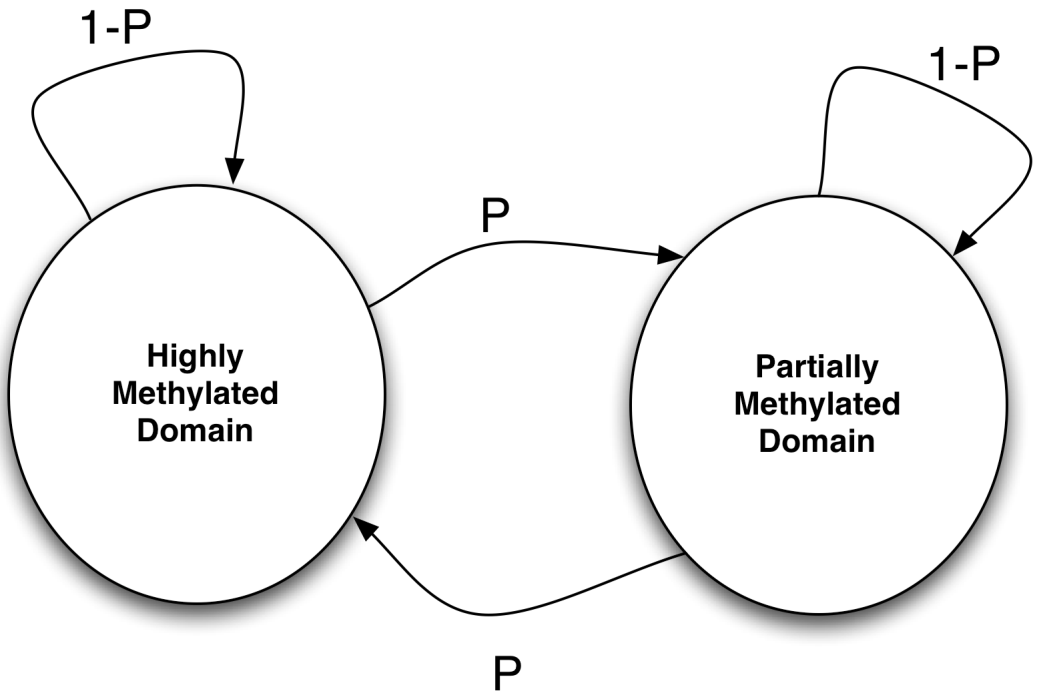
