

DNA Modification Study of Major Depressive Disorder: Beyond Locus-by-Locus Comparisons

Supplemental Information

MZ Twin Sample Selection.....	2
Bioinformatic Methods.....	4
Table S1. Sample information	7
Table S2. PCR conditions for pyrosequencing.....	8
Table S3. Attributes selected for the discrimination analysis.....	9
Table S4. Significant loci identified using locus-by-locus comparison	19
Table S5. Overlapping loci in SMRI brain.....	29
Table S6. Modification percentage of repetitive elements.....	30
Table S7. Top 10 GO enrichment terms.....	31
Table S8. Nodes overlapping with differentially methylated loci in brain and sperm	37
Table S9. Summary of discriminatory analysis using Weka	41
Supplemental References	42

MZ Twin Sample Selection

Australian participants were drawn from the Australian Twin Registry. Phenotypic information was collected as part of studies undertaken at the Queensland Institute of Medical Research. Depending on the study, depressive symptoms were evaluated with either the Semi-Structured Assessment for the Genetics of Alcoholism (SSAGA) (1), (adapted for use in Australia) or the Composite International Diagnostic Interview (CIDI) (2) questionnaires. A diagnosis of MDD was constructed from these questionnaires according to the DSM-IV criteria. A total of 17,958 individuals from 6,855 families have completed psychiatric interviews, including a total of 1,812 MZ twin pairs, of whom 261 were found to be discordant for MDD. The personality domains of neuroticism and extraversion were evaluated from the Eysenck Personality Questionnaire revised (3) or revised shortened versions. Normalized scores from the full study sample were calculated from the residuals after regression of arcsin transformed values on age and sex. Forty pairs (31 female) were selected for this study, prioritizing those with DNA collected at the same age, those where the MDD diagnosis was for recurrent MDD (16 pairs) and otherwise selected pairs matched for lifetime smoking and for whom a normalized score of neuroticism was more than 0.5 standard deviation units higher than its co-twin. The average age of the twins was 41.2 years, standard deviation 11.5 years. Zygosity was determined by genotyping nine independent microsatellite markers. All participants provided written informed consent under study protocols approved by the Queensland Institute of Medical Research Human Research Ethics Committee.

Phenotypic data for the Dutch twins were collected by the population-based Netherlands Twin Register (NTR: www.tweelingenregister.org). Depressive symptoms

were assessed in longitudinal surveys by the Beck Depression Inventory (BDI) short form, the anxious-depression scale of the Youth or Adult Self Report or the Composite International Diagnostic Interview (CIDI); data on neuroticism were assessed by standard personality questionnaires (see (4) for a description). For MZ twins who took part in the NTR biobank project (5), discordance was assessed based on all available data from these questionnaires and interviews. The average age of the twins was 38.4 years, standard deviation 12.7 years. Zygosity was confirmed by GWAS. All participants provided written informed consent under study protocols approved by the medical ethics board of the VUMC Amsterdam, the Netherlands.

UK participants were drawn from the St Thomas's Hospital and the Maudsley Hospital (London) Twin Registers and from ongoing studies of volunteer twins. Two of 15 pairs* were male and the mean age at the time of assessment was 53.7 years (range of 21 to 65 years). All subjects were interviewed for lifetime occurrence of psychiatric disorders using the Schedules for Clinical Assessment in Neuropsychiatry (SCAN) (6). Diagnoses were made according to DSM-IV. In addition to the SCAN interview, all study participants completed the Eysenck Personality Questionnaire (7). Zygosity was determined by questionnaire. (*After all analyses were performed, we learned that one set of twins, by mistake, was taken from the pool of dizygotic twins).

Bioinformatic Methods

Permutation analysis was performed by randomly sampling probes from a list of probes that were either directly overlapping or proximal to a gene based on distance cut-off; each sampling consisted of probes equal to the number of loci with nominal $p < 0.05$. The sampled probes were then compared to a second list of genes that were identified as differentially modified (nominal $p < 0.05$) in an independent study (8) and were represented in the CpG island arrays. The number of overlapping genes was recorded and this process was repeated 1,000,000 times to calculate the chance of overlap that is equal to or greater than the actual number of overlaps observed in the real dataset.

Analysis of sample heterogeneity, heteroscedasticity, and epigenetic outliers. Sample heterogeneity was tested by comparing the number of significant loci ($p < 0.05$), identified by *t*-test, between the twin groups separated by their degree of discordance (four groups, $n = 20/\text{group}$) (9). Fisher's exact test was used to detect differences in the number of significant loci between the groups. Heteroscedastic loci were identified using the *lmtest* package in R (10, 11). Harrison-McCabe test (12) was used to derive locus-by-locus p -values on a linearly fitted model for modC and age, followed by FDR correction ($q < 0.05$). Outlier analysis was performed using the Hampel identifier (13).

Median-based DNA modification differences. modC differences were identified by subtracting locus-wise median modification values for MDD patients and controls. Differentially modified probes were defined as those with median modC differences between affected and control individuals of more than 1.64 standard deviations (SD)

from the mean of the ^{mod}C differences (two-sided). Gene-Ontology (GO) enrichment analysis was performed using GO-Elite (14, 15).

Discriminant analysis was performed using Weka (16). The MDD and control samples were divided into two groups, learning or testing, using stratified random sampling. Informative attributes were selected from the learning group using Correlation-based Feature Subset Selection (CfsSubsetEval) and BestFirst greedy hill-climbing algorithm with backtracking (17). Using the informative attributes, a model to discriminate between affected and control samples was built using NaiveBayes and RandomForest; the model was evaluated using data from the testing group (18). The receiver operating characteristic (ROC) curve of the model was used to calculate the ROC area, where an area of 1 equals perfect classification while 0.5 and < 0.5 reflect random and reverse classification, respectively. All attributes selected for the discriminant analysis can be found in Table S3.

Weighted gene co-modification network analysis was performed using the WGCNA package for R (11, 19), and modules were generated by clustering loci with highly correlated ^{mod}C patterns (biweight midcorrelation). Outliers were eliminated by visually inspecting the dendrogram. The modules were then correlated with quantitative phenotypic trait measures. GO term enrichment analysis was then performed using GO-Elite on modules which showed significant correlation with phenotypic traits (14, 15).

Partial-correlation network analysis was performed as previously described (11, 20, 21). A connection was drawn between pairs of spots with *p*-values below the cutoff ($p < 3.0 \times 10^{-16}$). Modules in a network were identified by using a modularity

maximization algorithm (22). An implementation of this algorithm and plotting of the network can be found in the `inetwork` package for R (11).

Table S1. Sample information

SMRI Cortex			
	Controls	Depression	Depression w/ Psychosis
<i>n</i>	12	12	11
Mean Age	46.8 (24-63)	43.6 (24-62)	39.6 (28-63)
Sex	8M, 4F	6M, 6F	6M, 6F
Mean PMI	25.2	27.2	33.1

QSBB Cortex		
	Controls	Depression
<i>n</i>	18	18
Mean Age	41 (15-87)	43.6 (18-82)
Sex	18M	18M
Mean PMI	37.7	30.4

MZ twins WBC		
	Controls	Depression
<i>n</i>	97	103
Mean Age	40.8 (18-73)	40.8 (18-73)
Sex	20M, 77F	20M, 83F

Bipolar Germline		
	Controls	Bipolar
<i>n</i>	17	16
Mean Age	40.1 (16-60)	39.8 (23-55)
Sex	17M	16M

F, female; M, male; MZ, monozygotic; PMI, postmortem interval; QSBB, Quebec Suicide Brain Bank; SMRI, Stanley Medical Research Institute; WBC, white blood cell.

Table S2. PCR conditions for pyrosequencing

	Forward Primer	Reverse Primer	Sequencing Primer	Annealing Temp
<i>LIN28A</i>	GTAGGGAGGTGATTTTATGTGGTAG	CCCCTTAAATCCAATAACCAATTA	GTAGGGAGGTGATTTTATGTGGTAG	62
<i>LRRC41</i>	GGAAGGGTTAAGGAGGTTTTA	CCTCCAACCCAACCTAACC	GGAAGGGTTAAGGAGGTTTTA	62
<i>LINE-1</i>	TTTTGAGTTAGGTGTGGGATATA	AAAATCAAAAATTCCCTTTC	TTYGTGGTGTGTYGTTT	51.2
<i>NBL-2</i>	GGTTTTTTGGTTGATTTGT	CAAATCTAAACCCTAAACTC	GATAGTATTTTTTTT	51.2
<i>D4Z4</i>	GTGGTTTGGGTTAGGTATAGA	AACACTAACCAATCCCACAAC	GTGTATGGATTTTATTTTT	51.2

Table S3. Attributes selected for the discrimination analysis

UHNID	Genome Location	Within Distance	Gene Symbol	Upstream Distance	Gene Symbol	Downstream Distance	Gene Symbol	Type	CpG Island
SMRI and QSBB Cortex									
UHNhscpg0000491	chr6:15942026 7-159421444	0	RSPH3	141603	OSTCL	34584	TAGAP	Coding	YES
UHNhscpg0001906	chr15:6253654 2-62537390	0	FLJ38723	189	DQ571691	285	DQ596703	Coding	NO
UHNhscpg0002893	chr5:14872504 4-148725435	0	GRPEL2	3679	AFAP1L1	12134	PCYOX1L	Coding	YES
UHNhscpg0003336	chr8:76802162 -76802236	#N/A	#N/A	323103	HNF4G	720878	LOC100192378	Intergenic	NO
UHNhscpg0003441	chr18:3953504 6-39535321	0	PIK3C3	434485	KC6	69973	PIK3C3	Coding	NO
UHNhscpg0004402	chr10:1298422 61-129842403	0	PTPRE	151050	CLRN3	3409	PTPRE	Coding	NO
UHNhscpg0004492	chr17:4198474 7-41984929	0	MPP2	5817	MPP2	9646	APR2	Coding	YES
UHNhscpg0004603	chr7:14263820 3-142638652	0	KEL	246	C7orf34	84634	OR9A2	Coding	NO
UHNhscpg0004634	chr9:90112366 -90113736	0	DAPK1	231956	BC093087	148	DAPK1	Coding	YES
UHNhscpg0005234	chr2:85776232 -85776288	0	GGCX	3831	MAT2A	28359	VAMP8	Coding	NO
UHNhscpg0005249	chr3:15288045 2-152881307	0	RAP2B	324611	P2RY1	221415	AK092619	Coding	YES
UHNhscpg0005767	chr20:6252116 3-62521855	0	TPD52L2	26822	C20orf135	4662	DNAJC5	Coding	NO
UHNhscpg0005795	chr1:18044678 0-180446838	0	ACBD6	201421	AK096250	81271	BC036830	Coding	NO
UHNhscpg0005979	chr1:10246172 1-102462009	0	OLFM3	149121	OLFM3	880014	COL11A1	Coding	NO
UHNhscpg0006233	chr9:10173727 2-101737491	0	COL15A1	124914	GALNT12	129920	TGFBR1	Coding	NO
UHNhscpg0006767	chr2:15718873 2-157189019	0	NR4A2	4206	NR4A2	4064	AK128708	Coding	NO

