SUPPLEMENTAL METHODS

Characterization of the alcohol use disorder discordant monozygotic twin pairs

RAPI is a self-report measure of alcohol-related problems experienced during the previous 12 months. It measures negative consequences of drinking and does not provide quantity-frequency outputs. The original RAPI has 23 items, but in the twin study, the item on whether alcohol use interfered with school work or examination preparation was omitted, creating a 22-item Finnish adaptation of RAPI with 4 response options. The internal consistency of RAPI was excellent (coefficient $\alpha = 0.88$). The age 18.5 RAPI scores have been found to robustly predict alcohol DSM diagnoses in young adulthood(1). We also used DSM at age 25+ to confirm discordance (so we have two measures of alcohol problems/dependence) 7+ years apart with consistent discordance in the pairs we report upon. We have recently reported alcohol-related adverse health outcomes in AUD discordant twins selected on the basis of these criteria(2).

Genome-wide methylation assay

Briefly, 400 ng of gDNA from each individual (18 twin pairs) were restricted by the methylation-sensitive restriction enzymes (MSREs) *HpaII* (Fermentas GmbH, St. Leon-Rot, Germany) and *Hin6I* (Fermentas GmbH, St. Leon-Rot, Germany), producing DNA fragments with 5'-CpG protruding ends. Pooled gDNA from healthy individuals,

used as control DNA, was restricted in parallel to the gDNA from twins, together with fully methylated (HeLa; Sigma-Aldrich Company Ltd., Dorset, England), fully unmethylated (genome-wide amplified DNA) and negative water-controls to ensure that digestion ensued. 200 ng of the digested DNA were ligated to double-stranded adapters (Eurofins MWG Operon, Ebersberg, Germany) using T4 ligase (Fermentas GmbH, St. Leon-Rot, Germany). The template was then treated with *McrBC* endonuclease (New England Biolabs, Herts, England) to remove methylated cytosines. Following *McrBC* digestion, the gDNA was enriched by aminoallyl-PCR amplification using CG1b primers, complementary to the double-stranded adapters used during ligation. Cycling consisted of an initial cycle at 72°C for 5 minutes and 95°C for 1 minute, 25 cycles of 93°C for 40 seconds and 68°C for 2 minutes, and a final extension at 72°C for 5 minutes.

Enriched gDNA resulting from each individual and from the control gDNA was then purified using Qiagen MinElute PCR Purification Kit (Qiagen, West Sussex, England) and was labelled with reactive fluorescent dyes Cy3 and Cy5, respectively (GE Healthcare, Buckinghamshire, England). The Cy3-labelled DNA of each twin was hybridized with the Cy5-labelled control DNA to a NimbleGen DNA Methylation 385k Array (Roche NimbleGen Inc., Madison, USA). Following hybridization in a NimbleGen Hybridization System for 16 hours at 42°C, the arrays were washed to eliminate salts, dried and scanned in a GenePix 4000B scanner (Molecular Devices Ltd, Berkshire, England) to detect fluorescence from the cyanine dyes.

The microarray data was normalization using the software GenePix Pro 6.0 (Molecular Devices, Sunnyvale, USA). Normalization was performed using the software enabled global correction of unbalanced dye signals derived from slight differences in

the amount of dye with which each sample is labelled. To ensure comparable ratio between the two dyes manual normalization of intensity of the dyes were performed. This procedure involves adjusting histograms and ratio-signals from the two fluorescence-sensitive channels until a ratio-score of as close as possible to 1:1 is reached.

Analysis of the array

RP computes the pair-wise ratio (methylation difference) between each twin pair, ranks genes by increasing ratio, and calculates the rank products from all twin-pair comparisons. The method uses a 10.000 permutation-based estimation procedure to determine the likelihood of the observed RP value being better than the random RP value. The approximate RP value distribution was obtained by independently permuting methylation value within each sample relative to methylation probe and calculating the RP value as explained above. By counting how many random RP values were smaller than or equal to the observed RP value (x), the method calculates the average expected value (E = x/N, where N is the total number of permutations). Subsequently, it estimates the percentage of false-positives (PFP). If a probe is significantly differentially methylated ([q = E / rank(probe], rank(probe) denotes the position of the probe in a list of all probes sorted by increasing RP value).

