Supplemental Table 3. Replication sample subject demographics

Diagnosis	**Gender** (F/M)	**Mean RIN** (range)	**Mean Age** (range)
ASD	4/15	7.64 (6.2 to 9.1)	31 (2 to 60)
Control	5/23	7.50 (6.2 to 8.4)	36 (15 to 60)

This table indicates the demographic information for the subjects used in the replication analysis. ASD = autism spectrum disorder, F= female, M = male.

Supplemental Table 5. Significantly differentially expressed genes between ASD and Control subjects

Gene Symbol	Log2 Fold Change	p value	q value	Function
SNORA74A	-0.71	2.73E-07	0.002	RNA modification
SNORA53	-1.01	2.89E-07	0.002	RNA modification
SNORD17	-0.91	3.58E-06	0.016	RNA modification
TUBE1	0.18	7.26E-06	0.024	Microtubule organization
SNORA54	0.84	9.45E-06	0.025	RNA modification
SNORA74B	-0.46	1.46E-05	0.032	RNA modification
RP6-206I17.3	0.29	1.93E-05	0.036	Unknown
SNORD114-23	-0.54	2.93E-05	0.048	RNA modification

Statistical summary of the significantly differential expressed genes between autism spectrum disorder (ASD) and control subjects of the dorsolateral prefrontal cortex (DLPFC). Log fold change is indicated with the controls as the reference group; therefore, *TUBE1* is more highly expressed in ASD subjects than in control subjects. The p value indicates the significance for each gene in the differential expression analysis. The q value indicates the multiple testing corrected significance value for each gene.

Supplemental Table 6. Significantly differentially expressed genes between ASD and control subjects in the transcriptome-wide differential expression analysis using male subjects only

Gene Symbol	Log2 Fold Change	p value	q value	Function
SNORA54	1.01	1.17E-06	0.015	RNA modification
SNORA74A	-0.74	5.22E-06	0.034	RNA modification
SNORA53	-0.96	1.03E-05	0.045	RNA modification

Statistical summary of the significantly differential expressed genes between autism spectrum disorder (ASD) and control subjects of the dorsolateral prefrontal cortex (DLPFC) using only male subjects (10 ASD, 30 controls) out of 12 935 genes that passed the expression threshold. Log fold change is indicated with the controls as the reference group; therefore, *SNORA54* is more highly expressed in the male ASD subjects as compared to the male control subjects. The p value indicates the significance for each gene in the differential expression analysis. The q value indicates the multiple testing corrected significance value for each gene.

Supplemental Table 7. Statistical summary for the histaminergic genes of interest in the transcriptome-wide differential expression analysis using male subjects only

Gene Symbol	Log2 Fold Change	p value	q value	Control Mean Expression	ASD Mean Expression
HNMT	0.079	0.013	0.473	0.96	1.03
HRH1	0.054	0.396	0.873	0.58	0.64
HRH2	-0.002	0.964	0.996	0.87	0.87
HRH3	0.052	0.342	0.851	0.96	1.01

Statistical summary of the expression differences of the histamine genes of interest between autism spectrum disorder (ASD) and control subjects of the dorsolateral prefrontal cortex (DLPFC) using male subjects (10 ASD, 30 controls). Log fold change is indicated with the control subjects as the reference group; therefore, *HNMT* is more highly expressed in ASD subjects than in control subjects. The p value indicates the significance for each gene in the differential expression analysis. The q value indicates the multiple testing corrected significance value for each gene. Mean expression is expressed as log2(RPKM +1) values and reflect correction of confounding effects by principal components and covariates.

Supplemental Table 8. Statistical summary for the histaminergic genes of interest in the transcriptome-wide differential expression analysis using the replication data set

Gene Symbol	Log2 Fold Change	p value	q value	Control Mean Expression	ASD Mean Expression
HNMT	0.030	0.348	0.824	1.02	1.05
HRH1	-0.076	0.064	0.638	0.73	0.65
HRH2	-0.057	0.192	0.753	1.34	1.29
HRH3	0.136	0.013	0.536	-0.33	-0.19

Statistical summary of the histaminergic genes of interest that were sufficiently expressed for differential expression analysis between autism spectrum disorder (ASD) and control subjects of the replication data set. Log fold change is indicated with the controls as the reference group; therefore, *HNMT* is more highly expressed in ASD subjects than in control subjects. The p value indicates the significance for each gene in the differential expression analysis. The q value indicates the multiple testing corrected significance value for each gene. Mean expression is expressed as log2(RPKM+1) values and reflect correction of confounding effects by principal components.

