×10^{-5}	1.04×10^{-1}
GO:0004176	ATP-dependent peptidase activity	8.58×10^{-5}	2.25×10^{-1}
GO:0005858	axonemal dynein complex	9.47×10^{-5}	1.27×10^{-1}
GO:0044447	axoneme part	1.02×10^{-4}	1.20×10^{-1}
GO:0051890	regulation of cardioblast differentiation	2.19×10^{-4}	1.53×10^{-1}
GO:0001653	peptide receptor activity	2.90×10^{-4}	1.04×10^{-1}
GO:0019203	carbohydrate phosphatase activity	2.93×10^{-4}	1.43×10^{-1}
GO:0050308	sugar-phosphatase activity	3.32×10^{-4}	1.43×10^{-1}
GO:0000338	protein deneddylation	3.82×10^{-4}	2.04×10^{-1}
GO:0010388	cullin deneddylation	3.82×10^{-4}	2.04×10^{-1}
GO:0008528	G-protein coupled peptide receptor activity	3.91×10^{-4}	1.04×10^{-1}
GO:0006000	fructose metabolic process	4.59×10^{-4}	1.52×10^{-1}
GO:0004691	cAMP-dependent protein kinase activity	4.72×10^{-4}	1.76×10^{-1}
GO:0007340	acrosome reaction	4.93×10^{-4}	1.20×10^{-1}
GO:0045954	positive regulation of natural killer cell mediated cytotoxicity	6.27×10^{-4}	1.33×10^{-1}
GO:0044744	protein targeting to nucleus	6.36×10^{-4}	1.05×10^{-1}
GO:1902554	serine/threonine protein kinase complex	6.55×10^{-4}	1.09×10^{-1}
GO:0032852	#positive regulation of Ral GTPase activity	6.62×10^{-4}	2.15×10^{-1}
GO:0032859	#activation of GTPase activity	6.62×10^{-4}	2.15×10^{-1}
GO:0007338	single fertilization	6.73×10^{-4}	1.04×10^{-1}
GO:0030286	dynein complex	6.83×10^{-4}	1.08×10^{-1}
GO:0014855	striated muscle cell proliferation	7.21×10^{-4}	1.05×10^{-1}
GO:0005927	muscle tendon junction	8.08×10^{-4}	1.87×10^{-1}
GO:0000800	lateral element	8.74×10^{-4}	1.46×10^{-1}
GO:0002717	positive regulation of natural killer cell mediated immunity	8.89×10^{-4}	1.28×10^{-1}
GO:0006003	fructose 2,6-bisphosphate metabolic process	8.93×10^{-4}	1.52×10^{-1}
GO:0090090	negative regulation of canonical Wnt signalling pathway	9.53×10^{-4}	1.04×10^{-1}
GO:0032315	regulation of GTPase activity	9.84×10^{-4}	1.54×10^{-1}
GO:0032485	regulation of Ral protein signal transduction	9.84×10^{-4}	1.54×10^{-1}

***Benjamini-Hochberg corrected**

Supplementary table S5b: GO global test results of P-values < 10^{-3} for the second analysis, including age and sex

GO ID	GO Name	P-value	P-value*
GO:0090090	Wnt signalling pathway	1.57×10^{-4}	5.43×10^{-1}
GO:0050308	*sugar-phosphatase activity*	2.69×10^{-4}	5.43×10^{-1}
GO:0019203	*carbohydrate phosphatase activity*	2.75×10^{-4}	5.43×10^{-1}
GO:0004331	*fructose-2,6-bisphosphate 2-phosphatase activity*	3.16×10^{-4}	5.43×10^{-1}
GO:0003278	apoptotic process involved in heart morphogenesis	3.21×10^{-4}	5.43×10^{-1}
GO:0006003	*fructose 2,6-bisphosphate metabolic process*	5.85×10^{-4}	5.77×10^{-1}
GO:0060828	*regulation of canonical Wnt signalling pathway*	6.23×10^{-4}	5.77×10^{-1}
GO:0051890	*regulation of cardioblast differentiation*	7.02×10^{-4}	5.77×10^{-1}
GO:0004176	*ATP-dependent peptidase activity*	7.06×10^{-4}	5.77×10^{-1}
GO:0010827	*regulation of glucose transport*	7.91×10^{-4}	5.77×10^{-1}
GO:0046324	*regulation of glucose import*	8.36×10^{-4}	5.77×10^{-1}
GO:0046323	*glucose import*	8.87×10^{-4}	5.77×10^{-1}
GO:0006584	*catecholamine metabolic process*	9.54×10^{-4}	5.77×10^{-1}
GO:0009712	*catechol-containing compound metabolic process*	9.54×10^{-4}	5.77×10^{-1}

