Journal of Alzheimer's Disease 30 (2012) 1–7
IOS Press

Supplementary Data

Mitochondrial Dysfunction and Immune Activation are Detectable in Early Alzheimer's Disease Blood

Katie Lunnon[a,b,1], Zina Ibrahim[a,b,1], Petroula Proitsi[a,b], Anbarasu Lourdusamy[a,b], Stephen Newhouse[a,b], Martina Sattlecker[a,b], Simon Furney[a,b,3], Muzamil Saleem[a,b], Hilkka Soininen[c], Iwona Kłoszewska[d], Patrizia Mecocci[e], Magda Tsolaki[f], Bruno Vellas[g], Giovanni Coppola[h], Daniel Geschwind[h], Andrew Simmons[a,b], Simon Lovestone[a,b], Richard Dobson[a,b,2], Angela Hodges[a,b,2,*] and on Behalf of the AddNeuroMed Consortium

[a]*Institute of Psychiatry, King's Health Partners Centre for Neurodegeneration Research, King's College London, London, UK*

[b]*NIHR Biomedical Research Centre for Mental Health at South London and Maudsley NHS Foundation Trust and Institute of Psychiatry, King's College London, London, UK*

[c]*Department of Neurology, University of Eastern Finland and Kuopio University Hospital, Kuopio, Finland*

[d]*Medical University of Lodz, Lodz, Poland*

[e]*Institute of Gerontology and Geriatrics, University of Perugia, Perugia, Italy*

[f]*3rd Department of Neurology, "G. Papanicolaou" Hospital, Aristotle University of Thessaloniki, Thessaloniki, Greece*

[g]*INSERM U 558, University of Toulouse, Toulouse, France*

[h]*Department of Neurology, Programme in Neurogenetics, David Geffen School of Medicine, University of California at Los Angeles, CA, USA*

