Supplemental Data

# Genetic Control of Human Brain Transcript Expression in Alzheimer Disease

**Jennifer A. Webster, J. Raphael Gibbs, Jennifer Clarke, Monika Ray, Weixiong Zhang, Peter Holmans, Kristen Rohrer, Alice Zhao, Lauren Marlowe, Mona Kaleem, Donald S. McCorquodale III, Cindy Cuello, Doris Leung, Leslie Bryden, Priti Nath, Victoria L. Zismann, Keta Joshipura, Matthew J. Huentelman, Diane Hu-Lince, Keith D. Coon, David W. Craig, John V. Pearson, NACC-Neuropathology Group, Christopher B. Heward, Eric M. Reiman, Dietrich Stephan, John Hardy, and Amanda J. Myers**

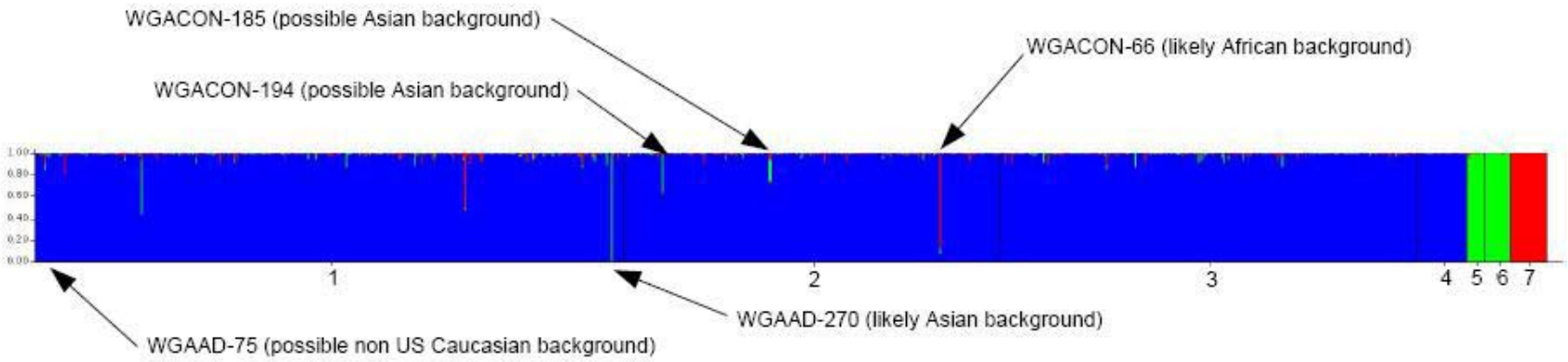
## Supplemental Acknowledgments

The directors, pathologists and technicians involved in this research include: **National Institute on Aging**: Ruth Seemann, **John Hopkins Alzheimer's Disease Research Center** (NIA grant AG05146): Juan C. Troncoso, MD, Dr. Olga Pletnikova, **University of California, Los Angeles** (NIA grant P50 AG16570): Harry Vinters, MD, Justine Pomakian, **The Kathleen Price Bryan Brain Bank, Duke University Medical Center** (NIA grant AG05128, NINDS grant NS39764, NIMH MH60451 also funded by Glaxo Smith Kline): Christine Hulette, MD, John F. Ervin, **Stanford University**: Dikran Horoupian, MD, Ahmad Salehi, MD, PhD, **New York Brain Bank, Taub Institute, Columbia University** (NYBB): Jean Paul Vonsattel, MD, Katerina Mancevska, **Massachusetts General Hospital:** E. Tessa Hedley-Whyte, MD, Karlotta Fitch, **University of Michigan** (NIH grant P50-AG08671): Roger Albin, MD, Andrew Lieberman MD, Lisa Bain, Eszter Gombosi, **University of Kentucky** (NIH grant AG05144): William Markesbery, MD, Sonya Anderson, **Mayo Clinic, Jacksonville**: Dennis W. Dickson, MD, Natalie Thomas, **University Southern California**: Caroll A. Miller, MD, Jenny Tang, M.S., Dimitri Diaz, **Washington University, St Louis Alzheimer's Disease Research Center** (NIH grant P50AG05681)**:** Dan McKeel, MD, John C. Morris, MD, Eugene Johnson, Jr., PhD, Virginia Buckles, PhD, Deborah Carter, **University of Washington, Seattle** (NIH grant P50 AG05136):Thomas Montine, MD, PhD, Aimee Schantz, MEd., **University of Pennsylvania School of Medicine, Alzheimer's Disease Research Center**: John Q Trojanowski, MD, Virginia M Lee, MD, Vivianna Van Deerlin, MD, Terry Schuck, **Boston University Alzheimer's Disease Research Center** (NIH grant P30-AG13846)**:** Ann C. McKee, MD, Carol Kubilus, **Sun Health Research Institute, Arizona** (NIA grant P30 AG19610**):** Joseph Rogers, PhD, Thomas G. Beach, MD, PhD, Lucia I. Sue, **Emory University**: Bruce H. Wainer, MD, PhD, Marla Gearing, PhD, **University of Texas, Southwestern Medical School:** Charles L. White, III, M.D., Roger Rosenberg, Marilyn Howell, Joan Reisch, **University of California, Davis:** William Ellis, MD, Mary Ann Jarvis, **Rush University Medical Center, Rush Alzheimer's Disease Center** (NIH grant AG10161)**:** David A. Bennett, M.D. Julie A. Schneider, MD, MS, Karen Skish, MS, PA (ASCP)MT, Wayne T Longman, **University of Miami/NPF Brain Endowment Bank:** Deborah C. Mash, MD, Margaret J Basile, Mitsuko Tanaka.