UHNID	Genome Location	Within Distance	Gene Symbol	Upstream Distance	Gene Symbol	Downstream Distance	Gene Symbol	Type	CpG Island
UHNhscpg0007207	chr1:11061089 3-110611161	0	ALX3	13469	FAM40A	43900	UBL4B	Coding	YES
UHNhscpg0007241	chr22:3431875 8-34318928	#N/A	#N/A	174	LARGE	1143200	ISX	Intergenic	NO
UHNhscpg0007469	chr15:8331607 1-83316383	0	CPEB1	17118	CPEB1	11650	AP3B2	Coding	YES
UHNhscpg0008224	chr18:1882318 1-18823415	0	GREB1L	131369	ROCK1	207568	GREB1L	Coding	YES
UHNhscpg0008601	chr19:3620526 2-36206171	0	ZBTB32	35897	UPK1A	2749	MLL4	Coding	NO
UHNhscpg0009228	chr17:1419832 9-14198542	#N/A	#N/A	58150	CDRT15	5963	HS3ST3B1	Intergenic	NO
UHNhscpg0010973	chr15:9960879 3-99609069	#N/A	#N/A	57769	LOC145814	36216	SYNM	Intergenic	NO
UHNhscpg0011168	chr3:19499230 3-194992512	#N/A	#N/A	408	C3orf21	2962	ACAP2	Intergenic	YES
UHNhscpg0013313	chr12:2930183 1-29302616	0	FAR2	598733	CCDC91	73981	FAR2	Coding	YES
UHNhscpg0013723	chr6:17706687 -17707405	0	NUP153	94738	FAM8A1	53179	KIN13A	Coding	YES
UHNhscpg0014137	chr19:3990249 2-39902782	#N/A	#N/A	2447	ZFP36	967	PLEKHG2	Intergenic	NO
UHNhscpg0014605	chr1:10059864 5-100598998	0	CCDC76	134	DKFZp761A078	15414	LRRC39	Coding	YES
UHNhscpg0014998	chr7:73587932 -73589384	0	RFC2	51078	LIMK1	34702	LAT2	Coding	YES
UHNhscpg0015431	chr8:67976415 -67976996	0	CSPP1	1844	COPS5	19086	CSPP1	Coding	YES
UHNhscpg0015972	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A
UHNhscpg0016627	chr10:1299253 28-129925524	#N/A	#N/A	860	MKI67	53875	AK124226	Intergenic	NO
UHNhscpg0017089	chr12:2192609 2-21926641	0	KCNJ8	115316	LDHB	23684	ABCC9	Coding	NO
UHNhscpg0017241	chr6:27513125 -27513486	#N/A	#N/A	72228	ZNF184	148378	BC035101	Intergenic	NO

UHNID	Genome Location	Within Distance	Gene Symbol	Upstream Distance	Gene Symbol	Downstream Distance	Gene Symbol	Type	CpG Island
UHNhscpg0017459	chr10:9606604 6-96066436	0	PLCE1	19219	CR599144	26552	NOC3L	Coding	NO
UHNhscpg0017872	chr5:17076418 4-170764320	#N/A	#N/A	25047	TLX3	49799	NPM1	Intergenic	NO
UHNhscpg0018030	chr2:17702736 9-177027588	#N/A	#N/A	9422	HOXD4	1216	HOXD3	Intergenic	NO
UHNhscpg0018204	chr4:94292597 -94292811	0	GRID2	88349	GRID2	457266	ATOH1	Coding	NO
UHNhscpg0018340	chr5:13541512 1-135415505	#N/A	#N/A	15615	TGFBI	49699	DAMS	Intergenic	YES
UHNhscpg0018461	chr22:3119857 0-31199230	0	OSBP2	133568	SLC35E4	19272	OSBP2	Coding	YES
UHNhscpg0018842	chr6:15046382 6-150464108	#N/A	#N/A	73543	alcan-gamma	79	PPP1R14C	Intergenic	YES
UHNhscpg0018915	chr22:3061893 6-30619246	#N/A	#N/A	45878	HORMAD2	17195	LIF	Intergenic	NO
UHNhscpg0019198	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A
UHNhscpg0019423	chr17:5079730 3-50797557	#N/A	#N/A	559926	CA10	1102681	KIF2B	Intergenic	NO
UHNhscpg0019951	chr2:12278972 5-122790282	#N/A	#N/A	264299	TSN	1979339	AX747402	Intergenic	NO
UHNhscpg0020108	chr6:12272028 6-122720877	0	HSF2	949414	GJA1	37056	BC022047	Coding	YES
UHNhscpg0020482	chr14:6095233 2-60953039	0	C14orf39	186529	PPM1A	22898	SIX6	Coding	YES
UHNhscpg0020874	chr6:11966941 1-119670551	0	MAN1A1	199053	FAM184A	103806	AK097101	Coding	YES
All Twins									
UHNhscpg0011063	chr17:4240132 2-42401903	0	SLC25A39	5285	RUNDC3A	20587	GRN	Coding	YES
UHNhscpg0015465	chr21:4784486 6-47845923	0	PCNT	75132	KIAA0402	28978	CR604262	Coding	YES
UHNhscpg0017424	chr12:6652400 1-66524250	0	LLPH	163933	HMGA2	6467	TMBIM4	Coding	NO

UHNID	Genome Location	Within Distance	Gene Symbol	Upstream Distance	Gene Symbol	Downstream Distance	Gene Symbol	Type	CpG Island
UHNhscpg0013480	chr15:4256493 7-42565851	0	TMEM87A	64435	VPS39	4	GANC	Coding	YES
UHNhscpg0018638	chr12:5449987 6-54500515	0	FLJ12825	3395	LOC100240734	19339	LOC400043	Coding	YES
UHNhscpg0017820	chr1:15667614 6-156676642	#N/A	#N/A	771	CRABP2	15770	ISG20L2	Intergenic	YES
UHNhscpg0000218	chr9:13702819 5-137029602	0	RNU6ATAC	3102	WDR5	179341	FLJ00318	Coding	YES
UHNhscpg0004772	chr5:15934318 5-159343665	#N/A	#N/A	449901	LOC285627	74	ADRA1B	Intergenic	YES
UHNhscpg0006370	chr2:20114887 -20114939	0	KIAA1336	13143	TTC32	19858	WDR35	Coding	NO
UHNhscpg0015679	chr10:1188936 48-118894667	0	VAX1	7551	KIAA1598	62332	KCNK18	Coding	YES
UHNhscpg0015941	chr6:10184669 4-101847129	0	GRIK2	517470	ASCC3	3328838	HACE1	Coding	YES
UHNhscpg0017563	chr2:23039176 2-230392172	0	DNER	255705	PID1	239757	TRIP12	Coding	NO
UHNhscpg0018281	chr2:31359764 -31360641	0	GALNT14	7274	GALNT14	35282	CAPN14	Coding	YES
UHNhscpg0019442	chr9:45202555 -45202993	#N/A	#N/A	203204	FAM27E3	238943	LOC100132167	Intergenic	NO
UHNhscpg0014854	chr8:33742842 -33743158	#N/A	#N/A	285403	DUSP26	1349816	UNC5D	Intergenic	NO
UHNhscpg0000946	chr6:15483096 6-154832122	0	CNKSR3	153066	IPCEF1	222389	RBM16	Coding	YES
UHNhscpg0016693	chr6:33129152 -33129769	#N/A	#N/A	32262	HLA-DPB1	699	COL11A2	Intergenic	YES

European Twins

UHNhscpg0002615	chr1:22105466 3-221055262	0	HLX	48895	BC045735	448007	C1orf140	Coding	YES
UHNhscpg0004022	chr3:25278397 -25278581	0	RARB	741944	NR1A2	146017	AK131021	Coding	NO
UHNhscpg0004403	chr10:1038802 75-103880427	#N/A	#N/A	65	LDB1	12359	PPRC1	Intergenic	YES

UHNID	Genome Location	Within Distance	Gene Symbol	Upstream Distance	Gene Symbol	Downstream Distance	Gene Symbol	Type	CpG Island
UHNhscpg0004747	chr17:6429933 9-64299822	0	PRKCA	73783	APOH	94366	BC033554	Coding	YES
UHNhscpg0004942	chr13:1132417 64-113242445	0	TUBGCP3	152763	C13orf28	58912	C13orf35	Coding	YES
UHNhscpg0005017	chrX:11836956 4-118370155	#N/A	#N/A	85022	KIAA1210	55	PGRMC1	Intergenic	YES
UHNhscpg0005105	chr16:5018629 9-50186697	#N/A	#N/A	46004	AF086132	131	PAPD5	Intergenic	NO
UHNhscpg0006054	chr10:1216395 2-12164124	0	DHTKD1	78929	UPF2	7515	SEC61A2	Coding	NO
UHNhscpg0006889	chr14:4507975 2-45080210	#N/A	#N/A	103253	FSCB	286296	C14orf28	Intergenic	NO
UHNhscpg0007691	chr6:43597645 -43597793	0	MAD2L1BP	709	GTPBP2	5783	MAD2L1BP	Coding	YES
UHNhscpg0007982	chr8:689089- 689250	0	BC022082	7863	ERICH1	143947	BC038783	Coding	YES
UHNhscpg0008439	chr15:8987569 8-89875949	0	POLG	15336	FANCI	28860	LOC254559	Coding	NO
UHNhscpg0008659	chr6:16289916 8-162899383	0	PARK2	34663	PARK2	248780	PACRG	Coding	NO
UHNhscpg0010604	chr1:22947832 9-229479169	0	C1orf96	37081	SPHAR	87825	ACTA1	Coding	YES
UHNhscpg0013846	chr1:10490766 -10491231	0	APITD1	10566	PGD	18739	CORT	Coding	YES
UHNhscpg0013859	chr1:35497024 -35497645	0	ZMYM6	46104	ZMYM6	47326	DKFZp313O192 5	Coding	YES
UHNhscpg0014044	chr3:6473366- 6473683	#N/A	#N/A	1211717	EDEM1	58482	AF279782	Intergenic	NO
UHNhscpg0014284	chr15:5286038 8-52860692	0	ARPP19	39141	MYO5A	12830	DKFZp666F111	Coding	NO
UHNhscpg0014854	chr8:33742842 -33743158	#N/A	#N/A	285403	DUSP26	1349816	UNC5D	Intergenic	NO
UHNhscpg0015254	chr6:27486860 -27487229	#N/A	#N/A	45963	ZNF184	174635	BC035101	Intergenic	YES
UHNhscpg0016673	chr19:5088743 9-50888058	0	POLD1	1173	NR1H2	14049	POLD1	Coding	YES

UHNID	Genome Location	Within Distance	Gene Symbol	Upstream Distance	Gene Symbol	Downstream Distance	Gene Symbol	Type	CpG Island
UHNhscpg0017059	chr16:5816238 2-58163309	0	C16orf80	81580	MMP15	28506	CSNK2A2	Coding	YES
UHNhscpg0017079	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A
UHNhscpg0017424	chr12:6652400 1-66524250	0	LLPH	163933	HMG2A	6467	TMBIM4	Coding	NO
UHNhscpg0017563	chr2:23039176 2-230392172	0	DNER	255705	PID1	239757	TRIP12	Coding	NO
UHNhscpg0018145	chr5:17416220 6-174162704	#N/A	#N/A	4305	MSX2	704971	DRD1	Intergenic	YES
UHNhscpg0019539	chr10:1454136 -1454463	0	ADARB2	207805	ADARB2	114361	NCRNA00168	Coding	NO
UHNhscpg0019690	chr2:44470094 -44470410	0	PPM1B	8353	PPM1B	32186	SLC3A1	Coding	NO
UHNhscpg0019748	chr7:36695243 -36696135	0	AOAH	201845	ANLN	197825	ELMO1	Coding	NO
UHNhscpg0020058	chr6:43612363 -43612939	0	RSPH9	3677	MAD2L1BP	26102	MRPS18A	Coding	YES