Supplemental Table 9. Statistical summary for a replication analysis of just the genes previously identified as significantly differentially expressed in the transcriptome-wide differential expression analysis using our data

Gene Symbol	Log2 Fold Change	p value	q value	Function
SNORA74A	0.11	0.518	0.893	RNA modification
SNORA53	-0.01	0.912	0.984	RNA modification
SNORD17	0.05	0.719	0.947	RNA modification
TUBE1	0.02	0.586	0.915	Microtubule organization
SNORA54	-0.01	0.891	0.980	RNA modification
SNORA74B	0.29	0.161	0.729	RNA modification
RP6-206I17.3	NA	NA	NA	Unknown
SNORD114-23	-0.02	0.798	0.965	RNA modification

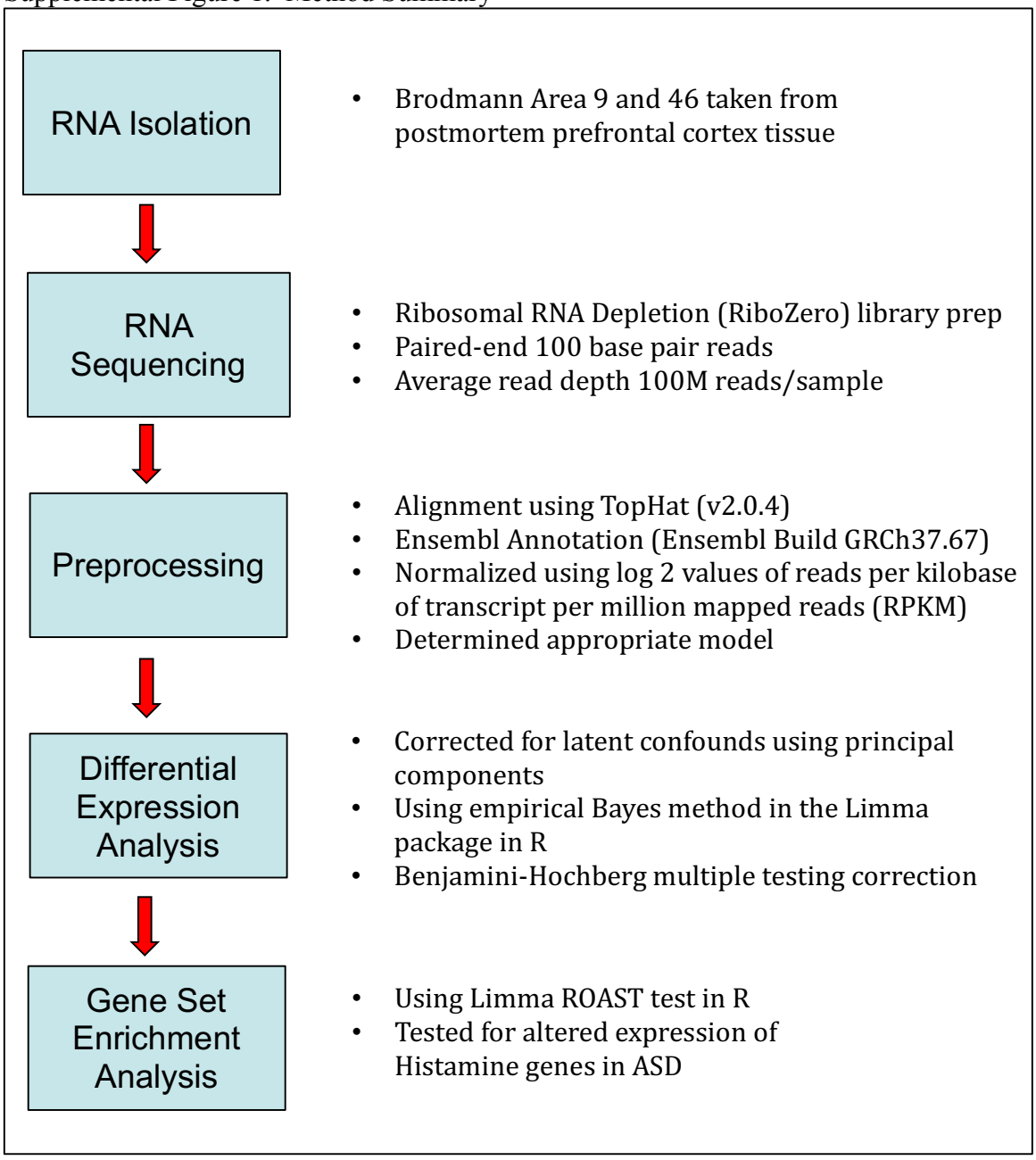
Statistical summary of the replication analysis evaluating only the previously identified significantly differential expressed genes between autism spectrum disorder (ASD) and control subjects of the dorsolateral prefrontal cortex (DLPFC). *RP6-206I17.3,* also called *HAS-LNCG003387,* was not analyzed in this analysis as the mean expression value was below the lower limit threshold, (log2 (RPKM+1) >0.5) in this data set. Log fold change is indicated with the controls as the reference group; therefore, *SNORD17* is more highly expressed in ASD subjects than in control subjects. The p value indicates the significance for each gene in the differential expression analysis. The q value indicates the multiple testing corrected significance value for each gene.