DNA methylation assay using the Sequonom EpiTYPER system

Five hundred ng of gDNA for each individual (36 individuals, 18 twin pairs) were treated with sodium bisulphite using the EZ DNA Methylation Kit (Zymo Research, Irvine, USA) according to the manufacturer's instructions. Sodium bisulphite treated

DNA was then amplified in duplicate; Positive controls and artificially unmethylated or fully methylated samples were included. The proportion of DNA methylation at defined CpG units within the amplified sequences was then determined using the MassCLEAVE procedure for MassARRAY (Sequenom Inc., Hamburg, Germany) according to the manufacturers recommended protocols, followed by analysis using EpiTyper 1.0 software (Sequenom Inc., Hamburg, Germany). Data was screened for outliers.

IMAGEN behavioural characterization

Data quality was controlled by context checks administered at the start of each task, which included information regarding confidentiality of the situation (i.e. was the adolescent alone); the noise level; their mood; time constraints and their tiredness. If any of the tasks were completed with dubious quality the participants were asked to recomplete the task at the study center.

Neuroimaging Analyses.

Stop signal task- The task was composed of Go trials and Stop trials. During Go trials (83%; 480 trials) participants were presented with arrows pointing either to the left or to the right. During these trials subjects were instructed to make a button response with their left or right index finger corresponding to the direction of the arrow. In the unpredictable Stop trials (17%; 80 trials), the arrows pointing left or right were followed (on average 300 ms later) by arrows pointing upwards; participants were instructed to inhibit their motor responses during these trials. A tracking algorithm changes the time interval between Go signal and Stop signal onsets according to each subject's performance on previous trials (average percentage of inhibition over previous Stop

trials, recalculated after each Stop trial), resulting in 50% successful and 50% unsuccessful inhibition trials. The inter-trial interval was 1,800 ms. the tracking algorithm of the task ensured that subjects were successful on 50% of Stop trials and worked at the edge of their own inhibitory capacity.

Functional magnetic resonance imaging data analysis- A one-way ANOVA was used to assess differences among scanners and acquisition sites on three measures of the reference Fusiform Face Area (FFA; X ¼ 42,Y ¼ _48, Z ¼ _23; ROI size 5 _ 5 _ 5 voxels): peak percent BOLD signal change (PBSC), mean percent BOLD signal change (PBSC) and normalized number of active voxels. No significant difference was observed per site and across all eight sites between Peak PBSC and Mean PBSC.

For each participant, Slice-time correction was conducted to adjust for time differences due to multislice imaging acquisition, all volumes were aligned to the first volume and non-linear warping was performed to an EPI template. Images were then smoothed with a Gaussian kernel of 5-mm full-width at half-maximum. Images with distorted magnetic field and structural abnormalities were excluded from the subsequent analyses. The contrast images were then entered into second-level random effect analyses to generate the statistical parametric maps (t-maps). High-resolution T1-weighted three-dimensional structural images were acquired for anatomical localization and coregistration with the functional time series. The functional T2-weighted images for the Stop Signal Task were acquired with a gradient-echo, echo-planar imaging (EPI) sequences. In total 444 volumes of scanned images per participant were acquired throughout the Stop Signal task. Each volume consisted of 40 slices aligned to the anterior commission/posterior commission line (2.4-mm slice thickness, 1-mm gap). The

echo time was optimized (echo time = 30 msec, repetition time = 2200 msec) to provide reliable imaging of subcortical areas.

Gene expression analysis

Total RNA was extracted from whole blood cells using the PAXgene Blood RNA Kit (QIAGEN Inc., Valencia, CA). Following quality control of the total RNA extracted, labeled complementary RNA (cRNA) was generated using the Illumina TotalPrep™ RNA Amplification kit (Applied Biosystems/Ambion, Austin, TX). Complementary RNA was purified and quantified using a Qubit 2.0 Fluorometer (Invitrogen, Paisly, UK). The size distributions of cRNA was determined through Bioanalyzer (Agilent Technologies, Santa Clara, CA) using the Eukaryotic mRNA Assay with smear analysis. Gene expression profiling was performed using Illumina HumanHT-12 v4 Expression BeadChips (Illumina Inc., San Diego, CA). Expression data was normalized using the mloess method (3).