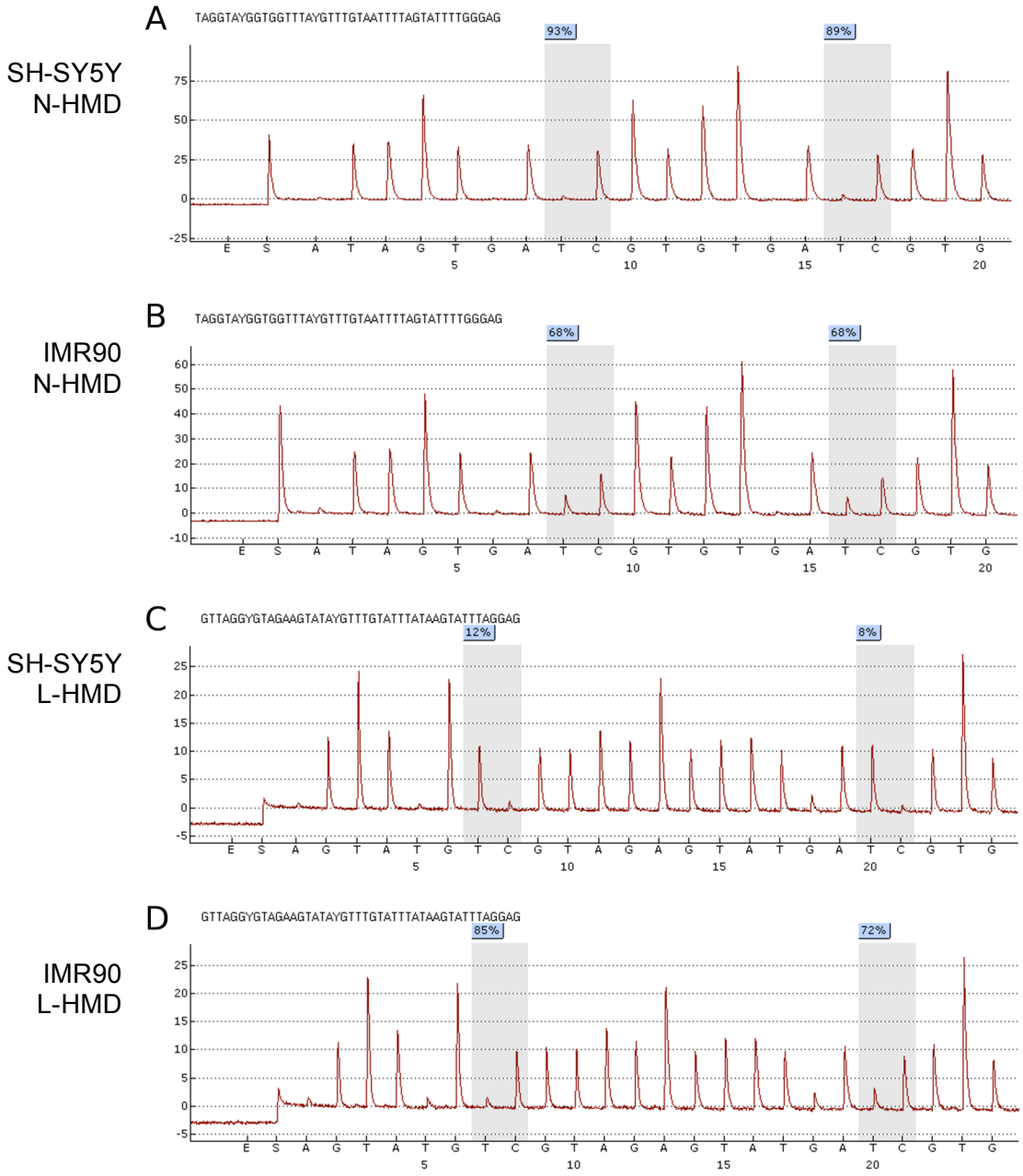
Supplementary Data

