doi:10.1093/brain/awy215

BRAIN
A JOURNAL OF NEUROLOGY

The human brainome: network analysis identifies *HSPA2* as a novel Alzheimer's disease target

Vladislav A. Petyuk,[1,*] Rui Chang,[2,*] Manuel Ramirez-Restrepo,[3,*] Noam D. Beckmann,[2,*] Marc Y. R. Henrion,[2] Paul D. Piehowski,[1] Kuixi Zhu,[2] Sven Wang,[2] Jennifer Clarke,[4] Matthew J. Huentelman,[5] Fang Xie,[1] Victor Andreev,[6] Anzhelika Engel,[3] Toumy Guettoche,[7] Loida Navarro,[7] Philip De Jager,[8,9,10] Julie A. Schneider,[11] Christopher M. Morris,[12] Ian G. McKeith,[13] Robert H. Perry,[14] Simon Lovestone,[15] Randall L. Woltjer,[16] Thomas G. Beach,[17] Lucia I. Sue,[17] Geidy E. Serrano,[17] Andrew P. Lieberman,[18] Roger L. Albin,[19,20] Isidre Ferrer,[21] Deborah C. Mash,[22] Christine M. Hulette,[23] John F. Ervin,[24] Eric M. Reiman,[25,26] John A. Hardy,[27] David A. Bennett,[11] Eric Schadt,[2] Richard D. Smith[1] and Amanda J. Myers[3,28,29]

*These authors contributed equally to this work.

Our hypothesis is that changes in gene and protein expression are crucial to the development of late-onset Alzheimer's disease. Previously we examined how DNA alleles control downstream expression of RNA transcripts and how those relationships are changed in late-onset Alzheimer's disease. We have now examined how proteins are incorporated into networks in two separate series and evaluated our outputs in two different cell lines. Our pipeline included the following steps: (i) predicting expression quantitative trait loci; (ii) determining differential expression; (iii) analysing networks of transcript and peptide relationships; and (iv) validating effects in two separate cell lines. We performed all our analysis in two separate brain series to validate effects. Our two series included 345 samples in the first set (177 controls, 168 cases; age range 65–105; 58% female; KRONOSII cohort) and 409 samples in the replicate set (153 controls, 141 cases, 115 mild cognitive impairment; age range 66–107; 63% female; RUSH cohort). Our top target is heat shock protein family A member 2 (HSPA2), which was identified as a key driver in our two datasets. HSPA2 was validated in two cell lines, with overexpression driving further elevation of amyloid-β_{40} and amyloid-β_{42} levels in APP mutant cells, as well as significant elevation of microtubule associated protein tau and phosphorylated-tau in a modified neuroglioma line. This work further demonstrates that studying changes in gene and protein expression is crucial to understanding late onset disease and further nominates HSPA2 as a specific key regulator of late-onset Alzheimer's disease processes.

1 Biological Sciences Division and Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, WA 99354, USA
2 Icahn Institute for Genomics and Multiscale Biology, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA
3 Department of Psychiatry and Behavioral Sciences, University of Miami Miller School of Medicine, Miami, FL 33136, USA
4 Food Science and Technology Department, University of Nebraska-Lincoln, Lincoln, NE 68588, USA
5 Neurogenomics Division, The Translational Genomics Research Institute, Phoenix, AZ 85004, USA
6 Arbor Research Collaborative for Health, 340 E Huron St # 300, Ann Arbor, MI 48104 USA
7 Roche Sequencing, 4300 Hacienda Drive, Pleasanton, CA 94588, USA

Received February 14, 2018. Revised April 20, 2018. Accepted June 22, 2018.