Australian Twins

UHNhscpg0000556	chr5:89769624 -89770410	0	MBLAC2	64021	CETN3	270	POLR3G	Coding	YES
UHNhscpg0000622	chr9:69197515 -69197884	#N/A	#N/A	4636	LOC440896	1595	FOXD4L6	Intergenic	YES
UHNhscpg0001156	chr19:9473507 -9474034	0	ZNF177	11669	ZNF559	49237	ZNF266	Coding	YES
UHNhscpg0001764	chr7:13276650 6-132767226	0	CHCHD3	45837	EEF1G	170596	EXOC4	Coding	YES
UHNhscpg0003223	chr19:3117400 2-31174292	#N/A	#N/A	125037	AK091070	223813	BC040412	Intergenic	NO
UHNhscpg0004336	chr16:3396360 7-33963696	#N/A	#N/A	177081	BC068290	440105	UBE2MP1	Intergenic	NO
UHNhscpg0005302	chr8:65494083 -65494444	0	LOC401463	4263	LOC401463	14084	CYP7B1	Coding	YES
UHNhscpg0006238	chr3:77733352 -77733413	#N/A	#N/A	36691	ROBO2	912974	ROBO1	Intergenic	NO

UHNID	Genome Location	Within Distance	Gene Symbol	Upstream Distance	Gene Symbol	Downstream Distance	Gene Symbol	Type	CpG Island
UHNhscpg0006569	chr11:1117495 20-111750323	0	ALG9	7215	ALG9	1176	C11orf1	Coding	YES
UHNhscpg0006727	chr13:8005496 9-80055953	0	NDFIP2	56503	BC034812	23555	KIAA1165	Coding	YES
UHNhscpg0006817	chr11:4359669 0-43596936	#N/A	#N/A	4747	BC031305	6007	DM004400	Intergenic	NO
UHNhscpg0007399	chr8:14292933 -14293021	0	SGCZ	867137	C8orf48	417925	DM004805	Coding	NO
UHNhscpg0007557	chr20:3965802 9-39658428	0	TOP1	4014	CR593014	8044	PRO0628	Coding	YES
UHNhscpg0008999	chr2:22008283 4-220084116	0	ABCB6	8465	ZFAND2B	10362	ANKZF1	Coding	YES
UHNhscpg0009514	chr1:11936735 6-119367510	#N/A	#N/A	639508	SPAG17	58155	TBX15	Intergenic	NO
UHNhscpg0009824	chr5:38188347 -38188479	#N/A	#N/A	348565	GDNF	70053	EGFLAM	Intergenic	NO
UHNhscpg0013426	chr4:37827816 -37829201	0	DKFZp566B152 4	139817	RELL1	63518	TBC1D1	Coding	YES
UHNhscpg0014371	chr15:1018743 10-101874797	0	PCSK6	38850	SNRPA1	12149	PCSK6	Coding	NO
UHNhscpg0016613	chr12:2165424 3-21654713	0	RECQL	30063	DKFZp762G094	24542	C12orf39	Coding	YES
UHNhscpg0017515	chr20:2274561 0-22746096	#N/A	#N/A	80096	AK055550	151089	CR618492	Intergenic	NO
UHNhscpg0017536	chr3:15632348 2-156324440	#N/A	#N/A	50547	SSR3	66523	LOC100287227	Intergenic	NO
UHNhscpg0017936	chr21:4632783 5-46328304	0	ITGB2	34094	PBF	12661	BC040064	Coding	NO
UHNhscpg0018420	chr1:86861379 -86862167	0	ODF2L	238933	COL24A1	27601	CLCA2	Coding	YES
UHNhscpg0018486	chr6:26596742 -26597919	0	ABT1	21817	BC033330	5041	BC043245	Coding	YES
UHNhscpg0018854	chr19:3734142 1-37342066	0	ZNF345	12137	ZNF790	40186	ZNF829	Coding	YES
UHNhscpg0019067	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A

UHNID	Genome Location	Within Distance	Gene Symbol	Upstream Distance	Gene Symbol	Downstream Distance	Gene Symbol	Type	CpG Island
Bipolar Germline									
UHNhscpg0001513	chr7:16684853 -16685241	0	ANKMY2	114648	SOSTDC1	517	BZW2	Coding	YES
UHNhscpg0002929	chr2:12145337 2-121454135	#N/A	#N/A	229447	FLJ14816	39063	hGli2	Intergenic	NO
UHNhscpg0005668	chr17:1762810 9-17629003	0	RAI1	133115	PEMT	50996	SMCR5	Coding	YES
UHNhscpg0008798	chr2:23844070 4-238441219	0	MLPH	117854	COL6A3	33997	PRLH	Coding	NO
UHNhscpg0014334	chr1:19789029 7-197891074	0	LHX9	3170	LHX9	235033	NEK7	Coding	YES
UHNhscpg0014952	chr1:15519708 4-155197908	0	GBA	8275	GBA	4892	MTX1	Coding	YES
UHNhscpg0016480	chr1:23676758 5-236768045	0	HEATR1	21089	HEATR1	81724	ACTN2	Coding	YES
UHNhscpg0017162	chr3:14141849 6-141418865	#N/A	#N/A	84311	CR620532	38185	RNF7	Intergenic	NO
UHNhscpg0018368	chr12:7223397 4-72234626	0	TBC1D15	52825	RAB21	7449	MRS2P2	Coding	NO
UHNhscpg0018474	chr1:17685451 3-176854910	0	ASTN1	39782	CR621436	50279	ASTN1	Coding	NO
UHNhscpg0019449	chr1:15130041 0-151300632	#N/A	#N/A	277	PI4KB	12483	RFX5	Intergenic	YES
UHNhscpg0005331	chr20:4980715 -4981294	0	SLC23A2	176424	RASSF2	67835	C20orf30	Coding	NO
UHNhscpg0020295	chr6:26234465 -26234971	0	HIST1H1D	7876	HIST1H3F	5682	HIST1H4F	Coding	NO
UHNhscpg0004710	chr4:14548340 8-145483723	#N/A	#N/A	421504	GYPA	80350	CR620567	Intergenic	NO
UHNhscpg0001970	chr1:20070856 2-200709206	0	CAMSAP1L1	69436	DDX59	132959	GPR25	Coding	YES
UHNhscpg0006664	chr17:2791679 1-27917267	0	ANKRD13B	181	GIT1	3259	ANKRD13B	Coding	YES
UHNhscpg0018000	chr18:7411091 3-74111185	0	ZNF516	18654	ZNF516	96291	FLJ44313	Coding	NO

UHNID	Genome Location	Within Distance	Gene Symbol	Upstream Distance	Gene Symbol	Downstream Distance	Gene Symbol	Type	CpG Island
UHNhscpg0000365	chr6:43682298 -43682469	#N/A	#N/A	26770	MRPS18A	55483	VEGFA	Intergenic	NO
UHNhscpg0001318	chr9:13277882 -13278712	0	MPDZ	27511	MPDZ	127749	AK123194	Coding	YES
UHNhscpg0017459	chr10:9606604 6-96066436	0	PLCE1	19219	CR599144	26552	NOC3L	Coding	NO
UHNhscpg0004551	chr10:1064408 77-106441606	0	SORCS3	14919	BC042079	495918	KIAA1059	Coding	NO
UHNhscpg0005929	chr16:7083500 4-70835231	0	VAC14	27859	hCG_1998535	6058	DKFZp434D0513	Coding	YES
UHNhscpg0013011	chr10:1040050 71-104005615	0	BC068543	3840	PITX3	148613	NFKB2	Coding	YES
UHNhscpg0015604	chr1:8276881- 8277598	#N/A	#N/A	190488	ERRFI1	106791	SLC45A1	Intergenic	YES
UHNhscpg0015683	chr3:16011779 7-160118272	0	IFT80	477	KIAA1374	74	SMC4	Coding	YES
UHNhscpg0017833	chr2:23858245 0-238583089	0	LRRFIP1	82714	RAB17	17717	LRRFIP1	Coding	NO
UHNhscpg0018956	chr4:10087024 5-100870545	0	H2AFZ	2366	UNQ9427	1090	LOC256880	Coding	YES
UHNhscpg0003453	chr9:8729315- 8729384	0	PTPRD	929516	C9orf123	128749	AK094342	Coding	NO
UHNhscpg0007719	chr5:12881231 7-128812426	0	ADAMTS19	362600	ISOC1	119825	ADAMTS19	Coding	NO
UHNhscpg0014552	chr9:5387049- 5387312	0	C9orf46	47176	RLN1	63246	CD274	Coding	NO
UHNhscpg0017283	chr3:14159427 1-141594631	#N/A	#N/A	58381	GRK7	838	ATP1B3	Intergenic	YES
UHNhscpg0014491	chr15:9011825 2-90118470	#N/A	#N/A	50987	LOC283761	347	C15orf42	Intergenic	NO
UHNhscpg0014730	chr6:30646820 -30647162	0	KIAA1949	5990	DHX16	8664	NRM	Coding	NO
UHNhscpg0017317	chr10:1193077 42-119308180	0	EMX2	3163	EMX2	456248	RAB11FIP2	Coding	NO
UHNhscpg0005689	chr6:13969682 4-139697362	#N/A	#N/A	1039	CITED2	92769	AK123801	Intergenic	NO

UHNID	Genome Location	Within Distance	Gene Symbol	Upstream Distance	Gene Symbol	Downstream Distance	Gene Symbol	Type	CpG Island
UHNhscpg0006522	chr7:10013645 7-100137224	0	AGFG2	44035	C7orf51	10794	AGFG2	Coding	YES
UHNhscpg0004515	chr20:3540194 0-35402589	0	DSN1	27459	NDRG3	3255	C20orf117	Coding	YES
UHNhscpg0010483	chr10:1227908 32-122790892	#N/A	#N/A	121797	KIAA1351	446952	FGFR2	Intergenic	NO
UHNhscpg0014149	chr5:103960- 104464	#N/A	#N/A	0	0	35908	PLEKHG4B	Intergenic	NO
UHNhscpg0010760	chr15:4121792 4-41218414	#N/A	#N/A	16389	BC031952	3132	DLL4	Intergenic	YES
UHNhscpg0006765	chrX:16370327 -16370471	#N/A	#N/A	198687	GRPR	235652	CTPS2	Intergenic	NO
UHNhscpg0015431	chr8:67976415 -67976996	0	CSPP1	1844	COPS5	19086	CSPP1	Coding	YES
UHNhscpg0009599	chr2:20032347 0-200324089	0	SATB2	651	SATB2	8731	FLJ32063	Coding	YES
UHNhscpg0016887	chr15:4321352 6-43213926	#N/A	#N/A	519	TTBK2	21176	UBR1	Intergenic	YES
UHNhscpg0000521	chr12:5441042 8-54410753	0	HOXC6	3885	HOXC8	11440	HOXC6	Coding	NO
UHNhscpg0004603	chr7:14263820 3-142638652	0	KEL	246	C7orf34	84634	OR9A2	Coding	NO
UHNhscpg0013506	chr8:61835534 -61836010	#N/A	#N/A	56071	CHD7	364514	RLBP1L1	Intergenic	YES
UHNhscpg0013740	chr12:7244598 -7245170	0	C1R	154	C1R	1980	C1RL	Coding	NO
UHNhscpg0018916	chr8:22547112 -22547408	0	EGR3	20451	BIN3	23356	PEBP4	Coding	NO
UHNhscpg0000054	chr19:4636603 2-46367007	0	SYMPK	6972	SYMPK	510	FOXA3	Coding	YES

The detailed information of location and the names of genes can be found at <http://data.microarrays.ca/cpg/>.

