Supplementary Text

Contents

Data retrieval

We retrieved gene expression datasets matching the keywords “autism” and “autistic” from the Gene Expression Omnibus (GEO `http://www.ncbi.nlm.nih.gov/geo/`) on September 10, 2012. There were no additional unique datasets found in ArrayExpress (`http://www.ebi.ac.uk/arrayexpress/`). Shortlisted datasets include human blood and brain expression profiling studies with case-control experiment designs only. A preliminary analysis was conducted on these datasets. Two datasets, GSE4187 and GSE26415 were deemed unsuitable for the analysis and thus removed, as described in the main paper.

Data Preprocessing

In the final twelve data sets, trancriptome analyses were performed on three common commercial platforms - Affymetrix, Illumina, Agilent, as well as one non-commercial platform, TIGR. Raw data (.CEL, .MEV etc.) are often processed with various methods provided by the vendors. Whenever possible, we downloaded raw data files for datasets on Affymetrix and Illumina platforms and preprocessed them locally. Affymetrix datasets were subjected to Robust Multi-array Analysis (RMA) from the *affy* [1] package in Bioconductor. Illumina datasets were quantile normalized and log_2 transformed using the *lumi* [2] package. Datasets on the TIGR platform were not locally preprocessed as the submitters' preprocessing methods are similar across the studies. There is only one dataset that uses Agilent arrays. Standard preprocessing is not necessary. Sample sources and tissue types for each dataset are specified in Supplementary Table 1.

Quality Control

The processed data were then subjected to an additional set of quality controls. We identified and excluded 17 samples that were used in more than one study, retaining data for the samples in the study with a smaller samples size. We removed eight samples from subjects with syndromic disorders of known genetic etiology (Fragile-X syndrome), nine non-ASD cases with mental retardation, as well as samples which were prepared differently than the rest of the samples (e.g. formalin fixed). Further details of exclusion are in Supplementary Table 3.

Sample Correlation Analysis

We removed sample outliers using a sample correlation analysis. Outlier samples were identified as those with correlation more than two standard deviations from the mean sample-to-sample expression profile correlation, and removed iteratively until no samples met the threshold for removal. This resulted in the removal of a toal of 54 samples, affecting seven studies. The remaining samples in the dataset were renormalized using quantile normalization.

Average Across Replicates

The identification of independent units is crucial for statistical analysis, because hidden correlations (Supplementary Figure 1) can lead to biases and inflate statistical significance [3]. As the question of interest in our study concerns a biological comparison between whole organisms, that is individuals with ASD and individuals

without ASD, we define independent biological units as unique individual subjects in each dataset. Multiple samples obtained from the same subject are regarded as technical replicates. Two of the studies included technical replicates for some specimens, in which case we computed the mean of the expression values to get a single expression profile for each subject.

Batch Effects Correction

We also looked at possible batch effects whenever batch information is available. Batch information was obtained by automated extraction of "scan dates" or "users" from raw data files, as well as supplementary texts and metadata provided by submitters. To detect possible batch effects, we compared the first two principal components to batch data, and visually identify groups that are separated by the principal components. The amount of variation explained by each principal component is also reported (Supplementary Figure 2). The *global test* method implemented in the bioconductor package of the same name is used as a secondary measure to obtain a significant value for the association between batch and the expression values (data not shown). In this analysis, we also discovered 34 samples in which batch effects were confounded with the case grouping. These samples were removed. In other datasets, we corrected for possible batch effects using ComBat [4] after discarding probes that are missing in more than 20% of the samples. Batch effects could not be corrected in GSE28521 due to small batch sizes, such that priors cannot be estimated in ComBat. Illumina slide numbers were used as batch information here.

Differential Expression Analysis

Differential expression analysis was conducted using analysis of variance (ANOVA) based on an empirical Bayes approach provided in the *limma* R package. For each dataset, we conducted a two group disease-control comparison for all probes. Phenotypic subgroups (Supplementary Table 2) were pooled into one generic autism disease group. To consider the direction of expression change in the meta-analyses, we computed one-tailed p-values from the resulting two-tailed p-values and t-statistics. Probes were annotated with platform specific annotations in Gemma [5], where gene assignments are made based on current genome annotations obtained via sequence analysis. Each data set was then collapsed to the gene level to allow cross-platform integration. This was done by taking the probe with the Bonferroni corrected minimum p-value (best scoring probe, $n \times min(p) < \alpha$), as the smallest p-value is least likely to occur by chance [6]. We excluded probes that map to multiple genes or do not map to a gene at all are from the analysis. The proportion of differentially expressed genes ($\pi_1 = 1 - \pi_0$) was estimated using the *qvalue* package in R. As the internal "bootstrap" method in the package does not return a standard error, we computed π_0 standard errors over one hundred bootstrap iterations locally. π_0 values from the "bootstrap" and "smoother" method were similar. We grouped π_0 values according to factors like sample size, tissue type, platform type, but there were no obvious trends.

We next compared the results from our re-analysis to that of previous publications, and evaluated the outcomes using the area under the receiver operating characteristic curve (AU-ROC, equivalent to the Wilcoxon rank-sum test), as well as average precision (AP, equivalent to the area under the precision-recall curve). The AU-ROC gives a probability for which a true positive is ranked higher than a false positive. AP gives us the amount of correct hits from the top N ranked genes, averaged across $N = \{1...n\}$, where n is the total number of genes in the rankings. While AU-ROC is an evaluation of where the true positives lie relative to the false positives, AP is sensitive to genes with higher rankings (top hits). The same threshold free evaluation measures are also used in later sections.

Meta Analysis of Gene Expression Datasets

We used a relative approach for this meta-analysis, such that each dataset was individually analyzed, and the results were subsequently combined. Two separate meta-analyses were conducted for the blood and brain data sets. We required each gene to be present in at least three datasets in each of the meta-analyses.

Fisher's Combined Probability Test

Fisher's combined probability test [7] takes the raw p-value calculated from the individual differential expression analysis for a gene across all datasets and combines them to generate a summary statistic (S) using the equation:

$$S_i = -2\sum_{j=1}^{k} log(p_i)$$

Using this method, we combined our results from multiple independent tests, all having the same null hypothesis (no difference between autism and control groups). Under the null hypothesis, the resulting test statistic has a χ^2 distribution. P-values for meta-analysis can then be obtained from the test-statistic using the χ^2 distribution with $2k$ degrees of freedom. We corrected the p-values for multiple testing using the Benjamini Hochberg [8] correction method. The resulting FDR represents the proportion of false positives among all the positive results returned at a given threshold.

Meta-rank Analysis

The meta-rank analysis is a rank aggregation strategy that involves using the average rank of the gene instead of the combination of p-values. For each individual gene, its rank was determined by their order of p-values. The smallest p-value in an experiment would have the highest rank. The meta-rank of gene i is computed by averaging the ranks across datasets 1 to k,

$$R_i = \frac{1}{k}\sum_{j=1}^{k} R_{ij}$$

The ranks of these averages were then computed for all genes. This method is less sensitive to possible dataset bias, where one potentially extremely low p-value could skew the results and have it dominate the combined result from Fisher's method. In other words, we can detect individual data sets that produce deviant significance values. However, unlike the Fisher's combined probability test, the distribution of ranks cannot be represented by a known statistical distribution.

In order to obtain a test statistic, we computed the permutation null distribution. We randomly permute the p-values in each study, and recalculate the metric R*. Repeating the process 10000 times gives an $M \times 10000$ matrix of permuted values, where M is the number of genes. The permutation null is then the empirical distribution of all values in that matrix. We can assess for significance by testing against the permutation null.

$F(x)$ is a function of the empirical cumulative distribution of the permutation null, where x is a random variable, which is, in this case, the metarank of gene i. This is computed for the number of studies, $k = 3, 4, 5, ..., 9$ for the null distribution of blood data and $k = 3$ for brain data. To ensure that the meta-signature genes are not sensitive to the choice of the meta-analysis method used, we reanalyzed both blood and brain datasets

using the method described. The rank correlations between these two methods are 0.80 and 0.61 for brain and blood datasets respectively (averaged over up-regulated and down-regulated lists), suggesting that there are some discrepancies between these methods. However, by quantifying the predictive power of this method with respect to meta-signatures from Fisher's method, we observed that the choice of method will not have a substantial effect on our selection of the top hits in the signatures (Supplementary Figure 4). Subsequent functional analyses were thus based on results yield from Fisher's method.

Sex-linked genes and other biological factors that might affect gene expression

Close inspection of differentially expressed genes revealed an abundance of highly differentially expressed sex-linked genes, such as USP9Y (Supplementary Figure 5. This can be attributed to the unbalanced study design where not all sibling pairs are sex-matched. Specifically, the propensity for the probands to be male in GSE37772 (OR=5.73, Supplementary Table 4) resulted in the surfacing of Y-linked genes. To eradicate the identified bias, we excluded Y-linked genes as well as autosomal genes showing evidence of sexual dimorphism in downstream analysis. As global expression profiles differ in the blood and brain, sex biased genes reported in Kang et al [9] were used to identify sexually dimorphic genes in the brain meta-signature, and those reported in Whitney et al [10] were used for the blood meta-signature. We also excluded genes that have been shown to escape X-inactivation [11]. Only those showing strong evidence (royal blue, violet and indigo; n = 61) in the cited study are selected for exclusion here.

A previous analysis has been done by Mistry et al to investigate gene expression changes affected by other biological factors in postmortem human brain, such as age, and postmortem interval (PMI) [12]. To check whether the results of our analysis are impacted by these factors, we overlapped signatures from the brain with the genes reported. In general, this meta-analysis is minimally affected as only three genes (PMI: HMOX1, HINT2; Age: FDX1L) were flagged. Generally these genes were not highly ranked in the meta signatures compared to other genes (AU-ROC).

The Jackknife Procedure

As the combined probability method applied has a propensity to skew towards a single deviant p-value for any gene in a study, we used a jackknife approach to further select for genes that are robust to the analysis [13]. The jackknife procedure implemented in the *bootstrap* R package involves repeating the meta-analysis k times, where k is the number of data sets. For each trial k, the k^{th} data set is left out. The agreement among these k jackknife meta-analyses was used as a basis for identifying a "core" signature that excludes genes appearing due to the influence of a single data set. Specifically, genes found in the top 200 across every core signature are considered robust.

Functional Enrichment Analysis

Gene set enrichment analysis was performed using ErmineJ Version 3.0 ([14] , `http://erminej.chibi.ubc.ca`), a software for determining enrichment of *Gene Ontology* (GO) [15] terms for a given gene list. GO terms represent a controlled vocabulary that links a certain molecular function, biological process, or cellular component to genes. We focused on terms under the "biological process" tree for our analysis. Significant enrichment of a specific GO term could suggest that the gene list is enriched for genes involved in a particular biological process. The PrecisionRecall (PR) method is used here. PR uses average precision

as a scoring function, thus it is sensitive to genes at the top of the rankings, without having to set a threshold. For each run, the negative log of Fisher's corrected p-values were used as the input, testing against gene sets within the $5 - 200$ size range over 500000 iterations [16]. We ran enrichment analysis on meta-analysis results of gene expression in blood and brain. Additionally, we performed meta-analyses on functional enrichment results of each individual data set, in both blood and brain. We used a jackknife analysis to identify concordantly dysregulated gene sets across all studies. We observed a relatively high enrichment of "cellular respiration" (GO:0045333) genes in brain expression data with both methods.

We also downloaded additional candidate gene sets from the Simons Foundation Autism Research Initiative (SFARI) database (`www.sfari.org`, retrieved in December 2012). Seven categories were established by SFARI based on their gene scoring syndrome - 1. High Confidence; 2. Strong Candidate; 3. Suggestive Evidence; 4. Minimal Evidence; 5. Hypothesized; 6. Not Supported; S Syndromic. We tested for enrichment of ASD candidate genes in Categories 2-5 and "S". As a control, we also tested Category 6. Category 1 was empty and thus excluded for analysis.

Literature Derived Candidate Genes

We performed a literature search for the candidate genes derived using Phenocarta [17](`http://phenocarta.chibi.ubc.ca`) as a main resource. Using this resource, we captured 807 genes under the umbrella term "autism spectrum disorder". Candidate genes from model organisms (mouse = 11,zebrafish = 1) were mapped to their human homologs using *HomoloGene* (`ftp://ftp.ncbi.nih.gov/pub/HomoloGene/`, build 67), resulting in 798 unique human genes. In addition to studies in Table 1, significantly differentially expressed genes from three additional expression profiling studies [18–20] for which data were not publicly available, as well as Hu et al.'s study that was excluded entirely due to duplicated samples (GSE4187, [21]) were obtained and compared to our results. There were no overlaps from Hu et al. and Glatt et al.'s studies. Because autism has been repeatedly associated with schizophrenia, we were also interested to see if there were similarities between the autism meta signatures and schizophrenia meta signatures. We obtained the list of differentially expressed genes reported in a genome-wide expression profiling study of combined cohorts in schizophrenia [13] and compared it with our brain meta-signatures. Results suggest that there are some similarities between autism and schizophrenia expression profiles in postmortem brain (Supplementary Figure 11).

CNV Enrichment Analysis

Copy number variation data was obtained from the Autism Chromosomal Rearrangement Database (ACRD, `http://projects.tcag.ca/autism/`) [22], Sanders et al [23] (Table S4 in original study) as well as Pinto et al [24] (Table S8 in original study), obtaining a total of 1023 CNVs. These variants are thought to be pathological, but to ensure uniformity, we computed their frequencies in the Database of Genomic Variants (DGV, retrieved in March 2013) [25] as described in Sanders et al. We identified seven common CNVs. After lifting the genomic coordinates genes over to hg18 (to match CNV data) using the UCSC liftOver tool [26] (`http://genome.ucsc.edu/cgi-bin/hgLiftOver`), Fisher's exact test was used to compute global enrichment of dysregulated genes in ASD-asociated rare CNV (n=1016). To reduce the amount of overlap in computing enrichment of dysregulated genes, we merged individual CNVs into CNV regions using a 90% reciprocal overlap. We also merged small CNVs that are completely nested within larger CNVs by taking the breakpoints of the larger CNV (union). The total number of merged CNV regions is 732 (Gain=385, Loss=340, Unknown=1).We

included all classes of CNV transmissions (inherited, de novo, unknown). Restricting our analysis to de novo CNVs did not substantially affect the results.

Single Nucleotide Polymorphisms (SNPs) in Candidate Genes

To see if common SNPs could potentially affect the expression levels of our candidate genes and lead to false positives, we looked for common SNPs (>=1% minor allele frequency, UCSC build 138, `http://hgdownload.cse.ucsc.edu/goldenpath/hg19/database/`) within probes used in various platforms analyzed in this study. Instead of using gene mappings provided by different platform manufacturers, we obtained probe alignments (hg19) from Gemma [5] as sequence analysis is uniformly performed across all platforms. The results are presented in Supplementary Tables 7- 10. Some of the probes (denoted NA) have been excluded for gene mapping in Gemma due to poor sequence alignment [27]. Several probes that target genes on sex chromosomes align to genomic regions in both chromosomes X and Y. Readers should be aware of this ambiguity when analyzing these genes, as with other sex-linked genes flagged in the results table.

Network Analysis

We conducted the network analysis on a human protein-protein interaction (PPIN) network. The PPIN network comprise data from the Human Protein Reference Database (HPRD) [28], Molecular Interaction Database(MINT) [29], Database of Interacting Proteins (DIP) [30], innateDB [31] and irefIndex [32]. With the aggregated network, we computed local network properties for the core candidate gene sets. 10000 random gene sets with similar size and node degree were sampled from the network to construct permutation distributions of the average shortest path length (Dijkstra's algorithm) and local clustering coefficient [13]. The results are in Supplementary Figure 10

Cell Type Subgroup Meta-analysis

We investigated the impact of cell type heterogeneity by conducting separate analyses on a homogenous subgroup comprising data sets with samples from LCL, and a non-LCL subgroup comprising datasets with samples from whole blood and lymphocytes. At an FDR threshold of 0.05, we observed 1 down regulated gene and 35 up-regulated genes in the LCL subgroup; there were 164 up-regulated genes and 143 down-regulated genes in the non-LCL subgroup. The large number of genes found in the non-LCL subgroup is likely due to the lack of independence between two data sets from the same study, GSE18123.1 and GSE18123.2 (Supplementary Figure 16). We applied the jackknife procedure to detect core signatures in each subgroup. As seen in Supplementary Figure 14, we found 17 up-regulated and 10 down-regulated core signature genes in the LCL subgroup. In the non-LCL subgroup, there were 9 up-regulated and 8 down-regulated core signature genes (Supplementary Figure 15). While additional genes were detected in this subgroup analysis, the trends were not strong across data sets in each subgroup (Supplementary Figure 13). Further investigations are required to determine the significance of these genes.

Supplementary Tables

List of Supplementary Tables

	Tissue Type	Tissue Source
Brain		
GSE28475	Dorsal lateral prefrontal cortex, middle frontal gyrus	HBTRC, NICHD
GSE28521	Frontal cortex (BA9), temporal cortex (BA41/42 or BA22)	ATP, HBB
GSE38322	BA19	ATP, HBTRC, NICHD
Blood		
GSE6575	Whole blood	CHARGE
GSE7329	Lymphoblastoid cell lines	AGRE
GSE15402	Lymphoblastoid cell lines	AGRE
GSE15451	Lymphoblastoid cell lines	AGRE
GSE18123.1	Whole blood	CHB, ACB
GSE18123.2	Whole blood	CHB, ACB
GSE25507	Peripheral blood lymphocytes	Phoenix
GSE32136	Lymphoblastoid cell lines	AGRE
GSE37772	Lymphoblastoid cell lines	Simon Simplex Collection

HBTRC: Harvard Brain Tissue Resource Centre.
NICHD: National Institute for Child Health and Human Development Brain and Tissue Bank.
ATP: Autism Tissue Program.
HBB: Harvard Brain Bank.
CHARGE: Childhood Autism Risks from Genetics and the Environment.
AGRE: Autism Genetic Resource Exchange.
CHB: Children's Hospital Boston.
ACB: Autism Consortium Boston.

Supplementary Table 1: Summary of tissue sources.

Supplementary Table 2: Summary of diagnosis criteria and ASD phenotypes in the original studies. Refer to Table 1 for study citations.

	Diagnosis Criteria	**Phenotypic descriptions**
Brain		
GSE28475	ADI-R, ADOS, TARF, medical records	Autism
GSE28521	Available upon request, includes ADI-R diagnostic scores, AN-Brain Bank Case Number	Autism
GSE38322	ADI-R	Autism
Blood		
GSE6575	DSM-IV,ADI-R, ADOS	Autism no regression, autism with regression
GSE7329	ADI-R, ADOS, Raven-IQ	ASD with dup(15q)
GSE15402	ADI-R, Raven's score, Peabody Picture Vocabulary Test	ASD[a]
GSE15451	ADI-R	ASD[b]
GSE18123.1	DSM-IV-TR,ADOS, ADI-R, comprehensive clinical testing	Autism, Asperger's Disorder, PDD-NOS
GSE18123.2	DSM-IV-TR,ADOS, ADI-R, comprehensive clinical testing	Autism, Asperger's Disorder, PDD-NOS
GSE25507	DSM-IV, ADOS, ADI-R	classical autism
GSE32136	-	ASD
GSE37772	Refer to the SFARI database for phenotype information	Autism

[a] Language, Mild, Savant (cluster analysis of ADI-R scores)
[b] severe language impairment (cluster analysis of ADI-R scores)
ADI-R: Autism Diagnostic Interview-Revised
ADOS: Autism Diagnostic Observation Schedule
TARF: The Autism Research Foundation
DSM-IV: Diagnostics and Statistical Manual of Mental Disorders IV, (TR: text revision)
SFARI: Simons Foundation Autism Research Initiative

(a)

Accession	Samples excluded
GSE6575	Removed non ASD subjects with mental retardation or developmental delay.
GSE7329	Removed samples with a Fragile X (FMR1-FM) mutation.Remove samples in 2005 batch (scan date).
GSE15402	Remove samples in batches not performed by user KyungS.
GSE15451	Removed tissue samples that overlap with GSE32136 [Blood ID: HI0779, HI2022, HI2772, HI3143, HI3914, HI2044, HI2769, HI3144, HI4360, HI0777], as well as a subject that overlap with GSE15402 [Subject ID: AU0325].
GSE18123.2	Excluded samples without an assigned batch.
GSE28475	Excluded seizure samples, IVT assays.Excluded formalin-fixed samples. Removed samples that are also present in GSE38322 [Subject ID: UMB4670, UMB1860].
GSE28521	Excluded samples from the cerebellum. Removed subjects that overlap with subjects in GSE38322 [Subject ID: AN19511, AN06420, AN08873, AN10833].
GSE32136	Excluded PPA/Propanol treated samples and samples in batch “bcmmes”.
GSE38322	Excluded samples from the cerebellum.
GSE37772	Excluded samples from mothers.