This research was supported in part by the Intramural Research Program of the NIH, National Institute on Aging (project Z01 AG000949-02). These studies were supported by Kronos Sciences Laboratory, the Verum foundation, the Bisgrove charitable donation, the NIH Neuroscience Blueprint (U24NS051872), the ENDGAME Consortium (UO1HL084744), the MRC, the Brain Research Trust, the Reta Lila Trust and the state of Arizona.

We also thank Dr. Edward B. Suh and Mr. James Lowey at TGen for the use of the ASU-TGen cluster-based supercomputer. AJM would like to thank the Johnnie B. Byrd Institute for support. None of the sponsors were involved in the design or conduct of the study, in the collection, analysis, and interpretation of the data, or in the preparation, review, or approval of the manuscript. There are no conflicts of interest for any of the authors.

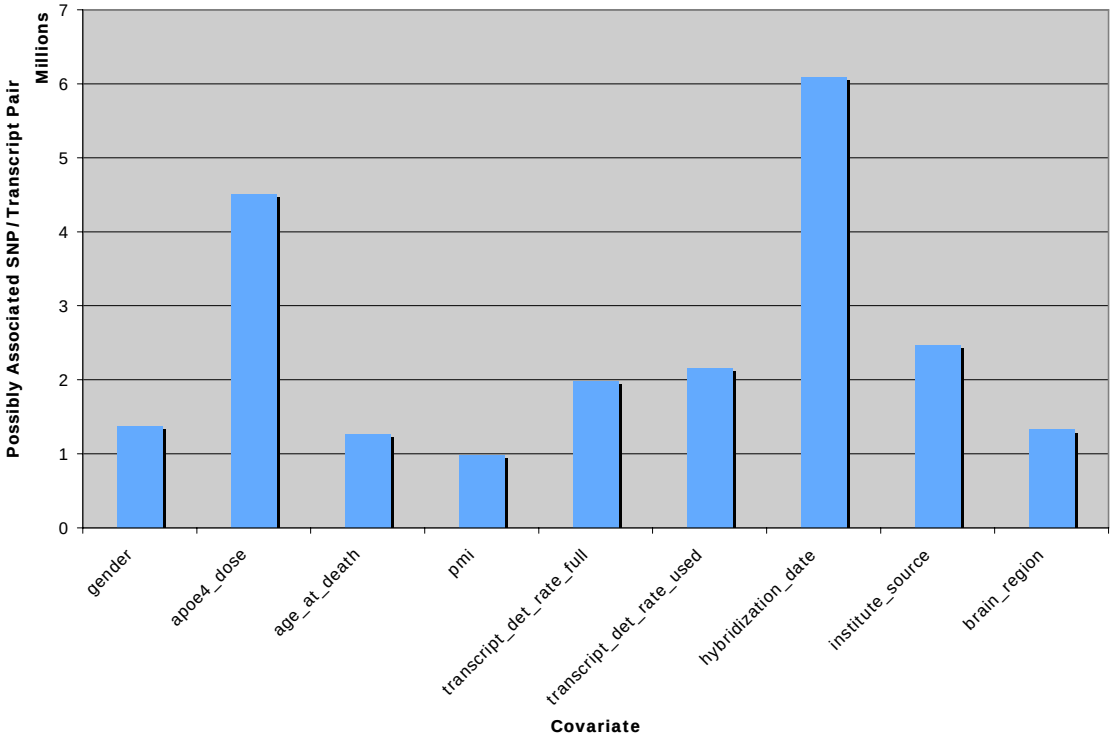
**Figure S1**. **Analysis of Population Structure within Cohort**