Accepted 25 February 2012

GLOSSARY FOR WGCNA ANALYSIS

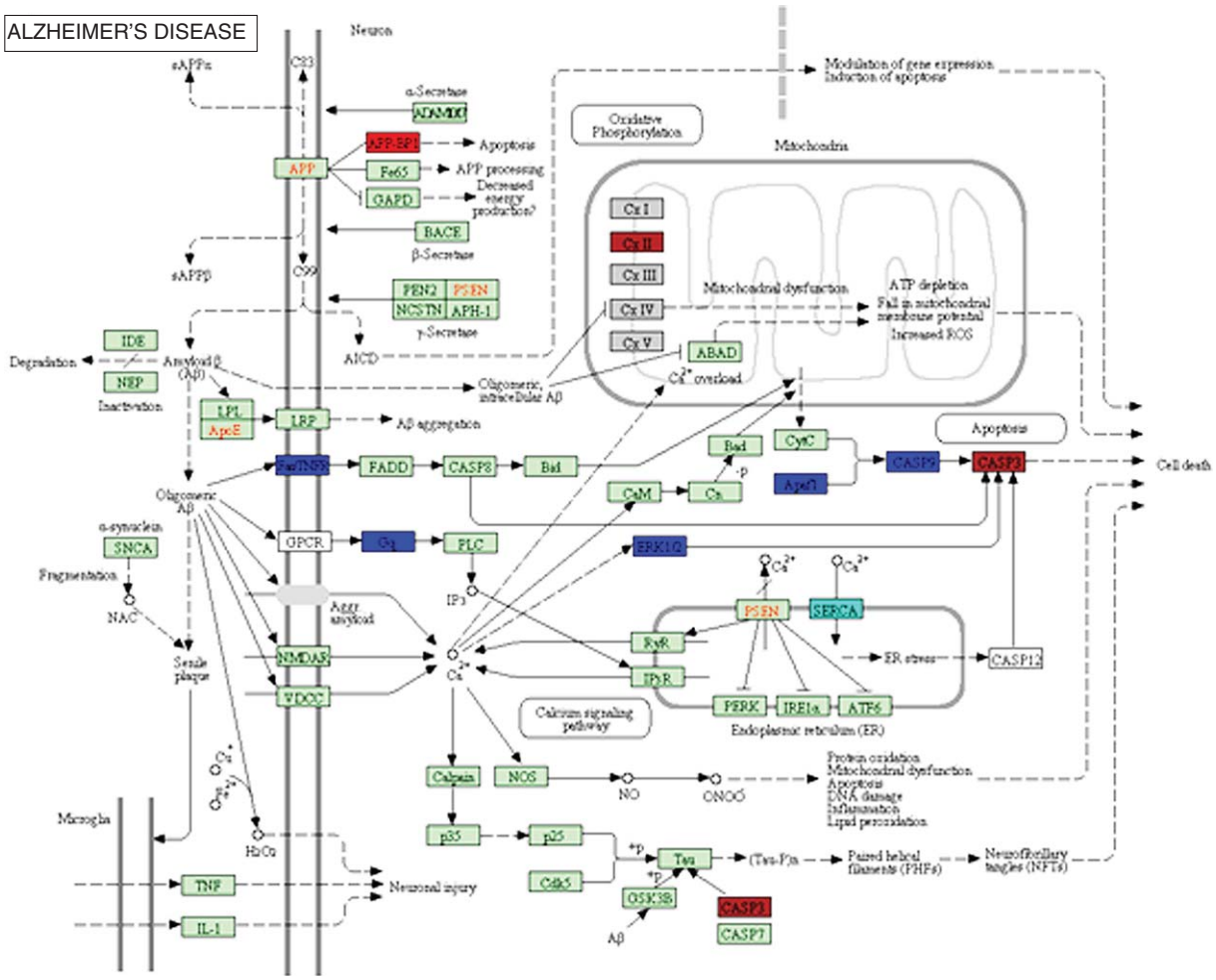
We have used gene expression data consisting of individual probes from microarrays to perform a Weighted Gene Co-expression Network Analysis (WGCNA). A gene co-expression network is a graphical representation of the relationship between genes according to the similarity of their expression profiles and thus potentially their biological relatedness. Within the network, a node represents a probe and an edge exists between two probes if they exhibit similar expression patterns across the samples, i.e., they are co-expressed. The following terms and definitions are used to represent different features of the network and associated analyses. For further details we refer readers to the glossary provided at the WGCNA web site: http://www.genetics.ucla.edu/labs/horvath/CoexpressionNetwork/Rpackages/WGCNA/Tutorials/Simulated-00-Background.pdf.

[1,2]These authors contributed equally to the work.

[3]Present address: The Institute of Cancer Research, London, UK.

*Correspondence to: Angela Hodges, King's College London, Institute of Psychiatry, De Crespigny Park, London, UK. Tel.: +44 207 848 0772; Fax: +44 207 848 0632; E-mail: angela.k.hodges@kcl.ac.uk.

ISSN 1387-2877/12/$27.50



Supplementary Figure 1. Over represented KEGG pathways ($p < 5 \times 10^{-4}$) identified by DAVID analysis (http://david.abcc.ncifcrf.gov/) [4] using probes with higher-than-median module membership and trait significance for each module (i.e., probes highlighted in green in Fig. 5 and listed in column Q, supplementary Tables 2–7). KEGG pathways with significantly over-represented modules include: Alzheimer's disease ($p = 1.8 \times 10^{-4}$; black module), Oxidative Phosphorylation ($p = 3.9 \times 10^{-5}$; black module), Ribosome ($p = 6.6 \times 10^{-5}$; black and $p = 1.7 \times 10^{-7}$; red module), Leukocyte Transendothelial Migration ($p = 1.2 \times 10^{-4}$; blue module). In each KEGG pathway probes with higher-than-median module membership and trait significance are indicated by a star with the color of the star indicating their assigned module, except the black module which is represented by grey.