Supplemental Figure 1. Methylation hidden Markov model



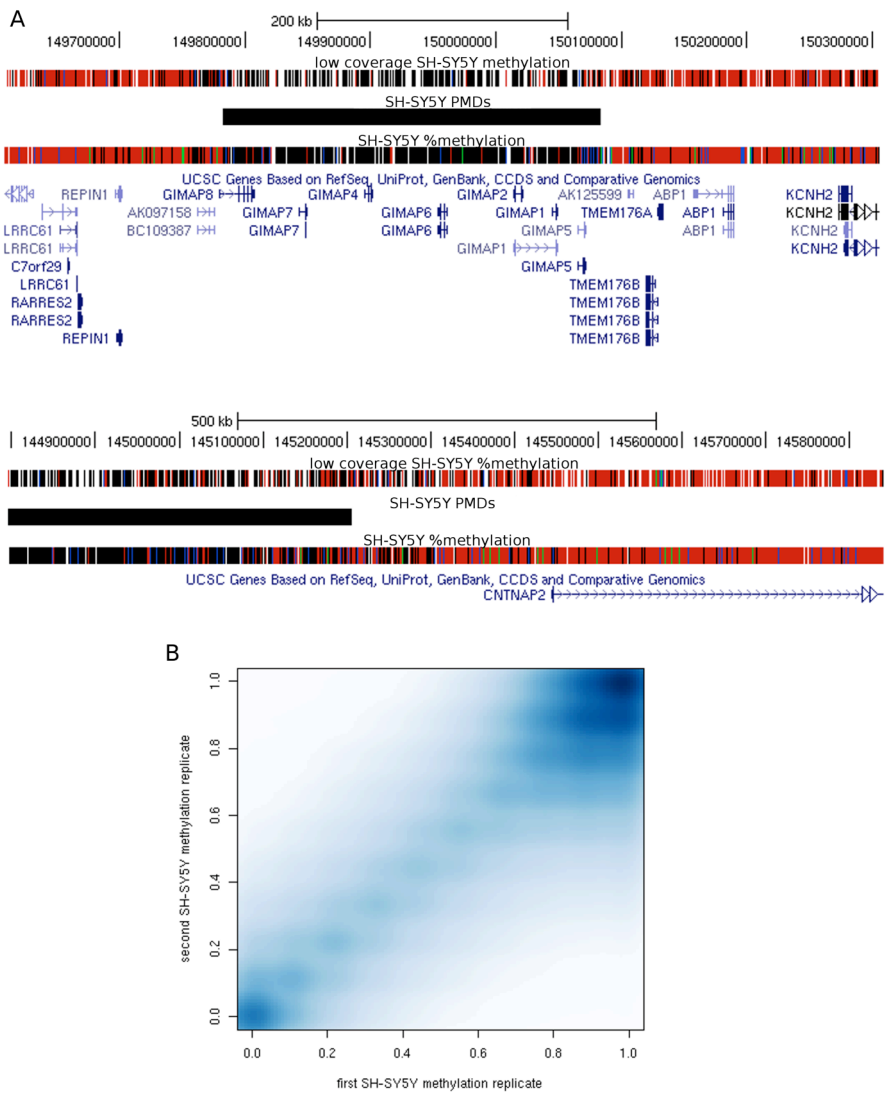
Circles represent the states of the model. Transitions are represented by the arrows. P is the transition probability between states (10^{-30}). While 1-P is the complement of P and represents the probability of the transition from the state to itself.

Supplemental Figure 2. Pyrosequencing validation of tissue-specific methylation differences of genetic loci within N-HMDs and L-HMDs