1) 383 US Caucasian AD Case Samples, our 181 is a subset of the group (Affy 500K)
2) 243 US Caucasian Control Samples, our 201 is a subset of this group (Affy 500K)
3) 271 US Caucasian Control from Coriell Cell Line Repositories (Illumina 550K)
4) 32 unrelated HapMap CEU samples (Illumina 550K)
5) 12 unrelated HapMap CHB samples (Illumina 550K)
6) 16 unrelated HapMap JPT samples (Illumina 550K)
7) 24 unrelated HapMap YRI samples (Illumina 550K)

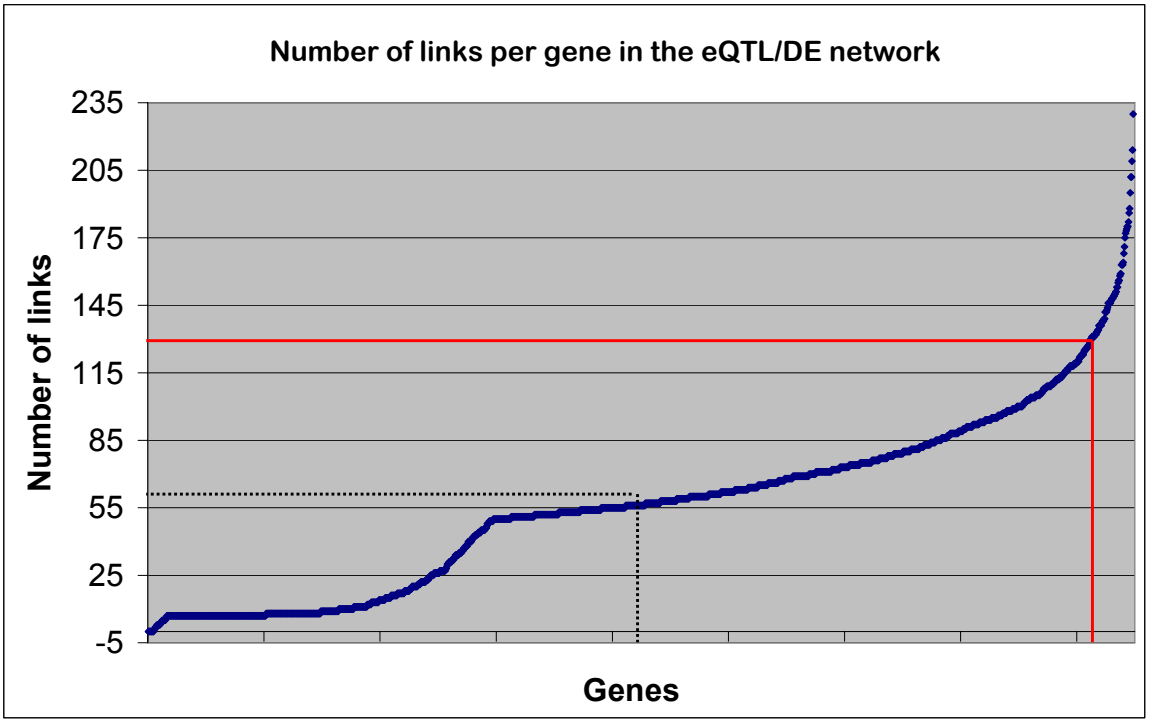
Clusters were calculated for genotype data using the program STRUCTURE[1,2] for 7 cohorts to examine ethnic bias within our series. On the figure blue corresponds to primarily European origin, green is primarily Asian descent and red corresponds to samples that are primarily African descent. As is seen on the figure there were 2 ethnic outliers within the LOAD population used for this study (WGAAD-75 and WGAAD-270, bins 1 and 2 on the figure) and 3 within the control population (WGCON-185, WGACON-194 and WGACON-66, bin 2 on the figure).

**Figure S2. Significant Transcript Effects for Each Covariate**




Graphed are the numbers of associated SNP/transcript pairs from the initial correction regression using R. Only pairs that reached possible significance (uncorrected p-value, $\alpha \leq 0.05$) for each covariate are graphed.