***Benjamini-Hochberg corrected**

Gene Ontology gene sets

None of the corrected P-values were significant. For GO, no permutation tests were performed due to its hierarchical structure.

The Wnt signalling pathway is a developmental pathway, ranking first in the age and sex corrected analysis of GO and on rank 30 without this correction. However, adult neurogenesis is also tightly regulated by multiple signalling pathways, including the canonical Wnt/β-catenin pathway [55]. Wnt glycoproteins activate several signalling pathways, and have key functions in midbrain dopaminergic neuron development [56].

Polygenic risk scores

a) b)

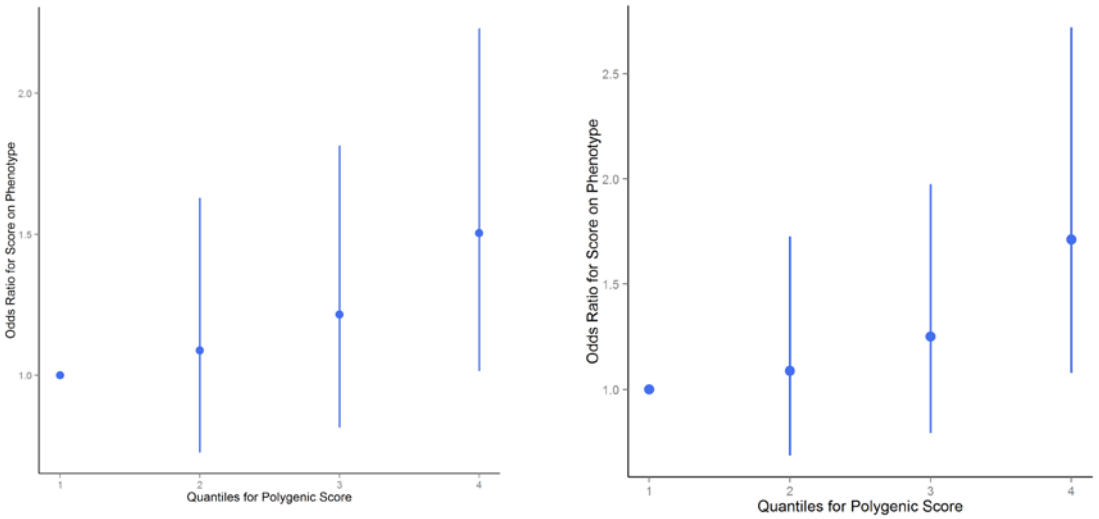

Figure S1a and b: Quartile plots for the polygenic risk score of alcohol dependence (including P-values < 0.5). Image S1a includes only PC1 to 5; b also includes age and sex correction. Polygenic risk scores were converted to quartiles, and quartile 1 was used as reference. Odds ratios and 9 % confidence intervals were estimated using logistic regression with five principal components to control for population stratification.

a) b)