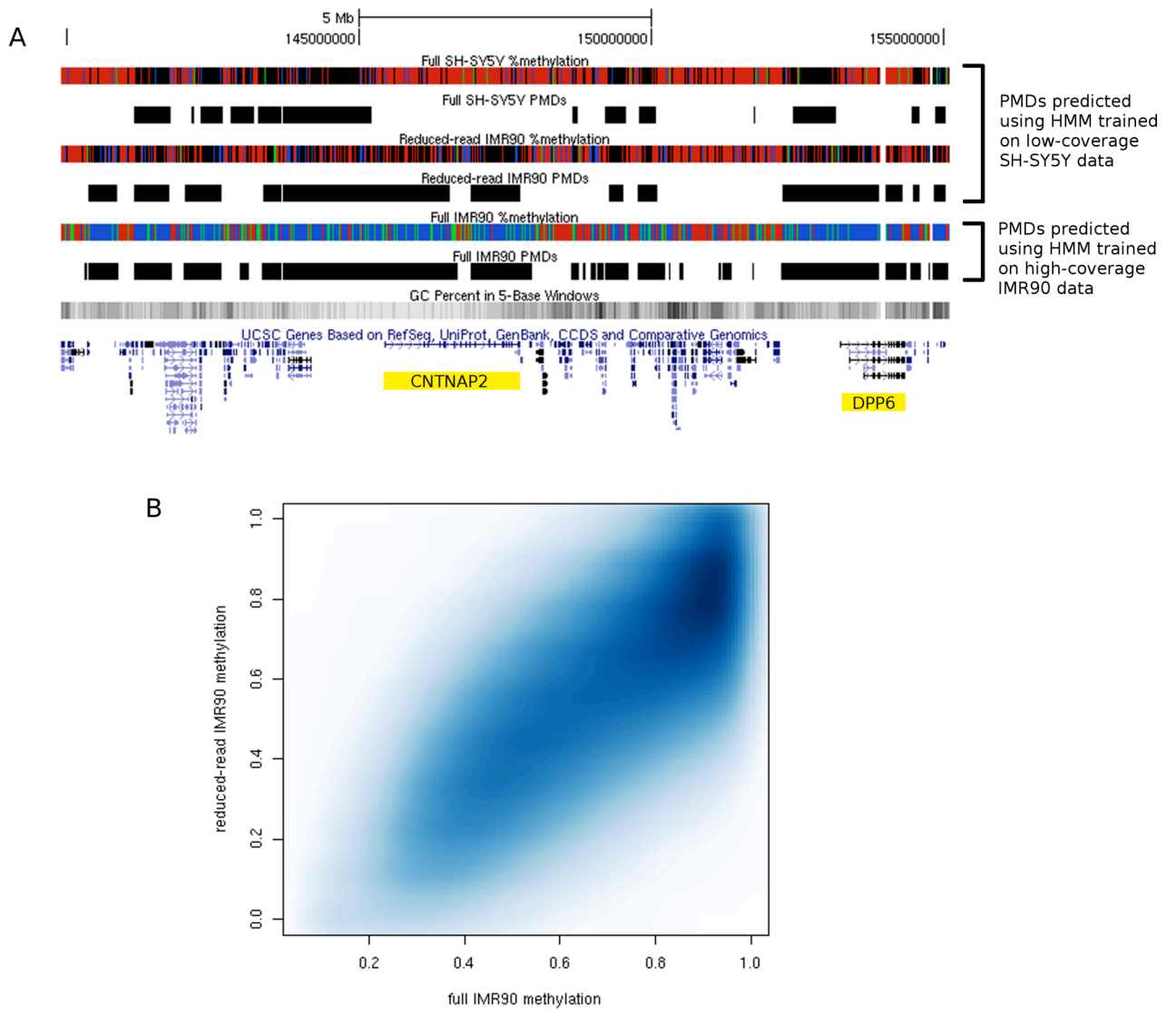
Pyrograms for DNA methylation at two genomic regions in SH-SY5Y and IMR90 cells. Gray shaded areas show CpG sites. Within the gray shaded boxes C denotes signal from methylated cytosines and T denotes signal from unmethylated cytosines. Blue boxes show percent methylation for that CpG site. (A-B) Pyrograms for N-HMD site just downstream of *CNTNAP2* (chr7:147,788,348-147,788,613). SH-SY5Y cells show 91% average methylation and IMR90 cells show 68% average methylation. (C-D) Pyrograms for L-HMD site between the *CAV1* and *MET* (chr7:116,073,633-116,073,874). SH-SY5Y cells show 10% average methylation and IMR90 cells show 79% average methylation.