**Figure S3. Determination of Hub Genes**




The graph plots the number of links for all of the genes within the DE/eQTL network we mapped (n=1697). The X-axis plots the genes in ascending order of the number of links (range=1-230). The Y-axis plots the number of links for each gene. The dashed line indicates the mean number of links, and the red line indicates the hub gene cutoff, with any gene having greater than 134 links defined as a hub gene.

**Table S1. Sample Description**

| | control cohort | case cohort |
|---|---|---|
| % female | 45% | 50% |
| average age | 81 | 84 |
| age range | 65-100 | 68-102 |
| % frontal cortex | 21% | 18% |
| % temporal cortex | 73% | 60% |
| % parietal cortex | 2% | 10% |
| % cerebellar cortex | 3% | 13% |
| # of different hybridization dates | 9 | 9 |
| average # of samples per hybridization | 21 | 19 |
| number of institute sources | 18 | 16 |
| average number of samples per institute | 11 | 11 |
| average PMI | 10 hrs | 9 hrs |
| average transcriptome detection rate per sample | 77% | 76% |

Listed are descriptive statistics for the samples used.

**Table S2. Top Ontology Groupings for Each Cluster within Our eQTL/DE Network**

Table S2A

| C | Score | Term | Process | Count | P_Value |
|---|---|---|---|---|---|
| 1 | 1.72 | GOTERM_BP_ALL | organ morphogenesis | 8 | 2.90E-03 |
| | | GOTERM_BP_ALL | developmental process | 31 | 6.40E-03 |
| | | GOTERM_BP_ALL | organ development | 14 | 1.10E-02 |
| | | GOTERM_BP_ALL | multicellular organismal development | 22 | 1.70E-02 |
| | | GOTERM_BP_ALL | system development | 18 | 2.40E-02 |
| | | GOTERM_BP_ALL | anatomical structure development | 20 | 4.60E-02 |
| | | GOTERM_BP_ALL | anatomical structure morphogenesis | 12 | 4.90E-02 |
| | | GOTERM_BP_ALL | multicellular organismal process | 24 | 9.10E-02 |
| 2 | 0.99 | INTERPRO | Zinc finger, C2H2-type | 8 | 4.10E-03 |
| | | GOTERM_MF_ALL | ion binding | 23 | 5.30E-03 |
| | | INTERPRO | KRAB box | 5 | 6.60E-03 |
| | | SMART | KRAB | 5 | 6.60E-03 |
| | | GOTERM_MF_ALL | cation binding | 21 | 7.00E-03 |
| | | GOTERM_MF_ALL | metal ion binding | 22 | 1.00E-02 |
| | | GOTERM_MF_ALL | zinc ion binding | 14 | 2.40E-02 |
| | | PIR_SUPERFAMILY | PIRSF005559:zinc finger protein ZFP-36 | 3 | 2.80E-02 |
| | | GOTERM_BP_ALL | RNA metabolic process | 16 | 4.40E-02 |
| | | COG_ONTOLOGY | General function prediction only | 8 | 5.30E-02 |
| | | INTERPRO | Zinc finger, C2H2-subtype | 4 | 5.40E-02 |
| | | SP_PIR_KEYWORDS | zinc-finger | 10 | 5.70E-02 |
| | | SP_PIR_KEYWORDS | metal-binding | 14 | 5.80E-02 |
| | | GOTERM_BP_ALL | nucleobase, nucleoside, nucleotide and nucleic acid metabolic process | 19 | 6.10E-02 |
| | | SP_PIR_KEYWORDS | zinc | 11 | 7.90E-02 |
| | | SP_PIR_KEYWORDS | Transcription regulation | 10 | 7.90E-02 |
| | | SP_PIR_KEYWORDS | dna-binding | 9 | 9.40E-02 |
| | | SP_PIR_KEYWORDS | Transcription | 10 | 9.70E-02 |
| | | GOTERM_MF_ALL | transition metal ion binding | 14 | 9.80E-02 |
| | | GOTERM_BP_ALL | gene expression | 17 | 1.10E-01 |
| | | GOTERM_BP_ALL | regulation of transcription, DNA-dependent | 11 | 1.40E-01 |
| | | SP_PIR_KEYWORDS | nucleus | 18 | 1.60E-01 |