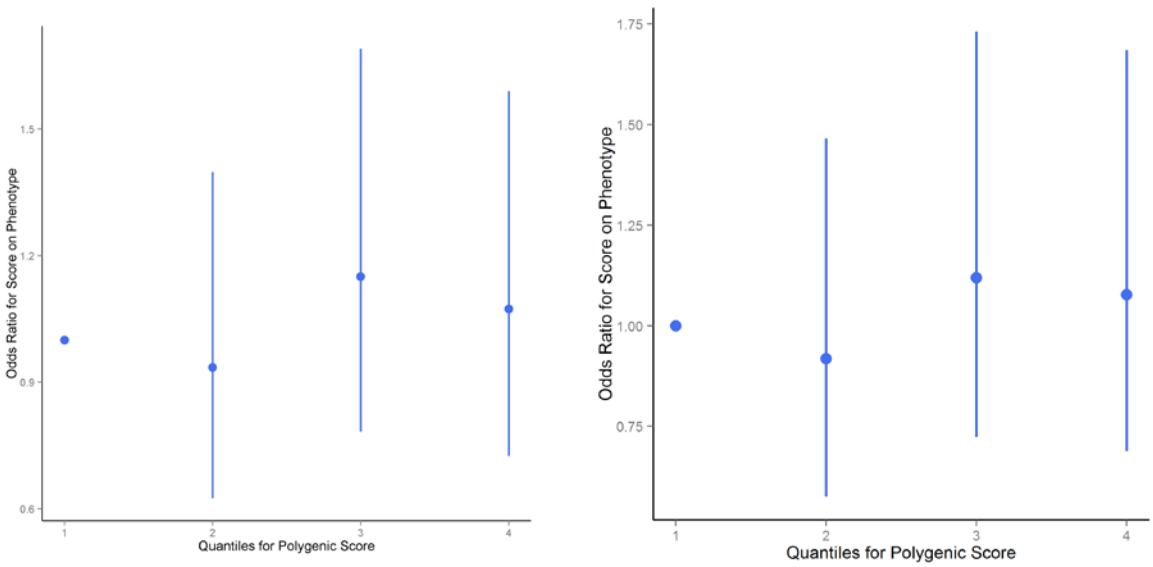

Figure S2a and b: Quartile plots for the polygenic risk score of disordered gambling (including P-values <0.5). Image S2a includes only PC1 to 5; b also includes age and sex correction. Polygenic risk scores were converted to quartiles, and quartile 1 was used as reference. Odds ratios and 95% confidence intervals were estimated using logistic regression with five principal components to control for population stratification.

References:

[1] Mishra A, Macgregor S. VEGAS2: Software for More Flexible Gene-Based Testing. Twin Res Hum Genet Off J Int Soc Twin Stud 2015;18:86–91. doi:10.1017/thg.2014.79.

[2] Liu JZ, Mcrae AF, Nyholt DR, Medland SE, Wray NR, Brown KM, et al. A Versatile Gene-Based Test for Genome-wide Association Studies. Am J Hum Genet 2010;87:139–45. doi:10.1016/j.ajhg.2010.06.009.

[3] Frank J, Cichon S, Treutlein J, Ridinger M, Mattheisen M, Hoffmann P, et al. Genome-wide significant association between alcohol dependence and a variant in the ADH gene cluster. Addict Biol 2012;17:171–80. doi:10.1111/j.1369-1600.2011.00395.x.

[4] Lind PA, Zhu G, Montgomery GW, Madden PAF, Heath AC, Martin NG, et al. Genome-wide association study of a quantitative disordered gambling trait. Addict Biol 2013;18:511–22. doi:10.1111/j.1369-1600.2012.00463.x.

[5] http://www.genome.jp/kegg/pathway.html. KEGG n.d.

[6] http://geneontology.org/page/go-database. Reactome n.d.

[7] www.geneontology.org. Geneontology n.d.

[8] Carson, M. org.Hs.eg.db: Genome wide annotation for Human. R package version 2.14.0. n.d.

[9] Juraeva D, Haenisch B, Zapatka M, Frank J, GROUP Investigators, PSYCH-GEMS SCZ Working Group, et al. Integrated pathway-based approach identifies association between genomic regions at CTCF and CACNB2 and schizophrenia. PLoS Genet 2014;10:e1004345. doi:10.1371/journal.pgen.1004345.

[10] Pruitt KD, Brown GR, Hiatt SM, Thibaud-Nissen F, Astashyn A, Ermolaeva O, et al. RefSeq: an update on mammalian reference sequences. Nucleic Acids Res 2014;42:D756–63. doi:10.1093/nar/gkt1114.