(b)

Datasets	Original Samples	Number Excluded	Number Failed QC	Remaining Samples
Brain				
GSE28475	123	63	14	34*
GSE28521	79	25	4	27*
GSE38322	36	26	0	10
Blood				
GSE6575	56	9	3	44
GSE7329	30	17	0	13
GSE15402	116	10	0	106
GSE15451	38	11	0	27
GSE18123.1	99	0	7	92
GSE18123.2	186	12	18	156
GSE25507	146	0	3	143
GSE32136	23	15	0	8
GSE37772	439	3	5	431

Supplementary Table 3: a) Samples excluded in each study; b) Summary statistics of samples removed. Asterisks denote data sets with technical replicates. The remaining samples comprise unique subjects only.

Supplementary Table 4: Demographics I - Gender. Gender imbalance is seen in some data sets, such as GSE37772. OR: Odds ratio.

Datasets	**ASD**		**Control**			
	Male	**Female**	**Male**	**Female**	**OR**	**Total**
Brain						
GSE28475	11	2	18	3	0.92	34
GSE28521	8	4	14	1	0.14	27
GSE38322	4	0	6	0	-	10
Blood						
GSE6575	28	5	8	3	2.10	44
GSE7329	7	0	6	0	-	13
GSE15402	77	0	29	0	-	106
GSE15451	15	0	12	0	-	27
GSE18123.1	64	0	28	0	-	92
GSE18123.2	72	21	30	33	3.80	156
GSE25507	80	0	63	0	-	143
GSE32136	4	0	4	0	-	8
GSE37772	198	34	105	94	5.20	431

Supplementary Table 5: Demographics II - Age, PMI and race of subjects in each study.

Datasets	Age range (years)		PMI		Race
	ASD	Control	ASD	Control	
Brain					
GSE28475	2-56	3-56	4-43.25	5-36	Caucasian, African American, unknown, mixed
GSE28521	5-51	6.75-43.25	4.75-32.92	16-56	Predominantly Caucasian, Asian
GSE38322	2-39	4-22.5	13-24.2	1-60	White, unknown
Blood					
GSE6575	matched	matched	-		
GSE7329	-	-	-	-	-
GSE15402	5-28	3-34	-	-	Predominantly White, Asian, multi
GSE15451	4-12	2-12	-	-	Predominantly White, unknown
GSE18123.1	3.4-17.5	2.8-16	-	-	Predominantly Caucasian, Asian, unknown, mixed
GSE18123.2	2-21	2.5-22	-	-	Predominantly Caucasian, Asian, unknown, mixed
GSE25507	2-14	3-11	-	-	Primarily Caucasian
GSE32136	-	-	-	-	-
GSE37772	4-17.7	3-23.8	-	-	Caucasian and non-caucasian

Supplementary Table 6: Overlap between results reported in the literature and individual re-analysis of differential expression. Per dataset significant probes reported at FDR < 0.05. Gene symbols are used as a proxy for probes in GSE18123.1; GenBank accessions are used in GSE15451 and GSE15402; Spot IDs are used for GSE7329. GSE25507 computed differences in expression variance instead of differential expression; GSE37772 reported outlier genes instead of differentially expressed genes; GSE32136 is not published.

Datasets	**Significant probes**	**Probes reported**	**Overlap**	**AUC**	**Precision(%)**
GSE15402	73	530	65	0.98	58.70
GSE15451	0	45	0	0.86	1.58
GSE18123.1	284	489	43	0.79	8.34
GSE18123.2	69	610	43	0.92	31.00
GSE25507	-	-	-	-	-
GSE28475	0	200	0	0.89	6.94
GSE28521	4	588	2	0.89	21.50
GSE32136	-	-	-	-	-
GSE37772	-	-	-	-	-
GSE38322	0	41	0	0.98	6.84
GSE6575	0	55	0	0.96	3.34
GSE7329	596	1281	339	0.95	44.10

See spreadsheets attached.

Supplementary Table 7: Up-regulated meta-signature in brain.

Supplementary Table 8: Down-regulated meta-signature in brain.

Supplementary Table 9: Up-regulated meta-signature in blood.

Supplementary Table 10: Down-regulated meta-signature in blood.

Symbol	Name	Meta p-value
UBE3A	ubiquitin protein ligase E3A	4.72E-06
CDKL5	cyclin-dependent kinase-like 5	1.43E-03
DMD	dystrophin	2.36E-03
SHANK3	SH3 and multiple ankyrin repeat domains 3	6.58E-03
HOXA1	homeobox A1	1.45E-02
PTEN	phosphatase and tensin homolog	2.34E-02
TSC1	tuberous sclerosis 1	3.09E-02
DHCR7	7-dehydrocholesterol reductase	3.99E-02
SCN1A	sodium channel, voltage-gated, type I, alpha subunit	7.29E-02
AHI1	Abelson helper integration site 1	9.63E-02
NF1	neurofibromin 1	9.67E-02
CACNA1C	calcium channel, voltage-dependent, L type, alpha 1C subunit	1.00E-01
RAI1	retinoic acid induced 1	1.52E-01
ALDH5A1	aldehyde dehydrogenase 5 family, member A1	1.66E-01
MECP2	methyl CpG binding protein 2 (Rett syndrome)	2.12E-01
ARX	aristaless related homeobox	2.21E-01
SLC9A6	solute carrier family 9, subfamily A (NHE6, cation proton antiporter 6), member 6	2.93E-01
ADSL	adenylosuccinate lyase	3.27E-01
DMPK	dystrophia myotonica-protein kinase	3.50E-01

Supplementary Table 11: All genes and respective p-values in the SFARI syndromic category.

Gene Symbol	Gene Name	Meta p-value
ATP5O	ATP synthase, H+ transporting, mitochondrial F1 complex, O subunit	1.83E-05
UQCRQ	ubiquinol-cytochrome c reductase, complex III subunit VII, 9.5kDa	5.45E-05
UQCRC1	ubiquinol-cytochrome c reductase core protein I	1.86E-04
CYC1	cytochrome c-1	2.90E-04
COX5B	cytochrome c oxidase subunit Vb	2.98E-04
NDUFA11	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 11, 14.7kDa	4.38E-04
ATP5L	ATP synthase, H+ transporting, mitochondrial Fo complex, subunit G	4.53E-04
UQCR10	ubiquinol-cytochrome c reductase, complex III subunit X	4.53E-04
UQCRC2	ubiquinol-cytochrome c reductase core protein II	5.25E-04
NDUFA13	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 13	5.35E-04
SLC25A12	solute carrier family 25 (aspartate/glutamate carrier), member 12	5.37E-04
FH	fumarate hydratase	7.55E-04
UQCR11	ubiquinol-cytochrome c reductase, complex III subunit XI	7.74E-04
NDUFS4	NADH dehydrogenase (ubiquinone) Fe-S protein 4, 18kDa (NADH-coenzyme Q reductase)	8.29E-04
IDH3A	isocitrate dehydrogenase 3 (NAD+) alpha	9.06E-04

Supplementary Table 12: Top genes in the “cellular respiration" GO category at a meta-analysis raw p-value threshold of 0.0001. There are a total of 116 genes in this functional group.

See spreadsheets attached.

Supplementary Table 13: a) Comparisons between core signatures and lists of differentially expressed genes from original studies; b) Pairwise comparisons of original hits.

Supplementary Table 14: Dysregulated genes (FDR<0.05, meta-signature) within ASD-associated CNV. Fisher's exact test was used to compute significance. NS: Not significant.

	Observed		**Expected**		**Total**	**p-value**
	n	%	n	%		
Brain						
Up	1	3.33	3	10.00	30	NS
Down	4	8.16	4	8.16	49	NS
Total	5	6.3	7	8.86	79	
Blood						
Up	22	13.75	17	10.63	160	0.12
Down	3	3.16	9	9.47	95	NS
Total	25	9.80	26	10.20	255	

Supplementary Table 15: Meta-signature genes that are also dysregulated in schizophrenia. FDR<0.1

	Genes	P-Value	FDR
Up-regulated			
	ABCA1	7.19e-05	4.16e-02
	P4HA1	6.83e-04	9.19e-02
Down-regulated			
	CCDC25	2.59e-05	3.27e-02
	LRRC17	3.06e-04	5.77e-02
	RMND5B	4.30e-04	6.09e-02
	SLC25A12	5.37e-04	6.57e-02
	FARSA	8.45e-04	7.78e-02
	PPA2	9.61e-04	8.26e-02
	APBA2	1.11e-03	9.01e-02

Data Set	Sample ID	Source
GSE28521	AN11989	ATP, HBB
GSE28521	AN12457	ATP, HBB
GSE28521	AN16115	ATP, HBB
GSE28521	AN16641	ATP, HBB
GSE28521	AN17138	ATP, HBB
GSE28521	AN17254	ATP, HBB
GSE28521	AN17777	ATP, HBB
GSE28521	AN01570	ATP, HBB
GSE28521	AN08166	ATP, HBB
GSE28521	AN08792	ATP, HBB
GSE28521	AN00493	ATP, HBB
GSE28521	AN00764	ATP, HBB
GSE28521	AN10028	ATP, HBB
GSE28521	AN12137	ATP, HBB
GSE28521	AN12240	ATP, HBB
GSE28521	AN14757	ATP, HBB
GSE28521	AN15566	ATP, HBB
GSE28521	AN17425	ATP, HBB
GSE28521	AN19442	ATP, HBB
GSE28521	AN19760	ATP, HBB
GSE28521	AN01125	ATP, HBB
GSE28521	AN01410	ATP, HBB
GSE28521	AN03217	ATP, HBB
GSE28521	AN04479	ATP, HBB
GSE28521	AN07176	ATP, HBB
GSE28521	AN00142	ATP, HBB
GSE28521	AN00544	ATP, HBB
GSE38322	AN10723	ATP, HBTRC, NICHD
GSE38322	AN06420	ATP, HBTRC, NICHD
GSE38322	AN03345	ATP, HBTRC, NICHD
GSE38322	AN08873	ATP, HBTRC, NICHD
GSE38322	AN10833	ATP, HBTRC, NICHD
GSE38322	UMB4670	ATP, HBTRC, NICHD
GSE38322	BTB1453	ATP, HBTRC, NICHD
GSE38322	BTB3228	ATP, HBTRC, NICHD
GSE38322	UMB4543	ATP, HBTRC, NICHD
GSE38322	AN19511	ATP, HBTRC, NICHD
GSE28475	UMB797	HBTRC, NICHD
GSE28475	UMB1674	HBTRC, NICHD
GSE28475	B6756	HBTRC, NICHD
GSE28475	B5873	HBTRC, NICHD
GSE28475	UMB1349	HBTRC, NICHD
GSE28475	UMB1796	HBTRC, NICHD
GSE28475	UMB4231	HBTRC, NICHD

Continued. . .

Data Set	Sample ID	Source
GSE28475	B5334	HBTRC, NICHD
GSE28475	B5251	HBTRC, NICHD
GSE28475	B5352	HBTRC, NICHD
GSE28475	B6736	HBTRC, NICHD
GSE28475	B7109	HBTRC, NICHD
GSE28475	B7085	HBTRC, NICHD
GSE28475	B5813	HBTRC, NICHD
GSE28475	B6677	HBTRC, NICHD
GSE28475	B4756	HBTRC, NICHD
GSE28475	B7079	HBTRC, NICHD
GSE28475	B6860	HBTRC, NICHD
GSE28475	B5666	HBTRC, NICHD
GSE28475	B6399	HBTRC, NICHD
GSE28475	UMB4849	HBTRC, NICHD
GSE28475	UMB4722	HBTRC, NICHD
GSE28475	UMB4899	HBTRC, NICHD
GSE28475	B1469	HBTRC, NICHD
GSE28475	UMB1185	HBTRC, NICHD
GSE28475	UMB1714	HBTRC, NICHD
GSE28475	UMB4787	HBTRC, NICHD
GSE28475	UMB1377	HBTRC, NICHD
GSE28475	UMB818	HBTRC, NICHD
GSE28475	UMB1649	HBTRC, NICHD
GSE28475	UMB1499	HBTRC, NICHD
GSE28475	UMB4898	HBTRC, NICHD
GSE28475	UMB4671	HBTRC, NICHD
GSE28475	UMB1650	HBTRC, NICHD

Supplementary Table 16: Brain samples included in the meta-analysis.

Data Set	Individual ID	Sample ID	Source
GSE7329	01-19	-	AGRE
GSE7329	02-07	-	AGRE
GSE7329	03-43	-	AGRE
GSE7329	98-19	-	AGRE
GSE7329	AU006504	-	AGRE
GSE7329	AU010603	-	AGRE
GSE7329	AU010604	-	AGRE
GSE7329	AU016703	-	AGRE
GSE7329	AU060003	-	AGRE
GSE7329	AU081206	-	AGRE
GSE7329	AU0943303	-	AGRE
GSE7329	AU0995302	-	AGRE
GSE7329	AU1038304	-	AGRE
GSE32136	-	HI0779	AGRE
GSE32136	-	HI2772	AGRE
GSE32136	-	HI3143	AGRE
GSE32136	-	HI3914	AGRE
GSE32136	-	HI2044	AGRE
GSE32136	-	HI2769	AGRE
GSE32136	-	HI3144	AGRE
GSE32136	-	HI4360	AGRE
GSE15451	AU014504	HI0487	AGRE
GSE15451	AU014506	HI0488	AGRE
GSE15451	AU014505	HI0489	AGRE
GSE15451	AU022703	HI0354	AGRE
GSE15451	AU022705	HI0355	AGRE
GSE15451	AU022704	HI0356	AGRE
GSE15451	AU032505	HI0364	AGRE
GSE15451	AU032504	HI0366	AGRE
GSE15451	AU1059302	HI3141	AGRE
GSE15451	AU010806	HI0735	AGRE
GSE15451	AU010803	HI0737	AGRE
GSE15451	AU046504	HI0781	AGRE
GSE15451	AU046503	HI0785	AGRE
GSE15451	AU055504	HI0802	AGRE
GSE15451	AU055506	HI3168	AGRE
GSE15451	AU080805	HI2019	AGRE
GSE15451	AU080804	HI2020	AGRE
GSE15451	AU0939302	HI2590	AGRE
GSE15451	AU0939303	HI2591	AGRE
GSE15451	AU1327305	HI3299	AGRE
GSE15451	AU1327303	HI3301	AGRE
GSE15451	AU1244302	HI3649	AGRE
GSE15451	AU1244301	HI3650	AGRE

Continued. . .

Data Set	Individual ID	Sample ID	Source
GSE15451	AU1181302	HI3913	AGRE
GSE15451	AU1443302	HI4107	AGRE
GSE15451	AU1443304	HI4109	AGRE
GSE15451	AU1234304	HI4359	AGRE
GSE15402	AU1156303	HI4004	AGRE
GSE15402	AU1417301	HI3673	AGRE
GSE15402	AU1204302	HI2983	AGRE
GSE15402	AU1339302	HI3369	AGRE
GSE15402	AU0883301	HI3323	AGRE
GSE15402	AU0905301	HI2427	AGRE
GSE15402	AU0997302	HI2779	AGRE
GSE15402	AU073107	HI1367	AGRE
GSE15402	AU0920301	HI2402	AGRE
GSE15402	AU1327304	HI3303	AGRE
GSE15402	AU1409301	HI3599	AGRE
GSE15402	AU072504	HI1657	AGRE
GSE15402	AU1196302	HI3522	AGRE
GSE15402	AU1405303	HI4516	AGRE
GSE15402	AU015003	HI0591	AGRE
GSE15402	AU062203	HI1943	AGRE
GSE15402	AU0941301	HI2532	AGRE
GSE15402	AU1078302	HI2687	AGRE
GSE15402	AU017504	HI0266	AGRE
GSE15402	AU028905	HI0110	AGRE
GSE15402	AU051504	HI0792	AGRE
GSE15402	AU079103	HI1861	AGRE
GSE15402	AU0934301	HI2536	AGRE
GSE15402	AU0997301	HI2781	AGRE
GSE15402	AU1047312	HI2683	AGRE
GSE15402	AU1261301	HI3118	AGRE
GSE15402	AU1270301	HI3290	AGRE
GSE15402	AU041904	HI0649	AGRE
GSE15402	AU025005	HI0050	AGRE
GSE15402	AU015903	HI0928	AGRE
GSE15402	AU1652301	HI4820	AGRE
GSE15402	AU1546302	HI4341	AGRE
GSE15402	AU1067301	HI3424	AGRE
GSE15402	AU1332302	HI3470	AGRE
GSE15402	AU1685302	HI4838	AGRE
GSE15402	AU1397301	HI3660	AGRE
GSE15402	AU1648301	HI4751	AGRE
GSE15402	AU053504	HI1495	AGRE
GSE15402	AU056604	HI1234	AGRE
GSE15402	AU1174301	HI2853	AGRE

Continued. . .

Data Set	Individual ID	Sample ID	Source
GSE15402	AU0943302	HI2576	AGRE
GSE15402	AU1047302	HI2680	AGRE
GSE15402	AU1097301	HI2905	AGRE
GSE15402	AU073804	HI2251	AGRE
GSE15402	AU1182302	HI2846	AGRE
GSE15402	AU024004	HI1364	AGRE
GSE15402	AU1056301	HI4420	AGRE
GSE15402	AU0962301	HI2609	AGRE
GSE15402	AU1069301	HI2828	AGRE
GSE15402	AU1182301	HI2835	AGRE
GSE15402	AU1196301	HI4870	AGRE
GSE15402	AU1520301	HI4002	AGRE
GSE15402	AU081203	HI2028	AGRE
GSE15402	AU1008302	HI4461	AGRE
GSE15402	AU016803	HI1492	AGRE
GSE15402	AU1025301	HI3379	AGRE
GSE15402	AU0955303	HI2791	AGRE
GSE15402	AU1102303	HI2824	AGRE
GSE15402	AU005215	HI1276	AGRE
GSE15402	AU083604	HI2163	AGRE
GSE15402	AU005214	HI1279	AGRE
GSE15402	AU029803	HI0624	AGRE
GSE15402	AU080403	HI2039	AGRE
GSE15402	AU048103	HI1555	AGRE
GSE15402	AU1048301	HI2679	AGRE
GSE15402	AU067703	HI2008	AGRE
GSE15402	AU1007302	HI2706	AGRE
GSE15402	AU1164302	HI2883	AGRE
GSE15402	AU069604	HI1428	AGRE
GSE15402	AU1498302	HI4185	AGRE
GSE15402	AU0917302	HI2530	AGRE
GSE15402	AU1048302	HI2677	AGRE
GSE15402	AU1102301	HI2815	AGRE
GSE15402	AU045004	HI0570	AGRE
GSE15402	AU1494302	HI4196	AGRE
GSE15402	AU007605	HI0217	AGRE
GSE15402	AU005303	HI1102	AGRE
GSE15402	AU070808	HI1911	AGRE
GSE15402	AU069606	HI4281	AGRE
GSE15402	AU0903303	HI2442	AGRE
GSE15402	AU1346304	HI3987	AGRE
GSE15402	AU083605	HI2162	AGRE
GSE15402	AU1429304	HI4090	AGRE
GSE15402	AU059407	HI1534	AGRE

Continued. . .

Data Set	Individual ID	Sample ID	Source
GSE15402	AU007505	HI0742	AGRE
GSE15402	AU1135203	HI2725	AGRE
GSE15402	AU1331304	HI3304	AGRE
GSE15402	AU045012	HI1866	AGRE
GSE15402	AU1007301	HI2705	AGRE
GSE15402	AU1283302	HI3216	AGRE
GSE15402	AU020105	HI0507	AGRE
GSE15402	AU059406	HI1537	AGRE
GSE15402	AU059405	HI1535	AGRE
GSE15402	AU061405	HI1545	AGRE
GSE15402	AU1135201	HI2723	AGRE
GSE15402	AU053304	HI1788	AGRE
GSE15402	AU057903	HI0813	AGRE
GSE15402	AU001504	HI1693	AGRE
GSE15402	AU059403	HI1539	AGRE
GSE15402	AU019404	HI1047	AGRE
GSE15402	AU032503	HI0365	AGRE
GSE15402	AU077503	HI1706	AGRE
GSE15402	AU0885304	HI2357	AGRE
GSE15402	AU0885301	HI2356	AGRE
GSE15402	AU005605	HI0614	AGRE
GSE15402	AU062903	HI1161	AGRE
GSE6575	whole blood_100683	-	CHARGE
GSE6575	whole blood_101291	-	CHARGE
GSE6575	whole blood_100923	-	CHARGE
GSE6575	whole blood_100652	-	CHARGE
GSE6575	whole blood_101584	-	CHARGE
GSE6575	whole blood_100676	-	CHARGE
GSE6575	whole blood_100812	-	CHARGE
GSE6575	whole blood_101007	-	CHARGE
GSE6575	whole blood_101611	-	CHARGE
GSE6575	whole blood_101832	-	CHARGE
GSE6575	whole blood_101094	-	CHARGE
GSE6575	whole blood_100537	-	CHARGE
GSE6575	whole blood_101369	-	CHARGE
GSE6575	whole blood_100603	-	CHARGE
GSE6575	whole blood_100309	-	CHARGE
GSE6575	whole blood_100476	-	CHARGE
GSE6575	whole blood_100247	-	CHARGE
GSE6575	whole blood_101592	-	CHARGE
GSE6575	whole blood_101025	-	CHARGE
GSE6575	whole blood_100214	-	CHARGE
GSE6575	whole blood_101060	-	CHARGE
GSE6575	whole blood_101062	-	CHARGE

Continued. . .