Supplemental Table 10. Meta-analysis of the genes and gene sets across the original and replication data sets

Meta-Analysis Test	chisq	meta p value	meta q value
Significant Genes in the LIBD Data Set			
SNORA74A	31.54	2.38E-06	3.62E-05
SNORA53	30.3	4.26E-06	3.62E-05
SNORD17	25.74	3.57E-05	1.52E-04
TUBE1	24.73	5.69E-05	1.93E-04
SNORA54	23.37	1.07E-04	3.03E-04
SNORA74B	25.92	3.29E-05	1.52E-04
RP6-206I17.3	NA	NA	NA
SNORD114-23	21.33	2.73E-04	6.63E-04
Histamine Genes of Interest			
HNMT	13.54	8.90E-03	0.012
HRH1	10.77	2.93E-02	0.033
HRH2	4.921	2.96E-01	0.296
HRH3	12.27	1.54E-02	0.019
Significant Gene Sets in the LIBD Data Set			
Hypothesis Driven Gene Set, up hypothesis	16.17	2.80E-03	0.004
GO: Histamine Receptor Activity, up hypothesis	9.882	4.25E-02	0.045
Hypothesis Driven Gene Set, mixed hypothesis	19.41	6.54E-04	0.001
GO: Histamine Receptor Activity, mixed hypothesis	14.38	6.19E-03	0.009
GeneRIF: Histaminergic, mixed hypothesis	19.05	7.68E-04	0.001
GeneRIF: Histamine, mixed hypothesis	16.31	2.63E-03	0.004