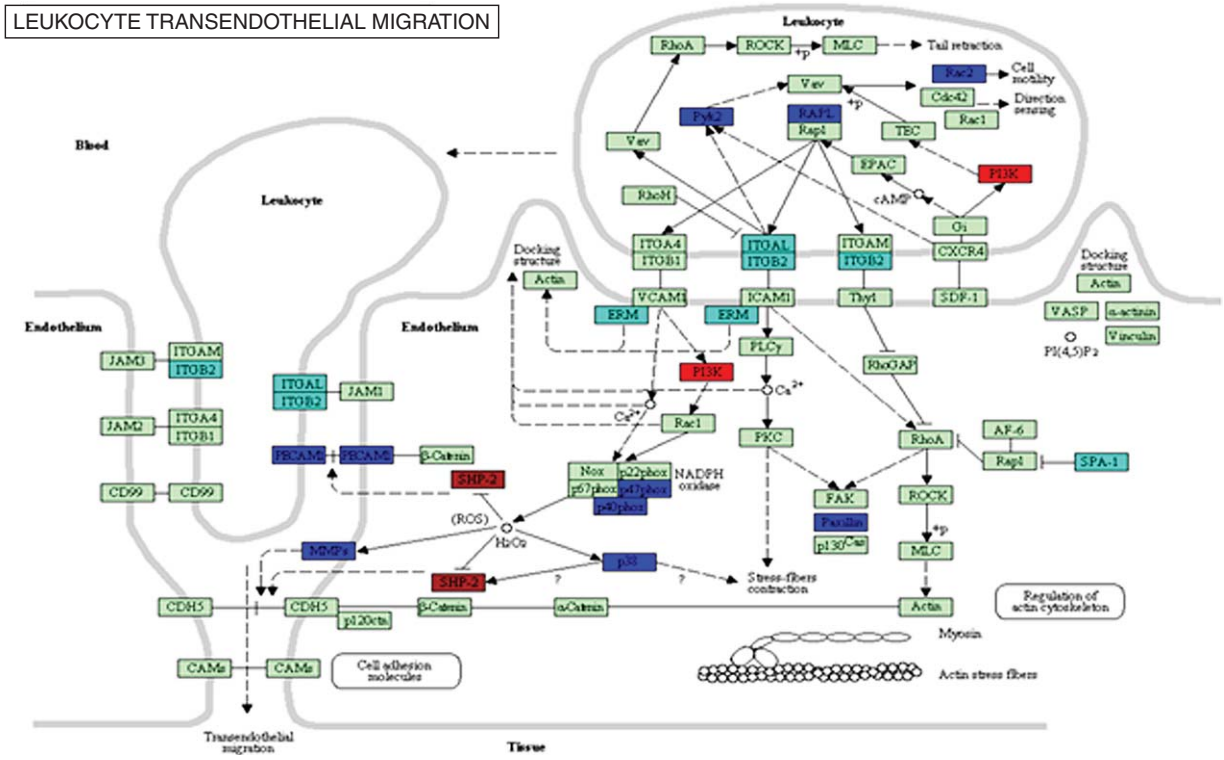
Probe: A probe assesses the expression levels of a particular gene within a given sample.

Connectivity: In its simplest form, the connectivity of a probe is computed as the number of neighbors it is connected to in a co-expression network, or:

$$\text{Connectivity}_i = k_i = \sum_{j \neq i} a_{ij}$$

In a weighted network, connectivity can be measured by different parameters, including topological overlap (see below). Probes with high connectivity values share a similar profile of gene expression with a relatively large number of other probes.

Topological Overlap (TO): Topological Overlap provides the score/weight for the edges in the co-expression network. To calculate the topological overlap for a pair of probes, their connections with all other probes in the network are compared. If the two probes show similar patterns of correlation with other probes, then they have a high topological overlap. Several studies have shown that probes showing high topological overlap are more likely to be functionally related than probes that do not. For two nodes



Supplementary Figure 1. (*continued*)

i and j, the topological overlap of the two nodes (t_{ij}) is computed as follows:

$$t_{ij} = (I_{jj} + a_{ij})/(\min\{k_i, k_j\}\_1 - a_{ij}) \quad \text{if } i \neq j$$
$$= 1 \quad \text{if } i = j$$

Where I_{jj} k_i and k_j are the connectivity measures of nodes i and j as defined earlier.