Genotyping method

DNA purification and genotyping was performed by the Centre National de Génotypage in Paris. DNA was purified from whole blood samples (~10 ml) preserved in BD Vacutainer EDTA tubes (Becton, Dickinson and Company) using Gentra Puregene Blood Kit (QIAGEN) according to the manufacturer's instructions. Genotype information was collected at 582,982 markers using the Illumina HumanHap610 Genotyping BeadChip.

SNP Imputation

Single nucleotide polymorphisms with call rates of < 98%, minor allele frequency < 1% or deviation from the Hardy-Weinberg equilibrium ($p \leq 1\times10\text{-}4$) were excluded from the analyses. Individuals with an ambiguous sex code, excessive missing genotypes (failure rate > 2%), and outlying heterozygosity (heterozygosity rate 3 standard deviations from the mean) were also excluded. Identity-by-state similarity was used to estimate cryptic relatedness for individual using PLINK software. Closely related individuals with identity-by-descent (IBD > 0.1875) were eliminated from the subsequent analysis. Population stratification for the genome-wide association study data was examined by principal component analysis (PCA) using EIGENSTRAT software. The four HapMap populations were used as reference groups in the PCA and individuals with divergent ancestry (from CEU) were also excluded.

A quality control step was used to select the SNPs and subjects retained for imputation. Shortly, the following thresholds were applied to select the SNPs (MAF=0.05, HWE pval=0.001, GENO=0.05). The data was checked for duplicates or outliers subjects. Homogenous Caucasian (CEU) ethnicity was studied with code Structure (4) and the reference set Hapmap.rel2.3. This control produces a genotyping dataset with 1835 subjects and 477,234 SNPs (autosome and X chromosomes) that entered the imputation.

We used the imputation protocol and setup defined for the ENIGMA2 projects. These are detailed at enigma.ini.usc.edu. From the 1KGP reference set (phase 1 release v3), monomorphic SNPs in Caucasian were filtered out to obtain a set of reference files that contain all 13,479,643 genetic variants observed more than once in

the European populations. After the ENIGMA2 protocol, we followed the three steps described in the manuscript from Howie and collaborators (5): (i) first check for ambiguous SNPs and potential strand flipping, (ii) phasing using *MaCH1* and (iii) imputation using *minimac.*

Association analysis:

Permutations with 10,000 iterations were used to control for the two brain areas tested in Stop Signal Task BOLD response, right subthalamic nucleus and inferior frontal gyrus and for externalizing behaviors and escalation of daily drinking. *P* values thus obtained are indicated as $p_{\text{corrected}}$. One tailed linear test was used to determine association between *PPM1G* methylation and its gene expression as we hypothesized that higher DNA methylation is associated with lower gene expression (6) and association between *PPM1G* methylation and escalation of daily drinking as we hypothesized that higher *PPM1G* methylation was associated to enhanced drinking behavior as shown in the alcohol use disorder discordant twins.

References

1. Dick DM, Aliev F, Viken R, Kaprio J, Rose RJ. Rutgers alcohol problem index scores at age 18 predict alcohol dependence diagnoses 7 years later. Alcoholism, clinical and experimental research. 2011;35:1011-1014.
2. Rose RJ, Winter T, Viken RJ, Kaprio J. Adolescent Alcohol Abuse and Adverse Adult Outcomes: Evaluating Confounds with Drinking-Discordant Twins. Alcoholism, clinical and experimental research. 2014.
3. Sasik R, Calvo E, Corbeil J. Statistical analysis of high-density oligonucleotide arrays: a multiplicative noise model. Bioinformatics. 2002;18:1633-1640.
4. Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. Genetics. 2000;155:945-959.
5. Howie B, Fuchsberger C, Stephens M, Marchini J, Abecasis GR. Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. Nature genetics. 2012;44:955-959.
6. Suzuki MM, Bird A. DNA methylation landscapes: provocative insights from epigenomics. Nature reviews Genetics. 2008;9:465-476.

SUPPLEMENTAL TABLES

TABLE S1. Classification of the 18 monozygotic twin pairs discordant for alcohol use disorder. Mean RAPI scores at age 18 and 25 years were significantly different between the AUD and non-AUD individuals among the 18 twin pairs at 18 years (t=-1.998; df=17; p=0.031) and at 25 years (t=-2.586, df=16, p=0.01).