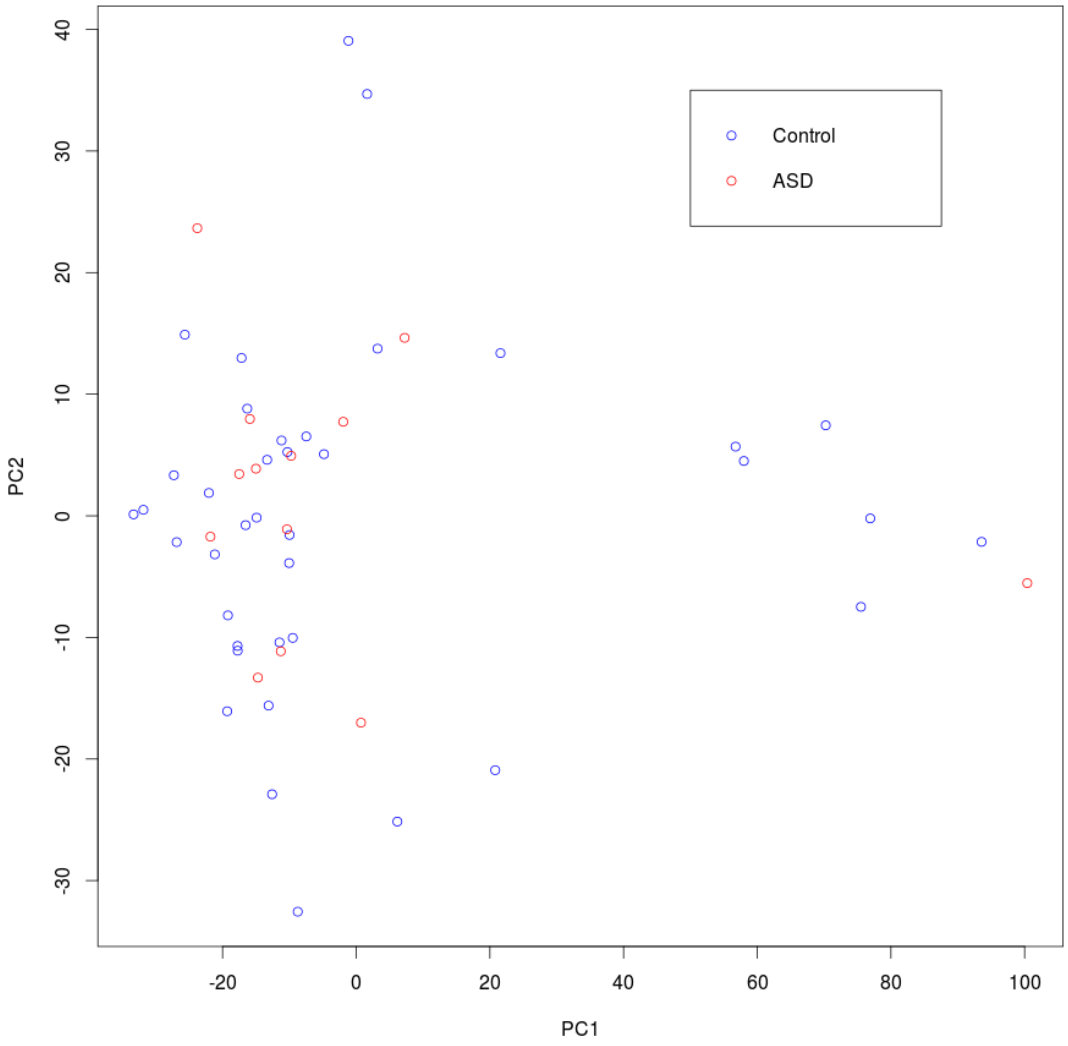
This table shows meta-analysis results using the Fisher's method of the p value statistics from both the original and replication data sets. These results were obtained using the sumlog() function of the "matap" package in R. These analyses could only be done for genes expressed in both datasets, therefore meta-analysis of *RP6-206I17.3* was not possible. Multiple testing correction was done using the Benjamini-Hochberg method.

Supplemental Figure 1. Method Summary



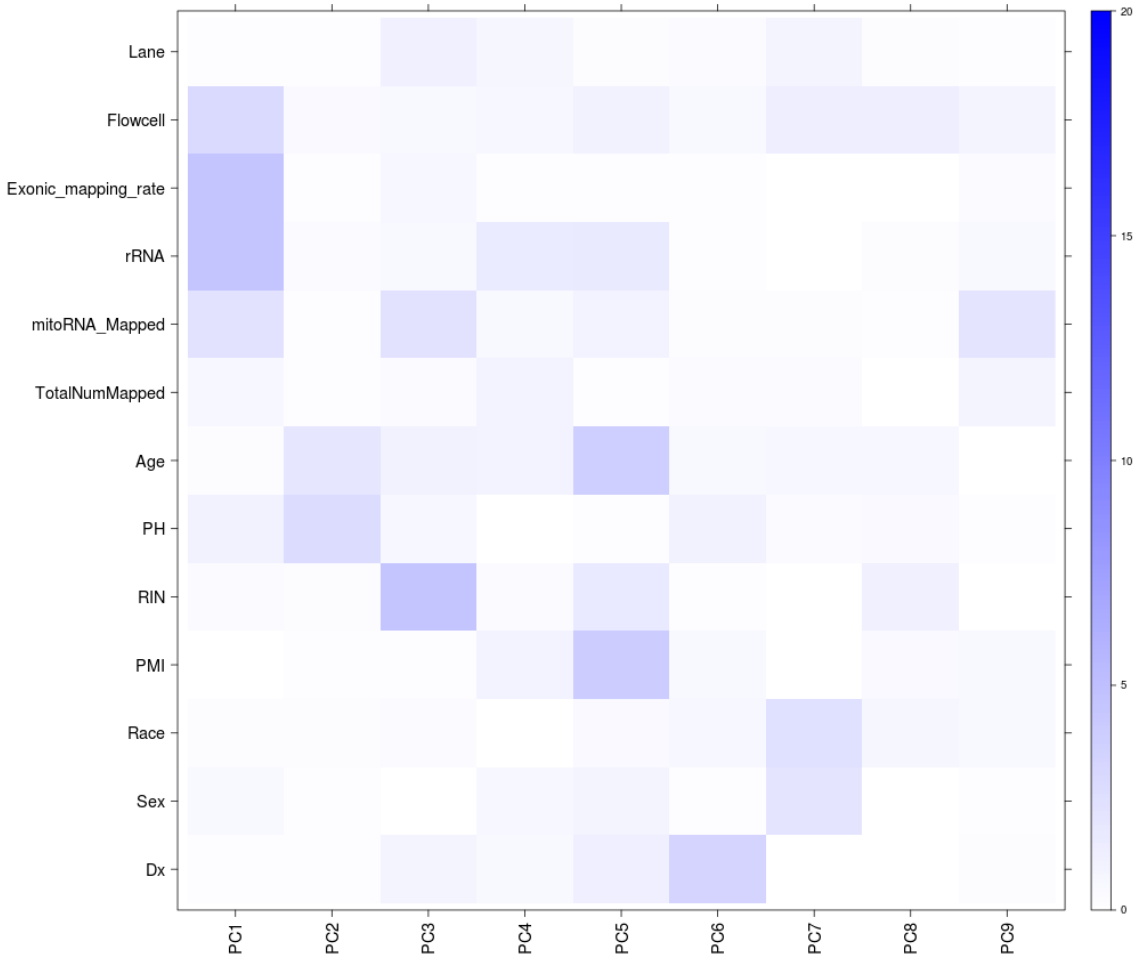
This flowchart depicts a summary of the methods used in this analysis of the expression of histamine genes in postmortem human brain samples of Autism Spectrum Disorder (ASD) subjects and healthy controls.