Data Set	Individual ID	Sample ID	Source
GSE6575	whole blood_101209	-	CHARGE
GSE6575	whole blood_100823	-	CHARGE
GSE6575	whole blood_101234	-	CHARGE
GSE6575	whole blood_100716	-	CHARGE
GSE6575	whole blood_101700	-	CHARGE
GSE6575	whole blood_102321	-	CHARGE
GSE6575	whole blood_101669	-	CHARGE
GSE6575	whole blood_100158	-	CHARGE
GSE6575	whole blood_101682	-	CHARGE
GSE6575	whole blood_100278	-	CHARGE
GSE6575	whole blood_100416	-	CHARGE
GSE6575	whole blood_100981	-	CHARGE
GSE6575	whole blood_101339	-	CHARGE
GSE6575	whole blood_101644	-	CHARGE
GSE6575	whole blood_100354	-	CHARGE
GSE6575	whole blood_101074	-	CHARGE
GSE6575	whole blood_101467	-	CHARGE
GSE6575	whole blood_101726	-	CHARGE
GSE6575	whole blood_100409	-	CHARGE
GSE6575	whole blood_101794	-	CHARGE
GSE6575	whole blood_100374	-	CHARGE
GSE6575	whole blood_100190	-	CHARGE
GSE37772	12032.p1	-	SSC
GSE37772	11201.s1	-	SSC
GSE37772	12219.p1	-	SSC
GSE37772	12117.s1	-	SSC
GSE37772	12007.p1	-	SSC
GSE37772	12235.p1	-	SSC
GSE37772	12295.s1	-	SSC
GSE37772	12457.s1	-	SSC
GSE37772	11041.p1	-	SSC
GSE37772	11435.p1	-	SSC
GSE37772	11696.p1	-	SSC
GSE37772	11129.p1	-	SSC
GSE37772	11090.p1	-	SSC
GSE37772	11433.p1	-	SSC
GSE37772	12647.p1	-	SSC
GSE37772	12435.p1	-	SSC
GSE37772	12736.p1	-	SSC
GSE37772	12512.p1	-	SSC
GSE37772	12984.p1	-	SSC
GSE37772	12958.p1	-	SSC
GSE37772	13128.p1	-	SSC
GSE37772	13195.p1	-	SSC

Continued. . .

Data Set	Individual ID	Sample ID	Source
GSE37772	11387.p1	-	SSC
GSE37772	11343.p1	-	SSC
GSE37772	11809.p1	-	SSC
GSE37772	11080.p1	-	SSC
GSE37772	11689.p1	-	SSC
GSE37772	11962.p1	-	SSC
GSE37772	11962.s1	-	SSC
GSE37772	11824.s1	-	SSC
GSE37772	12239.p1	-	SSC
GSE37772	12339.p1	-	SSC
GSE37772	12383.p1	-	SSC
GSE37772	12343.p1	-	SSC
GSE37772	12351.s1	-	SSC
GSE37772	12685.p1	-	SSC
GSE37772	12451.p1	-	SSC
GSE37772	12581.p1	-	SSC
GSE37772	12297.p1	-	SSC
GSE37772	12299.s1	-	SSC
GSE37772	11532.p1	-	SSC
GSE37772	11334.p1	-	SSC
GSE37772	11135.p1	-	SSC
GSE37772	11390.p1	-	SSC
GSE37772	11629.p1	-	SSC
GSE37772	11680.p1	-	SSC
GSE37772	12100.p1	-	SSC
GSE37772	12308.p1	-	SSC
GSE37772	12417.p1	-	SSC
GSE37772	12594.p1	-	SSC
GSE37772	11868.p1	-	SSC
GSE37772	11868.s1	-	SSC
GSE37772	12014.p1	-	SSC
GSE37772	12014.s1	-	SSC
GSE37772	12523.p1	-	SSC
GSE37772	12523.s1	-	SSC
GSE37772	12603.p1	-	SSC
GSE37772	12603.s1	-	SSC
GSE37772	11610.p1	-	SSC
GSE37772	11610.s1	-	SSC
GSE37772	11407.p1	-	SSC
GSE37772	11407.s1	-	SSC
GSE37772	11879.p1	-	SSC
GSE37772	11736.p1	-	SSC
GSE37772	11356.s1	-	SSC
GSE37772	11066.p1	-	SSC

Continued. . .

Data Set	Individual ID	Sample ID	Source
GSE37772	11066.s1	-	SSC
GSE37772	11254.p1	-	SSC
GSE37772	11254.s1	-	SSC
GSE37772	11484.p1	-	SSC
GSE37772	11484.s1	-	SSC
GSE37772	11480.p1	-	SSC
GSE37772	12048.s1	-	SSC
GSE37772	12184.s1	-	SSC
GSE37772	12399.s1	-	SSC
GSE37772	12385.p1	-	SSC
GSE37772	12385.s1	-	SSC
GSE37772	12420.s1	-	SSC
GSE37772	12327.p1	-	SSC
GSE37772	12420.p1	-	SSC
GSE37772	12279.p1	-	SSC
GSE37772	11948.s1	-	SSC
GSE37772	12346.s1	-	SSC
GSE37772	12279.s1	-	SSC
GSE37772	12297.s1	-	SSC
GSE37772	12048.p1	-	SSC
GSE37772	12399.p1	-	SSC
GSE37772	11998.p1	-	SSC
GSE37772	11998.s1	-	SSC
GSE37772	12096.p1	-	SSC
GSE37772	12096.s1	-	SSC
GSE37772	12015.p1	-	SSC
GSE37772	12015.s1	-	SSC
GSE37772	11839.p1	-	SSC
GSE37772	11718.p1	-	SSC
GSE37772	12078.s1	-	SSC
GSE37772	12073.p1	-	SSC
GSE37772	12073.s1	-	SSC
GSE37772	11979.p1	-	SSC
GSE37772	11979.s1	-	SSC
GSE37772	11990.p1	-	SSC
GSE37772	11990.s1	-	SSC
GSE37772	12044.p1	-	SSC
GSE37772	11540.p1	-	SSC
GSE37772	11540.s1	-	SSC
GSE37772	11267.p1	-	SSC
GSE37772	11274.p1	-	SSC
GSE37772	11274.s1	-	SSC
GSE37772	11247.p1	-	SSC
GSE37772	11247.s1	-	SSC

Continued...

Data Set	Individual ID	Sample ID	Source
GSE37772	11356.p1	-	SSC
GSE37772	11625.s1	-	SSC
GSE37772	11947.s1	-	SSC
GSE37772	11177.p1	-	SSC
GSE37772	11333.p1	-	SSC
GSE37772	11333.s1	-	SSC
GSE37772	11495.p1	-	SSC
GSE37772	11495.s1	-	SSC
GSE37772	11267.s1	-	SSC
GSE37772	12261.s1	-	SSC
GSE37772	12044.s1	-	SSC
GSE37772	12241.p1	-	SSC
GSE37772	12241.s1	-	SSC
GSE37772	12184.p1	-	SSC
GSE37772	11948.p1	-	SSC
GSE37772	12224.p1	-	SSC
GSE37772	12261.p1	-	SSC
GSE37772	11839.s1	-	SSC
GSE37772	11718.s1	-	SSC
GSE37772	11831.p1	-	SSC
GSE37772	11831.s1	-	SSC
GSE37772	12083.p1	-	SSC
GSE37772	12083.s1	-	SSC
GSE37772	12078.p1	-	SSC
GSE37772	11276.p1	-	SSC
GSE37772	11276.s1	-	SSC
GSE37772	11765.p1	-	SSC
GSE37772	11765.s1	-	SSC
GSE37772	11723.p1	-	SSC
GSE37772	11723.s1	-	SSC
GSE37772	11731.p1	-	SSC
GSE37772	11731.s1	-	SSC
GSE37772	11736.s1	-	SSC
GSE37772	11879.s1	-	SSC
GSE37772	11393.p1	-	SSC
GSE37772	11393.s1	-	SSC
GSE37772	11857.p1	-	SSC
GSE37772	11857.s1	-	SSC
GSE37772	11634.p1	-	SSC
GSE37772	11634.s1	-	SSC
GSE37772	11006.p1	-	SSC
GSE37772	11006.s1	-	SSC
GSE37772	11554.p1	-	SSC
GSE37772	11554.s1	-	SSC

Continued. . .

Data Set	Individual ID	Sample ID	Source
GSE37772	11342.p1	-	SSC
GSE37772	11342.s1	-	SSC
GSE37772	11113.s1	-	SSC
GSE37772	11113.p1	-	SSC
GSE37772	11220.s1	-	SSC
GSE37772	11114.s1	-	SSC
GSE37772	11220.p1	-	SSC
GSE37772	11114.p1	-	SSC
GSE37772	11489.s1	-	SSC
GSE37772	11482.p1	-	SSC
GSE37772	11076.s1	-	SSC
GSE37772	11076.p1	-	SSC
GSE37772	11089.p1	-	SSC
GSE37772	11550.s1	-	SSC
GSE37772	11156.p1	-	SSC
GSE37772	11399.p1	-	SSC
GSE37772	11323.p1	-	SSC
GSE37772	11284.s1	-	SSC
GSE37772	11325.s1	-	SSC
GSE37772	11284.p1	-	SSC
GSE37772	11325.p1	-	SSC
GSE37772	11193.p1	-	SSC
GSE37772	11193.s1	-	SSC
GSE37772	11428.s1	-	SSC
GSE37772	11328.s1	-	SSC
GSE37772	11046.s1	-	SSC
GSE37772	11502.p1	-	SSC
GSE37772	11091.s1	-	SSC
GSE37772	11208.s1	-	SSC
GSE37772	11004.s1	-	SSC
GSE37772	11502.s1	-	SSC
GSE37772	11376.p1	-	SSC
GSE37772	11265.p1	-	SSC
GSE37772	11376.s1	-	SSC
GSE37772	11291.p1	-	SSC
GSE37772	11285.s1	-	SSC
GSE37772	11197.p1	-	SSC
GSE37772	11479.p1	-	SSC
GSE37772	11411.s1	-	SSC
GSE37772	11216.p1	-	SSC
GSE37772	11216.s1	-	SSC
GSE37772	11057.p1	-	SSC
GSE37772	11005.p1	-	SSC
GSE37772	11499.p1	-	SSC

Continued. . .

Data Set	Individual ID	Sample ID	Source
GSE37772	11316.s1	-	SSC
GSE37772	11316.p1	-	SSC
GSE37772	11053.s1	-	SSC
GSE37772	11406.s1	-	SSC
GSE37772	11053.p1	-	SSC
GSE37772	11519.s1	-	SSC
GSE37772	11207.p1	-	SSC
GSE37772	11207.s1	-	SSC
GSE37772	11300.p1	-	SSC
GSE37772	11300.s1	-	SSC
GSE37772	11501.p1	-	SSC
GSE37772	11501.s1	-	SSC
GSE37772	11192.s1	-	SSC
GSE37772	11102.p1	-	SSC
GSE37772	11083.p1	-	SSC
GSE37772	11301.p1	-	SSC
GSE37772	11219.s1	-	SSC
GSE37772	11551.s1	-	SSC
GSE37772	11219.p1	-	SSC
GSE37772	11551.p1	-	SSC
GSE37772	11191.p1	-	SSC
GSE37772	11180.s1	-	SSC
GSE37772	11578.p1	-	SSC
GSE37772	11062.p1	-	SSC
GSE37772	11510.s1	-	SSC
GSE37772	11062.s1	-	SSC
GSE37772	11469.p1	-	SSC
GSE37772	11579.p1	-	SSC
GSE37772	11293.p1	-	SSC
GSE37772	11293.s1	-	SSC
GSE37772	11071.s1	-	SSC
GSE37772	11071.p1	-	SSC
GSE37772	11073.p1	-	SSC
GSE37772	11073.s1	-	SSC
GSE37772	11186.p1	-	SSC
GSE37772	11186.s1	-	SSC
GSE37772	11509.p1	-	SSC
GSE37772	11509.s1	-	SSC
GSE37772	11452.p1	-	SSC
GSE37772	11452.s1	-	SSC
GSE37772	11059.p1	-	SSC
GSE37772	11459.p1	-	SSC
GSE37772	11285.p1	-	SSC
GSE37772	11075.p1	-	SSC

Continued. . .

Data Set	Individual ID	Sample ID	Source
GSE37772	11557.s1	-	SSC
GSE37772	11075.s1	-	SSC
GSE37772	11546.p1	-	SSC
GSE37772	11546.s1	-	SSC
GSE37772	11577.p1	-	SSC
GSE37772	11577.s1	-	SSC
GSE37772	11417.p1	-	SSC
GSE37772	11411.p1	-	SSC
GSE37772	11197.s1	-	SSC
GSE37772	11625.p1	-	SSC
GSE37772	11233.p1	-	SSC
GSE37772	11461.s1	-	SSC
GSE37772	11418.s1	-	SSC
GSE37772	11345.s1	-	SSC
GSE37772	11466.s1	-	SSC
GSE37772	11466.p1	-	SSC
GSE37772	11178.s1	-	SSC
GSE37772	11178.p1	-	SSC
GSE37772	11244.s1	-	SSC
GSE37772	11244.p1	-	SSC
GSE37772	11189.s1	-	SSC
GSE37772	11189.p1	-	SSC
GSE37772	11420.s1	-	SSC
GSE37772	11420.p1	-	SSC
GSE37772	11424.s1	-	SSC
GSE37772	11424.p1	-	SSC
GSE37772	11410.s1	-	SSC
GSE37772	11555.p1	-	SSC
GSE37772	11511.p1	-	SSC
GSE37772	11511.s1	-	SSC
GSE37772	11329.p1	-	SSC
GSE37772	11098.p1	-	SSC
GSE37772	11154.p1	-	SSC
GSE37772	11519.p1	-	SSC
GSE37772	11007.s1	-	SSC
GSE37772	11457.p1	-	SSC
GSE37772	11482.s1	-	SSC
GSE37772	11457.s1	-	SSC
GSE37772	11327.p1	-	SSC
GSE37772	11327.s1	-	SSC
GSE37772	11489.p1	-	SSC
GSE37772	11490.s1	-	SSC
GSE37772	11102.s1	-	SSC
GSE37772	11146.s1	-	SSC

Continued. . .

Data Set	Individual ID	Sample ID	Source
GSE37772	11474.p1	-	SSC
GSE37772	11474.s1	-	SSC
GSE37772	11523.p1	-	SSC
GSE37772	11191.s1	-	SSC
GSE37772	11537.s1	-	SSC
GSE37772	11014.p1	-	SSC
GSE37772	11417.s1	-	SSC
GSE37772	11479.s1	-	SSC
GSE37772	11328.p1	-	SSC
GSE37772	11057.s1	-	SSC
GSE37772	11085.s1	-	SSC
GSE37772	11089.s1	-	SSC
GSE37772	11152.p1	-	SSC
GSE37772	11418.p1	-	SSC
GSE37772	11450.p1	-	SSC
GSE37772	11260.p1	-	SSC
GSE37772	11455.s1	-	SSC
GSE37772	11412.s1	-	SSC
GSE37772	11581.s1	-	SSC
GSE37772	11412.p1	-	SSC
GSE37772	11581.p1	-	SSC
GSE37772	11425.p1	-	SSC
GSE37772	11563.s1	-	SSC
GSE37772	11425.s1	-	SSC
GSE37772	11555.s1	-	SSC
GSE37772	11329.s1	-	SSC
GSE37772	11364.p1	-	SSC
GSE37772	11032.p1	-	SSC
GSE37772	11098.s1	-	SSC
GSE37772	11572.p1	-	SSC
GSE37772	11520.s1	-	SSC
GSE37772	11475.s1	-	SSC
GSE37772	11242.s1	-	SSC
GSE37772	11378.s1	-	SSC
GSE37772	11520.p1	-	SSC
GSE37772	11475.p1	-	SSC
GSE37772	11242.p1	-	SSC
GSE37772	11154.s1	-	SSC
GSE37772	11364.s1	-	SSC
GSE37772	11415.p1	-	SSC
GSE37772	11415.s1	-	SSC
GSE37772	11121.s1	-	SSC
GSE37772	11353.p1	-	SSC
GSE37772	11353.s1	-	SSC

Continued. . .

Data Set	Individual ID	Sample ID	Source
GSE37772	11000.p1	-	SSC
GSE37772	11579.s1	-	SSC
GSE37772	11233.s1	-	SSC
GSE37772	11473.s1	-	SSC
GSE37772	11445.p1	-	SSC
GSE37772	11445.s1	-	SSC
GSE37772	11469.s1	-	SSC
GSE37772	11007.p1	-	SSC
GSE37772	11480.s1	-	SSC
GSE37772	11168.s1	-	SSC
GSE37772	11030.p1	-	SSC
GSE37772	11030.s1	-	SSC
GSE37772	11149.p1	-	SSC
GSE37772	11149.s1	-	SSC
GSE37772	11382.p1	-	SSC
GSE37772	11382.s1	-	SSC
GSE37772	11265.s1	-	SSC
GSE37772	11028.p1	-	SSC
GSE37772	11046.p1	-	SSC
GSE37772	11383.s1	-	SSC
GSE37772	11383.p1	-	SSC
GSE37772	11083.s1	-	SSC
GSE37772	11303.s1	-	SSC
GSE37772	11301.s1	-	SSC
GSE37772	11410.p1	-	SSC
GSE37772	11029.s1	-	SSC
GSE37772	11029.p1	-	SSC
GSE37772	11121.p1	-	SSC
GSE37772	11291.s1	-	SSC
GSE37772	11429.p1	-	SSC
GSE37772	11572.s1	-	SSC
GSE37772	11406.p1	-	SSC
GSE37772	11180.p1	-	SSC
GSE37772	11335.p1	-	SSC
GSE37772	11461.p1	-	SSC
GSE37772	11533.s1	-	SSC
GSE37772	11337.s1	-	SSC
GSE37772	11499.s1	-	SSC
GSE37772	11458.s1	-	SSC
GSE37772	11458.p1	-	SSC
GSE37772	11533.p1	-	SSC
GSE37772	11345.p1	-	SSC
GSE37772	11379.s1	-	SSC
GSE37772	11490.p1	-	SSC

Continued. . .

Data Set	Individual ID	Sample ID	Source
GSE37772	11379.p1	-	SSC
GSE37772	11450.s1	-	SSC
GSE37772	11032.s1	-	SSC
GSE37772	11000.s1	-	SSC
GSE37772	11059.s1	-	SSC
GSE37772	11459.s1	-	SSC
GSE37772	11303.p1	-	SSC
GSE37772	11378.p1	-	SSC
GSE37772	11563.p1	-	SSC
GSE37772	11413.s1	-	SSC
GSE37772	11413.p1	-	SSC
GSE37772	11524.s1	-	SSC
GSE37772	11427.s1	-	SSC
GSE37772	11005.s1	-	SSC
GSE37772	11335.s1	-	SSC
GSE37772	11524.p1	-	SSC
GSE37772	11427.p1	-	SSC
GSE37772	11348.s1	-	SSC
GSE37772	11578.s1	-	SSC
GSE37772	11473.p1	-	SSC
GSE37772	11177.s1	-	SSC
GSE37772	11192.p1	-	SSC
GSE37772	11146.p1	-	SSC
GSE37772	11523.s1	-	SSC
GSE37772	11063.p1	-	SSC
GSE37772	11338.p1	-	SSC
GSE37772	11338.s1	-	SSC
GSE37772	11271.p1	-	SSC
GSE37772	11271.s1	-	SSC
GSE37772	11260.s1	-	SSC
GSE37772	11138.s1	-	SSC
GSE37772	11138.p1	-	SSC
GSE37772	11168.p1	-	SSC
GSE37772	11063.s1	-	SSC
GSE37772	11587.p1	-	SSC
GSE37772	11537.p1	-	SSC
GSE37772	11587.s1	-	SSC
GSE37772	11348.p1	-	SSC
GSE37772	11014.s1	-	SSC
GSE37772	11337.p1	-	SSC
GSE37772	11443.p1	-	SSC
GSE37772	11156.s1	-	SSC
GSE37772	11399.s1	-	SSC
GSE37772	11323.s1	-	SSC

Continued...