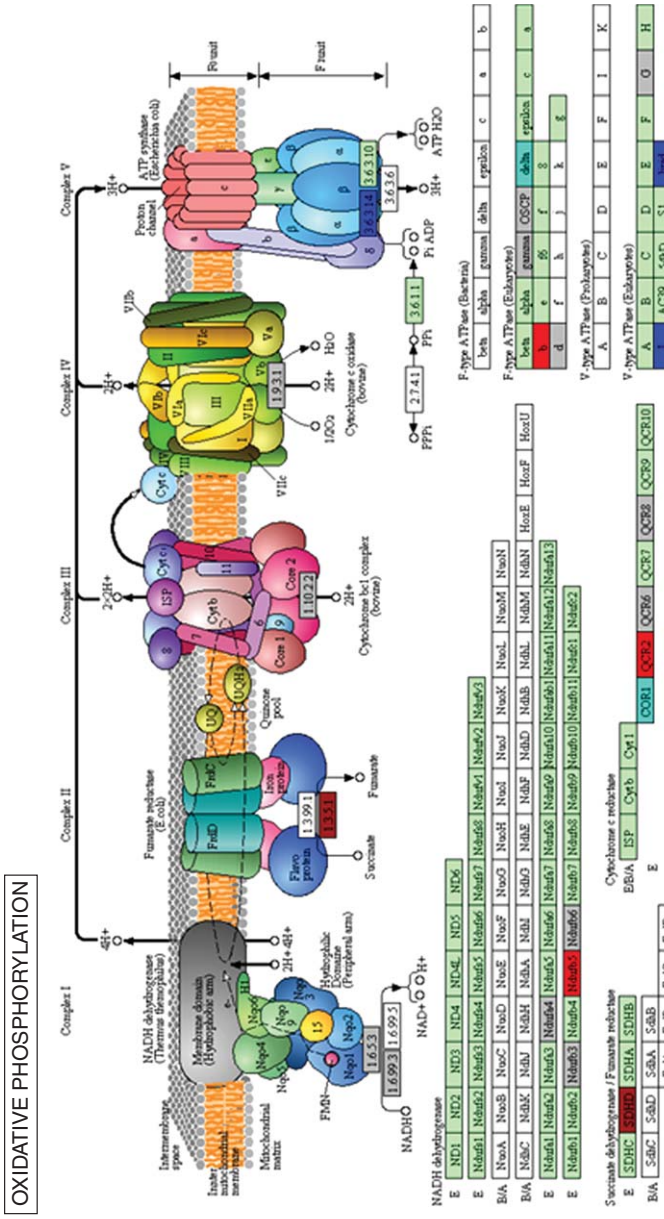
Topological Overlap Matrix (TOM): The Topological Overlap Matrix describes the pairwise TO between all probes in the network [1]. The numbers in the matrix measure similarity amongst the probes in the network. In this work, the TOM was used to define edges between probe pairs.

Modules: Modules are sub-networks of the larger network, comprising probes with similar expression patterns across samples. Probes belonging to the same module are thought to be functionally related, e.g., represent genes encoding a pathway or a protein complex or related biological function and are therefore considered biological important [2]. The biological characteristics and behavior of modules may reveal far more than only considering individual genes in isolation. Computationally, a network module is comprised of a set of probes which are closely connected according to a suitably defined measure of interconnectivity (TOM) and the set of samples from which the expression data is derived.