Supplemental Figure 2. Plot of the first two principal components of the gene expression data



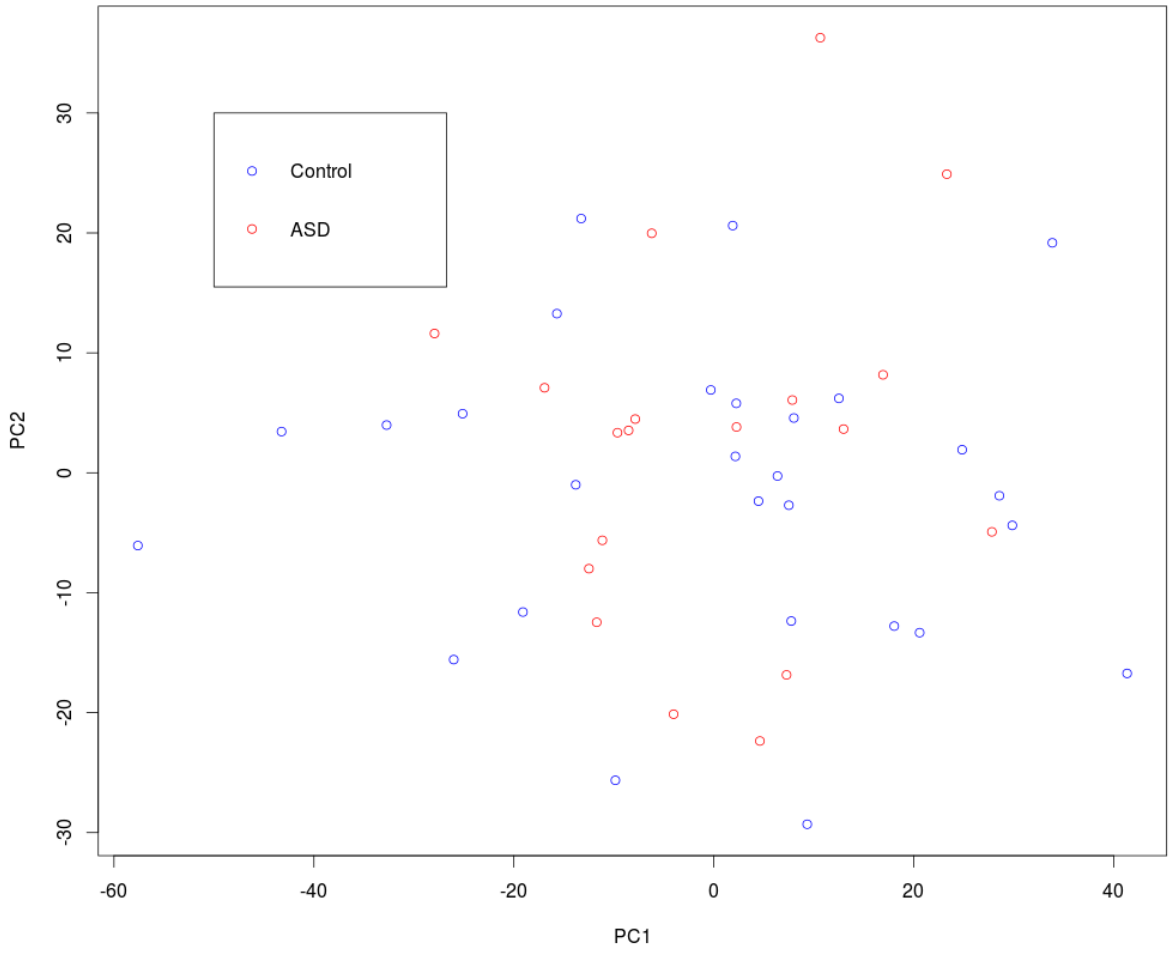
This figure shows the values of principal component 1 and 2 for each subject to indicate the general variability of the gene expression values across subjects.