Data Set	Individual ID	Sample ID	Source
GSE37772	11091.p1	-	SSC
GSE37772	11208.p1	-	SSC
GSE37772	11428.p1	-	SSC
GSE37772	11004.p1	-	SSC
GSE37772	11947.p1	-	SSC
GSE37772	11442.p1	-	SSC
GSE37772	11443.s1	-	SSC
GSE37772	11442.s1	-	SSC
GSE37772	11275.p1	-	SSC
GSE37772	11426.p1	-	SSC
GSE37772	11275.s1	-	SSC
GSE37772	11426.s1	-	SSC
GSE37772	11510.p1	-	SSC
GSE25507	0118-01-C	-	Phoenix
GSE25507	0120-01-C	-	Phoenix
GSE25507	0137-01-C	-	Phoenix
GSE25507	0147-01-C	-	Phoenix
GSE25507	0148-01-C	-	Phoenix
GSE25507	0152-01-C	-	Phoenix
GSE25507	0153-01-C	-	Phoenix
GSE25507	0155-01-c	-	Phoenix
GSE25507	0156-01-C	-	Phoenix
GSE25507	0164-01-c	-	Phoenix
GSE25507	0165-01-C	-	Phoenix
GSE25507	0167-01-C	-	Phoenix
GSE25507	0170-01-C	-	Phoenix
GSE25507	0175-01-C	-	Phoenix
GSE25507	0178-01-C	-	Phoenix
GSE25507	0180-01-C	-	Phoenix
GSE25507	0184-01-1	-	Phoenix
GSE25507	0191-01-C	-	Phoenix
GSE25507	0192-01-C	-	Phoenix
GSE25507	0193-01-C	-	Phoenix
GSE25507	0195-01-C	-	Phoenix
GSE25507	0196-01-C	-	Phoenix
GSE25507	0198-01-c	-	Phoenix
GSE25507	0201-01-C-RR	-	Phoenix
GSE25507	0205-01-C	-	Phoenix
GSE25507	0466-01-C	-	Phoenix
GSE25507	0510-01-C	-	Phoenix
GSE25507	0511-01-c	-	Phoenix
GSE25507	0514-01-C	-	Phoenix
GSE25507	0515-01-C	-	Phoenix
GSE25507	0516-01-C	-	Phoenix

Continued. . .

Data Set	Individual ID	Sample ID	Source
GSE25507	0523-01-C	-	Phoenix
GSE25507	0536-01-C	-	Phoenix
GSE25507	0537-01-C	-	Phoenix
GSE25507	0651-01-C	-	Phoenix
GSE25507	0652-01-C	-	Phoenix
GSE25507	0653-01-C	-	Phoenix
GSE25507	0671-01-C	-	Phoenix
GSE25507	0702-01-C-b	-	Phoenix
GSE25507	0966-01-C	-	Phoenix
GSE25507	0967-01-C	-	Phoenix
GSE25507	0968-01-C	-	Phoenix
GSE25507	0970-01-C	-	Phoenix
GSE25507	0972-01-C	-	Phoenix
GSE25507	0974-01-C	-	Phoenix
GSE25507	0975-01-C	-	Phoenix
GSE25507	0977-01-C	-	Phoenix
GSE25507	0978-01-C	-	Phoenix
GSE25507	0979-01-C	-	Phoenix
GSE25507	0982-01-c	-	Phoenix
GSE25507	0983-01-C	-	Phoenix
GSE25507	0986-01-C	-	Phoenix
GSE25507	0989-01-C	-	Phoenix
GSE25507	0990-01-C	-	Phoenix
GSE25507	0991-01-C	-	Phoenix
GSE25507	0992-01-c	-	Phoenix
GSE25507	0993-01-C	-	Phoenix
GSE25507	0998-01-C	-	Phoenix
GSE25507	1118-01-C	-	Phoenix
GSE25507	121-01-C	-	Phoenix
GSE25507	166-01-C	-	Phoenix
GSE25507	SARRC-0146-01-C	-	Phoenix
GSE25507	SARRC-0994-01-C	-	Phoenix
GSE25507	0105-01-A	-	Phoenix
GSE25507	0151-01-C	-	Phoenix
GSE25507	0199-01-C	-	Phoenix
GSE25507	0238-01-A	-	Phoenix
GSE25507	0301-01-A	-	Phoenix
GSE25507	0561-01-A	-	Phoenix
GSE25507	0562-01-A	-	Phoenix
GSE25507	0577-01-A	-	Phoenix
GSE25507	0705-01-A	-	Phoenix
GSE25507	0711-04-A	-	Phoenix
GSE25507	0790-01-A	-	Phoenix
GSE25507	0810-01-A	-	Phoenix

Continued. . .

Data Set	Individual ID	Sample ID	Source
GSE25507	0820-01-A	-	Phoenix
GSE25507	0833-01-C	-	Phoenix
GSE25507	0844-01-A	-	Phoenix
GSE25507	0922-01-A	-	Phoenix
GSE25507	0943-01-A	-	Phoenix
GSE25507	0947-01-A-b	-	Phoenix
GSE25507	0949-01-A	-	Phoenix
GSE25507	0950-01-A	-	Phoenix
GSE25507	0959-01-A	-	Phoenix
GSE25507	1005-01-A	-	Phoenix
GSE25507	1006-03-A	-	Phoenix
GSE25507	1009-03-A	-	Phoenix
GSE25507	1011-03-A-b	-	Phoenix
GSE25507	1015-01-A	-	Phoenix
GSE25507	1023-01-A-b	-	Phoenix
GSE25507	1030-03-A-b	-	Phoenix
GSE25507	1044-03-A	-	Phoenix
GSE25507	1049-03-A	-	Phoenix
GSE25507	1051-03-A-b	-	Phoenix
GSE25507	1052-03-A	-	Phoenix
GSE25507	1053-03-A	-	Phoenix
GSE25507	1056-03-A	-	Phoenix
GSE25507	1056-05-A	-	Phoenix
GSE25507	1061-03-A-b	-	Phoenix
GSE25507	1064-03-A	-	Phoenix
GSE25507	1065-03-C	-	Phoenix
GSE25507	1067-03-A	-	Phoenix
GSE25507	1076-03-A	-	Phoenix
GSE25507	1082-01-C	-	Phoenix
GSE25507	1083-03-A	-	Phoenix
GSE25507	1099-03-A	-	Phoenix
GSE25507	1106-03-A	-	Phoenix
GSE25507	1116-03-A	-	Phoenix
GSE25507	1119-03-A	-	Phoenix
GSE25507	1122-03-A	-	Phoenix
GSE25507	1124-02-A	-	Phoenix
GSE25507	1132-03-A	-	Phoenix
GSE25507	1138-03-A-b	-	Phoenix
GSE25507	1156-03-A	-	Phoenix
GSE25507	1163-03-A	-	Phoenix
GSE25507	1164-03-A	-	Phoenix
GSE25507	1186-03-A-b	-	Phoenix
GSE25507	1193-03-A-b	-	Phoenix
GSE25507	1194-03-A	-	Phoenix

Continued. . .

Data Set	Individual ID	Sample ID	Source
GSE25507	1195-03-A	-	Phoenix
GSE25507	1197-03-A	-	Phoenix
GSE25507	1198-03-A	-	Phoenix
GSE25507	1199-01-A	-	Phoenix
GSE25507	1200-03-A	-	Phoenix
GSE25507	1202-03-A	-	Phoenix
GSE25507	1203-03-A	-	Phoenix
GSE25507	1204-03-A-b	-	Phoenix
GSE25507	1205-01-A	-	Phoenix
GSE25507	1209-03-A	-	Phoenix
GSE25507	2095-03-A	-	Phoenix
GSE25507	2096-03-A	-	Phoenix
GSE25507	2097-03-A	-	Phoenix
GSE25507	2098-03-A	-	Phoenix
GSE25507	2103-03-A	-	Phoenix
GSE25507	2107-03-A	-	Phoenix
GSE25507	2112-01-A	-	Phoenix
GSE25507	2117-01-A	-	Phoenix
GSE25507	2119-03-A	-	Phoenix
GSE25507	2120-03-A	-	Phoenix
GSE25507	SARRC-0248-01-A	-	Phoenix
GSE25507	SARRC-1021-01-A	-	Phoenix
GSE25507	SARRC-1100-03-Aq	-	Phoenix
GSE25507	SARRC-2089-02-A	-	Phoenix
GSE18123.1	A-0006-P1	-	Boston
GSE18123.1	A-0008-P1	-	Boston
GSE18123.1	A-0010-P1	-	Boston
GSE18123.1	A-0016-P1	-	Boston
GSE18123.1	A-0021-P1	-	Boston
GSE18123.1	A-0022-P1	-	Boston
GSE18123.1	A-0023-P1	-	Boston
GSE18123.1	A-0025-P1	-	Boston
GSE18123.1	A-0028-P1	-	Boston
GSE18123.1	A-0030-P1	-	Boston
GSE18123.1	A-0031-P1	-	Boston
GSE18123.1	A-0032-P1	-	Boston
GSE18123.1	A-0034-P1	-	Boston
GSE18123.1	A-0041-P1	-	Boston
GSE18123.1	A-0042-P1	-	Boston
GSE18123.1	A-0044-P1	-	Boston
GSE18123.1	A-0050-P1	-	Boston
GSE18123.1	A-0052-P1	-	Boston
GSE18123.1	A-0053-P1	-	Boston
GSE18123.1	A-0055-A1	-	Boston

Continued. . .

Data Set	Individual ID	Sample ID	Source
GSE18123.1	A-0057-P1	-	Boston
GSE18123.1	A-0058-P1	-	Boston
GSE18123.1	A-0065-P1	-	Boston
GSE18123.1	A-0068-P1	-	Boston
GSE18123.1	A-0070-P1	-	Boston
GSE18123.1	A-0072-P1	-	Boston
GSE18123.1	A-0074-A1	-	Boston
GSE18123.1	A-0076-P1	-	Boston
GSE18123.1	A-0078-P1	-	Boston
GSE18123.1	A-0086-P1	-	Boston
GSE18123.1	A-0088-P1	-	Boston
GSE18123.1	A-0090-P1	-	Boston
GSE18123.1	A-0091-P1	-	Boston
GSE18123.1	A-0094-P1	-	Boston
GSE18123.1	A-0095-P1	-	Boston
GSE18123.1	A-0096-P1	-	Boston
GSE18123.1	A-0101-P1	-	Boston
GSE18123.1	A-0105-P1	-	Boston
GSE18123.1	A-0108-P1	-	Boston
GSE18123.1	A-0111-P1	-	Boston
GSE18123.1	A-0116-P1	-	Boston
GSE18123.1	A-0119-P1	-	Boston
GSE18123.1	A-0120-P1	-	Boston
GSE18123.1	A-0126-P1	-	Boston
GSE18123.1	A-0127-P1	-	Boston
GSE18123.1	A-0128-P1	-	Boston
GSE18123.1	A-0131-P1	-	Boston
GSE18123.1	A-0132-P1	-	Boston
GSE18123.1	A-0133-P1	-	Boston
GSE18123.1	A-0134-P1	-	Boston
GSE18123.1	AC02-1003-01	-	Boston
GSE18123.1	AC02-1004-01	-	Boston
GSE18123.1	AC02-1006-01	-	Boston
GSE18123.1	AC02-1007-01	-	Boston
GSE18123.1	AC02-1009-01	-	Boston
GSE18123.1	AC02-1113-01	-	Boston
GSE18123.1	AC02-1140-01	-	Boston
GSE18123.1	AC02-1141-01	-	Boston
GSE18123.1	AC02-1172-01	-	Boston
GSE18123.1	AC02-1178-01	-	Boston
GSE18123.1	AC02-1183-01	-	Boston
GSE18123.1	AC02-1189-01	-	Boston
GSE18123.1	AC02-1197-01	-	Boston
GSE18123.1	AC02-1211-01	-	Boston

Continued. . .

Data Set	Individual ID	Sample ID	Source
GSE18123.1	M-02-141	-	Boston
GSE18123.1	M-03-040	-	Boston
GSE18123.1	M-03-049	-	Boston
GSE18123.1	M-03-058	-	Boston
GSE18123.1	M-03-075	-	Boston
GSE18123.1	M-03-168	-	Boston
GSE18123.1	M-03-169	-	Boston
GSE18123.1	M-04-006	-	Boston
GSE18123.1	M-04-022	-	Boston
GSE18123.1	M-04-035	-	Boston
GSE18123.1	M-04-065	-	Boston
GSE18123.1	M-06-103	-	Boston
GSE18123.1	M-09-090	-	Boston
GSE18123.1	M-11-003	-	Boston
GSE18123.1	M-11-155	-	Boston
GSE18123.1	M-12-052	-	Boston
GSE18123.1	M-12-060	-	Boston
GSE18123.1	M-12-070	-	Boston
GSE18123.1	M-12-102	-	Boston
GSE18123.1	M-12-152	-	Boston
GSE18123.1	M-13-005	-	Boston
GSE18123.1	M-13-012	-	Boston
GSE18123.1	M-13-032	-	Boston
GSE18123.1	M-13-087	-	Boston
GSE18123.1	M-13-116	-	Boston
GSE18123.1	M-15-023	-	Boston
GSE18123.1	M-15-030	-	Boston
GSE18123.1	M-16-018	-	Boston
GSE18123.2	AC02-1253-01	-	Boston
GSE18123.2	AC02-1288-01	-	Boston
GSE18123.2	AC02-1289-01	-	Boston
GSE18123.2	AC02-1297-01	-	Boston
GSE18123.2	AC02-1302-01	-	Boston
GSE18123.2	AC02-1302-04	-	Boston
GSE18123.2	AC02-1317-01	-	Boston
GSE18123.2	AC02-1324-04	-	Boston
GSE18123.2	AC02-1333-01	-	Boston
GSE18123.2	AC02-1344-01	-	Boston
GSE18123.2	AC02-1351-01	-	Boston
GSE18123.2	AC02-1354-01	-	Boston
GSE18123.2	AC02-1293-01	-	Boston
GSE18123.2	AC02-1318-04	-	Boston
GSE18123.2	AC02-1336-04	-	Boston
GSE18123.2	A-0071-P1	-	Boston

Continued. . .

Data Set	Individual ID	Sample ID	Source
GSE18123.2	A-0074-P2	-	Boston
GSE18123.2	A-0075-P1	-	Boston
GSE18123.2	A-0076-A1	-	Boston
GSE18123.2	A-0077-P1	-	Boston
GSE18123.2	A-0081-P1	-	Boston
GSE18123.2	A-0082-P1	-	Boston
GSE18123.2	A-0085-P2	-	Boston
GSE18123.2	A-0098-P1	-	Boston
GSE18123.2	A-0103-P1	-	Boston
GSE18123.2	A-0114-P2	-	Boston
GSE18123.2	A-0119-A1	-	Boston
GSE18123.2	A-0119-A2	-	Boston
GSE18123.2	A-0125-A1	-	Boston
GSE18123.2	A-0125-A2	-	Boston
GSE18123.2	AC01-0004-01	-	Boston
GSE18123.2	AC01-0037-01	-	Boston
GSE18123.2	AC01-0061-01	-	Boston
GSE18123.2	AC01-0076-01	-	Boston
GSE18123.2	AC02-1007-04	-	Boston
GSE18123.2	AC02-1008-01	-	Boston
GSE18123.2	AC02-1115-01	-	Boston
GSE18123.2	AC02-1153-01	-	Boston
GSE18123.2	AC02-1168-01	-	Boston
GSE18123.2	AC02-1168-05	-	Boston
GSE18123.2	AC02-1181-01	-	Boston
GSE18123.2	AC02-1189-01	-	Boston
GSE18123.2	AC02-1198-01	-	Boston
GSE18123.2	AC02-1201-01	-	Boston
GSE18123.2	AC02-1206-01	-	Boston
GSE18123.2	AC02-1215-01	-	Boston
GSE18123.2	AC02-1217-01	-	Boston
GSE18123.2	AC02-1217-04	-	Boston
GSE18123.2	AC02-1224-01	-	Boston
GSE18123.2	AC02-1225-01	-	Boston
GSE18123.2	AC02-1226-01	-	Boston
GSE18123.2	AC02-1247-01	-	Boston
GSE18123.2	AC02-1274-01	-	Boston
GSE18123.2	AC02-1278-01	-	Boston
GSE18123.2	AC02-1278-04	-	Boston
GSE18123.2	AC03-1001-01	-	Boston
GSE18123.2	AC03-1002-01	-	Boston
GSE18123.2	AC03-1003-01	-	Boston
GSE18123.2	AC03-1009-01	-	Boston
GSE18123.2	AC04-0015-01	-	Boston

Continued. . .

Data Set	Individual ID	Sample ID	Source
GSE18123.2	AC04-0020-01	-	Boston
GSE18123.2	AC05-0002-01	-	Boston
GSE18123.2	AC05-0002-04	-	Boston
GSE18123.2	AC05-0070-01	-	Boston
GSE18123.2	AC05-0071-01	-	Boston
GSE18123.2	AC05-0078-01	-	Boston
GSE18123.2	AC05-1001-01	-	Boston
GSE18123.2	AC05-1003-04	-	Boston
GSE18123.2	AC05-1005-04	-	Boston
GSE18123.2	AC05-1009-01	-	Boston
GSE18123.2	AC05-1013-01	-	Boston
GSE18123.2	AC05-1013-04	-	Boston
GSE18123.2	F-03-046	-	Boston
GSE18123.2	F-03-066	-	Boston
GSE18123.2	F-04-081	-	Boston
GSE18123.2	F-04-162	-	Boston
GSE18123.2	F-04-172	-	Boston
GSE18123.2	F-04-202	-	Boston
GSE18123.2	F-04-207	-	Boston
GSE18123.2	F-05-188	-	Boston
GSE18123.2	F-05-239	-	Boston
GSE18123.2	F-06-098	-	Boston
GSE18123.2	F-09-071	-	Boston
GSE18123.2	F-11-093	-	Boston
GSE18123.2	F-11-192	-	Boston
GSE18123.2	F-11-214	-	Boston
GSE18123.2	F-12-174	-	Boston
GSE18123.2	F-12-241	-	Boston
GSE18123.2	F-13-101	-	Boston
GSE18123.2	F-13-112	-	Boston
GSE18123.2	F-15-059	-	Boston
GSE18123.2	F-16-115	-	Boston
GSE18123.2	F-16-149	-	Boston
GSE18123.2	M-02-147	-	Boston
GSE18123.2	M-03-175	-	Boston
GSE18123.2	M-03-203	-	Boston
GSE18123.2	M-03-230	-	Boston
GSE18123.2	M-04-201	-	Boston
GSE18123.2	M-04-240	-	Boston
GSE18123.2	M-06-217	-	Boston
GSE18123.2	M-07-194	-	Boston
GSE18123.2	M-12-148	-	Boston
GSE18123.2	M-12-186	-	Boston
GSE18123.2	M-13-157	-	Boston

Continued. . .