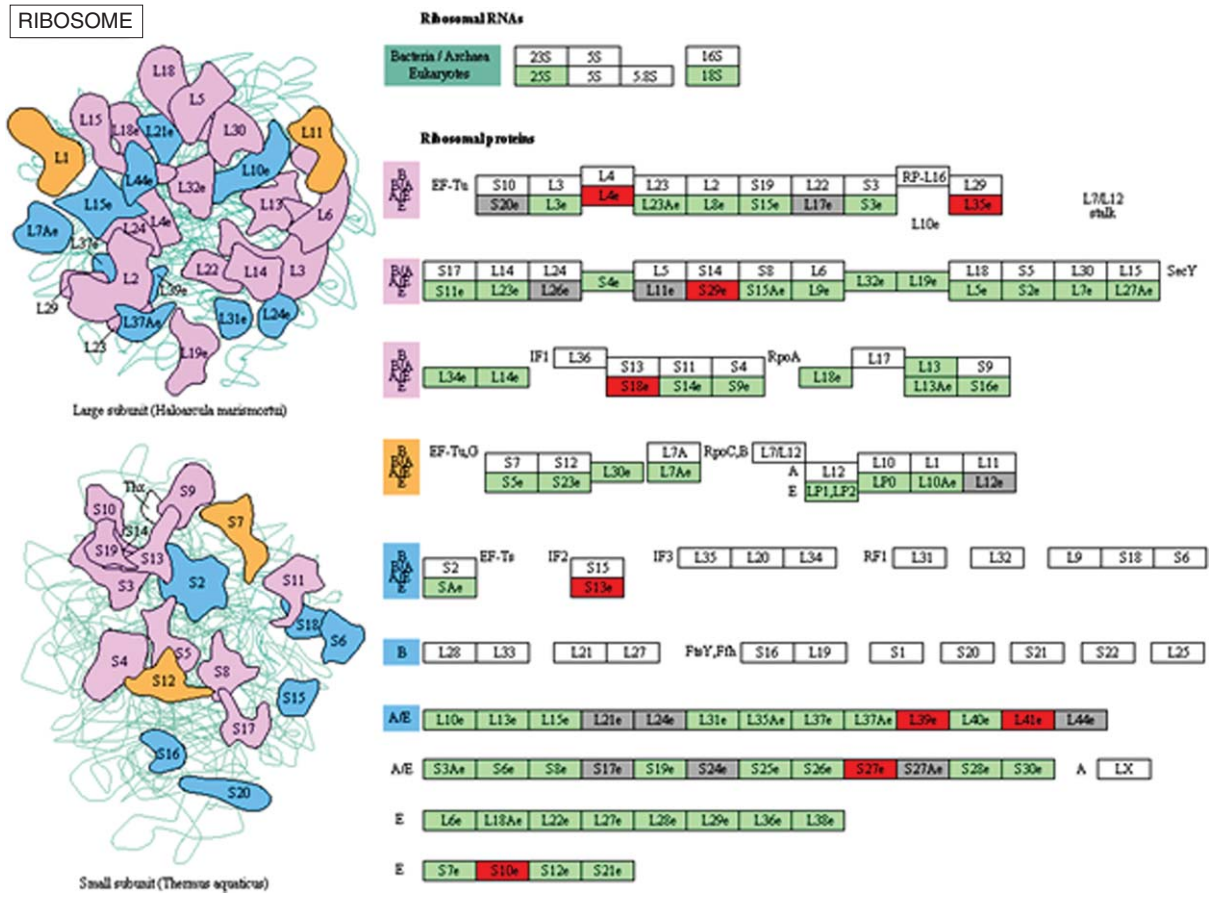
Module Eigengene (ME): A Module Eigengene is the expression profile chosen to represent that of the module. Module Eigengenes are important for establishing whether there are correlations between modules and clinical traits and each other. Mathematically, an eigengene is computed as the first eigenvector of the adjacency matrix of the module and represents the first-principal component of the genes within the module [2].

Module Membership (MM): Module Membership is a measure of the extent to which a probe conforms to the characteristics of the module it is assigned to. It is measured by the correlation between the expression profile of a probe and the Module Eigengene (ME) of the corresponding module to which the probe belongs.

Gene Significance (GS): A Gene Significance measure of a gene is used to assess the biological significance of a particular probe and therefore gene, with respect to a trait (e.g., disease severity). GS is defined as the correlation coefficient resulting from



Supplementary Figure 1. (*continued*)



Supplementary Figure 1. (*continued*)

correlating the outcome of the trait in question with the expression profile of the gene [3]. GS can take positive or negative values depending on the correlation relationship (a positive GS results from a positive correlation while a negative GS results from a negative correlation). A GS value of zero indicates no significance while higher absolute values indicate a higher significance of the gene to the trait [3].

Weighted Co-expression Network: A Weighted Co-expression network is a network in which the edges are annotated with numbers (weights) denoting the extent to which two nodes (probes) are similar. In this case the weights represent the topological overlap between nodes, i.e., the numbers represent the strength of the correlation of the expression profiles of the nodes connected via the edge.