Supplemental Figure 3. Heatmap for the association of known gene expression confounding variables and each principal component



The y-axis shows various demographic and technical variables. The x-axis shows the principal components included in the differential expression analysis model. The intensity of the heatmap represents the association of each variable with each PC, as measured by -log10 p values for these associations.

Supplemental Figure 4. Plot of the first two principal components of the gene expression data in the replication data set

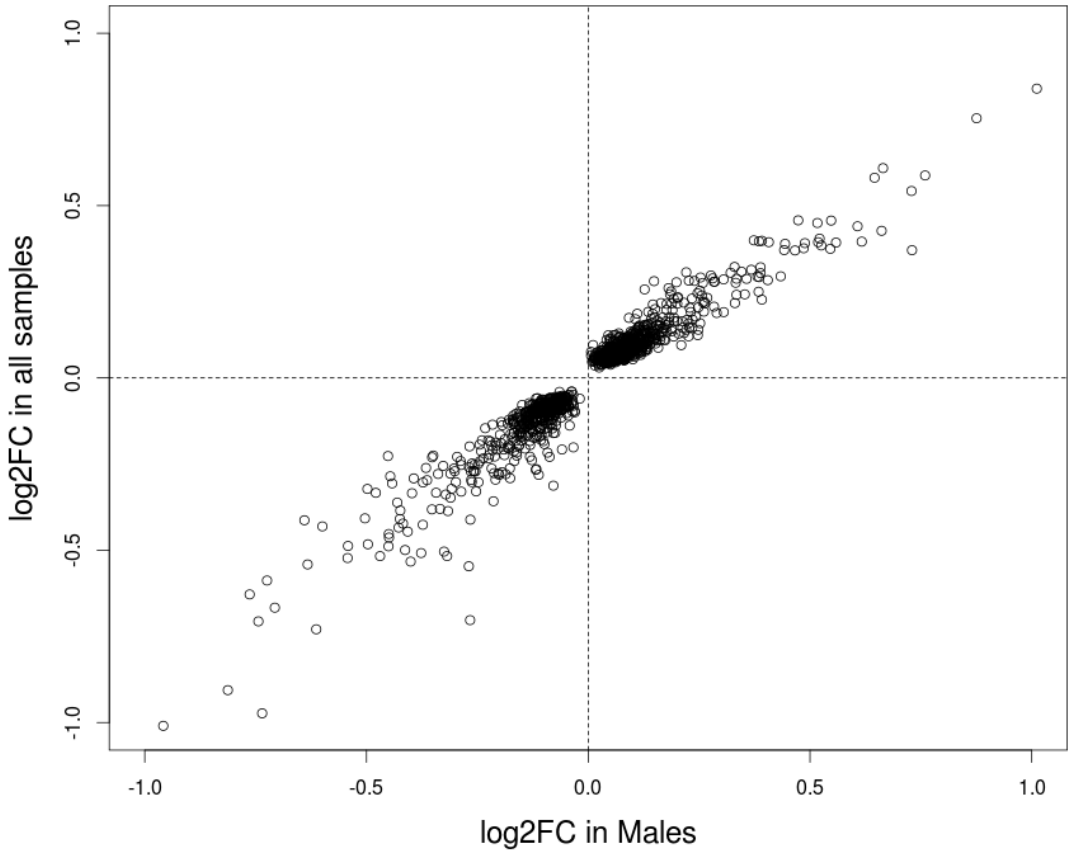


This figure shows the values of principal component 1 and 2 for each subject to indicate the general variability of the gene expression values across subjects in the replication data set.