Supplemental Figure 3. Reproducibility of MethylC-seq biological replicates



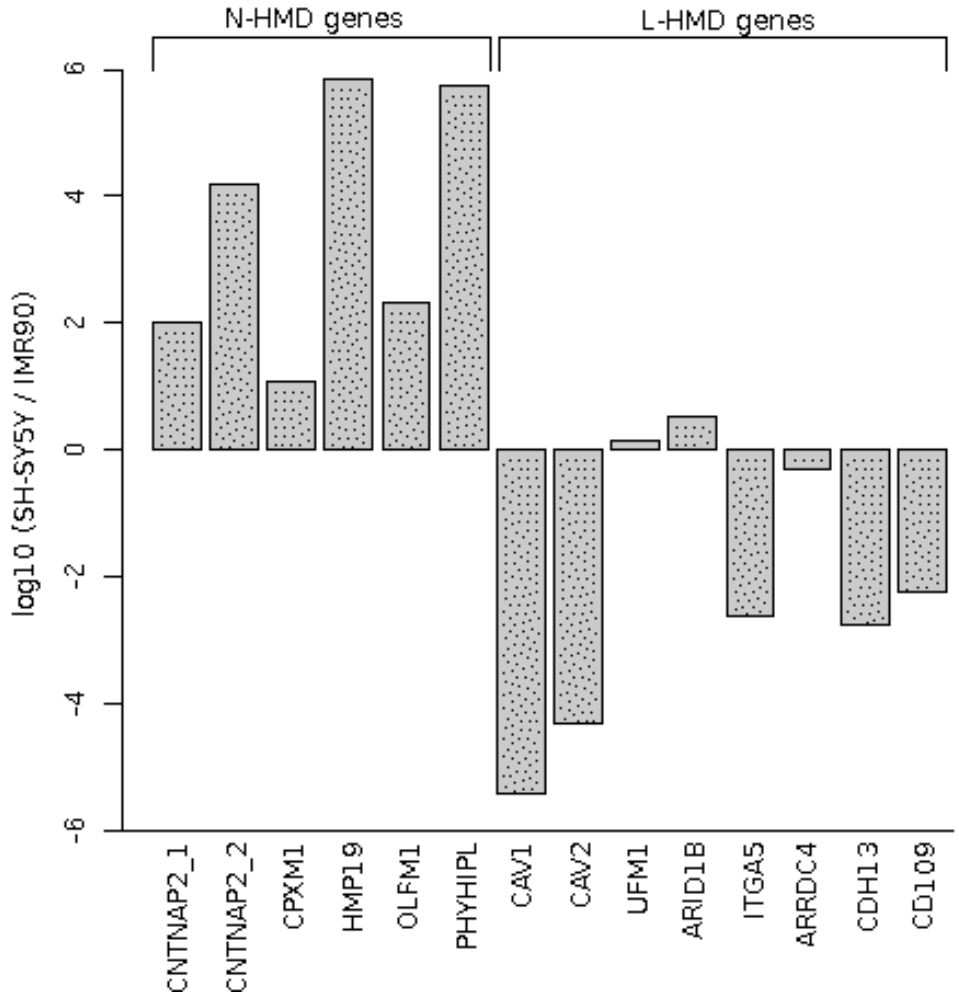
(A) Two examples on chromosome 7 showing that PMDs can be observed with very low sequencing coverage and the SH-SY5Y MethylC-seq data was reproducible. Top track shows MethylC-seq percent methylation data for SH-SY5Y cells using a single Illumina GAII sequencing lane on a biological replicate. The second track shows PMD as defined by the HMM on the original SH-SY5Y MethylC-seq data. The third track shows the original percent methylation data used in the study (from two Illumina GAII sequencing lanes). (B) Visualization of correlation between IMR90 methylation datasets after smoothing. For smoothing, means were calculated over running windows of nine CpG sites. Spearman rank correlation rho = .71. For unsmoothed datasets, Spearman rank correlation rho = .63. Since PMD detection by HMM uses binned methylation ranges rather than precise percentiles, replication of PMD locations is possible even without absolute correlation of percent methylation between the data sets.

Supplemental Figure 4. PMD detection in IMR90 reduced-read dataset and methylation correlation.



25.4 million alignable reads (similar to the 26.7 alignable reads in the SH-SY5Y full dataset) were randomly chosen from the full IMR90 dataset to look at the effects of reduced coverage on PMD detection and DNA methylation correlation. For the reduced-read IMR90 data, PMDs were detected using the HMM trained on the full SH-SY5Y dataset. 97% of the genomic length covered by IMR90 reduced-read PMDs were also in the full IMR90 PMDs. However, the sensitivity of PMD detection was reduced with only 66% of the genomic sequence covered by full IMR90 PMDs also covered by the reduced-read PMDs. The fact that the SH-SY5Y HMM was trained on PMDs with a lower average percent methylation than IMR90 PMDs (Figure 1B) could explain some of this lost sensitivity. (A) Browser view of the unsmoothed methylation levels across the CNTNAP2 locus in two IMR90 datasets. Below the methylation tracks are the corresponding PMDs detected by the HMMs. (B) Visualization of correlation between IMR90 methylation datasets after smoothing. For smoothing, means were calculated over running windows of nine CpG sites. Spearman rank correlation rho = .76. For unsmoothed datasets, Spearman rank correlation rho = .50.

Supplemental Figure 5. qRT-PCR validation of gene expression differences in N-HMDs and L-HMDs.