Data Set	Individual ID	Sample ID	Source
GSE18123.2	M-14-129	-	Boston
GSE18123.2	M-16-128	-	Boston
GSE18123.2	AC05-1047-01	-	Boston
GSE18123.2	AC01-1019-04	-	Boston
GSE18123.2	AC02-1253-04	-	Boston
GSE18123.2	AC02-1331-01	-	Boston
GSE18123.2	AC02-1334-01	-	Boston
GSE18123.2	AC02-1334-04	-	Boston
GSE18123.2	AC02-1360-01	-	Boston
GSE18123.2	AC02-1361-01	-	Boston
GSE18123.2	AC02-1364-01	-	Boston
GSE18123.2	AC02-1364-04	-	Boston
GSE18123.2	AC02-1370-01	-	Boston
GSE18123.2	AC02-1375-01	-	Boston
GSE18123.2	AC02-1378-01	-	Boston
GSE18123.2	AC02-1380-01	-	Boston
GSE18123.2	AC02-1382-01	-	Boston
GSE18123.2	AC02-1392-04	-	Boston
GSE18123.2	AC02-1336-05	-	Boston
GSE18123.2	AC02-1387-01	-	Boston
GSE18123.2	AC02-1387-04	-	Boston
GSE18123.2	A-0087-P2	-	Boston
GSE18123.2	A-0097-P2	-	Boston
GSE18123.2	AC01-0030-04	-	Boston
GSE18123.2	AC01-0069-01	-	Boston
GSE18123.2	AC02-1010-01	-	Boston
GSE18123.2	AC02-1173-01	-	Boston
GSE18123.2	AC02-1181-04	-	Boston
GSE18123.2	AC02-1184-01	-	Boston
GSE18123.2	AC02-1197-04	-	Boston
GSE18123.2	AC02-1211-04	-	Boston
GSE18123.2	AC02-1223-04	-	Boston
GSE18123.2	AC02-1238-04	-	Boston
GSE18123.2	AC02-1240-04	-	Boston
GSE18123.2	AC02-1248-05	-	Boston
GSE18123.2	AC02-1249-01	-	Boston
GSE18123.2	AC03-1009-04	-	Boston
GSE18123.2	AC05-0003-04	-	Boston
GSE18123.2	AC05-0023-04	-	Boston
GSE18123.2	AC05-1001-04	-	Boston
GSE18123.2	AC05-1004-04	-	Boston
GSE18123.2	AC05-1016-01	-	Boston
GSE18123.2	AC05-1024-01	-	Boston
GSE18123.2	AC05-1032-01	-	Boston

Continued. . .

Data Set	Individual ID	Sample ID	Source
GSE18123.2	AC05-1032-05	-	Boston
GSE18123.2	F-02-083	-	Boston
GSE18123.2	F-02-117	-	Boston
GSE18123.2	F-02-218	-	Boston
GSE18123.2	F-03-067	-	Boston
GSE18123.2	F-03-167	-	Boston
GSE18123.2	F-04-076	-	Boston
GSE18123.2	M-14-190	-	Boston

Supplementary Table 17: Blood samples included in the meta-analysis.

(a)

GSE28521	ASD	Control	Total
Number of samples	8	10	18
Age range (yrs)	5-39	16-56	
Gender (M:F)	6:2	All males	
PMI	6.75-43.25	4.75-28.92	
GSE38322	**ASD**	**Control**	**Total**
Number of samples	7	8	15
Age range (yrs)	2-39	1-60	
Gender (M:F)	All male	All male	
PMI	4.00-43.00	13.00-24.20	

(b)

	Differentially expressed genes at FDR<0.05	Up-regulated genes	Down-regulated genes
GSE28521	14	1	13
GSE38322	0	0	0

(c)

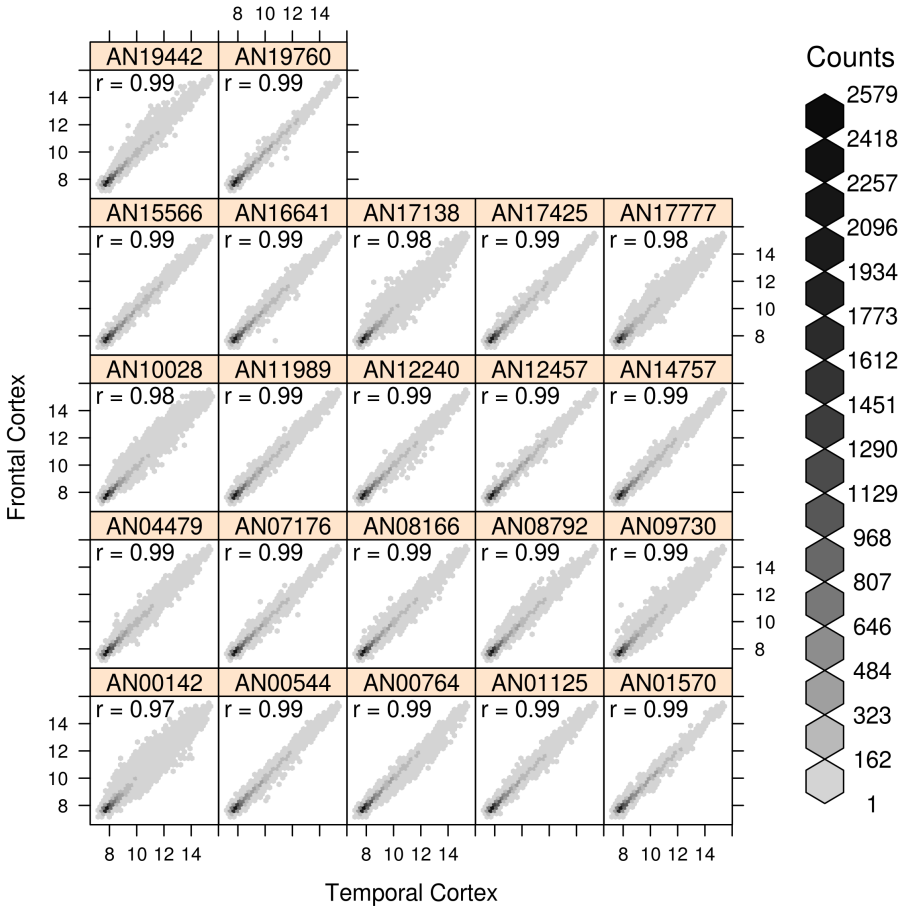
	GeneSymbol	Gene ID	p-value	FDR	Map loci
Up					
	LY6H	4062	2.63e-06	4.34e-02	8q24.3
Down					
	SGCA	6442	2.73e-06	4.51e-02	17q21
	ENTPD6	955	8.07e-06	4.56e-02	20p11.21
	COL11A2	1302	8.98e-06	4.56e-02	6p21.3
	KIAA0513	9764	1.19e-05	4.56e-02	16q24.1
	MEG3	55384	1.69e-05	4.56e-02	14q32
	CPLX3	594855	2.55e-05	4.56e-02	15q24.1
	FNDC9	408263	2.64e-05	4.56e-02	5q33.3
	MAP3K6	9064	2.80e-05	4.56e-02	1p36.11
	BRICD5	283870	2.84e-05	4.56e-02	16p13.3
	MORN1	79906	2.86e-05	4.56e-02	1p36.33-p36.32
	ULK3	25989	3.04e-05	4.56e-02	15q24.1
	SCN4B	6330	3.43e-05	4.71e-02	11q23.3
	SLC7A8	23428	3.75e-05	4.76e-02	14q11.2

Supplementary Table 18: a) Study design for samples from the cerebellum in GSE38322 and GSE28521. Samples that failed quality control tests were removed. We also collapsed technical replicates by obtaining the average expression of samples from the same individual.; b) Number of differentially expressed genes identified in the cerebellum. Sex-linked genes were removed from the GSE28521's results due to an imbalanced gender ratio in the study design; c) Differentially expressed genes in GSE28521.

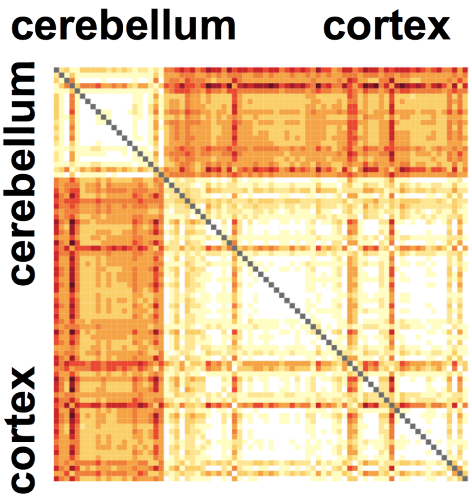
Supplementary Figures

List of Supplementary Figures

(a)

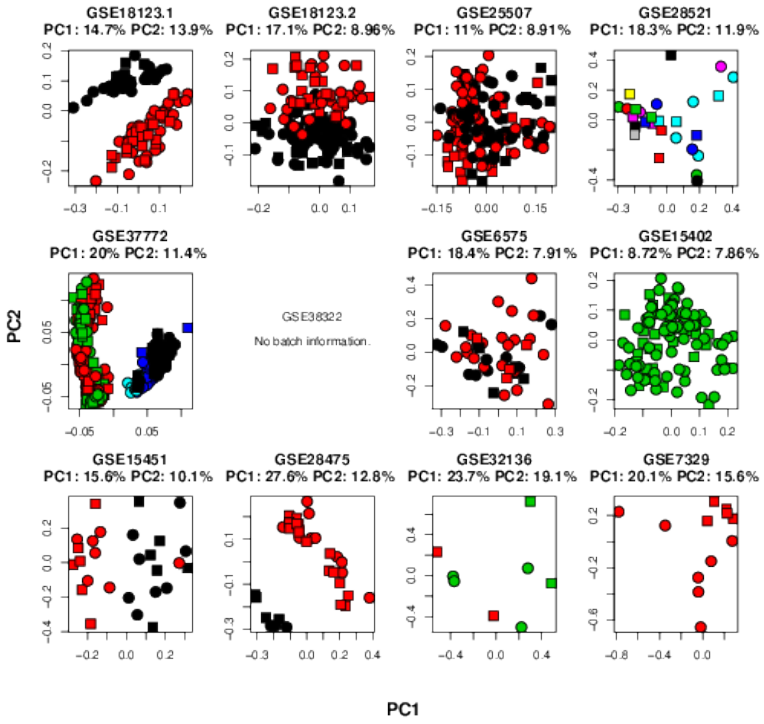


(b)

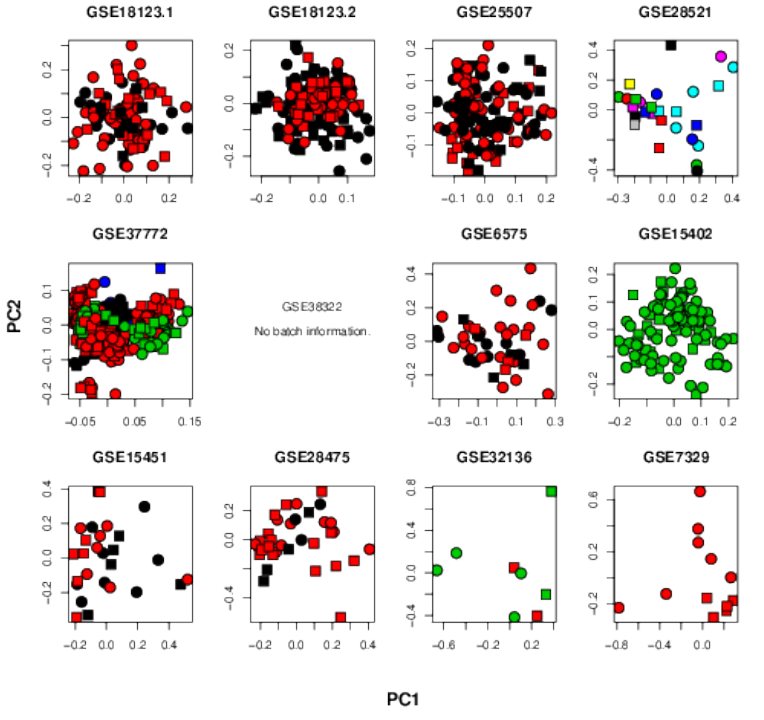


Supplementary Figure 1: Quality control of brain samples: a) Samples obtained from the temporal cortex and frontal cortex of the same individual exhibit highly correlated expression values. Data from GSE28521 shown here; b) Expression profiles of samples from the cerebellum differ from that of the cortex, as seen in the sample correlation matrix of GSE28521.

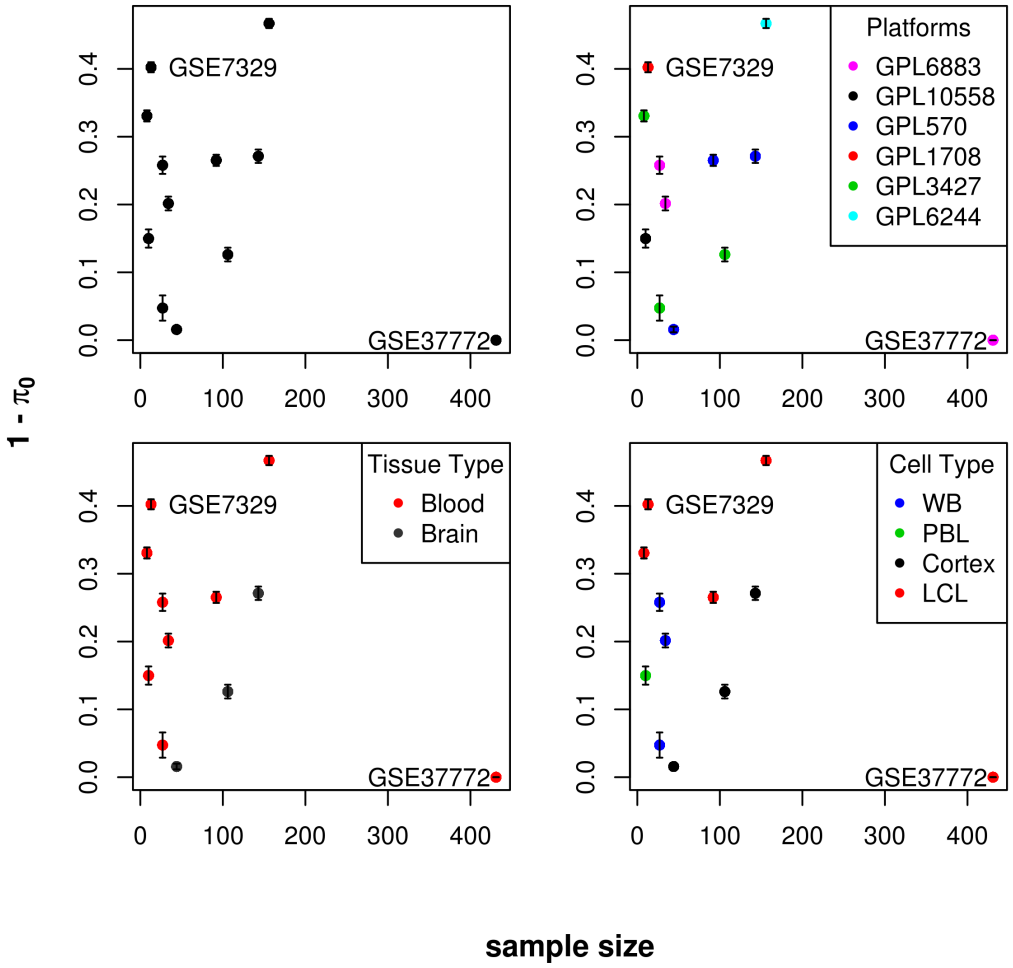
(a)



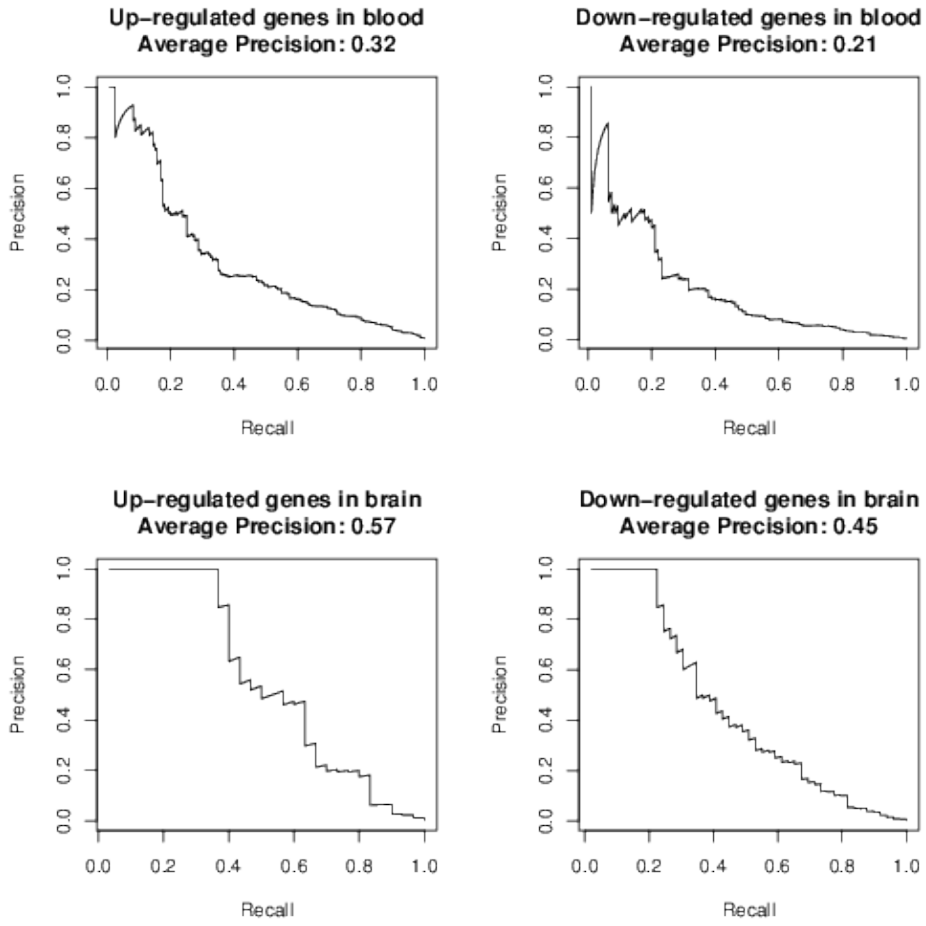
(b)



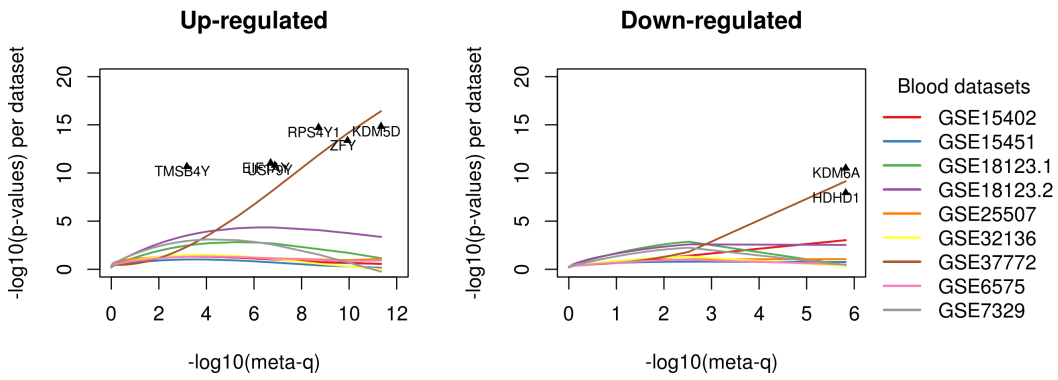
Supplementary Figure 2: Batch effects: a) Clustering of datapoints into distinct batches with respective percentage variances, suggesting the presence of batch effects. b) Batch effects were removed after batch correction (includes robust probes only).