Supplemental Figure 5. Heatmap for the association of known gene expression confounding variables and each principal component in the replication sample

The y-axis shows various demographic and technical variables. The x-axis shows the principal components included in the differential expression analysis model. The intensity of the heatmap represents the association of each variable with each PC, as measured by - log10 p values for these associations.

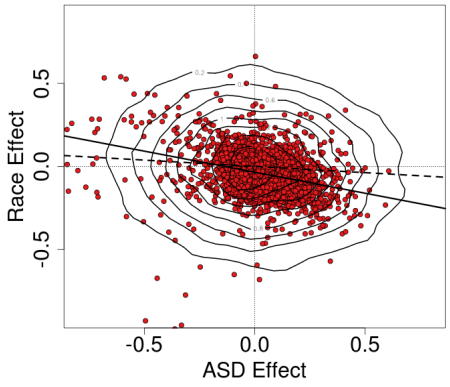
Supplemental Figure 6. Sensitivity analysis in male subjects vs full data set



The log2 fold changes of gene expression between Autism Spectrum Disorder (ASD) subjects and control subjects are shown on the y-axis for all Ensembl genes evaluated in this differential expression analysis for all marginally significant genes, while the x-axis shows the log2 fold change values of the same genes in an analysis using only the male subjects. Consistent log2 fold changes were identified, in line with the sex-balanced study design.

Supplementary Figure 7. Analysis of confounder correction using covariates and PCs

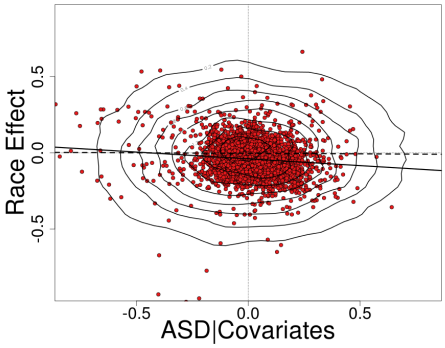
Explicitly including confounders in the model does not fully capture the effect of these confounders on the analysis. The use of PCs however more sufficiently corrects for the influence of known and unknown confounders.



Influence of race on uncorrected model

Plotting the log fold change of the univariate model for the effect of Race on gene expression on the y-axis and the univariate model for the effect of diagnosis on the x-axis demonstrates that Race greatly influences the effect of diagnosis on gene expression when there is no correction for confounds. The relationship between the two models is highly correlated (r = -0.33) and represented by the solid black line The dotted line and the contour lines represent the relationship expected by chance with random data.

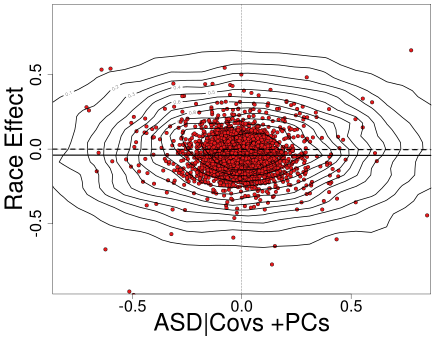
Model: Expression ~ Diagnosis



Influence of race on covariate corrected model

Modeling for diagnosis and many known confounding variables, the correlation between the corrected model and the effect of race is decreased but remains somewhat correlated (r = -0.11).