Signed Weighted Co-expression Network: Signed weighted co-expression network is a variant of weighted co-expression networks which attaches a sign to the weights assigned to its edges. The sign designates the direction of expression change among the expression profiles. Signed networks are thought to be more biologically relevant than unsigned networks whereby the modules are created based on absolute measures of correlation, i.e., genes assigned to the same module can have opposite directions of change in their gene expression profiles.

Table 1: Differential gene expression in blood samples from AD, MCI, and control subjects. A total of 2,908 significant differences were identified between the three groups (FDR corrected $p < 0.01$). Positive or negative fold-change indicates increased or decreased expression in MCI and/or AD with respect to control blood or AD with respect to MCI blood ($p < 0.001$ in post-hoc T-test).

Tables 2–7: A list of probes assigned to the disease-associated modules red, black, pink, brown, blue, and turquoise, respectively. Probe level associations with the diagnostic traits control, MCI-MCI, MCI-AD, AD, ALL AD, and disease severity are indicated. Gene

significance (GS) of a gene describes the strength and sign of the correlation between the probe and the trait in question, while the module membership score (MM) quantifies the extent to which a gene conforms to the characteristics of a module. The combination of MM and GS identifies genes which play important roles in a given network module and their significance for the clinical trait in question. Probes with higher-than-median module membership and trait significance for each module are indicated in column Q.

Table 8: Compiled gene lists comprising top Alzheimer's GWAS genes, other candidate genes thought to be associated with Alzheimer's, OXPHOS genes, MRP genes, and immune genes. Genes are annotated with their module membership.

Table 9: Test for over-representation of MCI-associated (A) and AD-associated (B) gene expression changes in specific blood cell populations in blood samples from AD patients or normal elderly controls. A total of 19,161 probes were used in the analysis (see methods) of which some had significantly altered expression in MCI ($n = 1{,}999$ with FDR <0.01) and/or AD ($n = 1{,}319$ with FDR <0.01). These were mapped to a set of probes previously reported to be enriched in particular blood cell types by Watkins et al. [5] using RNA from blood analyzed with the same arrays. Over-representation of cell-type enriched transcripts was examined using Chi-square or the Fisher's exact test if the number of probes was less than 10 (*). To increase confidence in our results, we also tested whether more cell lineage probes attained a given p-value than would be expected by chance by randomly selecting 1,319 or 1,999 of the 19,161 used in the analysis and repeating the analysis for each cell-type enriched probe list for 10,000 permutations. We further tested for over- rather than under-representation of significantly altered probes in particular blood cells in AD blood by a hypergeometric probability test.

Table 10: We tested the blood modules (column A) for enrichment using a large pre-defined collection of brain-related gene sets (column C) [6–27]. Classification categories and functional annotation for each test dataset were defined by the individual study (column G). Significance was computed using a hypergeometric test. Each dataset is identifiable by the publication first describing each study (column B).

REFERENCES

[1] Yip AM, Horvath S (2007) Gene network interconnectedness and the generalized topological overlap measure. *BMC Bioinformatic* **8**, 22.

[2] Zhang B, Horvath S (2005) A general framework for weighted gene co-expression network analysis. *Stat Appl Genet Mol Biol* **4**, Article 17.

[3] Presson AP, Sobel EM, Papp JC, Suarez CJ, Whistler T, Rajeevan MS, Vernon SD, Horvath S (2008) Integrated weighted gene co-expression network analysis with an application to chronic fatigue syndrome. *BMC Syst Biol* **2**, 95.

[4] Huang DW, Sherman BT, Lempicki RA (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Prot* **4**, 44-57.

[5] Watkins NA, Gusnanto A, de Bono B, De S, Miranda-Saavedra D, Hardie DL, Angenent WGJ, Attwood AP, Ellis PD, Erber W, Foad NS, Garner SF, Isacke CM, Jolley J, Koch K, Macaulay IC, Morley SL, Rendon A, Rice KM, Taylor N, Thijssen-Timmer DC, Tijssen MR, van der Schoot CE, Wernisch L, Winzer T, Dudbridge F, Buckley CD, Langford CF, Teichmann S, Gottgens B, Ouwehand WH (2009) A HaemAtlas: Characterizing gene expression in differentiated human blood cells. *Blood* **113**, E1-E9.