ID	RAPI 18 y means	RAPI 25 y means	DSM3R Alcohol Dep.	DSM3R Alcohol Abuse	RAPI 18 y items	RAPI 25 y items
Pair 1	0.6234	1.8182	1	0	.	9
Pair 2	0.5909	0.6364	1	0	1	4
Pair 3	0.5455	0.2273	1	0	8	4
Pair 4	0.5	0	1	0	9	1
Pair 5	0.2727	0.4091	1	0	5	8
Pair 6	0.2273	0.5	1	0	5	7
Pair 7	0.1818	0.2727	1	0	4	4
Pair 8	0.1472	0.3182	1	0	.	7
Pair 9	0.1364	0.1818	1	0	3	2
Pair 10	0.1364	0.3182	0	1	2	4
Pair 11	0.1364	0.0455	1	0	3	1
Pair 12	0.0909	0.5455	1	0	1	6
Pair 13	0.0909	0.2727	1	0	1	6
Pair 14	0.0909	.	0	1	2	.
Pair 15	0.0455	0.1364	1	0	1	3
Pair 16	0.0455	0.0455	1	0	0	1
Pair 17	0	0.0455	1	0	2	1
Pair 18	0	0.0909	1	0	0	2

Abbreviations: AD, alcohol dependence; AA, alcohol abuse; Δ, difference of alcohol use disorder affected twin - non affected twin; RAPI, Rudger Alcohol Problem Index.

TABLE S2. Primers and PCR conditions for the verification phase

Gene	Primers	Annealing T (C°)	Fragment length (bp)	CpG sites present /sites detected
OPRL1	GATTGGTTTTTTTGGGGGAT	66	333	3/3
	AACAACACAAAACTTATATCCACA			
	TTGGAGGGATTATTTTATTTTTGGT	64	478	5/4
	AACCTCAATACACAAATCCTAACCC			
AIM1	TTTTAGGAGATATAAGAAGTTTAGGTTTTT	64	610	39/30
	CATATAACCATAAAACAACAATAAACACAC			
	TTGTGGTTGTGTAGTAGTGTTATGAAA	66	323	17/14
	AAACCCCAAAAAACCTTCTTC			
FMN1	GGTAGGGTTGAAGATTAGTGGAAA	66	286	22/21
	CCCCATCTAAACAAATCCAAAA			
SEPHS2	GGGATTTTGTAGTTTTAGTTTTTT	64	666	45/38
	TCAAATAACCTTTAAAACAATCCT			
SLC6A3	TTAGGGAGAAGTATATTTGGGTGGT	64	284	7/5
	AAAAATCCCAAATCAATAAACACAA			
	AGATTTATTTATATGTTGTATAATGGTGG	66	176	4/4
	AAAACTAAACCCTAATCAAAAAAAA			
PPM1G	TTGGGAAGGATAGTAGAGGTTTT	64	300	6/4
	AAACTCCAACCAAAAAATAACAAAC			
INS_IGF2	GGGATTTTTTTAGGGTTAAGGG	62	329	15/13
	AAATCACCTATTCAAACTCCCA			
PIPOX	GAGATTAGTTTGGTTAATATGGTGAAA	64	284	7/5
	CCATAACACTTTCTAACACAACCCT			
	TTTGAGTTTAGGAAGTGGAGGTT	62	434	7/7
	CCTCTCCTCAACCACAAAAAA			

TABLE S3. DMRs in the AUD twin compared to the unaffected twin ($p<1x10^{-5}$)