Table S4. Significant loci identified using locus-by-locus comparison

UHNID	Genome Location	Within Distance	Gene Symbol	Upstream Distance	Gene Symbol	Downstream Distance	Gene Symbol	Type	CpG island
<i>SMRI Cortex (p < 0.01)</i>									
UHNhscpg0019055	chr5:58120723 -58121002	0	RAB3C	328538	GAPT	32126	AB074188	Coding	NO
UHNhscpg0018579	chr14:9696808 5-96968781	0	PAPOLA	12973	AK7	90288	BC035096	Coding	YES
UHNhscpg0009466	chr2:45162966 -45163278	#N/A	#N/A	163237	C2orf34	5758	SIX3	Intergenic	YES
UHNhscpg0013186	chrX:25023665 -25024257	0	ARX	8565	POLA1	1132205	MAGEB18	Coding	YES
UHNhscpg0019316	chr2:23379157 2-233791933	0	NGEF	25422	NGEF	105448	NEU2	Coding	YES
UHNhscpg0019423	chr17:5079730 3-50797557	#N/A	#N/A	559926	CA10	1102681	KIF2B	Intergenic	NO
UHNhscpg0018193	chr14:3073980 5-30740102	#N/A	#N/A	342906	PRKD1	288226	KIAA1333	Intergenic	NO
UHNhscpg0000519	chr8:29567585 -29568036	#N/A	#N/A	359400	DUSP4	10741	C8orf75	Intergenic	NO
UHNhscpg0013502	chr17:5586059 4-55861067	#N/A	#N/A	103295	MSI2	55776	MRPS23	Intergenic	NO
UHNhscpg0011427	chr6:74447715 -74447811	0	CD109	41972	AK124950	331355	AF086303	Coding	NO
UHNhscpg0013340	chr2:19829942 8-198299981	0	SF3B1	123933	ANKRD44	18249	COQ10B	Coding	YES
UHNhscpg0006806	chr11:4233354 -4233561	#N/A	#N/A	9595	RRM1	155021	OR52B4	Intergenic	NO
UHNhscpg0013489	chr5:68462950 -68463665	0	CCNB1	37071	SLC30A5	13859	BC112270	Coding	YES
UHNhscpg0010138	chr12:1219503 25-121950517	0	KDM2B	2483	KDM2B	113937	ORAI1	Coding	NO
UHNhscpg0011152	chr6:11998921 3-119989303	#N/A	#N/A	176746	AK097101	1411323	C6orf170	Intergenic	NO
UHNhscpg0010271	chr16:5157878 1-51579065	#N/A	#N/A	393598	SALL1	537911	AK098676	Intergenic	NO

UHNID	Genome Location	Within Distance	Gene Symbol	Upstream Distance	Gene Symbol	Downstream Distance	Gene Symbol	Type	CpG island
UHNhscpg0019629	chr16:2980265 1-29802822	0	KIF22	10812	ZG16	5391	KIF22	Coding	NO
UHNhscpg0013430	chr1:15667681 7-156676922	#N/A	#N/A	1442	CRABP2	15490	ISG20L2	Intergenic	NO
UHNhscpg0004565	chr1:63787138 -63787404	#N/A	#N/A	4237	BC040309	1325	FOXD3	Intergenic	YES
UHNhscpg0004621	chr13:9433264 7-94332823	0	GPC6	813162	GPC5	759019	DCT	Coding	NO
UHNhscpg0005863	chr22:3174047 9-31741739	0	PATZ1	51959	PIK3IP1	1132	Em:AC005003.4	Coding	YES
UHNhscpg0008904	chr11:1113370 35-111337089	#N/A	#N/A	9705	AB231705	1168	BTG4	Intergenic	NO
UHNhscpg0019368	chr11:4811167 2-48112256	0	PTPRJ	241615	NUP160	44986	PTPRJ	Coding	NO
UHNhscpg0015028	chr2:12363257 5-123633124	#N/A	#N/A	1107149	TSN	1136497	AX747402	Intergenic	NO
UHNhscpg0013862	chr2:11303275 0-113033389	0	ZC3H6	20086	ZC3H8	24036	KIAA2035	Coding	YES
UHNhscpg0016149	chr10:1242344 58-124234929	0	HTRA1	17591	ARMS2	85251	DMBT1	Coding	NO
UHNhscpg0005051	chr14:3705356 8-37053835	#N/A	#N/A	1782	NKX2-8	72946	PAX9	Intergenic	YES
UHNhscpg0013618	chr20:4442028 2-44420887	0	WFDC3	65947	SPINT4	20367	UBE2C	Coding	YES
UHNhscpg0020697	chr8:10850942 5-108509966	0	ANGPT1	160682	ANGPT1	401578	RSPO2	Coding	NO
UHNhscpg0018135	chr4:18590515 9-185905589	#N/A	#N/A	84544	AY429595	34405	HELT	Intergenic	NO
UHNhscpg0011200	chr2:15701145 -15701536	0	NBAS	136553	NBAS	30233	DDX1	Coding	YES
UHNhscpg0014607	chr2:17274594 6-172746152	0	SLC25A12	141027	DYNC112	32782	HAT1	Coding	NO
UHNhscpg0007784	chr8:21861286 -21861615	0	XPO7	27254	KIAA0745	20020	NPM2	Coding	NO
UHNhscpg0014405	chr3:18385307 8-183853678	0	EIF2B5	28296	HTR3E	7285	EIF2B5	Coding	YES

UHNID	Genome Location	Within Distance	Gene Symbol	Upstream Distance	Gene Symbol	Downstream Distance	Gene Symbol	Type	CpG island
UHNhscpg0010375	chr2:85197520 -85197654	#N/A	#N/A	63721	TMSB10	576	KCMF1	Intergenic	YES
UHNhscpg0004741	chr6:37467161 -37467926	0	C6orf129	17879	FTSJD2	132357	MDGA1	Coding	YES
UHNhscpg0013603	chr6:80714127 -80714799	0	TTK	56812	ELOVL4	101544	BCKDHB	Coding	YES
UHNhscpg0000745	chr18:3386098 9-33861362	#N/A	#N/A	12305	MOCOS	16339	FHOD3	Intergenic	NO
UHNhscpg0015471	chr16:8558947 2-85589875	#N/A	#N/A	443358	FAM92B	55153	KIAA0182	Intergenic	YES
UHNhscpg0015961	chr19:4761624 5-47616792	0	ZC3H4	759	KIAA1064	17287	SAE1	Coding	YES
QSBB Cortex ($p < 0.01$)									
UHNhscpg0002274	chr15:4124564 3-41246148	0	CHAC1	14406	DLL4	24932	INO80	Coding	YES
UHNhscpg0011787	chr6:10644224 8-106442372	#N/A	#N/A	591279	PREP	91822	PRDM1	Intergenic	YES
UHNhscpg0015683	chr3:16011779 7-160118272	0	IFT80	477	KIAA1374	74	SMC4	Coding	YES
UHNhscpg0014863	chr9:5832920- 5833466	0	ERMP1	56366	KIAA1432	57442	MLANA	Coding	YES
UHNhscpg0013780	chr2:12216720 7-122167620	0	KIAA0622	182	KIAA0622	120835	RNU4ATAC	Coding	NO
UHNhscpg0009771	chr5:84158463 -84158717	#N/A	#N/A	477852	EDIL3	1419544	NBPF22P	Intergenic	NO
UHNhscpg0007053	chr10:9837407 8-98374233	0	PIK3AP1	27269	TM9SF3	217805	LCOR	Coding	NO
UHNhscpg0017489	chr3:17246830 8-172469056	0	ECT2	39300	NCEH1	3190	ECT2	Coding	YES
UHNhscpg0007708	chr5:15700306 7-157003202	#N/A	#N/A	299	ADAM19	49484	SOX30	Intergenic	YES
UHNhscpg0015943	chr7:14270497 2-142705297	#N/A	#N/A	45469	KEL	17989	OR9A2	Intergenic	NO
UHNhscpg0019050	chr11:3558496 8-35585208	#N/A	#N/A	33120	PAMR1	54526	FJX1	Intergenic	NO

UHNID	Genome Location	Within Distance	Gene Symbol	Upstream Distance	Gene Symbol	Downstream Distance	Gene Symbol	Type	CpG island
UHNhscpg0000825	chr11:6235865 8-62359223	0	TUT1	13895	FLJ00132	1451	MTA2	Coding	YES
UHNhscpg0000210	chr17:5953160 5-59532218	0	TBX4	40964	C17orf82	1588	TBX4	Coding	YES
UHNhscpg0017459	chr10:9606604 6-96066436	0	PLCE1	19219	CR599144	26552	NOC3L	Coding	NO
UHNhscpg0016480	chr1:23676758 5-236768045	0	HEATR1	21089	HEATR1	81724	ACTN2	Coding	YES
UHNhscpg0017432	chr15:4195392 3-41954225	0	MGA	82397	TYRO3	72494	MGA	Coding	YES
UHNhscpg0013446	chr15:5928026 5-59281002	0	RNF111	54413	SLTM	42000	RNF111	Coding	YES
UHNhscpg0000904	chr4:8604897- 8605570	0	GPR78	15379	GPR78	263202	HMX1	Coding	NO
UHNhscpg0005794	chr2:70475858 -70475935	#N/A	#N/A	79	TIA1	8906	PCYOX1	Intergenic	YES
UHNhscpg0018958	chr22:5024817 1-50248477	0	ZBED4	26991	BRD1	48378	ALG12	Coding	NO
UHNhscpg0007521	chr18:4772142 1-47721819	0	MYO5B	97433	MYO5B	31743	CCDC11	Coding	YES
UHNhscpg0002001	chr1:18315504 2-183155730	0	LAMC2	40316	LAMC1	61648	KIAA0479	Coding	NO
UHNhscpg0005969	chr3:11227999 8-112280145	0	ATG3	61590	BTLA	749	SLC35A5	Coding	YES
UHNhscpg0011223	chr2:17693332 1-176933646	#N/A	#N/A	65807	KIAA1715	11190	EVX2	Intergenic	YES
UHNhscpg0013147	chr2:64872176 -64872585	0	SERTAD2	28562	BC041356	200678	AK097952	Coding	NO
UHNhscpg0013016	chr16:6775260 8-67753145	0	GFOD2	41221	AX747090	3860	RANBP10	Coding	YES
UHNhscpg0005977	chr1:19873320 3-198733283	#N/A	#N/A	6658	PTPRC	94718	DM004432	Intergenic	NO
UHNhscpg0011454	chr16:3112872 7-31129395	0	MYST1	4617	BCKDK	7559	MYST1	Coding	YES
UHNhscpg0000347	chr6:85483901 -85484420	#N/A	#N/A	10002	TBX18	612518	AK024998	Intergenic	YES