| | | | | | |
|---|---|---|---|---|---|
| | | GOTERM_BP_ALL | RNA biosynthetic process | 11 | 1.70E-01 |
| | | GOTERM_BP_ALL | transcription, DNA-dependent | 11 | 1.70E-01 |
| | | GOTERM_BP_ALL | regulation of transcription | 11 | 2.00E-01 |
| | | GOTERM_BP_ALL | regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process | 11 | 2.20E-01 |
| | | GOTERM_BP_ALL | transcription | 11 | 2.40E-01 |
| | | GOTERM_BP_ALL | regulation of metabolic process | 12 | 2.60E-01 |
| | | GOTERM_BP_ALL | metabolic process | 32 | 2.70E-01 |
| | | GOTERM_BP_ALL | regulation of gene expression | 11 | 2.70E-01 |
| | | GOTERM_BP_ALL | regulation of cellular metabolic process | 11 | 3.30E-01 |
| | | GOTERM_MF_ALL | nucleic acid binding | 14 | 3.80E-01 |
| | | GOTERM_MF_ALL | DNA binding | 9 | 4.30E-01 |
| | | GOTERM_BP_ALL | macromolecule metabolic process | 25 | 4.40E-01 |
| | | GOTERM_BP_ALL | cellular metabolic process | 28 | 4.50E-01 |
| | | GOTERM_BP_ALL | biopolymer metabolic process | 19 | 4.50E-01 |
| | | GOTERM_BP_ALL | primary metabolic process | 28 | 4.80E-01 |
| | | GOTERM_BP_ALL | regulation of cellular process | 14 | 6.30E-01 |
| | | GOTERM_BP_ALL | regulation of biological process | 15 | 6.40E-01 |
| | | GOTERM_BP_ALL | biological regulation | 16 | 6.60E-01 |
| | | GOTERM_BP_ALL | cellular process | 38 | 6.90E-01 |
| 3 | 1.12 | GOTERM_BP_ALL | cell growth | 5 | 2.40E-02 |
| | | GOTERM_BP_ALL | regulation of cell size | 5 | 2.60E-02 |
| | | GOTERM_BP_ALL | growth | 5 | 5.20E-02 |
| | | GOTERM_BP_ALL | regulation of cell growth | 4 | 5.80E-02 |
| | | GOTERM_BP_ALL | anatomical structure morphogenesis | 9 | 8.30E-02 |
| | | GOTERM_BP_ALL | regulation of growth | 4 | 8.40E-02 |
| | | GOTERM_BP_ALL | cellular structure morphogenesis | 6 | 1.10E-01 |
| | | GOTERM_BP_ALL | cell morphogenesis | 6 | 1.10E-01 |
| | | GOTERM_BP_ALL | regulation of biological quality | 7 | 2.00E-01 |
| | | GOTERM_BP_ALL | anatomical structure development | 13 | 2.00E-01 |
| 4 | 2.05 | GOTERM_BP_ALL | negative regulation of transcription, DNA-dependent | 13 | 2.40E-03 |
| | | GOTERM_BP_ALL | negative regulation of transcription | 16 | 2.50E-03 |
| | | SP_PIR_KEYWORDS | repressor | 17 | 2.70E-03 |
| | | GOTERM_BP_ALL | negative regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process | 16 | 4.60E-03 |
| | | GOTERM_MF_ALL | transcription repressor activity | 13 | 5.10E-03 |