[11] Efron B, Tibshirani R. On testing the significance of sets of genes. Ann Appl Stat 207AD;1:107–29.

[12] Lind PA, Macgregor S, Vink JM, Pergadia ML, Hansell NK, de Moor MHM, et al. A genomewide association study of nicotine and alcohol dependence in Australian and Dutch populations. Twin Res Hum Genet Off J Int Soc Twin Stud 2010;13:10–29. doi:10.1375/twin.13.1.10.

[13] Rungta RL, Choi HB, Tyson JR, Malik A, Dissing-Olesen L, Lin PJC, et al. The cellular mechanisms of neuronal swelling underlying cytotoxic edema. Cell 2015;161:610–21. doi:10.1016/j.cell.2015.03.029.

[14] Zuo L, Lu L, Tan Y, Pan X, Cai Y, Wang X, et al. Genome-wide association discoveries of alcohol dependence: GWAS of Alcohol Dependence. Am J Addict 2014;23:526–39. doi:10.1111/j.1521-0391.2014.12147.x.

[15] Santos CRA, Martinho A, Quintela T, Gonçalves I. Neuroprotective and neuroregenerative properties of metallothioneins. IUBMB Life 2012;64:126–35. doi:10.1002/iub.585.

[16] Li C-Y, Mao X, Wei L. Genes and (common) pathways underlying drug addiction. PLoS Comput Biol 2008;4:e2. doi:10.1371/journal.pcbi.0040002.

[17] Hirose Y, Iwamoto Y, Sakuraba K, Yunokuchi I, Harada F, Ohkuma Y. Human phosphorylated CTD-interacting protein, PCIF1, negatively modulates gene expression by RNA polymerase II. Biochem Biophys Res Commun 2008;369:449–55. doi:10.1016/j.bbrc.2008.02.042.

[18] Pussinen PJ, Metso J, Malle E, Barlage S, Palosuo T, Sattler W, et al. The role of plasma phospholipid transfer protein (PLTP) in HDL remodeling in acute-phase patients. Biochim Biophys Acta 2001;1533:153–63.

[19] Huuskonen J, Olkkonen VM, Jauhiainen M, Ehnholm C. The impact of phospholipid transfer protein (PLTP) on HDL metabolism. Atherosclerosis 2001;155:269–81.

[20] http://www.ncbi.nlm.nih.gov/gene/128497. SPATA25 n.d.

[21] Ko KK, Powell MS, Hogarth PM. ZSWIM1: a novel biomarker in T helper cell differentiation. Immunol Lett 2014;160:133–8. doi:10.1016/j.imlet.2014.01.016.

[22] Loges NT, Olbrich H, Fenske L, Mussaffi H, Horvath J, Fliegauf M, et al. DNAI2 mutations cause primary ciliary dyskinesia with defects in the outer dynein arm. Am J Hum Genet 2008;83:547–58. doi:10.1016/j.ajhg.2008.10.001.

[23] Luciano M, Hansell NK, Lahti J, Davies G, Medland SE, Räikkönen K, et al. Whole genome association scan for genetic polymorphisms influencing information processing speed. Biol Psychol 2011;86:193–202. doi:10.1016/j.biopsycho.2010.11.008.

[24] Bell RL, Kimpel MW, McClintick JN, Strother WN, Carr LG, Liang T, et al. Gene expression changes in the nucleus accumbens of alcohol-preferring rats following chronic ethanol consumption. Pharmacol Biochem Behav 2009;94:131–47. doi:10.1016/j.pbb.2009.07.019.

[25] Kakiuchi C, Ishiwata M, Nanko S, Kunugi H, Minabe Y, Nakamura K, et al. Functional polymorphisms of HSPA5: Possible association with bipolar disorder. Biochem Biophys Res Commun 2005;336:1136–43. doi:10.1016/j.bbrc.2005.08.248.