UHNID	Genome Location	Within Distance	Gene Symbol	Upstream Distance	Gene Symbol	Downstream Distance	Gene Symbol	Type	CpG island
UHNhscpg0015824	chr7:87258078-87258585	0	ABCB1	28572	ABCB1	29825	RUNDC3B	Coding	YES
UHNhscpg0001118	chr1:23116557-3-231165949	0	FAM89A	29102	ARV1	132724	TRIM67	Coding	NO
UHNhscpg0003946	chr11:4382193-5-43822139	0	HSD17B12	44855	HSD17B12	29119	AK094325	Coding	NO
UHNhscpg0011252	chr18:2700634-2700695	0	SMCHD1	84001	NDC80	37	KIAA0650	Coding	NO
UHNhscpg0000186	chr2:17697662-0-176977907	#N/A	#N/A	2306	HOXD11	3584	HOXD10	Intergenic	YES
UHNhscpg0008770	chr6:26271172-26271840	0	HIST1H3G	18870	HIST1H2BH	545	BC079832	Coding	YES
WBC ($p < 0.01$)									
UHNhscpg0004801	chr11:1225263-51-122527098	0	UBASH3B	287884	LOC399959	182156	CRTAM	Coding	YES
UHNhscpg0014074	chr4:15360080-1-153601606	0	TMEM154	9542	CR599924	88901	TIGD4	Coding	NO
UHNhscpg0014839	chr11:1075695-4-10757727	#N/A	#N/A	41419	MRV11	15083	CTR9	Intergenic	NO
UHNhscpg0016645	chr4:15968961-1-159689927	#N/A	#N/A	45059	PPID	254	KIAA1450	Intergenic	YES
UHNhscpg0020074	chr19:4622097-2-46221275	0	FBXO46	333	AX746967	46768	SIX5	Coding	NO
UHNhscpg0005279	chr22:2946858-0-29469209	0	KREMEN1	10673	C22orf31	21036	KREMEN1	Coding	YES
UHNhscpg0013972	chr16:2512279-2-25123483	0	LCMT	38950	DQ577693	28414	AK056728	Coding	YES
UHNhscpg0018184	chrX:9838698-9838906	0	SHROOM2	104781	GPR143	41505	SHROOM2	Coding	NO
UHNhscpg0016512	chr7:39988730-39989046	#N/A	#N/A	93871	CR606158	912	CDC2L5	Intergenic	YES
UHNhscpg0001374	chr16:5168470-4-51685461	#N/A	#N/A	499521	SALL1	431515	AK098676	Intergenic	NO
UHNhscpg0016132	chr11:6360736-5-63607942	0	MARK2	12176	C11orf84	48044	MARK2	Coding	YES

UHNID	Genome Location	Within Distance	Gene Symbol	Upstream Distance	Gene Symbol	Downstream Distance	Gene Symbol	Type	CpG island
UHNhscpg0001429	chr1:24268658 6-242686962	0	PLD5	73821	PLD5	532649	AL137733	Coding	YES
UHNhscpg0004749	chr5:34007580 -34008345	0	AMACR	22800	SLC45A2	9618	C1QTNF3	Coding	YES
UHNhscpg0004539	chr9:12937473 8-129375867	#N/A	#N/A	101039	FLJ00001	880	LMX1B	Intergenic	YES
UHNhscpg0017868	chr2:12828349 6-128283906	0	IWS1	23757	BC022892	9471	MYO7B	Coding	YES
UHNhscpg0010932	chr3:17403486 4-174035019	#N/A	#N/A	33748	NLGN1	121343	NAALADL2	Intergenic	NO
UHNhscpg0006063	chr5:17741172 6-177412139	#N/A	#N/A	26832	AK126616	7096	PROP1	Intergenic	YES
UHNhscpg0001527	chr20:9048179 -9048734	#N/A	#N/A	182634	PLCB1	966	PLCB4	Intergenic	NO
UHNhscpg0013885	chr9:10820984 0-108210678	0	FSD1L	9058	SLC44A1	109732	FKTN	Coding	YES
UHNhscpg0004863	chr9:34588073 -34588820	0	CNTFR	5003	LOC415056	21672	C9orf23	Coding	NO
UHNhscpg0015565	chr5:15656934 8-156569947	0	MED7	33210	HAVCR2	19403	FAM71B	Coding	YES
UHNhscpg0006281	chr11:1346534 21-134653794	#N/A	#N/A	278404	BC112333	0	0	Intergenic	NO
UHNhscpg0016340	chr8:20054471 -20055392	0	ATP6V1B2	13754	SLC18A1	20778	AK127443	Coding	YES
UHNhscpg0001872	chr12:1335625 55-133563481	0	ZNF26	29687	ZNF605	50396	ZNF84	Coding	YES
UHNhscpg0016621	chr5:17005217 2-170052833	0	KCNIP1	235534	KCNMB1	157889	GABRP	Coding	NO
UHNhscpg0014613	chr17:3588366 1-35884305	0	SYNRG	10073	DUSP14	16449	SYNRG	Coding	NO
UHNhscpg0005354	chr7:43945525 -43945726	0	URG4	16292	URG4	20320	UBE2D4	Coding	YES
UHNhscpg0005011	chr11:1192274 87-119227764	0	USP2	10104	MFRP	24755	CR610619	Coding	YES
UHNhscpg0005880	chr19:5444684 6-54447081	#N/A	#N/A	578	CACNG7	19212	CACNG8	Intergenic	NO

UHNID	Genome Location	Within Distance	Gene Symbol	Upstream Distance	Gene Symbol	Downstream Distance	Gene Symbol	Type	CpG island
UHNhscpg0018446	chr21:4662963 7-46630263	0	ADARB1	33058	ADARB1	24003	C21orf89	Coding	NO
UHNhscpg0009719	chr16:1289728 6-12898039	0	CPPED1	98358	FLJ11151	97437	SHISA9	Coding	YES
UHNhscpg0014397	chr4:80412227 -80412950	#N/A	#N/A	82855	GK2	796	AK123913	Intergenic	NO
UHNhscpg0016186	chr4:11355806 6-113558444	0	C4orf21	48981	DKFZp434C0927	167	LARP7	Coding	YES
UHNhscpg0010652	chr10:3707708 2-37077248	#N/A	#N/A	1146720	FZD8	337536	ANKRD30A	Intergenic	NO
UHNhscpg0013881	chr16:1621394 1-16214788	0	ABCC1	63524	ABCC1	28634	MRP6	Coding	NO
UHNhscpg0019770	chr1:19999850 8-199999376	0	NR5A2	1008342	BC040869	9281	NR5A2	Coding	NO
UHNhscpg0013854	chr11:1285566 18-128556937	0	AX747861	99165	ETS1	4638	BC039676	Coding	NO
UHNhscpg0018753	chr10:3051397 -3052217	#N/A	#N/A	994855	AK097474	57534	PFKP	Intergenic	NO
UHNhscpg0000289	chr14:4560310 0-45603973	0	FKBP3	18298	PRPF39	1162	FANCM	Coding	YES
UHNhscpg0004672	chr3:11078961 6-110790285	#N/A	#N/A	177293	LOC151760	579	PVRL3	Intergenic	YES
UHNhscpg0001753	chr11:1277166 9-12771897	0	TEAD1	220259	PARVA	77770	TEAD1	Coding	NO
UHNhscpg0006286	chr5:18147047 -18147145	#N/A	#N/A	216450	BC028204	1326011	CDH18	Intergenic	NO
UHNhscpg0000030	chr2:44222763 -44223773	0	LRPPRC	117159	ABCG8	172226	PPM1B	Coding	YES
UHNhscpg0019153	chr2:22615535 9-226155615	#N/A	#N/A	248029	DOCK10	109986	KIAA1486	Intergenic	NO
Sperm ($p < 0.01$)									
UHNhscpg0009813	chr1:23231108 1-232311281	#N/A	#N/A	134065	DISC1	222432	SIPA1L2	Intergenic	NO
UHNhscpg0016173	chr14:6808770 3-68087829	0	ARG2	20686	PIGH	30039	VTI1B	Coding	NO

UHNID	Genome Location	Within Distance	Gene Symbol	Upstream Distance	Gene Symbol	Downstream Distance	Gene Symbol	Type	CpG island
UHNhscpg0019304	chr3:14712508 0-147125502	#N/A	#N/A	673	ZIC4	1678	ZIC1	Intergenic	NO
UHNhscpg0010998	chr4:62648453 -62648602	0	LPHN3	49181	LPHN3	126675	LPHN3	Coding	NO
UHNhscpg0007944	chr13:8832587 4-88326027	0	SLITRK5	2656	AF339814	2819514	BC038529	Coding	NO
UHNhscpg0006769	chr1:1240173- 1241428	0	ACAP3	5838	ACAP3	2565	PUSL1	Coding	NO
UHNhscpg0013635	chr16:8440190 3-84402563	0	ATP2C2	38457	WFDC1	35469	ATP2C2	Coding	YES
UHNhscpg0015033	chr15:6739583 0-67396381	0	SMAD3	236470	BX538221	21672	SMAD3	Coding	NO
UHNhscpg0006255	chr19:4761592 1-47616230	0	ZC3H4	435	KIAA1064	17849	SAE1	Coding	YES
UHNhscpg0016201	chr15:4274192 5-42742350	0	ZFP106	1022	ZFP106	45484	SNAP23	Coding	NO
UHNhscpg0013559	chr6:1379825- 1380247	#N/A	#N/A	64833	FOXQ1	9821	FOXF2	Intergenic	YES
UHNhscpg0005803	chr7:20369436 -20369688	#N/A	#N/A	112423	MACC1	636	ITGB8	Intergenic	NO
UHNhscpg0018455	chr1:17012273 -17012819	#N/A	#N/A	35359	MST1	4893	ESPN	Intergenic	NO
UHNhscpg0010233	chr12:9568826 0-95688349	0	VEZT	24296	DKFZp761C241	13846	DM119494	Coding	NO
UHNhscpg0009977	chr10:9444867 3-94449184	#N/A	#N/A	33523	KIF11	496	HHEX	Intergenic	YES
UHNhscpg0005920	chr1:6663435- 6663620	#N/A	#N/A	506	KLHL21	10135	PHF13	Intergenic	YES
UHNhscpg0008560	chr11:1737478 9-17375244	0	DKFZp686I21167	3268	AK096475	31553	KCNJ11	Coding	YES
UHNhscpg0007336	chr1:11690237 4-116902454	#N/A	#N/A	224513	C1orf161	12549	ATP1A1	Intergenic	NO
UHNhscpg0006928	chr12:1033449 46-103345236	#N/A	#N/A	33565	PAH	6215	ASCL1	Intergenic	NO
UHNhscpg0011627	chr10:1203042 54-120304390	#N/A	#N/A	202415	C10orf84	48525	PRLHR	Intergenic	NO