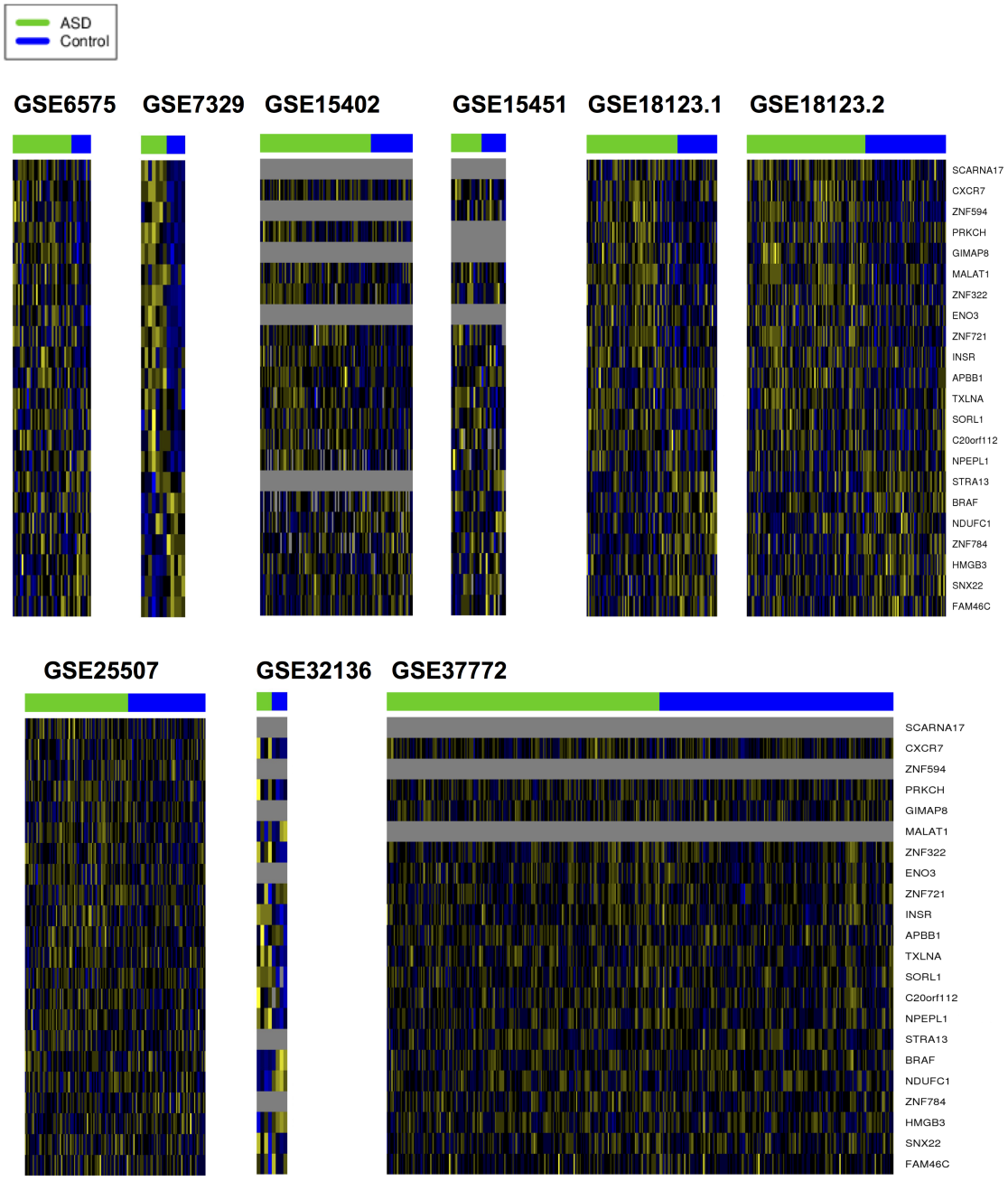
Supplementary Figure 3: $1 - \pi_0$ values for each study against sample size. Error bars denote standard errors for 100 bootstrap iterations. PBL: peripheral blood lymphocytes; LCL: lymphoblastoid cell lines; WB: whole blood.



Supplementary Figure 4: Comparison between results obtained from the Fisher's method and the Meta-Rank method. The peaks on the left suggests that genes ranked at the top are similar for both methods.

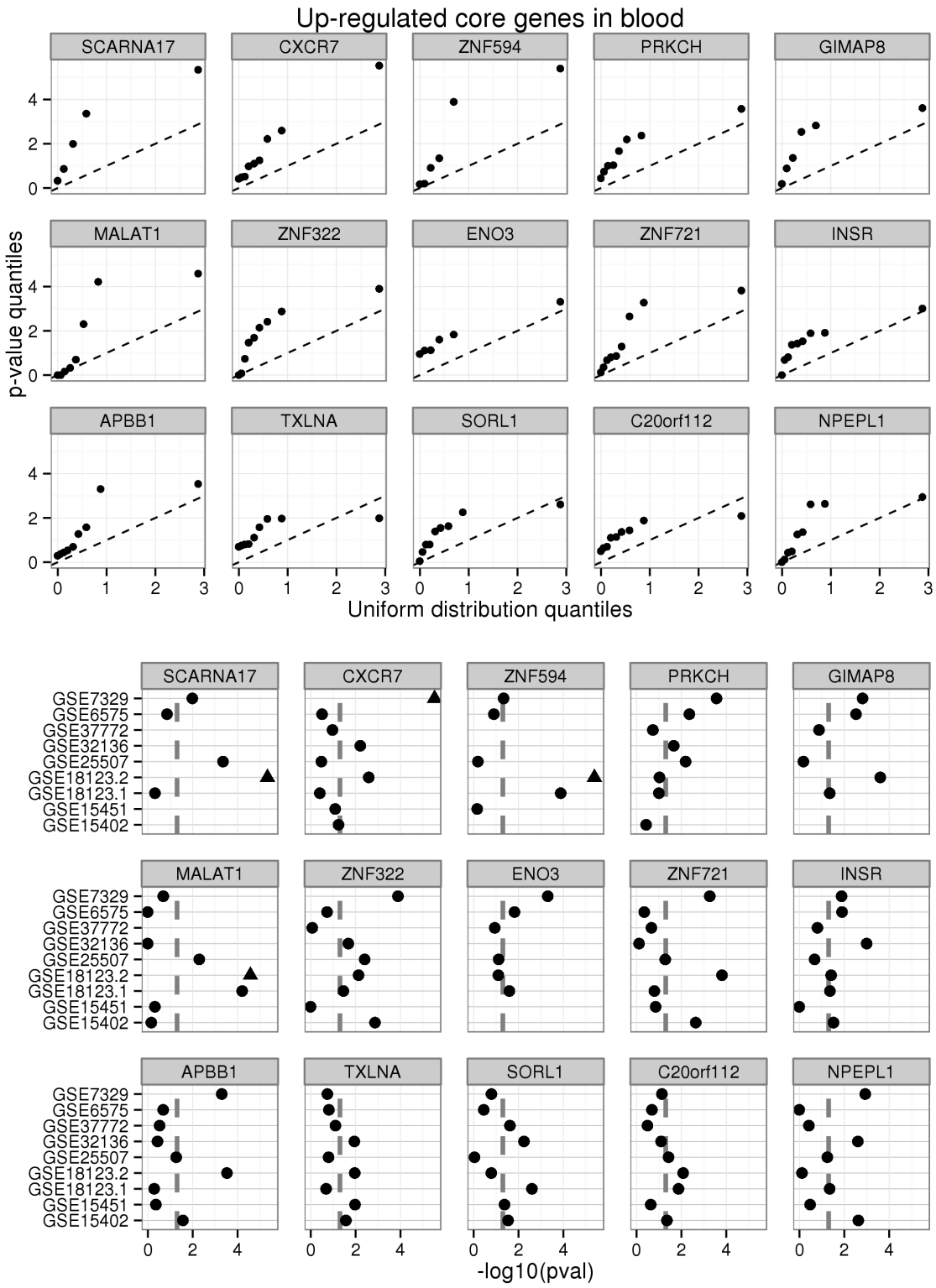


Supplementary Figure 5: Sex-linked genes, typically from studies with a gender-imbalanced study design, skew meta-profiles.

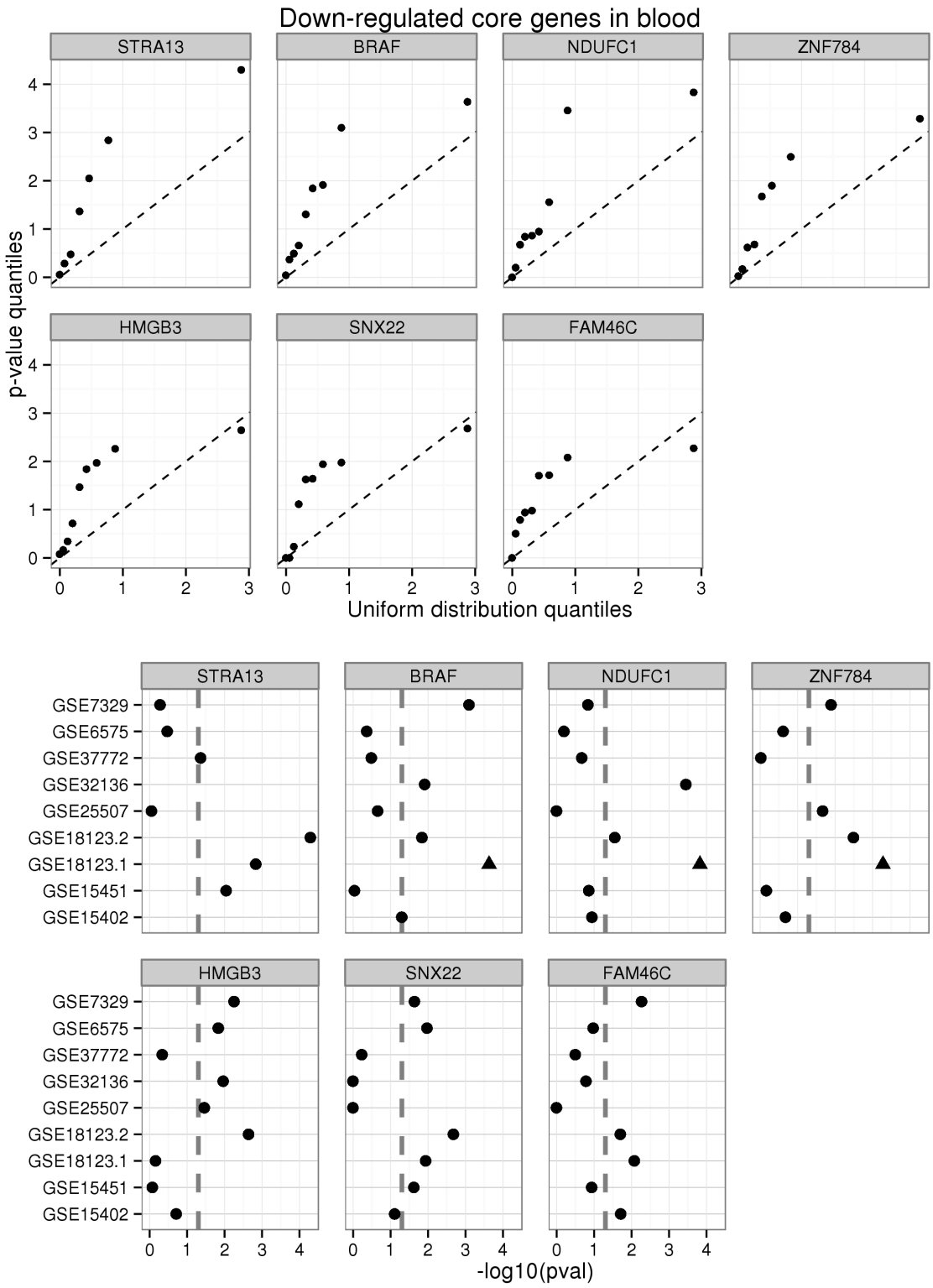


Supplementary Figure 6: Gene expression levels of core blood signature genes. Grey area: missing values.

(a)

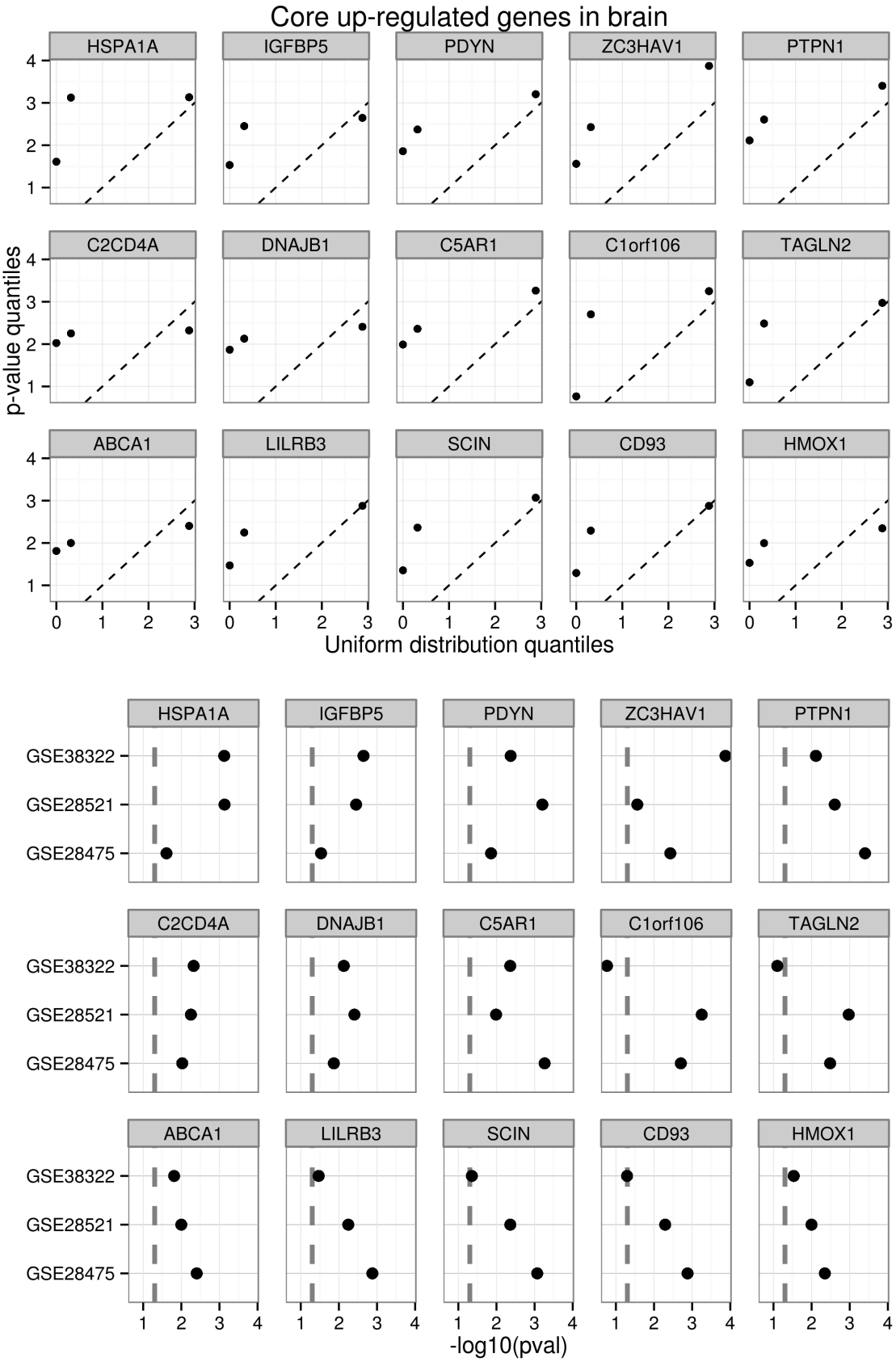

Up-regulated core genes in blood
SCARNA17
CXCR7
ZNF594
PRKCH
GIMAP8
MALAT1
ZNF322
ENO3
ZNF721
INSR
APBB1
TXLNA
SORL1
C20orf112
NPEPL1
p-value quantiles
Uniform distribution quantiles
GSE7329
GSE6575
GSE37772
GSE32136
GSE25507
GSE18123.2
GSE18123.1
GSE15451
GSE15402
-log10(pval)

(b)

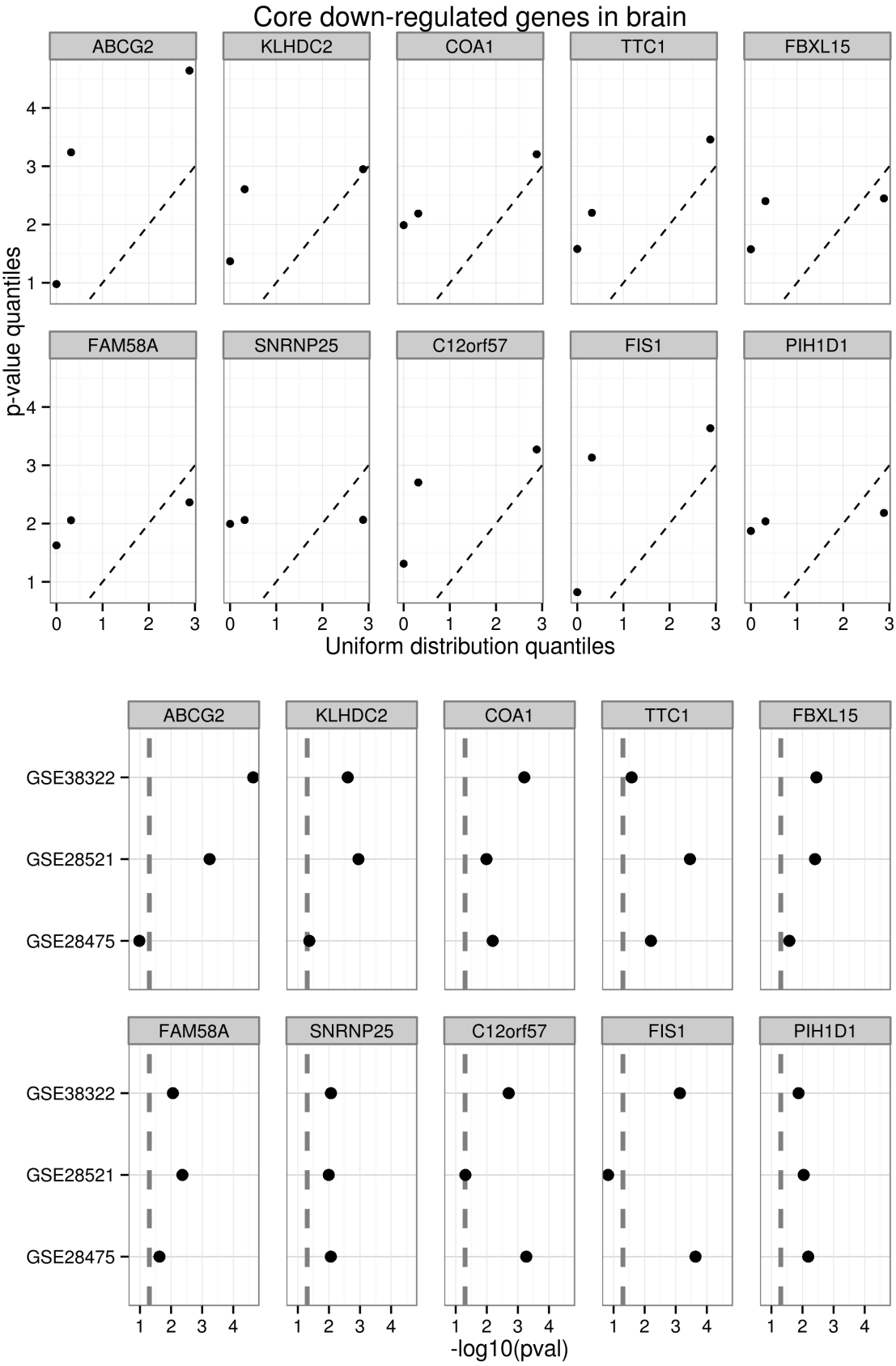


Supplementary Figure 7: P-values of core blood signatures in individual studies. Top: Quantile-quantile plots. Deviation from the diagonal for indicates that the gene is robust. Bottom: Raw p-values in each individual re-analysis, marked with a triangle if it meets an FDR threshold of 0.05 in that data set. a) Up-regulated genes. b) Down-regulated genes.