Model: Expression ~ Diagnosis + Race + RIN +Age + Sex + Exonic Mapping Rate

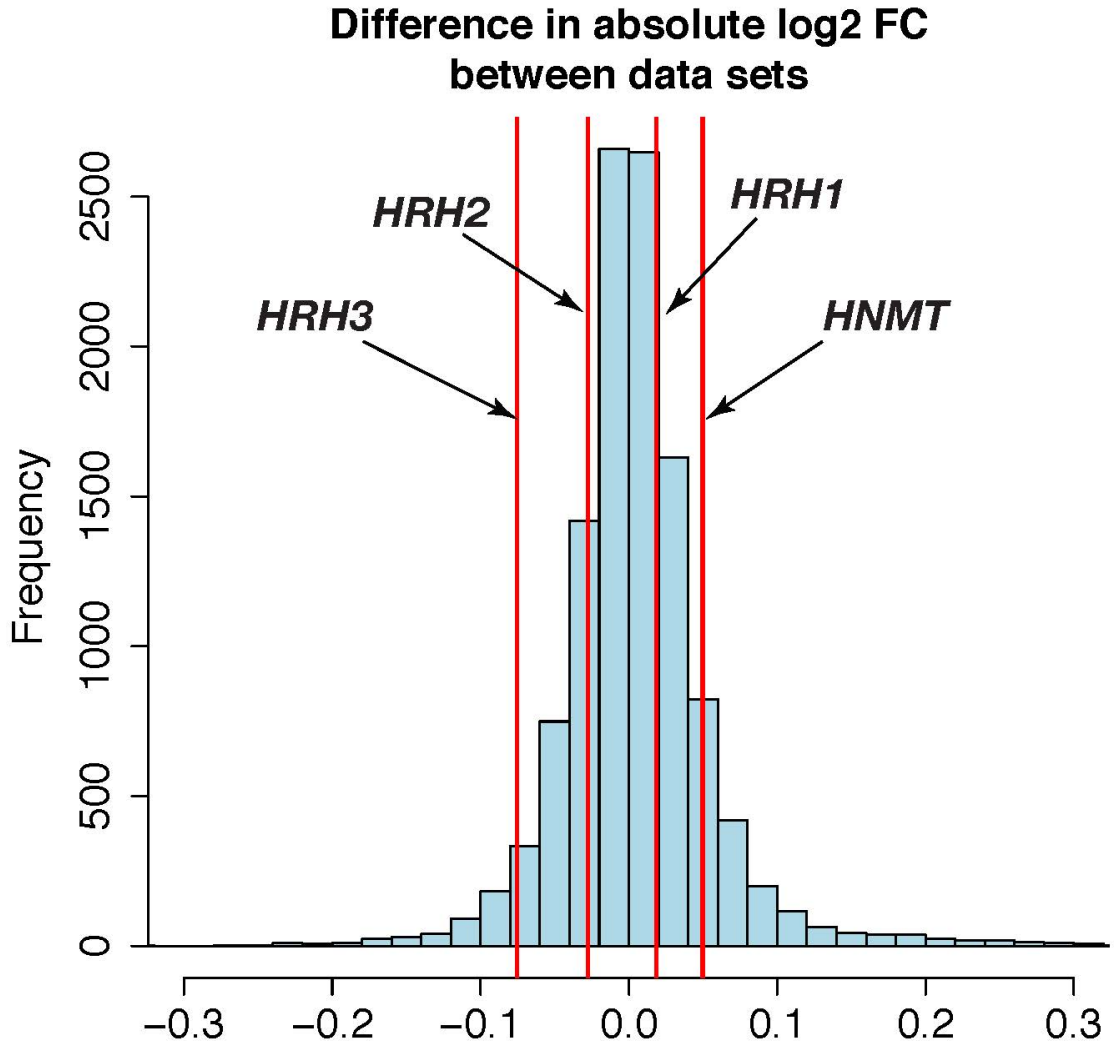


Influence of race on diagnosis model with number of PCs determined by num.sv and covariate correction

Modeling for the effect of diagnosis, covariates, and PCs, greatly reduces the correlation of the corrected model and the general effect of race on gene expression. (r = -0.001)

Model: Expression ~ Diagnosis + Race + RIN +Age +Sex + Exonic Mapping Rate +$\sum_i^9 PC_i$

Supplementary Figure 8. Difference in the absolute log2 fold changes of gene expression between the original data set and the replication data set



The difference in the identified absolute log2 fold changes between our data set and the replication data set for all Ensembl genes evaluated within both differential expression analyses are shown on the x-axis, while the y-axis shows the frequency, or number of genes with each log2 fold change value. The x-axis shown is limited to changes between -0.3 to 0.3 to allow easier evaluation of the genes of interest. The red lines indicate the difference in the absolute log2 fold changes identified for the genes of interest.