UHNID	Genome Location	Within Distance	Gene Symbol	Upstream Distance	Gene Symbol	Downstream Distance	Gene Symbol	Type	CpG island
UHNhscpg0015850	chr12:7207969 4-72080471	0	TMEM19	5277	THAP2	68186	RAB21	Coding	YES
UHNhscpg0005356	chr6:10132870 5-101329013	0	ASCC3	415900	SIM1	517891	GRIK2	Coding	YES
UHNhscpg0004538	chr5:14521444 1-145214708	0	PRELID2	1357497	KCTD16	24587	GRXCR2	Coding	YES
UHNhscpg0015239	chr15:5307595 6-53076170	0	ONECUT1	105003	KIAA1370	729767	WDR72	Coding	NO
UHNhscpg0018685	chr4:15959210 9-159592543	0	C4orf46	18847	RXFP1	733	ETFDH	Coding	NO
UHNhscpg0005201	chr15:3717488 2-37175149	0	LOC145845	64175	CSNK1A1P	8083	MEIS2	Coding	NO
UHNhscpg0005251	chr2:12795548 4-127955659	0	CYP27C1	90620	BIN1	59206	ERCC3	Coding	NO
UHNhscpg0011152	chr6:11998921 3-119989303	#N/A	#N/A	176746	AK097101	1411323	C6orf170	Intergenic	NO
UHNhscpg0007801	chr19:1418513 8-14185802	0	BC000492	15167	PALM3	13427	SAMD1	Coding	YES
UHNhscpg0014229	chr2:19808111 1-198081590	0	ANKRD44	18349	ANKRD44	175109	SF3B1	Coding	NO
UHNhscpg0016558	chr8:73163281 -73163691	0	AK309726	175462	TRPA1	285934	KCNB2	Coding	YES
UHNhscpg0016032	chr7:99258470 -99258781	0	FLJ00284	28440	ZNF498	9760	CYP3A5	Coding	NO
UHNhscpg0011724	chr1:23672458 1-236724700	0	HEATR1	11585	LGALS8	41432	HEATR1	Coding	NO
UHNhscpg0006964	chr11:1080930 68-108093311	0	NPAT	33004	NPAT	247	ATM	Coding	YES
WBC overlap with Cortex ($p < 0.05$)									
UHNhscpg0018546	chr1:61376616 -61377126	#N/A	#N/A	85360	AK097193	28790	BC030753	Intergenic	NO
UHNhscpg0015649	chr12:5825931 6-58259622	#N/A	#N/A	18569	CTDSP2	65685	CR602022	Intergenic	YES
UHNhscpg0018171	chr8:77593109 -77593505	0	LOC100192378	1114050	HNF4G	9	ZFH4	Coding	NO

UHNID	Genome Location	Within Distance	Gene Symbol	Upstream Distance	Gene Symbol	Downstream Distance	Gene Symbol	Type	CpG island
UHNhscpg0019016	chr2:54682720 -54682992	#N/A	#N/A	94006	C2orf73	461	SPTBN1	Intergenic	YES
UHNhscpg0015912	chr1:26738228 -26738564	0	LIN28	38962	ZNF683	20237	DHDDS	Coding	NO
UHNhscpg0014696	chr9:10074581 0-100746063	0	ANP32B	38676	HEMGN	61825	NANS	Coding	YES
UHNhscpg0004888	chr4:10374764 8-103747857	0	UBE2D3	30382	CR618043	1433	AK093356	Coding	YES
UHNhscpg0020874	chr6:11966941 1-119670551	0	MAN1A1	199053	FAM184A	103806	AK097101	Coding	YES
UHNhscpg0010993	chr1:46766929 -46768004	0	LRRC41	22784	RAD54L	1375	UQCRH	Coding	YES
UHNhscpg0017059	chr16:5816238 2-58163309	0	C16orf80	81580	MMP15	28506	CSNK2A2	Coding	YES
UHNhscpg0007942	chr1:54122058 -54122354	0	GLIS1	188900	DMRTB1	111035	TMEM48	Coding	NO
UHNhscpg0013497	chr6:28574154 -28575048	#N/A	#N/A	14630	AK056211	252353	AK291391	Intergenic	NO
UHNhscpg0010615	chr6:15651900 3-156519136	#N/A	#N/A	741966	NOX3	579949	ARID1B	Intergenic	NO

The detailed information of location and the names of genes can be found at <http://data.microarrays.ca/cpg/>.

Table S5. Overlapping loci in SMRI brain

Overlapping loci at nominal $p < 0.05$					
UHNID	Genome Location	Distance	Gene Symbol	Type	CpG island
UHNhscpg0000446	chr6:43214575-43215208	0	TTBK1	Coding	YES
UHNhscpg0013432	chr12:126018128-126018404	0	TMEM132B	Coding	YES
UHNhscpg0005870	chr9:103235957-103236387	0	TMEFF1	Coding	YES
UHNhscpg0005863	chr22:31740479-31741739	0	PATZ1	Coding	YES
UHNhscpg0005389	chr3:158520219-158520467	0	MFSD1	Coding	YES
UHNhscpg0015172	chr15:68125341-68125738	0	LBXCOR1	Coding	NO
UHNhscpg0014696	chr9:100745810-100746063	0	ANP32B	Coding	YES
UHNhscpg0014103	chr15:61521868-61522136	366	RORA	Intergenic	NO
UHNhscpg0004564	chr11:2292611-2292760	429	ASCL2	Intergenic	YES
UHNhscpg0001435	chr5:72740629-72740925	1161	FOXD1	Intergenic	YES
UHNhscpg0000203	chr17:36859818-36860097	1775	MLLT6	Intergenic	YES
UHNhscpg0001474	chr8:53326184-53326779	3745	ST18	Intergenic	NO
UHNhscpg0017826	chr13:112627415-112628118	93794	SOX1	Intergenic	YES
UHNhscpg0003336	chr8:76802162-76802236	323103	HNF4G	Intergenic	NO

The detailed information of location and the names of genes can be found at <http://data.microarrays.ca/cpg/>.

Table S6. Modification percentage of repetitive elements

Tissues	Phenotypes	LINE-1	NBL-2	D4Z4
Cortex (mean \pm SD% ^{mod} C)	MDD	53.8 \pm 4.2%	69.6 \pm 4.6%	67.7 \pm 10.6%
	Controls	55.5 \pm 3.5%	69.8 \pm 5.6%	68.2 \pm 9.7%
WBC (mean \pm SD% ^{mod} C)	MDD	57.2 \pm 3.0%	73.2 \pm 2.6%	68.1 \pm 6.3%
	Controls	53.9 \pm 5.6%	73.0 \pm 3.5%	67.1 \pm 5.9%

Table S7. Top 10 GO enrichment terms

GO Name(GOID)	Number Changed	Number Measured	Number in GO	Percent Changed	Percent Present	Z Score	PermuteP	Gene Symbols
<i>Methylation difference across different tissues</i>								
eukaryotic cell surface binding(GO:0043499)	3	5	18	60.0	27.8	6.6	0.000	APOH ATP5B PCSK6
cell proliferation in forebrain(GO:0021846)	3	5	8	60.0	62.5	6.6	0.001	EMX2 LHX5 ZEB2
regulation of cytokine biosynthetic process(GO:0042035)	5	24	100	20.8	24.0	4.4	0.001	CEBPG IGF2BP1 IL1A INHBB MAS T2
regulation of postsynaptic membrane potential(GO:0060078)	3	10	31	30.0	32.3	4.4	0.004	GRID2 GRIK2 NLGN1
negative regulation of translation(GO:0017148)	3	11	44	27.3	25.0	4.1	0.007	CAPRIN1 EIF2AK3 IGF2BP1
cell-cell signaling involved in cell fate commitment(GO:0045168)	3	12	24	25.0	50.0	3.9	0.009	FGF10 SIX3 SOX2
fat cell differentiation(GO:0045444)	4	20	73	20.0	27.4	3.8	0.005	EIF2AK3 FGF10 INHBB WNT5B
tube closure(GO:0060606)	4	21	49	19.0	42.9	3.7	0.007	COBL FGF10 ZEB2 ZIC2
cell surface(GO:0009986)	9	80	299	11.3	26.8	3.6	0.003	APOH ATP5B DMD FGF10 IL1A IR AK2 NLGN1 PCSK6 WNT5B
glutamate signaling pathway(GO:0007215)	3	14	26	21.4	53.8	3.5	0.014	GNAQ GRID2 GRIK2
<i>WGCNA network - Dutch twins (module 1)</i>								
carbon-oxygen lyase activity(GO:0016835)	3	15	58	20.0	25.9	6.1	0.002	HSD17B4 NEIL1 OGG1
response to insulin stimulus(GO:0032868)	4	59	248	6.8	23.8	3.5	0.012	FGF5 GPAM INSIG2 STRADA
regulation of lipid metabolic process(GO:0019216)	3	37	165	8.1	22.4	3.5	0.017	GPAM INSIG2 STRADA
transmembrane receptor protein tyrosine kinase signaling pathway(GO:0007169)	7	159	497	4.4	32.0	3.2	0.010	AC069542.2 CHUK FGF12 FGF5 R PS6KA1 STAM STRADA VAV1

GO Name(GOID)	Number Changed	Number Measured	Number in GO	Percent Changed	Percent Present	Z Score	PermuteP	Gene Symbols
glycerolipid metabolic process(GO:0046486)	3	47	213	6.4	22.1	2.9	0.034	FDXACB1 GPAM INSIG2
regulation of protein transport(GO:0051223)	3	52	208	5.8	25.0	2.7	0.043	GPAM OGG1 WWTR1
negative regulation of catalytic activity(GO:0043086)	5	126	425	4.0	29.6	2.5	0.041	NEIL1 NOSIP NR2F2 RPS6KA1 WTR1
cell proliferation(GO:0008283)	6	172	503	3.5	34.2	2.3	0.044	CHUK FGF5 IRF2 LHX9 STAT4 ZEB2
WGCNA network - Dutch twins (module 2)								
postsynaptic density(GO:0014069)	3	42	96	7.1	43.8	3.9	0.008	ANKS1B GRIA2 GRID2
negative regulation of cellular component organization(GO:0051129)	4	84	244	4.8	34.4	3.3	0.009	KIAA0319 MID1IP1 PAX5 SMAD6
postsynaptic membrane(GO:0045211)	3	56	166	5.4	33.7	3.2	0.018	ANKS1B GRIA2 GRID2
multicellular organismal development(GO:0007275)	9	324	965	2.8	33.6	3.1	0.004	EN1 EVX2 KIAA0319 LBH PAX5 PAX9 SERPINE2 SIM1 SPATA24
membrane fraction(GO:0005624)	7	228	931	3.1	24.5	3.0	0.008	CHPT1 DPAGT1 GRIA2 GRID2 LAMP1 RGS16 SERPINE2
synapse(GO:0045202)	4	105	333	3.8	31.5	2.8	0.021	ANKS1B GRIA2 GRID2 SERPINE2
lipid binding(GO:0008289)	4	108	474	3.7	22.8	2.7	0.020	ANXA4 CHPT1 HADHB SGIP1
neuronal cell body(GO:0043025)	3	74	219	4.1	33.8	2.5	0.037	GNB2L1 GRIA2 SERPINE2
negative regulation of cellular protein metabolic process(GO:0032269)	3	76	266	3.9	28.6	2.5	0.041	PAX5 SERPINE2 SMAD6
regulation of cell growth(GO:0001558)	3	77	258	3.9	29.8	2.4	0.041	CHPT1 NDRG3 SERPINE2
WGCNA network - UK twins (module 1)								
negative regulation of immune response(GO:0050777)	3	6	44	50.0	13.6	5.0	0.002	DUSP10 PSMB4 TGFB2
generation of neurons(GO:0048699)	3	7	11	42.9	63.6	4.5	0.004	NTF3 PTF1A TGFB2