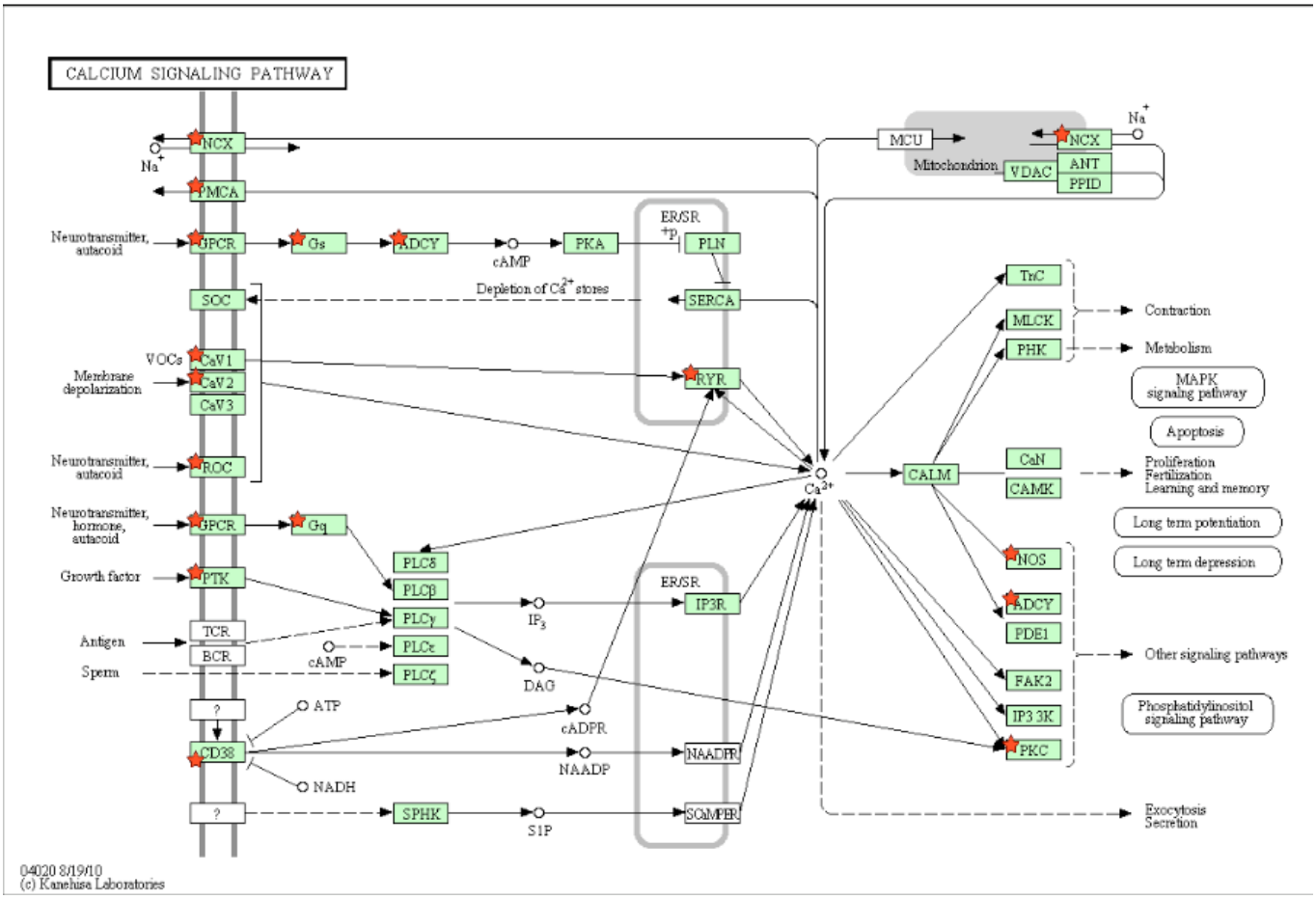


Gene expression was assessed for *CAV1*, *CAV2*, and two primer pairs to *CNTNAP2*. In addition we chose a random set of N-HMD and L-HMD genes that 1) had consistent microarray probe values across the gene in both tissues, 2) had microarray probe values above 6.1 (the 75th percentile) in at least one of the cell lines, and 3) were differentially expressed in the two cell lines after Benjamini correction for multiple hypothesis testing.

qRT-PCR results are the average of triplicates and were normalized in both cell lines to *GAPDH*.