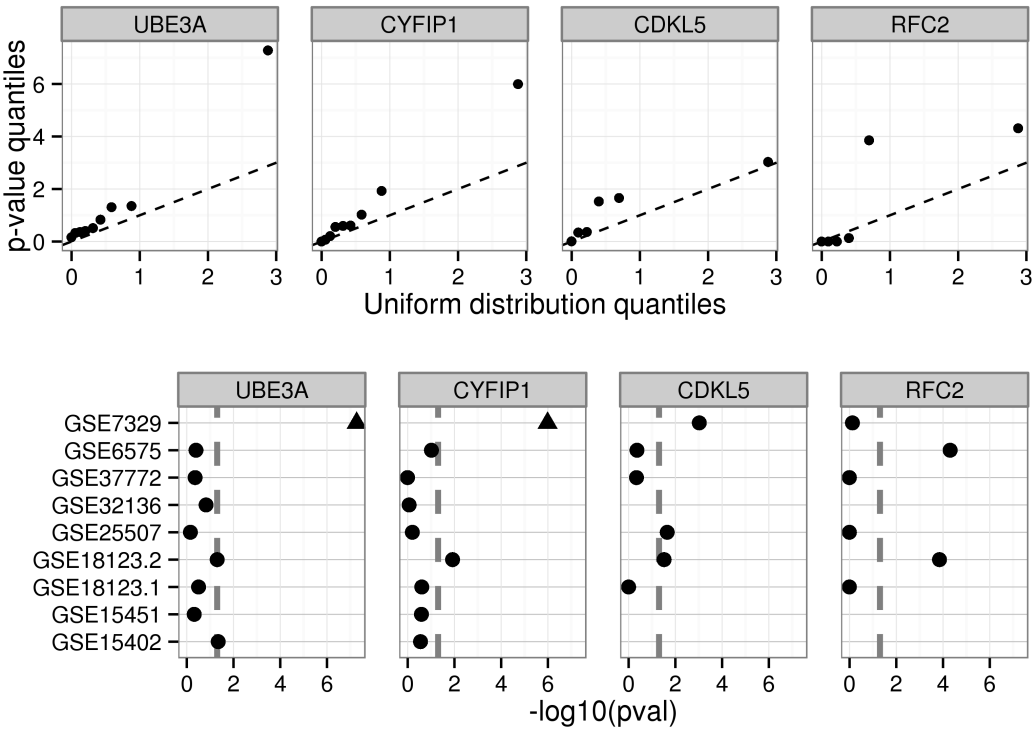
(a)


Core up-regulated genes in brain
HSPA1A
IGFBP5
PDYN
ZC3HAV1
PTPN1
C2CD4A
DNAJB1
C5AR1
C1orf106
TAGLN2
ABCA1
LILRB3
SCIN
CD93
HMOX1
p-value quantiles
Uniform distribution quantiles
GSE38322
GSE28521
GSE28475
-log10(pval)

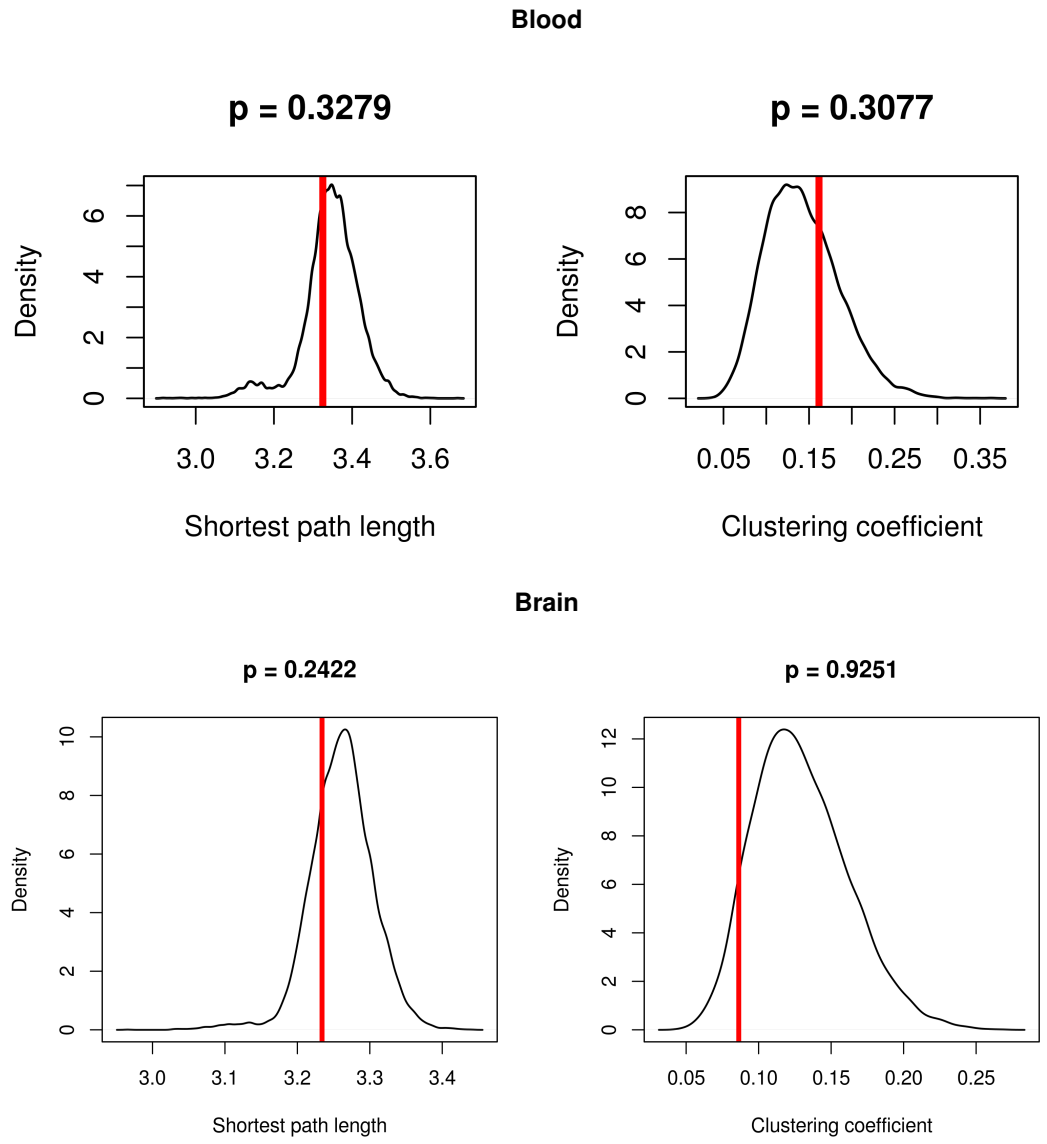
(b)



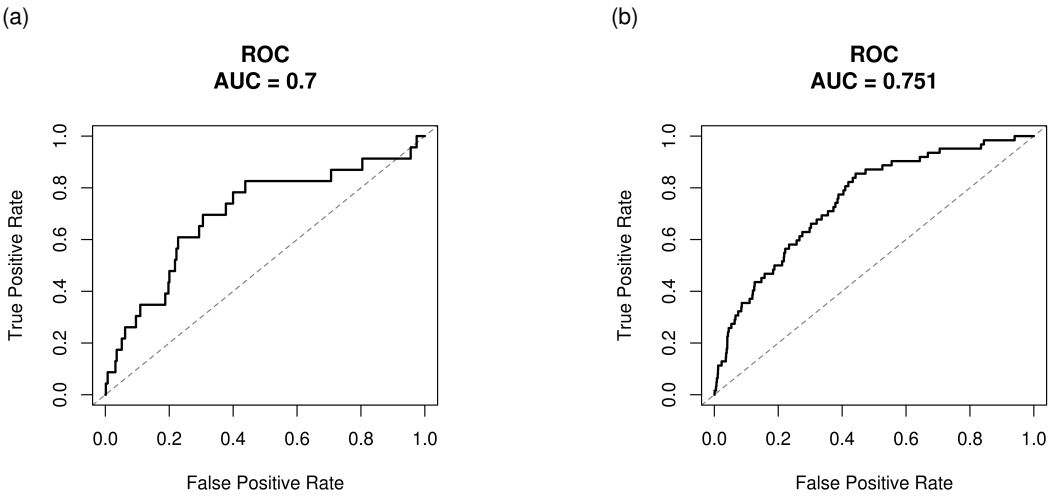
Supplementary Figure 8: P-values of core brain signatures in individual studies. Top: Quantile-quantile plots. Deviation from the diagonal for indicates that the gene is robust. Bottom: Raw p-values in each individual re-analysis, marked with a triangle if it meets an FDR threshold of 0.05 in that data set. a) Up-regulated genes. b) Down-regulated genes.



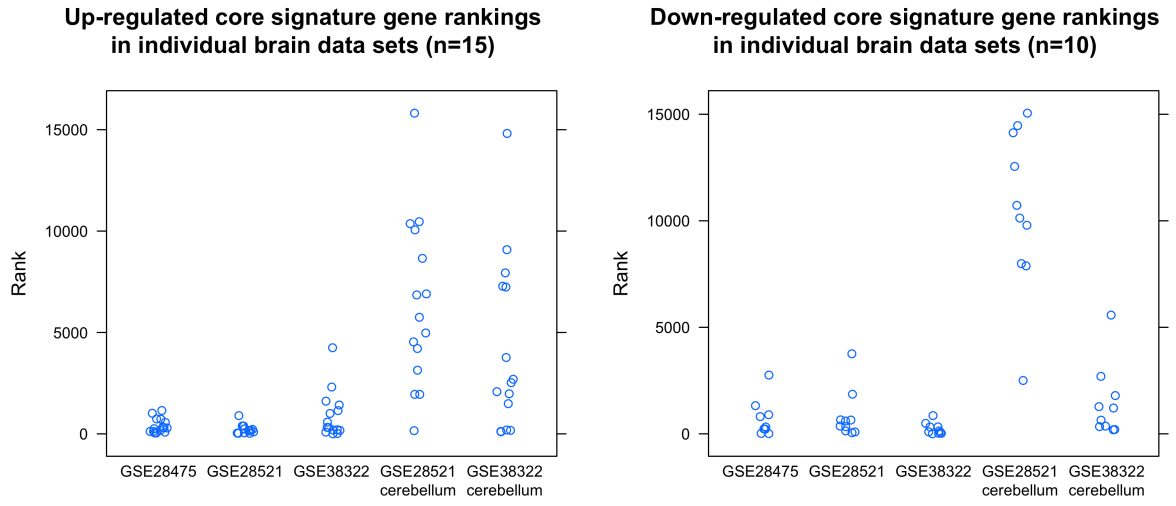
Supplementary Figure 9: Raw p-values of genes located in 15q11-13 (UBE3A, CYFIP1) Xp22 (CDKL5), and 7q11.23 (RFC2). Top(Q-Q plots): The lack of an overall deviation from the uniform diagonal suggests that the signals are skewed. Bottom: Per-dataset p-value with a p-value threshold of 0.05 (dashed grey); genes that meet an FDR threshold of 0.05 in the dataset are marked with a triangle. Compare with core blood signatures provided in the supplement.



Supplementary Figure 10: PPIN network properties of core candidate genes in the blood and brain. The properties of the genes were not significantly different from random gene sets.

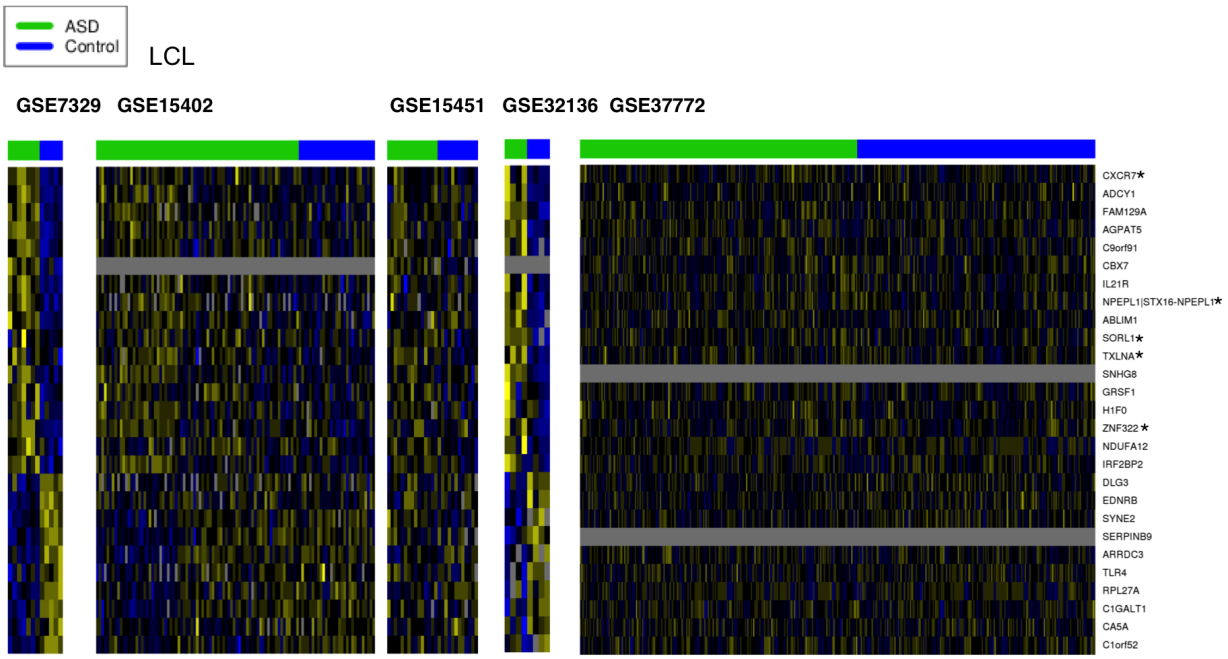


Supplementary Figure 11: Common brain meta-signatures between the autism (current study) and schizophrenia meta-analyses by Mistry et al. a) Up-regulated genes; b) Down-regulated genes.

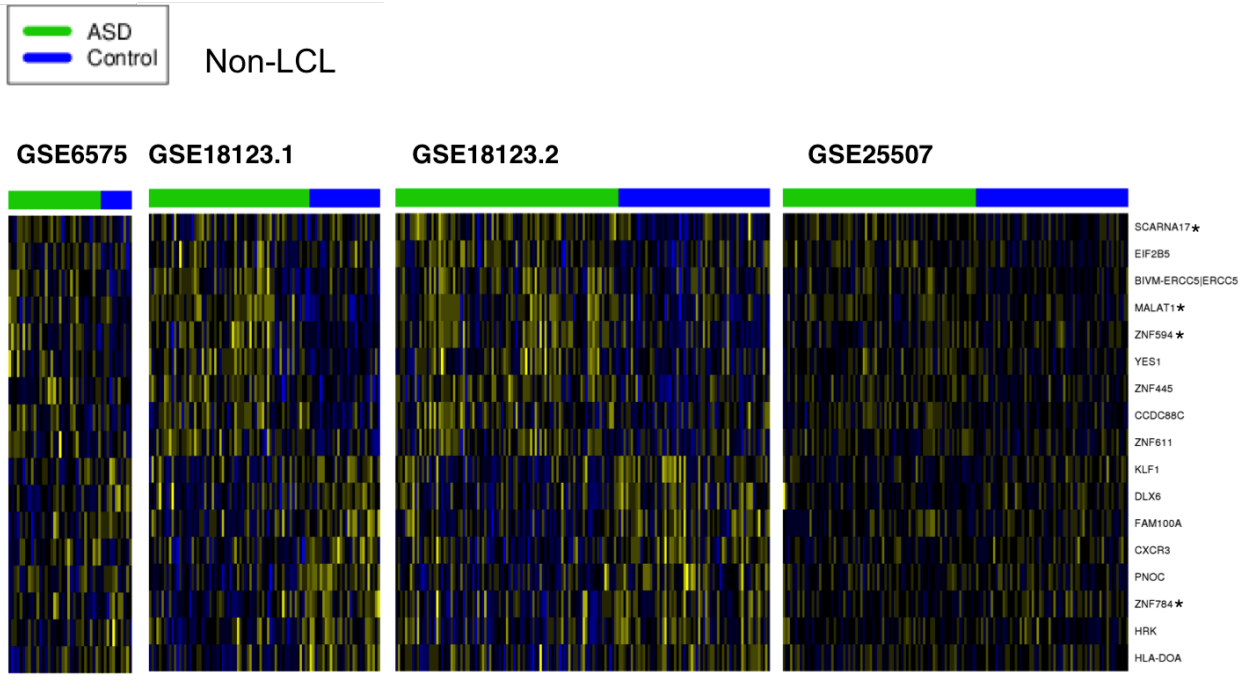


Supplementary Figure 12: Core signature genes are highly ranked (differentially expressed) in the cortex. The patterns differ in the cerebellum, suggesting that the cortex and cerebellum do not share similar molecular signatures in autism.

(a)

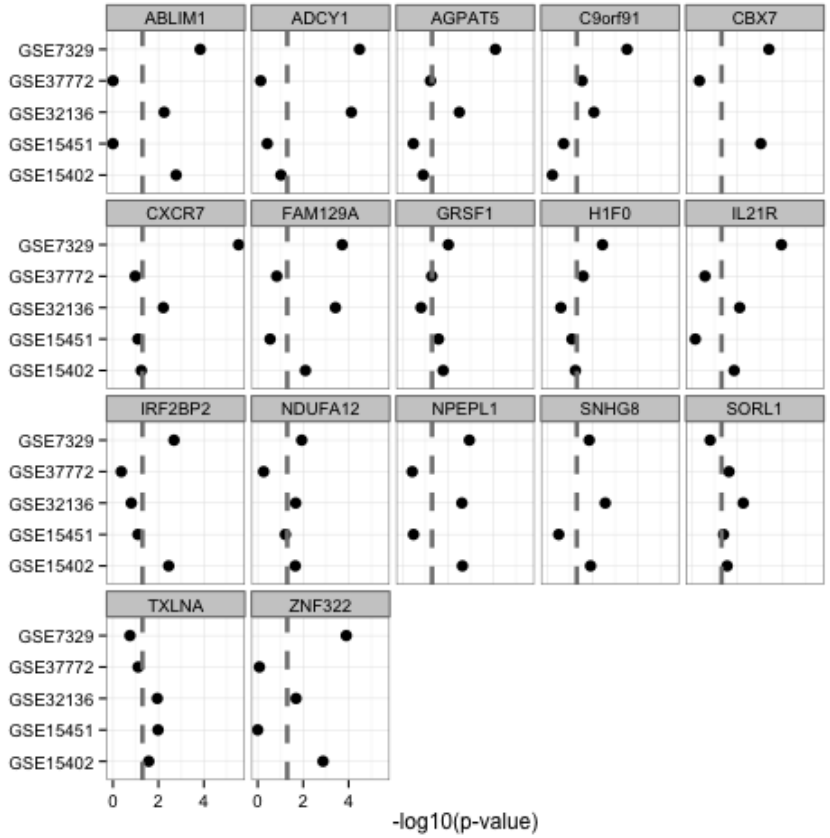


(b)

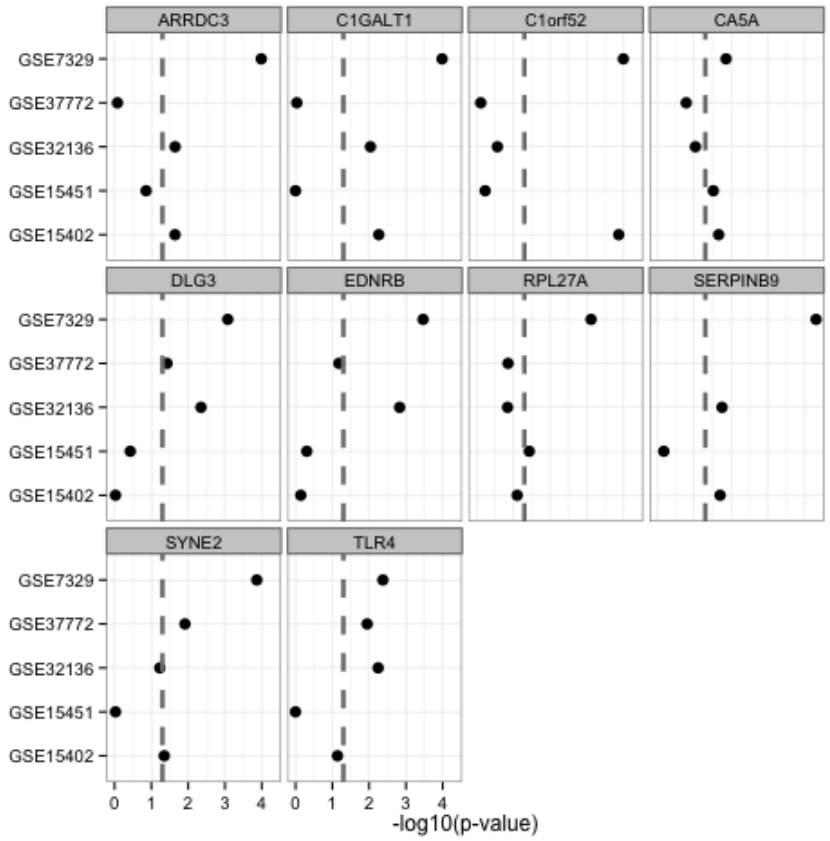


Supplementary Figure 13: Gene expression levels of core blood signature genes from a) LCL subgroup b) non-LCL subgroup. Grey area: missing values. Genes marked with an asterisk were also found in the analysis with all blood data sets.