GO Name(GOID)	Number Changed	Number Measured	Number in GO	Percent Changed	Percent Present	Z Score	PermuteP	Gene Symbols
cardiac cell differentiation(GO:0035051)	4	12	27	33.3	44.4	4.4	0.002	BMP2 CACYPB TBX3 TGFB2
hindbrain development(GO:0030902)	4	12	19	33.3	63.2	4.4	0.002	EN1 HNF1B PTF1A TGFB2
heart morphogenesis(GO:0003007)	5	18	39	27.8	46.2	4.4	0.001	BMPR1A HAND2 TBX3 TGFB2 TTN
branching involved in ureteric bud morphogenesis(GO:0001658)	5	19	38	26.3	50.0	4.2	0.002	BMP2 MYC PAX2 SALL1 WT1 BCOR BMP2 BMPR1A CACYPB FBN1 GATA4 HAND2 SALL1 TGFB2 TTN WT1
heart development(GO:0007507)	11	71	178	15.5	39.9	4.0	0.000	GATA4 HAND2 STIL TBX3 CIDEA MAPK14 PSMB4 PSMB7 SMC1A UACA UBC
heart looping(GO:0001947)	4	14	37	28.6	37.8	4.0	0.004	GATA4 HAND2 STIL TBX3 CIDEA MAPK14 PSMB4 PSMB7 SMC1A UACA UBC
signal transduction in response to DNA damage(GO:0042770)	7	36	121	19.4	29.8	3.9	0.003	UACA UBC
regulation of epidermis development(GO:0045682)	3	9	33	33.3	27.3	3.8	0.008	MAFG NAB2 TGFB2
WGCNA network - UK twins (module 2)								
methionine metabolic process(GO:0006555)	3	4	14	75.0	28.6	7.6	0.000	ENOPH1 MSRA MTHFD1
cell proliferation in forebrain(GO:0021846)	3	5	8	60.0	62.5	6.7	0.001	ARX EMX2 WNT7A
forebrain cell migration(GO:0021885)	3	7	26	42.9	26.9	5.6	0.001	ARX DAB1 EMX2
oxidoreductase activity, acting on a sulfur group of donors(GO:0016667)	3	8	53	37.5	15.1	5.1	0.003	GLRX MSRA TMX1
histone mRNA metabolic process(GO:0008334)	3	9	31	33.3	29.0	4.8	0.003	ERI1 LSM10 SSB
forebrain development(GO:0030900)	6	32	67	18.8	47.8	4.6	0.001	APLP2 ARX DLC1 EMX2 OTX2 ZIC5
regulation of transforming growth factor beta receptor signaling pathway(GO:0017015)	4	26	66	15.4	39.4	3.2	0.013	ONECUT1 ONECUT2 PRDM16 ZNF703

GO Name(GOID)	Number Changed	Number Measured	Number in GO	Percent Changed	Percent Present	Z Score	PermuteP	Gene Symbols
axonogenesis(GO:0007409)	6	50	118	12.0	42.4	3.2	0.008	KALRN MAPT PAX2 PRKCA SLITRK1 VAX2
cell fate commitment(GO:0045165)	6	51	106	11.8	48.1	3.1	0.009	DLX1 ERBB4 ONECUT1 ONECUT2 OTX2 PAX7
negative regulation of protein transport(GO:0051224)	3	18	77	16.7	23.4	3.0	0.025	DAB2IP DPH3 NDFIP2
WGCNA network - UK twins (module 3)								
forebrain anterior/posterior pattern formation(GO:0021797)	4	4	6	100.0	66.7	4.8	0.000	FEZF2 PAX6 SIX3 WNT1
hexose metabolic process(GO:0019318)	16	42	203	38.1	20.7	4.3	0.000	CALM1 CALM2 CALM3 DLAT GBE1 GMDS GYS1 IRS2 LDHB LRP5 MDH1 ONECUT1 PDX1 PGM1 PGM2L1 RPE TKTL1 TSTA3
protein polyubiquitination(GO:0000209)	12	28	102	42.9	27.5	4.2	0.000	BARD1 CDC16 FBXW11 MGRN1 NEDD4 PARK2 RNF111 RNF41 RNF6 UBE2K UBE2Q2 WWP2
endothelial cell differentiation(GO:0045446)	3	3	12	100.0	25.0	4.2	0.003	JAG1 PROX1 S1PR1
regulation of translational elongation(GO:0006448)	3	3	9	100.0	33.3	4.2	0.003	AARSD1 EIF5A2 RPS5
lens morphogenesis in camera-type eye(GO:0002089)	4	5	11	80.0	45.5	4.1	0.002	MEIS1 PROX1 PVRL3 SHROOM2
regulation of transcription involved in cell fate commitment(GO:0060850)	5	8	9	62.5	88.9	3.8	0.002	LBX1 NKX6-1 NR2F2 PAX6 PROX1
detection of chemical stimulus(GO:0009593)	5	8	41	62.5	19.5	3.8	0.002	CYB5R4 LBP NKX6-1 PDX1 SYT1
endocytic vesicle membrane(GO:0030666)	7	14	43	50.0	32.6	3.7	0.002	AP2S1 CACNG8 GRIA2 GRIA3 LAMP2 RILP SYT1
cell fate determination(GO:0001709)	9	21	49	42.9	42.9	3.6	0.002	DLL1 JAG1 LBX1 PAX2 PAX6 POU6F2 PROX1 TBX2 WNT1
WGCNA network - Australian twins (module 1)								
hair cycle process(GO:0022405)	5	15	61	33.3	24.6	6.6	0.000	ALX4 BARX2 FGF10 KRT27 LDB1

GO Name(GOID)	Number Changed	Number Measured	Number in GO	Percent Changed	Percent Present	Z Score	PermuteP	Gene Symbols
negative regulation of proteolysis(GO:0045861)	3	10	41	30.0	24.4	4.8	0.004	SERPINE2 SERPING1 UBXN1
Rho GTPase binding(GO:0017048)	3	12	44	25.0	27.3	4.2	0.007	DAAM1 DIAPH3 VCL
transferase activity, transferring pentosyl groups(GO:0016763)	3	12	50	25.0	24.0	4.2	0.007	MTAP SIRT2 XYLT1
positive regulation of peptidyl-tyrosine phosphorylation(GO:0050731)	3	14	73	21.4	19.2	3.8	0.011	FGF10 MTOR THBS4
phosphoric diester hydrolase activity(GO:0008081)	4	24	92	16.7	26.1	3.7	0.009	CHRM3 PDE7A PLCB4 TDP1
negative regulation of catabolic process(GO:0009895)	3	17	65	17.6	26.2	3.3	0.018	MTOR SERPINE2 UBXN1
response to unfolded protein(GO:0006986)	3	18	83	16.7	21.7	3.2	0.023	DERL1 HERPUD2 HSPA4L
sulfur compound metabolic process(GO:0006790)	4	30	143	13.3	21.0	3.1	0.020	GSTA4 MSRA SLC19A2 SULT2A1
DNA bending activity(GO:0008301)	3	20	54	15.0	37.0	3.0	0.031	FOXD4L6 FOXG1 FOXJ2

Partial correlation network - common nodes

negative regulation of lipid metabolic process(GO:0045833)	3	9	40	33.3	22.5	4.0	0.008	BRCA1 CNR1 PROX1
cell morphogenesis involved in differentiation(GO:0000904)	5	23	64	21.7	35.9	3.9	0.004	BMP2 FOXF2 PAX2 PROX1 SLC1A3
ethanolamine-containing compound metabolic process(GO:0042439)	3	10	39	30.0	25.6	3.8	0.010	BCHE ENPP6 SLC5A7
acetylglucosaminyltransferase activity(GO:0008375)	3	10	39	30.0	25.6	3.8	0.011	EXT2 GCNT2 XYLT1
response to alkaloid(GO:0043279)	4	17	61	23.5	27.9	3.7	0.007	BCHE CNR1 DPYSL2 NKX6-1
cell-matrix adhesion(GO:0007160)	4	18	69	22.2	26.1	3.5	0.010	CD44 COL5A3 NID1 VCL
lysosomal membrane(GO:0005765)	5	29	117	17.2	24.8	3.2	0.012	BLOC1S1 GPR143 ITM2C M6PR SLC29A3

GO Name(GOID)	Number Changed	Number Measured	Number in GO	Percent Changed	Percent Present	Z Score	PermuteP	Gene Symbols
condensed chromosome(GO:0000793)	3	13	42	23.1	31.0	3.1	0.022	BRCA1 SC65 SMC2
negative regulation of sequence-specific DNA binding transcription factor activity(GO:0043433)	4	22	77	18.2	28.6	3.0	0.019	ID1 IRAK2 PROX1 ZCCHC11
digestive tract development(GO:0048565)	3	15	33	20.0	45.5	2.8	0.033	ALX4 FGF9 FOXF2

Table S8. Nodes overlapping with differentially methylated loci in brain and sperm

UHNID	Genome Location	Within Distance	Gene Symbol	Upstream Distance	Gene Symbol	Downstream Distance	Gene Symbol	Type	CpG island
<i>SMRI Cortex (p < 0.01)</i>									
UHNhscpg0019055	chr5:58120723 -58121002	0	RAB3C	328538	GAPT	32126	AB074188	Coding	NO
UHNhscpg0018579	chr14:9696808 5-96968781	0	PAPOLA	12973	AK7	90288	BC035096	Coding	YES
UHNhscpg0009466	chr2:45162966 -45163278	#N/A	#N/A	163237	C2orf34	5758	SIX3	Intergenic	YES
UHNhscpg0013186	chrX:25023665 -25024257	0	ARX	8565	POLA1	1132205	MAGEB18	Coding	YES
UHNhscpg0019316	chr2:23379157 2-233791933	0	NGEF	25422	NGEF	105448	NEU2	Coding	YES
UHNhscpg0019423	chr17:5079730 3-50797557	#N/A	#N/A	559926	CA10	1102681	KIF2B	Intergenic	NO
UHNhscpg0018193	chr14:3073980 5-30740102	#N/A	#N/A	342906	PRKD1	288226	KIAA1333	Intergenic	NO
UHNhscpg0000519	chr8:29567585 -29568036	#N/A	#N/A	359400	DUSP4	10741	C8orf75	Intergenic	NO
UHNhscpg0013502	chr17:5586059 4-55861067	#N/A	#N/A	103295	MSI2	55776	MRPS23	Intergenic	NO
UHNhscpg0011427	chr6:74447715 -74447811	0	CD109	41972	AK124950	331355	AF086303	Coding	NO
UHNhscpg0013340	chr2:19829942 8-198299981	0	SF3B1	123933	ANKRD44	18249	COQ10B	Coding	YES
UHNhscpg0006806	chr11:4233354 -4233561	#N/A	#N/A	9595	RRM1	155021	OR52B4	Intergenic	NO
UHNhscpg0013489	chr5:68462950 -68463665	0	CCNB1	37071	SLC30A5	13859	BC112270	Coding	YES
UHNhscpg0010138	chr12:1219503 25-121950517	0	KDM2B	2483	KDM2B	113937	ORAI1	Coding	NO
UHNhscpg0011152	chr6:11998921 3-119989303	#N/A	#N/A	176746	AK097101	1411323	C6orf170	Intergenic	NO
UHNhscpg0010271	chr16:5157878 1-51579065	#N/A	#N/A	393598	SALL1	537911	AK098676	Intergenic	NO