[6] Albright AV, Gonzalez-Scarano F (2004) Microarray analysis of activated mixed glial (microglia) and monocyte-derived macrophage gene expression. *J Neuroimmunol* **157**, 27-38.

[7] Bachoo RM, Kim RS, Ligon KL, Maher EA, Brennan C, Billings N, Chan S, Li C, Rowitch DH, Wong WH, DePinho RA (2004) Molecular diversity of astrocytes with implications for neurological disorders. *Proc Natl Acad Sci U S A* **101**, 8384-8389.

[8] Blalock EM, Geddes JW, Chen KC, Porter NM, Markesbery WR, Landfield PW (2004) Incipient Alzheimer's disease: Microarray correlation analyses reveal major transcriptional and tumor suppressor responses. *Proc Natl Acad Sci U S A* **101**, 2173-2178.

[9] Bult CJ, Eppig JT, Kadin JA, Richardson JE, Blake JA (2008) The Mouse Genome Database (MGD): Mouse biology and model systems. *Nucl Acids Res* **36**, D724-D728.

[10] Cahoy JD, Emery B, Kaushal A, Foo LC, Zamanian JL, Christopherson KS, Xing Y, Lubischer JL, Krieg PA, Krupenko SA, Thompson WJ, Barres BA (2008) A transcriptome database for astrocytes, neurons, and oligodendrocytes: A new resource for understanding brain development and function. *J Neurosci* **28**, 264-278.

[11] Colangelo V, Schurr J, Ball MJ, Pelaez RP, Bazan NG, Lukiw WJ (2002) Gene expression profiling of 12633 genes in Alzheimer hippocampal CA1: Transcription and neurotrophic factor down-regulation and up-regulation of apoptotic and pro-inflammatory signaling. *J Neurosci Res* **70**, 462-473.

[12] Foster LJ, de Hoog CL, Zhang YL, Zhang Y, Xie XH, Mootha VK, Mann M (2006) A mammalian organelle map by protein correlation profiling. *Cell* **125**, 187-199.

[13] Gan L, Ye SM, Chu A, Anton K, Yi SL, Vincent VA, von Schack D, Chin D, Murray J, Lohr S, Patthy L, Gonzalez-Zulueta M, Nikolich K, Urfer R (2004) Identification of cathepsin B as a mediator of neuronal death induced by A beta-activated microglial cells using a functional genomics approach. *J Biol Chem* **279**, 5565-5572.

[14] Ginsberg SD, Che SL (2005) Expression profile analysis within the human hippocampus: Comparison of CA1 and CA3 pyramidal neurons. *J Comp Neurol* **487**, 107-118.

[15] it-Ghezala G, Mathura VS, Laporte V, Quadros A, Paris D, Patel N, Volmar CH, Kolippakkam D, Crawford F, Mullan M (2005) Genomic regulation after CD40 stimulation in microglia: Relevance to Alzheimer's disease. *Mol Brain Res* **140**, 73-85.

[16] Lein ES, Zhao XY, Gage FH (2004) Defining a molecular atlas of the hippocampus using DNA microarrays and high-

throughput in situ hybridization. *J Neurosci* **24**, 3879-3889.