| | | | | | |
|---|---|---|---|---|---|
| | | GOTERM_BP_ALL | negative regulation of cellular metabolic process | 17 | 9.90E-03 |
| | | GOTERM_BP_ALL | negative regulation of cellular process | 36 | 1.40E-02 |
| | | GOTERM_BP_ALL | negative regulation of metabolic process | 18 | 1.70E-02 |
| | | GOTERM_BP_ALL | negative regulation of biological process | 36 | 2.10E-02 |
| | | GOTERM_BP_ALL | negative regulation of transcription from RNA polymerase II promoter | 6 | 1.80E-01 |
| 5 | 2.06 | GOTERM_BP_ALL | protein kinase cascade | 9 | 2.20E-03 |
| | | GOTERM_BP_ALL | regulation of signal transduction | 10 | 2.80E-03 |
| | | GOTERM_BP_ALL | MAPKKK cascade | 5 | 1.80E-02 |
| | | GOTERM_BP_ALL | positive regulation of signal transduction | 4 | 5.20E-02 |
| 6 | 2.11 | INTERPRO | KRAB box | 9 | 5.50E-05 |
| | | PIR_SUPERFAMILY | PIRSF005559:zinc finger protein ZFP-36 | 6 | 6.20E-05 |
| | | SMART | KRAB | 9 | 1.50E-04 |
| | | SP_PIR_KEYWORDS | dna-binding | 22 | 3.30E-04 |
| | | UP_SEQ_FEATURE | zinc finger region:C2H2-type 11 | 6 | 1.20E-03 |
| | | UP_SEQ_FEATURE | zinc finger region:C2H2-type 10 | 6 | 1.90E-03 |
| | | UP_SEQ_FEATURE | zinc finger region:C2H2-type 9 | 6 | 3.30E-03 |
| | | INTERPRO | Zinc finger, C2H2-type | 11 | 3.70E-03 |
| | | INTERPRO | Zinc finger, C2H2-subtype | 7 | 4.30E-03 |
| | | UP_SEQ_FEATURE | zinc finger region:C2H2-type 12 | 5 | 5.00E-03 |
| | | UP_SEQ_FEATURE | zinc finger region:C2H2-type 8 | 6 | 5.30E-03 |
| | | UP_SEQ_FEATURE | zinc finger region:C2H2-type 7 | 6 | 6.70E-03 |
| | | COG_ONTOLOGY | General function prediction only | 16 | 7.50E-03 |
| | | UP_SEQ_FEATURE | domain:KRAB | 5 | 9.70E-03 |
| | | UP_SEQ_FEATURE | zinc finger region:C2H2-type 6 | 6 | 1.10E-02 |
| | | UP_SEQ_FEATURE | zinc finger region:C2H2-type 1; degenerate | 3 | 1.60E-02 |
| | | UP_SEQ_FEATURE | zinc finger region:C2H2-type 4 | 6 | 2.90E-02 |
| | | UP_SEQ_FEATURE | zinc finger region:C2H2-type 2 | 6 | 4.30E-02 |
| | | UP_SEQ_FEATURE | zinc finger region:C2H2-type 3 | 6 | 4.30E-02 |
| | | INTERPRO | Zinc finger, C2H2-type/integrase, DNA-binding | 5 | 5.80E-02 |
| | | UP_SEQ_FEATURE | zinc finger region:C2H2-type 5 | 5 | 7.10E-02 |
| | | UP_SEQ_FEATURE | zinc finger region:C2H2-type 13 | 3 | 9.70E-02 |
| | | INTERPRO | Zinc finger, C2H2-like | 5 | 1.40E-01 |
| | | SMART | ZnF_C2H2 | 5 | 2.00E-01 |
| | | UP_SEQ_FEATURE | zinc finger region:C2H2-type 1 | 3 | 5.30E-01 |
| 7 | 1.78 | GOTERM_MF_ALL | ligase activity, forming carbon-nitrogen bonds | 12 | 6.40E-03 |

| | | | | | |
|---|---|---|---|---|---|
| | | GOTERM_MF_ALL | ubiquitin-protein ligase activity | 10 | 9.00E-03 |
| | | GOTERM_MF_ALL | small protein conjugating enzyme activity | 10 | 9.60E-03 |
| | | GOTERM_BP_ALL | protein ubiquitination | 7 | 1.10E-02 |
| | | GOTERM_BP_ALL | ubiquitin cycle | 19 | 1.10E-02 |
| | | GOTERM_MF_ALL | small conjugating protein ligase activity | 10 | 1.20E-02 |
| | | GOTERM_BP_ALL | protein modification by small protein conjugation | 7 | 1.30E-02 |
| | | GOTERM_MF_ALL | acid-amino acid ligase activity | 10 | 1.90E-02 |
| | | SP_PIR_KEYWORDS | Ubl conjugation pathway | 16 | 2.40E-02 |
| | | GOTERM_MF_ALL | ligase activity | 14 | 6.20E-02 |
| | | SP_PIR_KEYWORDS | ligase | 10 | 9.40E-02 |
| 8 | 3.02 | SP_PIR_KEYWORDS | mitochondrion | 33 | 8.20E-06 |
| | | SP_PIR_KEYWORDS | transit peptide | 19 | 2.50E-03 |
| | | UP_SEQ_FEATURE | transit peptide:Mitochondrion | 12 | 4.10E-02 |
| 9 | 3.42 | SP_PIR_KEYWORDS | glycoprotein | 49 | 3.90E-06 |
| | | UP_SEQ_FEATURE | glycosylation site:N-linked (GlcNAc...) | 38 | 1.60E-04 |
| | | SP_PIR_KEYWORDS | signal | 36 | 4.00E-04 |
| | | UP_SEQ_FEATURE | signal peptide | 32 | 6.30E-04 |
| | | SP_PIR_KEYWORDS | Secreted | 19 | 7.80E-04 |
| | | UP_SEQ_FEATURE | disulfide bond | 21 | 2.40E-02 |