UHNID	Genome Location	Within Distance	Gene Symbol	Upstream Distance	Gene Symbol	Downstream Distance	Gene Symbol	Type	CpG island
UHNhscpg0019629	chr16:2980265 1-29802822	0	KIF22	10812	ZG16	5391	KIF22	Coding	NO
UHNhscpg0013430	chr1:15667681 7-156676922	#N/A	#N/A	1442	CRABP2	15490	ISG20L2	Intergenic	NO
UHNhscpg0004565	chr1:63787138 -63787404	#N/A	#N/A	4237	BC040309	1325	FOXD3	Intergenic	YES
UHNhscpg0004621	chr13:9433264 7-94332823	0	GPC6	813162	GPC5	759019	DCT	Coding	NO
UHNhscpg0005863	chr22:3174047 9-31741739	0	PATZ1	51959	PIK3IP1	1132	Em:AC005003.4	Coding	YES
UHNhscpg0008904	chr11:1113370 35-111337089	#N/A	#N/A	9705	AB231705	1168	BTG4	Intergenic	NO
UHNhscpg0019368	chr11:4811167 2-48112256	0	PTPRJ	241615	NUP160	44986	PTPRJ	Coding	NO
UHNhscpg0015028	chr2:12363257 5-123633124	#N/A	#N/A	1107149	TSN	1136497	AX747402	Intergenic	NO
UHNhscpg0013862	chr2:11303275 0-113033389	0	ZC3H6	20086	ZC3H8	24036	KIAA2035	Coding	YES
UHNhscpg0016149	chr10:1242344 58-124234929	0	HTRA1	17591	ARMS2	85251	DMBT1	Coding	NO
UHNhscpg0005051	chr14:3705356 8-37053835	#N/A	#N/A	1782	NKX2-8	72946	PAX9	Intergenic	YES
UHNhscpg0013618	chr20:4442028 2-44420887	0	WFDC3	65947	SPINT4	20367	UBE2C	Coding	YES
UHNhscpg0020697	chr8:10850942 5-108509966	0	ANGPT1	160682	ANGPT1	401578	RSPO2	Coding	NO
UHNhscpg0018135	chr4:18590515 9-185905589	#N/A	#N/A	84544	AY429595	34405	HELT	Intergenic	NO
UHNhscpg0011200	chr2:15701145 -15701536	0	NBAS	136553	NBAS	30233	DDX1	Coding	YES
UHNhscpg0014607	chr2:17274594 6-172746152	0	SLC25A12	141027	DYNC112	32782	HAT1	Coding	NO
UHNhscpg0007784	chr8:21861286 -21861615	0	XPO7	27254	KIAA0745	20020	NPM2	Coding	NO
UHNhscpg0014405	chr3:18385307 8-183853678	0	EIF2B5	28296	HTR3E	7285	EIF2B5	Coding	YES

UHNID	Genome Location	Within Distance	Gene Symbol	Upstream Distance	Gene Symbol	Downstream Distance	Gene Symbol	Type	CpG island
UHNhscpg0010375	chr2:85197520 -85197654	#N/A	#N/A	63721	TMSB10	576	KCMF1	Intergenic	YES
UHNhscpg0004741	chr6:37467161 -37467926	0	C6orf129	17879	FTSJD2	132357	MDGA1	Coding	YES
UHNhscpg0013603	chr6:80714127 -80714799	0	TTK	56812	ELOVL4	101544	BCKDHB	Coding	YES
UHNhscpg0000745	chr18:3386098 9-33861362	#N/A	#N/A	12305	MOCOS	16339	FHOD3	Intergenic	NO
UHNhscpg0015471	chr16:8558947 2-85589875	#N/A	#N/A	443358	FAM92B	55153	KIAA0182	Intergenic	YES
UHNhscpg0015961	chr19:4761624 5-47616792	0	ZC3H4	759	KIAA1064	17287	SAE1	Coding	YES
QSBB Cortex ($p < 0.01$)									
UHNhscpg0002274	chr15:4124564 3-41246148	0	CHAC1	14406	DLL4	24932	INO80	Coding	YES
UHNhscpg0011787	chr6:10644224 8-106442372	#N/A	#N/A	591279	PREP	91822	PRDM1	Intergenic	YES
UHNhscpg0015683	chr3:16011779 7-160118272	0	IFT80	477	KIAA1374	74	SMC4	Coding	YES
UHNhscpg0014863	chr9:5832920- 5833466	0	ERMP1	56366	KIAA1432	57442	MLANA	Coding	YES
UHNhscpg0013780	chr2:12216720 7-122167620	0	KIAA0622	182	KIAA0622	120835	RNU4ATAC	Coding	NO
UHNhscpg0009771	chr5:84158463 -84158717	#N/A	#N/A	477852	EDIL3	1419544	NBPF22P	Intergenic	NO
UHNhscpg0007053	chr10:9837407 8-98374233	0	PIK3AP1	27269	TM9SF3	217805	LCOR	Coding	NO
UHNhscpg0017489	chr3:17246830 8-172469056	0	ECT2	39300	NCEH1	3190	ECT2	Coding	YES
UHNhscpg0007708	chr5:15700306 7-157003202	#N/A	#N/A	299	ADAM19	49484	SOX30	Intergenic	YES
UHNhscpg0015943	chr7:14270497 2-142705297	#N/A	#N/A	45469	KEL	17989	OR9A2	Intergenic	NO
UHNhscpg0019050	chr11:3558496 8-35585208	#N/A	#N/A	33120	PAMR1	54526	FJX1	Intergenic	NO

UHNID	Genome Location	Within Distance	Gene Symbol	Upstream Distance	Gene Symbol	Downstream Distance	Gene Symbol	Type	CpG island
UHNhscpg0000825	chr11:6235865 8-62359223	0	TUT1	13895	FLJ00132	1451	MTA2	Coding	YES
UHNhscpg0000210	chr17:5953160 5-59532218	0	TBX4	40964	C17orf82	1588	TBX4	Coding	YES
UHNhscpg0017459	chr10:9606604 6-96066436	0	PLCE1	19219	CR599144	26552	NOC3L	Coding	NO
UHNhscpg0016480	chr1:23676758 5-236768045	0	HEATR1	21089	HEATR1	81724	ACTN2	Coding	YES
UHNhscpg0017432	chr15:4195392 3-41954225	0	MGA	82397	TYRO3	72494	MGA	Coding	YES
UHNhscpg0013446	chr15:5928026 5-59281002	0	RNF111	54413	SLTM	42000	RNF111	Coding	YES
UHNhscpg0000904	chr4:8604897- 8605570	0	GPR78	15379	GPR78	263202	HMX1	Coding	NO
UHNhscpg0005794	chr2:70475858 -70475935	#N/A	#N/A	79	TIA1	8906	PCYOX1	Intergenic	YES
UHNhscpg0018958	chr22:5024817 1-50248477	0	ZBED4	26991	BRD1	48378	ALG12	Coding	NO
UHNhscpg0007521	chr18:4772142 1-47721819	0	MYO5B	97433	MYO5B	31743	CCDC11	Coding	YES

The detailed information of location and the names of genes can be found at <http://data.microarrays.ca/cpg/>.

Table S9. Summary of discriminatory analysis using Weka

Model	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area
SMRI and QSBB Cortex						
NaiveBayes	0.81	0.2	0.82	0.81	0.807	0.773
RandomForest	0.81	0.191	0.81	0.81	0.81	0.818
All Twins						
NaiveBayes	0.538	0.46	0.55	0.538	0.541	0.6
RandomForest	0.385	0.548	0.403	0.385	0.355	0.545
European Twins						
NaiveBayes	0.704	0.32	0.703	0.704	0.699	0.625
RandomForest	0.815	0.198	0.815	0.815	0.814	0.844
Australian Twins						
NaiveBayes	0.438	0.533	0.456	0.438	0.431	0.46
RandomForest	0.25	0.742	0.258	0.25	0.25	0.238
Bipolar Germline model on SMRI and QSBB Cortex						
NaiveBayes	0.556	0.336	0.781	0.556	0.489	0.612
RandomForest	0.708	0.346	0.717	0.708	0.694	0.765
Bipolar Germline model on QSBB cortex						
NaiveBayes	0.722	0.278	0.821	0.822	0.699	0.784
RandomForest	0.75	0.25	0.833	0.75	0.733	0.951

Supplemental References

1. Bucholz KK, Cadoret R, Cloninger CR, Dinwiddie SH, Hesselbrock VM, Nurnberger JI, Jr., *et al.* (1994): A new, semi-structured psychiatric interview for use in genetic linkage studies: a report on the reliability of the SSAGA. *J Stud Alcohol.* 55:149-158.
2. Kessler RC, Ustun TB (2004): The World Mental Health (WMH) Survey Initiative Version of the World Health Organization (WHO) Composite International Diagnostic Interview (CIDI). *Int J Methods Psychiatr Res.* 13:93-121.
3. Eysenck SBG, Eysenck HJ, Barrett P (1985): A revised version of the psychoticism scale. *Pers Individ Dif* 6:21-29.
4. Boomsma DI, Beem AL, van den Berg M, Dolan CV, Koopmans JR, Vink JM, *et al.* (2000): Netherlands twin family study of anxious depression (NETSAD). *Twin Res.* 3:323-334.
5. Willemsen G, de Geus EJ, Bartels M, van Beijsterveldt CE, Brooks AI, Estourgie-van Burk GF, *et al.* (2010): The Netherlands Twin Register biobank: a resource for genetic epidemiological studies. *Twin Res Hum Genet.* 13:231-245.
6. Wing JK, Babor T, Brugha T, Burke J, Cooper JE, Giel R, *et al.* (1990): SCAN. Schedules for Clinical Assessment in Neuropsychiatry. *Arch Gen Psychiatry.* 47:589-593.
7. Eysenck HJ, Eysenck SBG (1975): Manual of the Eysenck Personality Questionnaire (Junior and Adult).
8. Sabunciyan S, Aryee MJ, Irizarry RA, Rongione M, Webster MJ, Kaufman WE, *et al.* (2012): Genome-Wide DNA Methylation Scan in Major Depressive Disorder. *PLoS One.* 7:e34451.
9. Kirk KM, Birley AJ, Statham DJ, Haddon B, Lake RI, Andrews JG, *et al.* (2000): Anxiety and depression in twin and sib pairs extremely discordant and concordant for neuroticism: prodromus to a linkage study. *Twin Res.* 3:299-309.
10. Zeileis A, Hothorn T (2002): Diagnostic Checking in Regression Relationships. *R News.* 2:7-10.
11. R Development Core Team (2011): *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing.
12. Harrison MJ, McCabe BPM (1979): A test for heteroscedasticity based on ordinary least squares residuals. *J Am Stat Assoc.* 74:494-499.
13. Davies L, Gather U (1993): The identification of multiple outliers. *J Am Stat Assoc.* 88:782-792.
14. Salomonis N, Nelson B, Vranizan K, Pico AR, Hanspers K, Kuchinsky A, *et al.* (2009): Alternative splicing in the differentiation of human embryonic stem cells into cardiac precursors. *PLoS Comp Biol.* 5:e1000553.

15. Salomonis N, Schlieve CR, Pereira L, Wahlquist C, Colas A, Zambon AC, *et al.* (2010): Alternative splicing regulates mouse embryonic stem cell pluripotency and differentiation. *Proc Natl Acad Sci U S A.* 107:10514-10519.
16. Hall M, Frank E, Holmes G, Pfahringer B, Reutemann P, Witten IH (2009): The WEKA data mining software: an update. *SIGKDD Explor Newsl.* 11:10-18.
17. Hall M (1999): *Correlation-based Feature Subset Selection for Machine Learning.* Department of Computer Science, University of Waikato.
18. Breiman L (2001): Random Forests. *Machine Learning.* 45:5-32.
19. Langfelder P, Horvath S (2008): WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics.* 9:559.
20. Mill J, Tang T, Kaminsky Z, Khare T, Yazdanpanah S, Bouchard L, *et al.* (2008): Epigenomic profiling reveals DNA-methylation changes associated with major psychosis. *Am J Hum Genet.* 82:696-711.
21. Opgen-Rhein R, Strimmer K (2007): From correlation to causation networks: a simple approximate learning algorithm and its application to high-dimensional plant gene expression data. *BMC Syst Biol.* 1:37.
22. Newman ME (2006): Modularity and community structure in networks. *Proc Natl Acad Sci U S A.* 103:8577-8582.