[17] Lein ES, Hawrylycz MJ, Ao N, Ayres M, Bensinger A, Bernard A, Boe AF, Boguski MS, Brockway KS, Byrnes EJ, Chen L, Chen L, Chen TM, Chin MC, Chong J, Crook BE, Czaplinska A, Dang CN, Datta S, Dee NR, Desaki AL, Desta T, Diep E, Dolbeare TA, Donelan MJ, Dong HW, Dougherty JG, Duncan BJ, Ebbert AJ, Eichele G, Estin LK, Faber C, Facer BA, Fields R, Fischer SR, Fliss TP, Frensley C, Gates SN, Glattfelder KJ, Halverson KR, Hart MR, Hohmann JG, Howell MP, Jeung DP, Johnson RA, Karr PT, Kawal R, Kidney JM, Knapik RH, Kuan CL, Lake JH, Laramee AR, Larsen KD, Lau C, Lemon TA, Liang AJ, Liu Y, Luong LT, Michaels J, Morgan JJ, Morgan RJ, Mortrud MT, Mosqueda NF, Ng LL, Ng R, Orta GJ, Overly CC, Pak TH, Parry SE, Pathak SD, Pearson OC, Puchalski RB, Riley ZL, Rockett HR, Rowland SA, Royall JJ, Ruiz MJ, Sarno NR, Schaffnit K, Shapovalova NV, Sivisay T, Slaughterbeck CR, Smith SC, Smith KA, Smith BI, Sodt AJ, Stewart NN, Stumpf KR, Sunkin SM, Sutram M, Tam A, Teemer CD, Thaller C, Thompson CL, Varnam LR, Visel A, Whitlock RM, Wohnoutka PE, Wolkey CK, Wong VY, Wood M, Yaylaoglu MB, Young RC, Youngstrom BL, Yuan XF, Zhang B, Zwingman TA, Jones AR (2007) Genome-wide atlas of gene expression in the adult mouse brain. *Nature* **445**, 168-176.

[18] Liang WS, Dunckley T, Beach TG, Grover A, Mastroeni D, Ramsey K, Caselli RJ, Kukull WA, McKeel D, Morris JC, Hulette CM, Schmechel D, Reiman EM, Rogers J, Stephan DA (2008) Altered neuronal gene expression in brain regions differentially affected by Alzheimer's disease: A reference data set. *Physiol Genomics* **33**, 240-256.

[19] Miller JA, Horvath S, Geschwind DH (2010) Divergence of human and mouse brain transcriptome highlights Alzheimer disease pathways. *Proc Natl Acad Sci U S A* **107**, 12698-12703.

[20] Morciano M, Burre J, Corvey C, Karas M, Zimmermann H, Volknandt W (2005) Immunoisolation of two synaptic vesicle pools from synaptosomes: A proteomics analysis. *J Neurochem* **95**, 1732-1745.

[21] Newrzella D, Pahlavan PS, Krueger C, Boehm C, Sorgenfrei O, Schroeck H, Eisenhardt G, Bischoff N, Vogt G, Wafzig O, Rossner M, Maurer MH, Hiemisch H, Bach A, Kuschinsky W, Schneider A (2007) The functional genome of CA1 and CA3 neurons under native conditions and in response to ischemia. *BMC Genomics* **8**, 370.

[22] Oldham MC, Horvath S, Geschwind DH (2006) Conservation and evolution of gene coexpression networks in human and chimpanzee brains. *Proc Natl Acad Sci U S A* **103**, 17973-17978.

[23] Oldham MC, Konopka G, Iwamoto K, Langfelder P, Kato T, Horvath S, Geschwind DH (2008) Functional organization of the transcriptome in human brain. *Nat Neurosci* **11**, 1271-1282.

[24] Parachikova A, Agadjanyan M, Cribbs D, Blurton-Jones M, Perreau V, Rogers J, Beach T, Cotman C (2007) Inflammatory changes parallel the early stages of Alzheimer disease. *Neurobiol Aging* **28**, 1821-1833.

[25] Thomas DM, Francescutti-Verbeem DM, Kuhn DM (2006) Gene expression profile of activated microglia under conditions associated with dopamine neuronal damage. *FASEB J* **20**, 515-517.

[26] Torres-Munoz JE, Van Waveren C, Keegan MG, Bookman RJ, Petito CK (2004) Gene expression profiles in microdissected neurons from human hippocampal subregions. *Mol Brain Res* **127**, 105-114.

[27] Winden KD, Oldham MC, Mirnics K, Ebert PJ, Swan CH, Levitt P, Rubenstein JL, Horvath S, Geschwind DH (2009) The organization of the transcriptional network in specific neuronal classes. *Mol Syst Biol* **5**, 291.