Gene (Nimblegen ID)	Gene symbol	pfp	Location	CpG context
Protein phosphatase 1G (CHR02P027457864)	*PPM1G*	$1x10^{-5}$	Body	Shore
Selenophosphate synthetase 2 (CHR16P030364503)	*SEPHS2*	$1x10^{-5}$	Body	Shelf
NA(CHR02P095678989)	NA	$1x10^{-5}$	Intergenic	Shore
Insulin- insulin-like growth factor 2 (CHR11P002137905)	*INS-IGF2*	$1x10^{-5}$	Body	Shore
Neurotrophic tyrosine kinase, receptor, type 1 (CHR01P155095309)	*NTRK1*	$1x10^{-5}$	Body	Shore
Formin 1 (CHR15P031274371)	*FMN1*	$1x10^{-5}$	Promoter	Island
Myocyte enhancer factor 2D (CHR05P064022468)	*MEF2D*	$3x10^{-4}$	Body	Island
NA (CHR01P154736441)	NA	$3x10^{-4}$	Intergenic	Island
6-phosphofructo-2-kinase/fructose-2,6-biphosphatase 3 (CHR10P006245627)	*PFKFB3*	$3x10^{-4}$	Body	Island
L-pipecolic acid oxidase (CHR08P142298123)	*PIPOX*	$4x10^{-4}$	Promoter	Open sea
Solute carrier family 45, member 4 (CHR17P024393350)	*SLC45A4*	$3x10^{-4}$	Body	Island
Fibroblast growth factor (acidic) intracellular binding protein (CHR11P065412364)	*FIBP*	$2.4x10^{-3}$	Body	Island
MIRLET7B host gene (non-protein coding) (CHR22P044859950)	*MIRLET7BHG*	$2.9x10^{-3}$	Body	Shore
Testis expressed 33 (CHR22P035733887)	*TEX33*	$3.7x10^{-3}$	Promoter	Open sea
nucleosome assembly protein 1-like 2 (CHRXP072351512)	*NAP1L2*	$3.8x10^{-3}$	Promoter	Open sea
Solute carrier family 6 (neurotransmitter transporter, dopamine), member 3 (CHR05P001469753)	*SLC6A3*	$4.5x10^{-3}$	Body	Shore
Hypothetical protein LOC149840 (CHR20P005679048)	C20orf196	$4.6x10^{-3}$	Promoter	Open sea
ubiquitin-conjugating enzyme E2M pseudogene 1 (CHR16P034261914)	*UBE2MP1*	$6.5x10^{-3}$	Body	Shore
Absent in melanoma 1 (CHR06P107065946)	*AIM1*	$7x10^{-3}$	Promoter	Shore
NA (CHR09P069075112)	NA	$7.4x10^{-3}$	Intergenic	Island
THUMP domain containing 1 (CHR16P020660421)	*THUMPD1*	$7.7x10^{-3}$	Body	Island
Opiate receptor-like 1 (CHR20P062188944)	*OPRL1*	$7.8x10^{-3}$	Body	Island
Replication termination factor 2 domain containing 1 (CHR20P054477313)	*RTFDC1*	$8.1x10^{-3}$	Body	Shore
NA (CHR12P006270027)	NA	$1x10^{-2}$	Intergenic	Island
DNA damage-inducible transcript 4 protein (CHR10P073704396)	*DDIT4*	$1.2x10^{-2}$	Body	Island
Laminin, gamma 3 precursor (CHR09P132891524)	*LAMC3*	$1.3x10^{-2}$	Body	Shore
olfactory receptor, family 1, subfamily I, member 1 (CHR19P015058994)	*OR1I1*	$1.4x10^{-2}$	Body	Open sea
NA (CHR12P065749657)	NA	$1.4x10^{-2}$	Intergenic	Island
Lipoprotein Lp(a) precursor (CHR06P160991416)	*LPA*	$1.4x10^{-2}$	Body	Open sea
Coenzyme A synthase (CHR17P037967818)	*COASY*	$1.5x10^{-2}$	Body	Island
Centrin 3 (CHR05P089741971)	*CETN3*	$1.5x10^{-2}$	Promoter	Shore
Centrin 3 (CHR05P089741771)	*CETN3*	$1.5x10^{-2}$	Promoter	Shore
NA (CHR08P144860147)	NA	$1.5x10^{-2}$	Intergenic	Shore