(a)

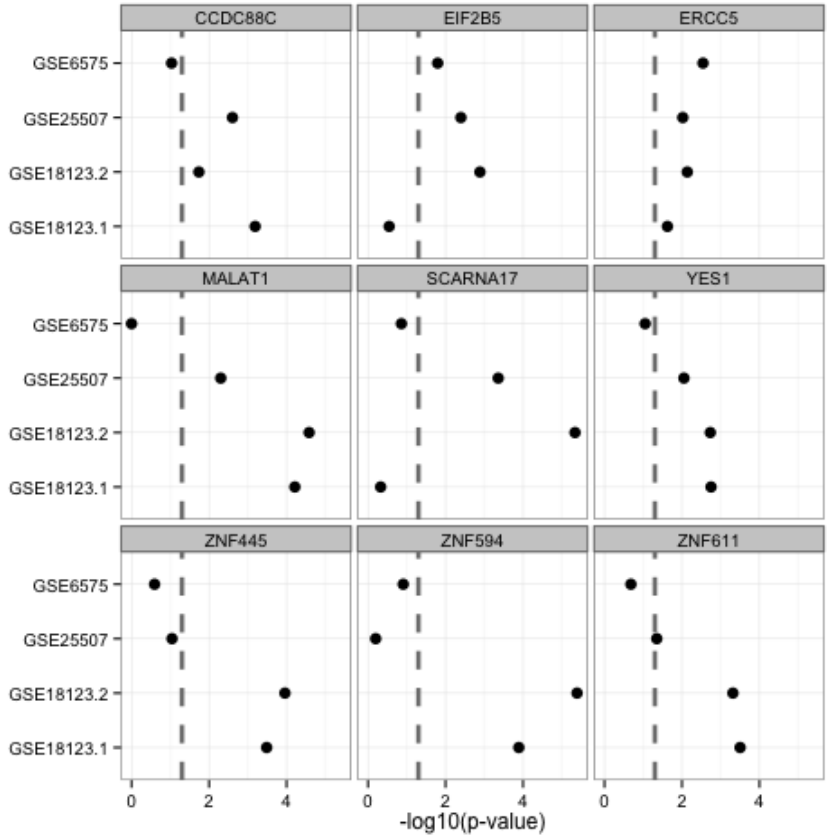


(b)

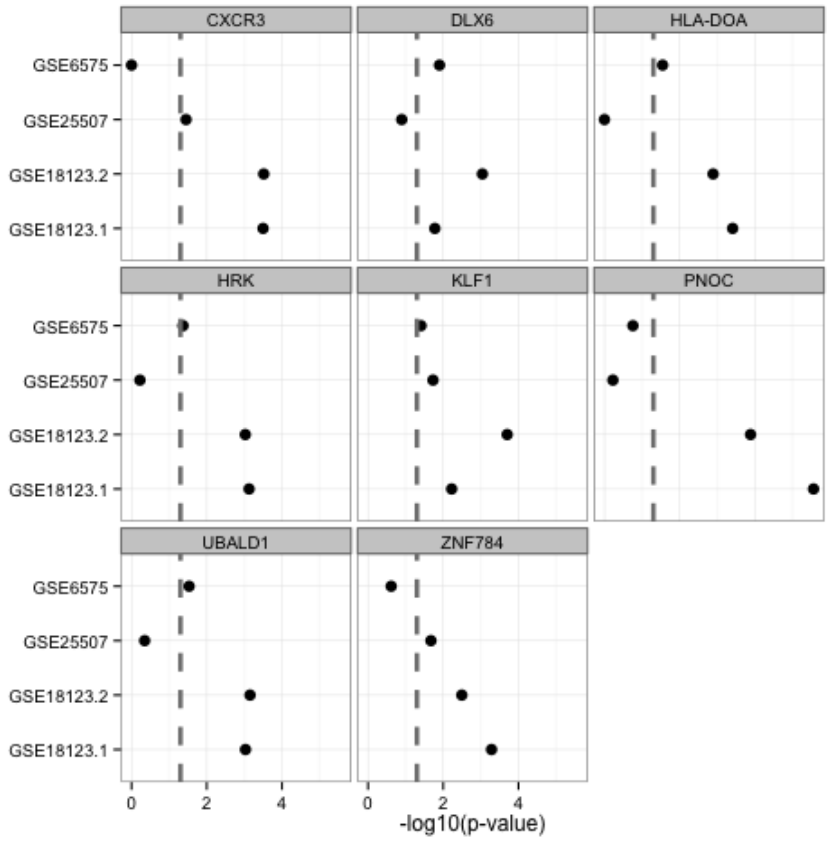


Supplementary Figure 14: P-values of core LCL signatures in individual studies. Dotted lines indicate p-value threshold of 0.05. a) Up-regulated genes. b) Down-regulated genes.

(a)

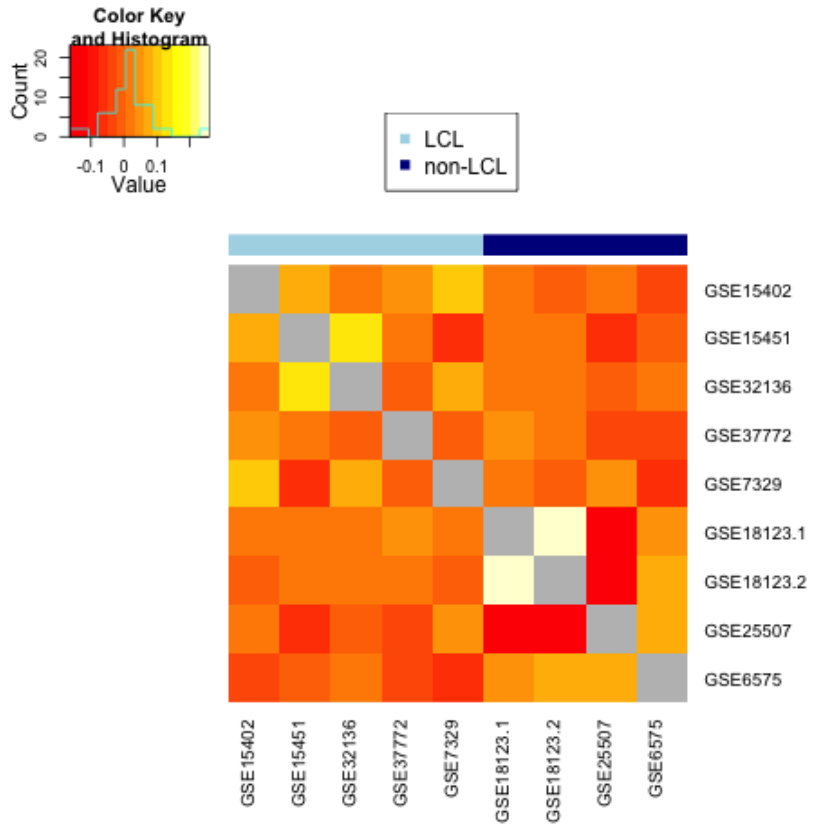


(b)

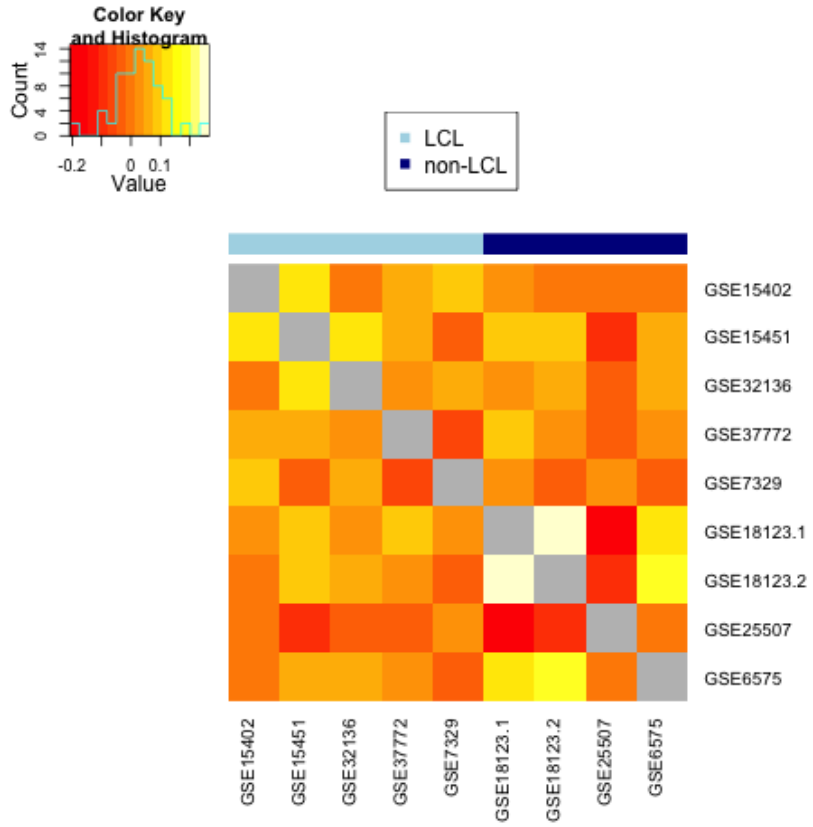


Supplementary Figure 15: P-values of core non-LCL signatures in individual studies. Dotted lines indicate p-value threshold of 0.05. a) Up-regulated genes. b) Down-regulated genes.

(a)



(b)



Supplementary Figure 16: Heatmaps illustrating spearman ranked correlations (scaled) of p-values among blood datasets.

References

[1] Gautier, L., Cope, L., Bolstad, B. M., and Irizarry, R. A. (2004) *Bioinformatics* **20(3)**, 307–315.

[2] Du, P., Kibbe, W. A., and Lin, S. M. July 2008 *Bioinformatics* **24(13)**, 1547–1548 PMID: 18467348.

[3] Churchill, G. A. December 2002 *Nat Genet* **32**, 490–495.

[4] Johnson, W. E., Li, C., and Rabinovic, A. January 2007 *Biostat* **8(1)**, 118–127 PMID: 16632515.

[5] Zoubarev, A., Hamer, K. M., Keshav, K. D., McCarthy, E. L., Santos, J. R. C., Van Rossum, T., McDonald, C., Hall, A., Wan, X., Lim, R., Gillis, J., and Pavlidis, P. July 2012 *Bioinformatics* **28(17)**, 2272–3 PMID: 22782548.

[6] Ramasamy, A., Mondry, A., Holmes, C. C., and Altman, D. G. September 2008 *PLoS Med* **5(9)**, e184.

[7] Fisher, R. (1948) *American Statistician* **2:30.**

[8] Benjamini, Y. and Hochberg, Y. (1995) *Journal of the Royal Statistical Society Series* B(57), 289–300.

[9] Kang, H. J., Kawasawa, Y. I., Cheng, F., Zhu, Y., Xu, X., Li, M., Sousa, A. M. M., Pletikos, M., Meyer, K. A., Sedmak, G., Guennel, T., Shin, Y., Johnson, M. B., Krsnik, u., Mayer, S., Fertuzinhos, S., Umlauf, S., Lisgo, S. N., Vortmeyer, A., Weinberger, D. R., Mane, S., Hyde, T. M., Huttner, A., Reimers, M., Kleinman, J. E., and Å estan, N. October 2011 *Nature* 478(7370), 483–489.

[10] Whitney, A. R., Diehn, M., Popper, S. J., Alizadeh, A. A., Boldrick, J. C., Relman, D. A., and Brown, P. O. February 2003 *PNAS* 100(4), 1896–1901 PMID: 12578971.

[11] Carrel, L. and Willard, H. F. March 2005 *Nature* 434(7031), 400–404.

[12] Mistry, M. and Pavlidis, P. (2010) *Neuroscience* 167, 384–95 2.

[13] Mistry, M., Gillis, J., and Pavlidis, P. February 2013 *Mol Psychiatry* 18(2), 215–225.

[14] Lee, H. K., Braynen, W., Keshav, K., and Pavlidis, P. (2005) *BMC Bioinformatics* 6, 269 PMID: 16280084.

[15] Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T., Harris, M. A., Hill, D. P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J. C., Richardson, J. E., Ringwald, M., Rubin, G. M., and Sherlock, G. May 2000 *Nat. Genet.* 25(1), 25–29 PMID: 10802651.

[16] Gillis, J., Mistry, M., and Pavlidis, P. June 2010 *Nat. Protocols* 5(6), 1148–1159.

[17] Portales-Casamar, E., Ch'ng, C., Lui, F., St-Georges, N., Zoubarev, A., Lai, A. Y., Lee, M., Kwok, C., Kwok, W., Tseng, L., and Pavlidis, P. February 2013 *BMC Genomics* 14(1), 129 PMID: 23442263.

[18] Garbett, K., Ebert, P. J., Mitchell, A., Lintas, C., Manzi, B., Mirnics, K., and Persico, A. M. June 2008 *Neurobiol Dis* 30(3), 303–311 PMID: 18378158 PMCID: PMC2693090.

[19] Purcell, A. E., Jeon, O. H., Zimmerman, A. W., Blue, M. E., and Pevsner, J. November 2001 *Neurology* 57(9), 1618–1628 PMID: 11706102.

[20] Glatt, S. J., Tsuang, M. T., Winn, M., Chandler, S. D., Collins, M., Lopez, L., Weinfeld, M., Carter, C., Schork, N., Pierce, K., and Courchesne, E. September 2012 *J Am Acad Child Adolesc Psychiatry* 51(9), 934–944.e2 PMID: 22917206.

[21] Hu, V. W., Frank, B. C., Heine, S., Lee, N. H., and Quackenbush, J. (2006) *BMC Genomics* 7, 118 PMID: 16709250.

[22] Marshall, C. R., Noor, A., Vincent, J. B., Lionel, A. C., Feuk, L., Skaug, J., Shago, M., Moessner, R., Pinto, D., Ren, Y., Thiruvahindrapduram, B., Fiebig, A., Schreiber, S., Friedman, J., Ketelaars, C. E. J., Vos, Y. J., Ficicioglu, C., Kirkpatrick, S., Nicolson, R., Sloman, L., Summers, A., Gibbons, C. A., Teebi, A., Chitayat, D., Weksberg, R., Thompson, A., Vardy, C., Crosbie, V., Luscombe, S., Baatjes, R., Zwaigenbaum, L., Roberts, W., Fernandez, B., Szatmari, P., and Scherer, S. W. February 2008 *Am. J. Hum. Genet.* 82(2), 477–488 PMID: 18252227.

[23] Sanders, S. J., Ercan-Sencicek, A. G., Hus, V., Luo, R., Murtha, M. T., Moreno-De-Luca, D., Chu, S. H., Moreau, M. P., Gupta, A. R., Thomson, S. A., Mason, C. E., Bilguvar, K., Celestino-Soper, P. B. S., Choi, M., Crawford, E. L., Davis, L., Wright, N. R. D., Dhodapkar, R. M., DiCola, M., DiLullo, N. M., Fernandez, T. V., Fielding-Singh, V., Fishman, D. O., Frahm, S., Garagaloyan, R., Goh, G. S., Kammela, S., Klei, L., Lowe, J. K., Lund, S. C., McGrew, A. D., Meyer, K. A., Moffat, W. J., Murdoch, J. D., O'Roak, B. J., Ober, G. T., Pottenger, R. S., Raubeson, M. J., Song, Y., Wang, Q., Yaspan, B. L., Yu, T. W., Yurkiewicz, I. R., Beaudet, A. L., Cantor, R. M., Curland, M., Grice, D. E., GÃ¼nel, M., Lifton, R. P., Mane, S. M., Martin, D. M., Shaw, C. A., Sheldon, M., Tischfield, J. A., Walsh, C. A., Morrow, E. M., Ledbetter, D. H., Fombonne, E., Lord, C., Martin, C. L., Brooks, A. I., Sutcliffe, J. S., Cook, Edwin H, J., Geschwind, D., Roeder, K., Devlin, B., and State, M. W. June 2011 *Neuron* 70(5), 863–885 PMID: 21658581.

[24] Pinto, D., Pagnamenta, A. T., Klei, L., Anney, R., Merico, D., Regan, R., Conroy, J., Magalhaes, T. R., Correia, C., Abrahams, B. S., Almeida, J., Bacchelli, E., Bader, G. D., Bailey, A. J., Baird, G., Battaglia, A., Berney, T., Bolshakova, N., Bölte, S., Bolton, P. F., Bourgeron, T., Brennan, S., Brian, J., Bryson, S. E., Carson, A. R., Casallo, G., Casey, J., Chung, B. H. Y., Cochrane, L., Corsello, C., Crawford, E. L., Crossett, A., Cytrynbaum, C., Dawson, G., deJonge, M., Delorme, R., Drmic, I., Duketis, E., Duque, F., Estes, A., Farrar, P., Fernandez, B. A., Folstein, S. E., Fombonne, E., Freitag, C. M., Gilbert, J., Gillberg, C., Glessner, J. T., Goldberg, J., Green, A., Green, J., Guter, S. J., Hakonarson, H., Heron, E. A., Hill, M., Holt, R., Howe, J. L., Hughes, G., Hus, V., Igliozzi, R., Kim, C., Klauck, S. M., Kolevzon, A., Korvatska, O., Kustanovich, V., Lajonchere, C. M., Lamb, J. A., Laskawiec, M., Leboyer, M., Le Couteur, A., Leventhal, B. L., Lionel, A. C., Liu, X.-Q., Lord, C., Lotspeich, L., Lund, S. C., Maestrini, E., Mahoney, W., Mantoulan, C., Marshall, C. R., McConachie, H., McDougle, C. J., McGrath, J., McMahon, W. M., Merikangas, A., Migita, O., Minshew, N. J., Mirza, G. K., Munson, J., Nelson, S. F., Noakes, C., Noor, A., Nygren, G., Oliveira, G., Papanikolaou, K., Parr, J. R., Parrini, B., Paton, T., Pickles, A., Pilorge, M., Piven, J., Ponting, C. P., Posey, D. J., Poustka, A., Poustka, F., Prasad, A., Ragoussis, J., Renshaw, K., Rickaby, J., Roberts, W., Roeder, K., Roge, B., Rutter, M. L., Bierut, L. J., Rice, J. P., Salt, J., Sansom, K., Sato, D., Segurado, R., Sequeira, A. F., Senman, L., Shah, N., Sheffield, V. C., Soorya, L., Sousa, I., Stein, O., Sykes, N., Stoppioni, V., Strawbridge, C., Tancredi, R., Tansey, K., Thiruvahindrapduram, B., Thompson, A. P., Thomson, S., Tryfon, A., Tsiantis, J., Van Engeland, H., Vincent, J. B., Volkmar, F., Wallace, S., Wang, K., Wang, Z., Wassink, T. H., Webber, C., Weksberg, R., Wing, K., Wittemeyer, K., Wood, S., Wu, J., Yaspan, B. L., Zurawiecki, D., Zwaigenbaum, L., Buxbaum, J. D., Cantor, R. M., Cook, E. H., Coon, H., Cuccaro, M. L., Devlin, B., Ennis, S., Gallagher, L., Geschwind, D. H., Gill, M., Haines, J. L.,

Hallmayer, J., Miller, J., Monaco, A. P., Nurnberger Jr, J. I., Paterson, A. D., Pericak-Vance, M. A., Schellenberg, G. D., Szatmari, P., Vicente, A. M., Vieland, V. J., Wijsman, E. M., Scherer, S. W., Sutcliffe, J. S., and Betancur, C. July 2010 *Nature* 466(7304), 368–372.

[25] Iafrate, A. J., Feuk, L., Rivera, M. N., Listewnik, M. L., Donahoe, P. K., Qi, Y., Scherer, S. W., and Lee, C. September 2004 *Nat Genet* 36(9), 949–951.

[26] Hinrichs, A. S., Karolchik, D., Baertsch, R., Barber, G. P., Bejerano, G., Clawson, H., Diekhans, M., Furey, T. S., Harte, R. A., Hsu, F., Hillman-Jackson, J., Kuhn, R. M., Pedersen, J. S., Pohl, A., Raney, B. J., Rosenbloom, K. R., Siepel, A., Smith, K. E., Sugnet, C. W., Sultan-Qurraie, A., Thomas, D. J., Trumbower, H., Weber, R. J., Weirauch, M., Zweig, A. S., Haussler, D., and Kent, W. J. January 2006 *Nucleic Acids Res* 34(Database issue), D590–D598 PMID: 16381938 PMCID: PMC1347506.

[27] Barnes, M., Freudenberg, J., Thompson, S., Aronow, B., and Pavlidis, P. (2005) *Nucleic Acids Res* 33, 5914–23 18.

[28] Prasad, T. S. K., Kandasamy, K., and Pandey, A. (2009) *Methods Mol. Biol.* 577, 67–79 PMID: 19718509.

[29] Chatr-aryamontri, A., Ceol, A., Palazzi, L. M., Nardelli, G., Schneider, M. V., Castagnoli, L., and Cesareni, G. January 2007 *Nucl. Acids Res.* 35(suppl 1), D572–D574 PMID: 17135203.

[30] Xenarios, I., Rice, D. W., Salwinski, L., Baron, M. K., Marcotte, E. M., and Eisenberg, D. January 2000 *Nucleic Acids Res* 28(1), 289–291 PMID: 10592249 PMCID: PMC102387.

[31] Lynn, D. J., Winsor, G. L., Chan, C., Richard, N., Laird, M. R., Barsky, A., Gardy, J. L., Roche, F. M., Chan, T. H. W., Shah, N., Lo, R., Naseer, M., Que, J., Yau, M., Acab, M., Tulpan, D., Whiteside, M. D., Chikatamarla, A., Mah, B., Munzner, T., Hokamp, K., Hancock, R. E. W., and Brinkman, F. S. L. September 2008 *Mol Syst Biol* 4(1).

[32] Razick, S., Magklaras, G., and Donaldson, I. M. September 2008 *BMC Bioinformatics* 9(1), 405 PMID: 18823568.