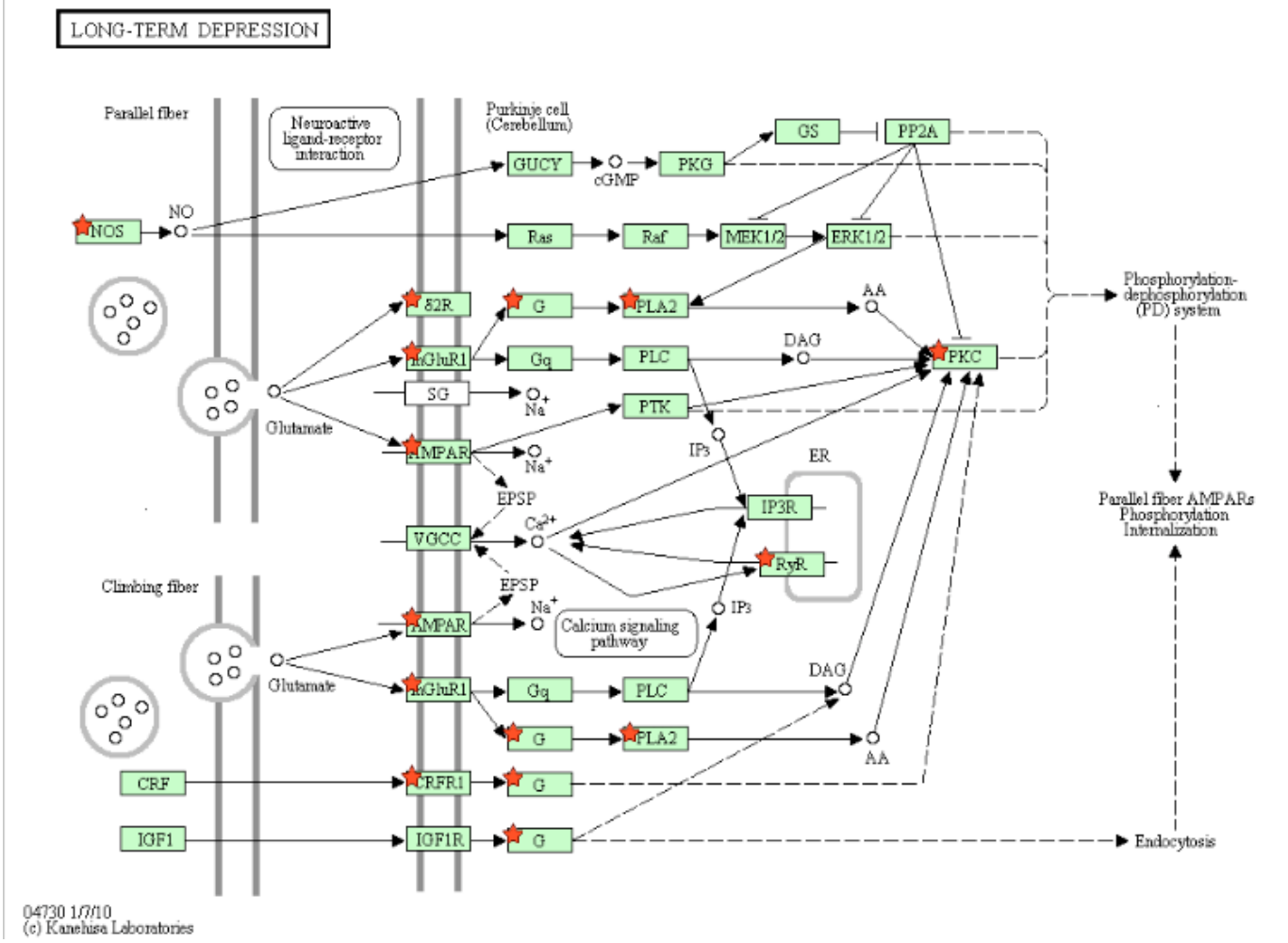
Bars above 0 show higher expression in SH-SY5Y cells and bars below 0 show higher expression in IMR90 cells.