[26] Kress W, Schropp C, Lieb G, Petersen B, Büsse-Ratzka M, Kunz J, et al. Saethre-Chotzen syndrome caused by TWIST 1 gene mutations: functional differentiation from Muenke coronal synostosis syndrome. Eur J Hum Genet EJHG 2006;14:39–48. doi:10.1038/sj.ejhg.5201507.

[27] http://www.uniprot.org/uniprot/Q2TAC6. KIF19 n.d.

[28] Mantha K, Laufer BI, Singh SM. Molecular Changes during Neurodevelopment following Second-Trimester Binge Ethanol Exposure in a Mouse Model of Fetal Alcohol Spectrum Disorder: From Immediate Effects to Long-Term Adaptation. Dev Neurosci 2014;36:29–43. doi:10.1159/000357496.

[29] Choi J, Ababon MR, Soliman M, Lin Y, Brzustowicz LM, Matteson PG, et al. Autism Associated Gene, ENGRAILED2, and Flanking Gene Levels Are Altered in Post-Mortem Cerebellum. PLoS ONE 2014;9:e87208. doi:10.1371/journal.pone.0087208.

[30] http://www.ncbi.nlm.nih.gov/gene/5539. PPY n.d.

[31] Horsnell H, Baldock PA. Osteoblastic Actions of the Neuropeptide Y System to Regulate Bone and Energy Homeostasis. Curr Osteoporos Rep 2016. doi:10.1007/s11914-016-0300-9.

[32] Kim YJ, Bi S. Knockdown of neuropeptide Y in the dorsomedial hypothalamus reverses high-fat diet-induced obesity and impaired glucose tolerance in rats. Am J Physiol Regul Integr Comp Physiol 2016;310:R134–42. doi:10.1152/ajpregu.00174.2015.

[33] Inaba A, Komori Y, Muroi Y, Kinoshita K, Ishii T. Neuropeptide Y signaling in the dorsal raphe nucleus inhibits male sexual behavior in mice. Neuroscience 2016;320:140–8. doi:10.1016/j.neuroscience.2016.01.069.

[34] Vähätalo LH, Ruohonen ST, Ailanen L, Savontaus E. Neuropeptide Y in noradrenergic neurons induces obesity in transgenic mouse models. Neuropeptides 2016;55:31–7. doi:10.1016/j.npep.2015.11.088.

[35] Marchler-Bauer A, Derbyshire MK, Gonzales NR, Lu S, Chitsaz F, Geer LY, et al. CDD: NCBI's conserved domain database. Nucleic Acids Res 2015;43:D222–6. doi:10.1093/nar/gku1221.

[36] SH2D7, NCBI website n.d.

[37] del Castillo FJ, Cohen-Salmon M, Charollais A, Caille D, Lampe PD, Chavrier P, et al. Consortin, a trans-Golgi network cargo receptor for the plasma membrane targeting and recycling of connexins. Hum Mol Genet 2010;19:262–75. doi:10.1093/hmg/ddp490.

[38] CTSA n.d.

[39] Lobo DSS, Souza RP, Tong RP, Casey DM, Hodgins DC, Smith GJ, et al. Association of functional variants in the dopamine D2-like receptors with risk for gambling behaviour in healthy Caucasian subjects. Biol Psychol 2010;85:33–7. doi:10.1016/j.biopsycho.2010.04.008.

[40] Lobo DSS, Aleksandrova L, Knight J, Casey DM, el-Guebaly N, Nobrega JN, et al. Addiction-related genes in gambling disorders: new insights from parallel human and pre-clinical models. Mol Psychiatry 2015;20:1002–10. doi:10.1038/mp.2014.113.

[41] da Silva Lobo DS, Vallada HP, Knight J, Martins SS, Tavares H, Gentil V, et al. Dopamine genes and pathological gambling in discordant sib-pairs. J Gambl Stud Co-Spons Natl Counc Probl Gambl Inst Study Gambl Commer Gaming 2007;23:421–33. doi:10.1007/s10899-007-9060-x.