Mitochondrial carrier triple repeat 2 (CHR18P027594617)	*MCART2*	$1.5x10^{-2}$	Body	Open sea
Dihydrouridine synthase 4-like (CHR07P106992170)	*DUS4L*	$1.5x10^{-2}$	Body	Shore
Transmembrane protein 160 (CHR19P052243750)	*TMEM160*	$1.6x10^{-2}$	Promoter	Island
Homo sapiens cDNA FLJ44416 fis, clone UTERU2000649 (CHR11P001957498)	AK126380	$1.7x10^{-2}$	Body	Open sea
Histone cluster 1, H1a (CHR06P026126303)	*HIST1H1A*	$1.7x10^{-2}$	Promoter	Shore
Basic leucine zipper nuclear factor 1 (CHR01P167604439)	*BLZF1*	$1.8x10^{-2}$	Body	Open sea
Amiloride-sensitive cation channel 1, neuronal (CHR17P028644764)	*ACCN1*	$2x10^{-2}$	Body	Shore
ORM1-like 3 (CHR17P035337074)	*ORMDL3*	$2.1x10^{-2}$	Body	Island
Solute carrier family 4 member 11 (CHR20P003166898)	*SLC4A11*	$2.2x10^{-2}$	Body	Island
hydroxysteroid (17-beta) dehydrogenase 4 (CHR05P118815438)	*HSD17B4*	$2.2x10^{-2}$	Promoter	Shore
SAMD15 sterile alpha motif domain containing 15 (CHR14P076913514)	*SAMD15*	$2.6x10^{-2}$	Promoter	Shore
Mitochondrial carrier triple repeat 2 (CHR18P027594722)	*MCART2*	$2.6x10^{-2}$	Body	Open sea
Developmental pluripotency associated 5 (CHR06P074120552)	*DPPA5*	$2.7x10^{-2}$	Body	Island
POU class 4 homeobox 2 (CHR04P147779280)	*POU4F2*	$2.7x10^{-2}$	Promoter	Island
Solute carrier family 6 (neurotransmitter transporter, dopamine), member 3 (CHR05P001469658)	*SLC6A3*	$2.7x10^{-2}$	Body	Shore
Lipase maturation factor 1 (CHR16P000912779)	*LMF1*	$2.7x10^{-2}$	Body	Island
Septin 4 (CHR17P053962067)	*SEP4*	$2.7x10^{-2}$	Body	Shore
Poly(A) binding protein, cytoplasmic 1-like (CHRXP072215343)	*PABPC1L2*	$2.8x10^{-2}$	Body	Shore
Zinc finger protein 773 (CHR19P062702331)	*ZNF773*	$2.8x10^{-2}$	Promoter	Shore
Solute carrier family 6 (neurotransmitter transporter, dopamine), member 3 (CHR05P000590263)	*SLC6A3*	$2.8x10^{-2}$	Body	Shore
N-methyl-D-aspartate receptor subunit 2D (CHR19P053594050)	*GRIN2D*	$2.9x10^{-2}$	Body	Shore
Zinc finger protein 304 (CHR19P062553802)	*ZNF304*	$2.9x10^{-2}$	Promoter	Shore
NA (CHR08P001342531)	NA	$2.9x10^{-2}$	Intergenic	Island
Homo sapiens, clone IMAGE:4852110 (CHR19P047328773)	*BC042152*	$3.2x10^{-2}$	Body	Shore
Ganglioside induced differentiation associated (CHR01P118273135)	*GDAP2*	$3.5x10^{-2}$	Body	Shore
Methyltransferase 10 domain containing (CHR17RP002317336)	*METT10D*	$3.7x10^{-2}$	Body	Open sea
Mitochondrial ribosomal protein L43 (CHR10P102736720)	*MRPL43*	$3.7x10^{-2}$	Body	Island
TP53 target 3 (CHR16P033113392)	*TP53TG3*	$4x10^{-2}$	Body	Island
Cutaneous T-cell lymphoma-associated antigen 1 (CHR10P000383303)	*CTAGE1*	$4x10^{-2}$	Promoter	Open sea
DIP2 disco-interacting protein 2 homolog C (CHR18P018251891)	*DIP2C*	$4x10^{-2}$	Body	Island
Family with sequence similarity 27, member E3 (CHR07P027195856)	*FAM27E3*	$4.1x10^{-2}$	Body	Shore
NA (CHR09P067376330)	NA	$4.1x10^{-2}$	Intergenic	Shore
Apolipoprotein A-IV precursor (CHR11P065412364)	*APOA4*	$4.2x10^{-2}$	Promoter	Shore