Supplemental Figure 6. Calcium signaling KEGG pathway



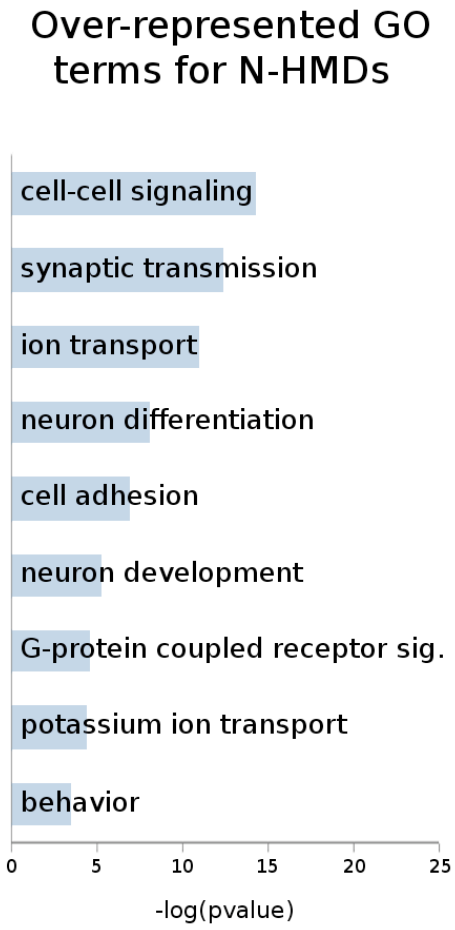
Red stars represent genes present in the N-HMD gene list. Figure is from KEGG and DAVID.

Supplemental Figure 7. Long-term depression KEGG pathway



Red stars represent genes present in the N-HMD gene list. Figure is from KEGG and DAVID.

Supplemental Figure 8. Compared to neuronally expressed genes, N-HMDs are enriched for genes with specific functions in neuronal differentiation and synaptic function.



Human cortex expression data was downloaded from GEO and 7,556 neuronally expressed genes were extracted by selecting all genes with expression levels above the 75th percentile. These were used as a background gene set in DAVID to find GO terms over-represented in the N-HMD genes.

Supplemental Table 1. GO terms for genes in domains

(Excel file SuppTable1.xls)

Supplemental Table 2. List of autism implicated and candidate genes found in N-HMDs from Table 1, grouped by function

Neuronal adhesion	*ASTN1, ASTN2, CDH22, CNTN4, CNTNAP2, NRXN1*
Receptor/drug binding	*CNR1, DRD3, ESRRB, GABRA5, GABRB3, GRID1, GRIK2, GRIN2A, GRM8, HTR2A, HTR3A, NTRK1, NTRK3, OPRM1, SLC1A1*
Ion channel	*CACNA1D, SCN2A*
Axon guidance/migration	*DAB1, NTNG1, RELN, ROBO1*
Involved in neuronal structure	*DLGAP2, MAP2, MARK1, SHANK1, SHANK2*
Other signal transduction	*GNAS*
Other synaptic transport	*MYO16*
Other transmembrane protein	*DPP10, DPP6, SYT17, TMEM195*
Other synapse-localized protein	*RIMS3*
Kinase	*DAPK1, FRK, NTRK1, NTRK3, PAK7, PRKAG2, PRKCB*
Metabolism enzyme	*ALDH5A1, CYP11B1, FHIT, GALNT13, TPH2*
Other enzyme	*PAPPA2, PARK2, PON1*
Other or unknown	*AUTS2, CCDC64, DLX6, DUOXA1, FBXP40*

Genes only in Pinto et al. 2010 autism candidate and implicated lists

Genes only in SFARI database

Autism genes in both

Supplemental Table 3. List of genes in N-HMDs, L-HMDs, and B-PMDs

(Excel file SuppTable3.xls)