Table S2B

| Term | Process | C1 | Ct1 | PV1 | C4 | Ct4 | PV4 |
|---|---|---|---|---|---|---|---|
| SP_PIR_KEYWORDS | activator | 3.74 | 14 | 1.30E-05 | 2.61 | 19 | 1.10E-03 |
| GOTERM_BP_ALL | biological regulation | 3.74 | 53 | 1.50E-05 | 2.61 | 124 | 3.90E-04 |
| GOTERM_BP_ALL | biopolymer metabolic process | 3.74 | 54 | 7.20E-05 | 2.61 | 110 | 3.00E-01 |
| GOTERM_BP_ALL | cellular metabolic process | 3.74 | 64 | 1.90E-02 | 2.61 | 154 | 8.00E-01 |
| GOTERM_BP_ALL | cellular process | 3.74 | 81 | 1.50E-01 | 2.61 | 231 | 3.80E-01 |
| GOTERM_MF_ALL | DNA binding | 3.74 | 31 | 1.50E-04 | 2.61 | 59 | 1.30E-03 |
| SP_PIR_KEYWORDS | dna-binding | 3.74 | 26 | 3.30E-05 | 2.61 | 48 | 9.80E-05 |
| GOTERM_BP_ALL | gene expression | 3.74 | 40 | 5.40E-04 | 2.61 | 89 | 1.90E-02 |
| GOTERM_BP_ALL | macromolecule metabolic process | 3.74 | 62 | 1.40E-03 | 2.61 | 142 | 4.80E-01 |
| GOTERM_BP_ALL | metabolic process | 3.74 | 69 | 1.60E-02 | 2.61 | 169 | 8.20E-01 |
| GOTERM_MF_ALL | nucleic acid binding | 3.74 | 44 | 6.20E-05 | 2.61 | 83 | 7.00E-03 |
| GOTERM_BP_ALL | nucleobase, nucleoside, nucleotide and nucleic acid metabolic process | 3.74 | 42 | 5.80E-04 | 2.61 | 85 | 2.30E-01 |
| SP_PIR_KEYWORDS | nucleus | 3.74 | 48 | 3.10E-04 | 2.61 | 119 | 3.80E-04 |
| GOTERM_BP_ALL | primary metabolic process | 3.74 | 62 | 6.50E-02 | 2.61 | 160 | 5.90E-01 |
| GOTERM_BP_ALL | regulation of biological process | 3.74 | 50 | 1.90E-05 | 2.61 | 117 | 2.40E-04 |
| GOTERM_BP_ALL | regulation of cellular metabolic process | 3.74 | 37 | 2.70E-06 | 2.61 | 77 | 1.10E-04 |
| GOTERM_BP_ALL | regulation of cellular process | 3.74 | 49 | 5.40E-06 | 2.61 | 113 | 6.50E-05 |
| GOTERM_BP_ALL | regulation of gene expression | 3.74 | 35 | 7.40E-06 | 2.61 | 74 | 1.20E-04 |
| GOTERM_BP_ALL | regulation of metabolic process | 3.74 | 37 | 8.60E-06 | 2.61 | 78 | 3.00E-04 |
| GOTERM_BP_ALL | regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process | 3.74 | 33 | 2.10E-05 | 2.61 | 71 | 1.40E-04 |
| GOTERM_BP_ALL | regulation of transcription | 3.74 | 32 | 3.30E-05 | 2.61 | 70 | 1.00E-04 |
| GOTERM_BP_ALL | regulation of transcription, DNA-dependent | 3.74 | 31 | 2.40E-05 | 2.61 | 63 | 7.70E-04 |
| GOTERM_BP_ALL | RNA biosynthetic process | 3.74 | 31 | 4.60E-05 | 2.61 | 64 | 1.10E-03 |
| GOTERM_BP_ALL | RNA metabolic process | 3.74 | 37 | 7.70E-05 | 2.61 | 69 | 1.00E-01 |
| SP_PIR_KEYWORDS | Transcription | 3.74 | 30 | 8.50E-06 | 2.61 | 56 | 2.70E-05 |
| GOTERM_BP_ALL | transcription | 3.74 | 32 | 9.10E-05 | 2.61 | 71 | 2.90E-04 |
| GOTERM_MF_ALL | transcription factor activity | 3.74 | 11 | 5.80E-02 | 2.61 | 24 | 2.10E-02 |
| GOTERM_BP_ALL | transcription from RNA polymerase II promoter | 3.74 | 11 | 2.10E-02 | 2.61 | 26 | 4.20E-03 |
| SP_PIR_KEYWORDS | Transcription regulation | 3.74 | 30 | 3.80E-06 | 2.61 | 55 | 1.70E-05 |
| GOTERM_MF_ALL | transcription regulator activity | 3.74 | 17 | 3.00E-02 | 2.61 | 40 | 3.00E-03 |
| GOTERM_BP_ALL | transcription, DNA-dependent | 3.74 | 31 | 4.60E-05 | 2.61 | 64 | 1.10E-03 |