Fibroblast growth factor (acidic) intracellular binding protein (CHR11P116199645)	*FIBP*	$4.2x10^{-2}$	Body	Island
Solute carrier family 6 (neurotransmitter transporter, dopamine), member 3 (CHR05P001469968)	*SLC6A3*	$4.2x10^{-2}$	Promoter	Island
Retinoid X receptor, alpha (CHR09P136358106)	*RXRA*	$4.2x10^{-2}$	Promoter	Shore
Resistin like beta (CHR03P109958940)	*RETNLB*	$4.2x10^{-2}$	Promoter	Open sea
Thioredoxin domain containing 9 (CHR02P099319238)	*TXNDC9*	$4.3x10^{-2}$	Body	Island
Hypothetical protein LOC126868 (CHR01P116455306)	C1orf161	$4.3x10^{-2}$	Promoter	Open sea
Centrin 3 (CHR05P089741871)	*CETN3*	$4.3x10^{-2}$	Promoter	Shore
Junctophilin 1 (CHR08P075396907)	*JPH1*	$4.3x10^{-2}$	Promoter	Shore
Homo sapiens cDNA: FLJ22202 fis (CHR20P059954441)	AK025855	$4.6x10^{-2}$	Promoter	Island
Rho guanine nucleotide exchange factor (GEF) 17 (CHR11P072751030)	*ARHGEF17*	$4.7x10^{-2}$	Body	Island
Serum amyloid A2 isoform a (CHR11P018226667)	*SAA2*	$4.8x10^{-2}$	Body	Open sea

TABLE S4. Association analyses of the alcohol dependence-related SNP rs2384629 and relative haplotypes to *PPM1G* methylation, escalation of daily drinking, impulsivity and BOLD response of right subthalamic nucleus during the Stop Signal Task

	rs2384629		**Hap 1**		**Hap 2**		**Hap 3**		**Hap 4**		**Hap 5**	
Variable	t	*p*	t	*p*	t	*p*	t	*p*	t	*p*	t	*p*
PPM1G methylation	1.205	0.299	1.69	0.090	-1.15	0.248	-1.01	0.159	-1.63	0.103	1.12	0.264
Escalation of daily drinking	-1.434	0.152	-0.85	0.393	-0.36	0.712	-0.44	0.658	0.89	0.375	0.91	0.362
Impulsivity	0.571	0.568	0.35	0.726	0.81	0.421	2.12	0.034	1.48	0.139	-0.92	0.358
BOLD response of right subthalamic nucleus	-0.311	0.783	1.71	0.089	-0.45	0.649	0.70	0.484	-0.58	0.557	-0.69	0.492

The general linear model was used and gender, puberty score and recruitment center were used as covariates. Handedness was used as additional covariate when analyzing the right subthalamic nucleus

TABLE S5. Association analyses of *PPM1G* methylation in peripheral blood and quantity and frequency of drinking in the last 12 months and in the last month in 14 year old

Variable	t	*p*
Quantity of drinking	-0.45	0.652
Frequency of drinking in the last 12 months	-0.15	0.880
Frequency of drinking in the last month	-0.86	0.388

The general linear model was used and gender, puberty score and recruitment center were used as covariates.

SUPPLEMENTAL FIGURES

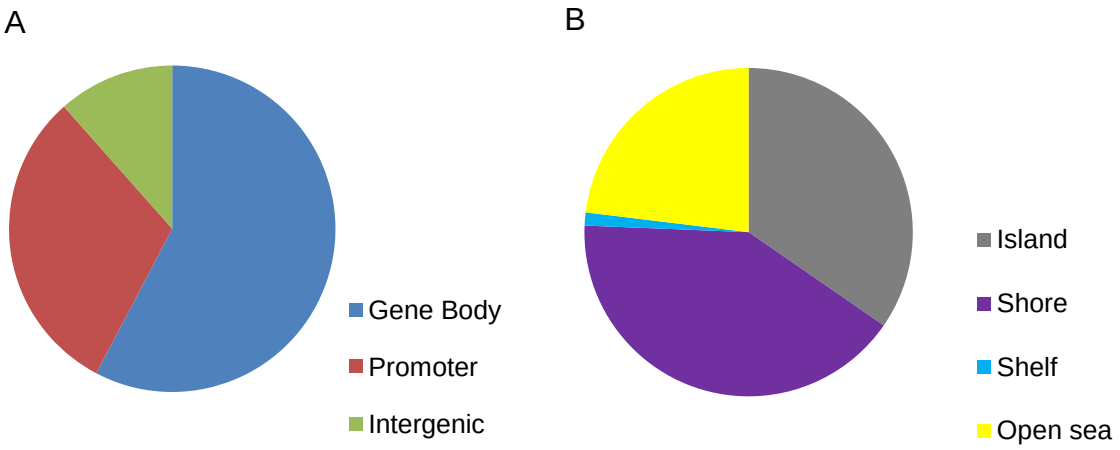

FIGURE S1. Genomic distribution of the alcohol use disorder differentially methylated regions regarding their position in the gene (**A**) and their CpG context (**B**).