[42] Lim S, Ha J, Choi S-W, Kang S-G, Shin Y-C. Association study on pathological gambling and polymorphisms of dopamine D1, D2, D3, and D4 receptor genes in a Korean population. J Gambl Stud Co-Spons Natl Counc Probl Gambl Inst Study Gambl Commer Gaming 2012;28:481–91. doi:10.1007/s10899-011-9261-1.

[43] Wilson D, da Silva Lobo DS, Tavares H, Gentil V, Vallada H. Family-based association analysis of serotonin genes in pathological gambling disorder: evidence of vulnerability risk in the 5HT-2A receptor gene. J Mol Neurosci MN 2013;49:550–3. doi:10.1007/s12031-012-9846-x.

[44] Comings DE, Gade-Andavolu R, Gonzalez N, Wu S, Muhleman D, Chen C, et al. The additive effect of neurotransmitter genes in pathological gambling. Clin Genet 2001;60:107–16.

[45] Chou S-Y, Lee Y-C, Chen H-M, Chiang M-C, Lai H-L, Chang H-H, et al. CGS21680 attenuates symptoms of Huntington's disease in a transgenic mouse model. J Neurochem 2005;93:310–20. doi:10.1111/j.1471-4159.2005.03029.x.

[46] Krishan S, Richardson DR, Sahni S. Adenosine monophosphate-activated kinase and its key role in catabolism: structure, regulation, biological activity, and pharmacological activation. Mol Pharmacol 2015;87:363–77. doi:10.1124/mol.114.095810.

[47] Ye W, Shimamura K, Rubenstein JL, Hynes MA, Rosenthal A. FGF and Shh signals control dopaminergic and serotonergic cell fate in the anterior neural plate. Cell 1998;93:755–66.

[48] Wechsler-Reya RJ, Scott MP. Control of neuronal precursor proliferation in the cerebellum by Sonic Hedgehog. Neuron 1999;22:103–14.

[49] Rondini TA, Baddini SP, Sousa LF, Bittencourt JC, Elias CF. Hypothalamic cocaine- and amphetamine-regulated transcript neurons project to areas expressing gonadotropin releasing hormone immunoreactivity and to the anteroventral periventricular nucleus in male and female rats. Neuroscience 2004;125:735–48. doi:10.1016/j.neuroscience.2003.12.045.

[50] Grant JE, Potenza MN, Weinstein A, Gorelick DA. Introduction to Behavioral Addictions. Am J Drug Alcohol Abuse 2010;36:233–41. doi:10.3109/00952990.2010.491884.

[51] Huang S, Czech MP. The GLUT4 Glucose Transporter. Cell Metab 2007;5:237–52. doi:10.1016/j.cmet.2007.03.006.

[52] Poletto AC, David-Silva A, Yamamoto AP de M, Machado UF, Furuya DT. Reduced Slc2a4/GLUT4 expression in subcutaneous adipose tissue of monosodium glutamate obese mice is recovered after atorvastatin treatment. Diabetol Metab Syndr 2015;7. doi:10.1186/s13098-015-0015-6.

[53] Chiang M-C, Chen C-M, Lee M-R, Chen H-W, Chen H-M, Wu Y-S, et al. Modulation of energy deficiency in Huntington's disease via activation of the peroxisome proliferator-activated receptor gamma. Hum Mol Genet 2010;19:4043–58. doi:10.1093/hmg/ddq322.

[54] Mochel F, Durant B, Meng X, O'Callaghan J, Yu H, Brouillet E, et al. Early Alterations of Brain Cellular Energy Homeostasis in Huntington Disease Models. J Biol Chem 2012;287:1361–70. doi:10.1074/jbc.M111.309849.

[55] Varela-Nallar L, Inestrosa NC. Wnt signaling in the regulation of adult hippocampal neurogenesis. Front Cell Neurosci 2013;7. doi:10.3389/fncel.2013.00100.

[56] Arenas E. Wnt signaling in midbrain dopaminergic neuron development and regenerative medicine for Parkinson's disease. J Mol Cell Biol 2014;6:42–53. doi:10.1093/jmcb/mju001.