Listed are the full ontologies for the top grouping from our analysis of each cluster using functional annotation clustering from the DAVID suite of programs[3,4]. Table S2A shows the cluster number (C) enrichment score (Score), ontology keyword (Term), ontology process (Process), number of transcripts (Count) and p-values for the EASE scores for each process (p-value). Table S2B shows the ontologies for top enrichment group for clusters 1 and 4 both of which had the same members with C1 and C4 being the enrichment score for cluster 1 and cluster 4, Ct1 and Ct4 the number of transcripts with that score in cluster 1 and cluster 4, PV1 and PV4 the p-values for those terms in cluster 1 and cluster 4

**Table S3. Summary Statistics for Original Cohort with Control Sample Generated from Bootstrap Analysis**

Original cohort: LOAD and control combined n = 364. Control sample: n = 364. Shown are the sample statistics comparing the original dataset (total cohort) with a permuted dataset (bootstrap cohort) to determine whether we had effects that were specific to disease tissue or whether novel effects were found simply because we had a larger sample.

| | total cohort | bootstrap cohort |
|---|---|---|
| % female | 48% | 50% |
| average age | 82 | 81 |
| age range | 65-102 | 65-100 |
| % frontal cortex | 20% | 18% |
| % temporal cortex | 67% | 72% |
| % parietal cortex | 6% | 2% |
| % cerebellar cortex | 8% | 3% |
| # of different hybridization dates | 9 | 9 |
| average # of samples per hybridization | 20 | 21 |
| number of institute sources | 18 | 18 |
| average number of samples per institute | 11 | 11 |
| average PMI | 10 hrs | 10 hrs |
| average transcriptome detection rate per sample | 77% | 77% |

**Table S4. Detection of Disease Effects in Entire Cohort Compared with Bootstrap Sample Based on Control Data Only**

| ALPHA | INTERACTION | | NO INTERACTION | |
|---|---|---|---|---|
| | number det | proportion det | number det | proportion det |
| α= 0.05 | 551 | 83% | 1820 | 100% |
| α= 0.01 | 495 | 74% | 1820 | 100% |
| α= 0.001 | 436 | 66% | 1820 | 100% |
| α= 0.0001 | 390 | 59% | 1815 | 100% |
| α= 0.00001 | 345 | 52% | 1794 | 97% |
| **α= 0.000001** | **303** | **46%** | **1737** | **95%** |

Shown are the numbers (number det) and the percentages (proportion det) of the SNP-transcript effects we found in the total cohort that were recovered in the bootstrap sample based just on our control dataset at different false positive stringencies. To survive multiple testing correction, p-values in our analysis had to be less that α= 0.000001. Note that at this level, only 46% of the SNP-transcript relationships which have a significant interaction with diagnosis (putative LOAD effects) were recovered, whereas 95% of the relationships without a significant interaction term were recovered. This demonstrates that most of the novel effects are likely due to our specific sampling of disease tissue and not an increase in cohort size.

**Supplemental References**

1. Pritchard JK, Stephens M, Rosenberg NA, Donnelly P. Inference of population structure using multilocus genotype data. Genetics. 2000 Jun;155 (2):945-59.
2. Falush D, Stephens M, Pritchard JK. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics. 2003 Aug;164(4):1567-87.
3. Dennis G Jr, Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, Lempicki RA. DAVID: Database for Annotation, Visualization, and Integrated Discovery. Genome Biology 2003;4(5):P3.
4. Douglas A Hosack, Glynn Dennis Jr, Brad T Sherman, HClifford Lane, Richard A Lempicki. Identifying Biological Themes within Lists of Genes with EASE. Genome Biology 2003 4(6):P4.