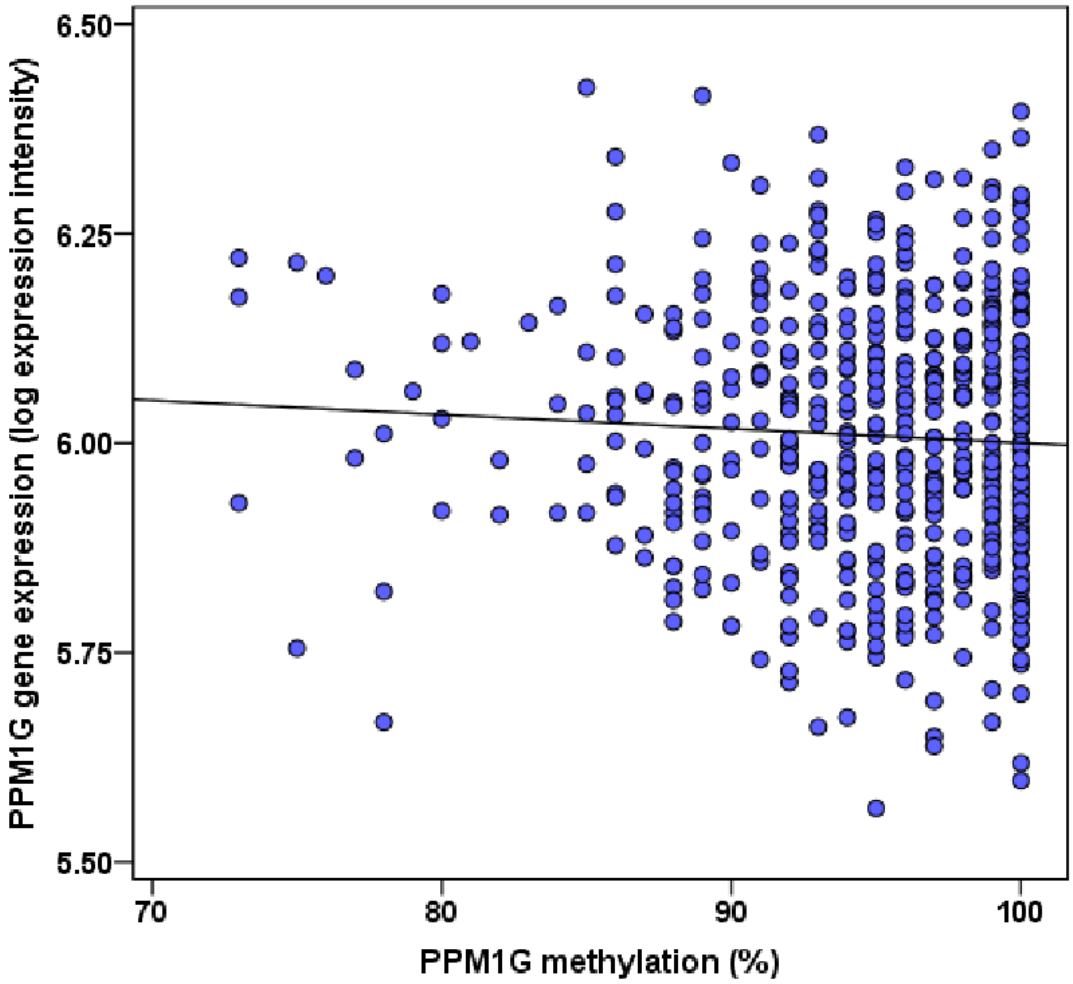

FIGURE S2. *PPM1G* gene expression (log expression intensity) and *PPM1G* methylation (%) in 14 year old peripheral blood show a negative association (p=0.027).