8 Center for Translational and Computational Neuroimmunology, Department of Neurology, Columbia University Medical Center, New York, NY 10032, USA
9 New York Genome Center, New York NY 10013, USA
10 Program in Medical and Population Genetics, Broad Institute, Cambridge, MA 02142, USA
11 Rush Alzheimer's Disease Center, Rush University Medical Center, Chicago, IL 60612, USA
12 Newcastle Brain Tissue Resource, Institute of Neuroscience, Campus for Ageing and Vitality, Newcastle University, Newcastle upon Tyne, NE4 5PJ, UK
13 NIHR Biomedical Research Centre, Institute for Ageing and Health, Newcastle University, Campus for Ageing and Vitality, Newcastle upon Tyne, NE4 5PJ, UK
14 Neuropathology and Cellular Pathology, Royal Victoria Infirmary, Queen Victoria Road, Newcastle upon Tyne, NE1 4LP, UK
15 University of Oxford, Medical Sciences Division, Department of Psychiatry, Warneford Hospital, Oxford OX3 7JX, UK
16 Neuropathology Core of the Layton Aging and Alzheimer's Disease Center, Oregon Health and Science University, Portland, OR 97239, USA
17 Banner Sun Health Research Institute, Sun City, AZ 85351, USA
18 Department of Pathology, University of Michigan, Ann Arbor, MI 48109, USA
19 Department of Neurology, University of Michigan, Ann Arbor, MI 48109, USA
20 Geriatrics Research, Education, and Clinical Center, VAAAHS, Ann Arbor, MI 48105, USA
21 Department of Pathology and Experimental Therapeutics, University of Barcelona; CIBERNED; Hospitalet de Llobregat 08007, Spain
22 Department of Neurology, University of Miami Miller School of Medicine, Miami, FL 33136, USA
23 Department of Pathology, Division of Neuropathology, Duke University Medical Center, Durham, NC 27710, USA
24 Kathleen Price Bryan Brain Bank, Department of Medicine, Division of Neurology, Duke University, Durham, NC 27710, USA
25 The Arizona Alzheimer's Consortium, Phoenix, Arizona 85014, USA
26 Banner Alzheimer's Institute, Phoenix, Arizona 85006, USA
27 Department of Molecular Neuroscience and Reta Lila Research Laboratories, University College London Institute of Neurology, London WC1E 6BT, UK
28 Interdepartmental Program in Neuroscience, University of Miami Miller School of Medicine, Miami, FL 33136, USA
29 Interdepartmental Program in Human Genetics and Genomics, University of Miami Miller School of Medicine, Miami, FL 33136, USA

Correspondence to: Amanda J. Myers
Rosenstiel Medical Sciences Building, room 7011A, D-106
1600 NW 10th Ave, Miami, FL 33136, USA
E-mail: profmyersUM@gmail.com

Keywords: genetic network; Alzheimer's disease; dementia; transcriptomics; proteomics

Introduction

Along with other groups, we have proposed that systems approaches to finding novel genes involved in disease pathways can be more powerful than DNA-only approaches (Myers, 2012, 2013, 2014). We have previously examined genotype–transcript relationships via expression quantitative trait loci analysis and constructed regulatory networks (Myers *et al.*, 2007*a*; Webster *et al.*, 2009; Zhang *et al.*, 2013). In this report, we extend the original work using two independent datasets and integrated mass spectrometry proteomics (Piehowski *et al.*, 2013). In addition, by experimentally validating the top replicated key drivers using two independent cell-based models, we directly tested predictions from the network models regarding their impact on late-onset Alzheimer's disease pathology.

Proteins are the primary effectors of human phenotypes, so it is crucial to understand protein expression in the context of gene variation and transcript expression. Prior work to integrate proteomic data into the analysis of biological networks has been successfully performed in humans (Garge *et al.*, 2010; Portelli *et al.*, 2014; Stark *et al.*, 2014), yeast (Foss *et al.*, 2007; Wu *et al.*, 2008) and rodents (Fei *et al.*, 2011; Ghazalpour *et al.*, 2011). Most of the human data to date have been collected from lymphoblasts (Garge *et al.*, 2010; Stark *et al.*, 2014), which are highly subject to *de novo* mutation. Such cell-based systems are not ideal for constructing models of human disease, especially as many targets found in lymphoblast screens do not replicate in brain tissues (Hong *et al.*, 2008).

In this report, we present the first replication 'omics screen that includes DNA variation, RNA expression, and tandem mass spectrometry proteome profiles in two series of human brains, ~50% of which are pathologically confirmed late-onset Alzheimer's disease tissues and one of which includes samples collected worldwide (Supplementary Fig. 1). These two sets were analysed independently to determine relationships between DNA, RNA and protein. Our pipeline (Fig. 1) involved testing for single effects, such as differential expression of both transcripts and peptides as well as expression quantitative trait loci to examine single relationships between transcript and peptide abundances and allele content.

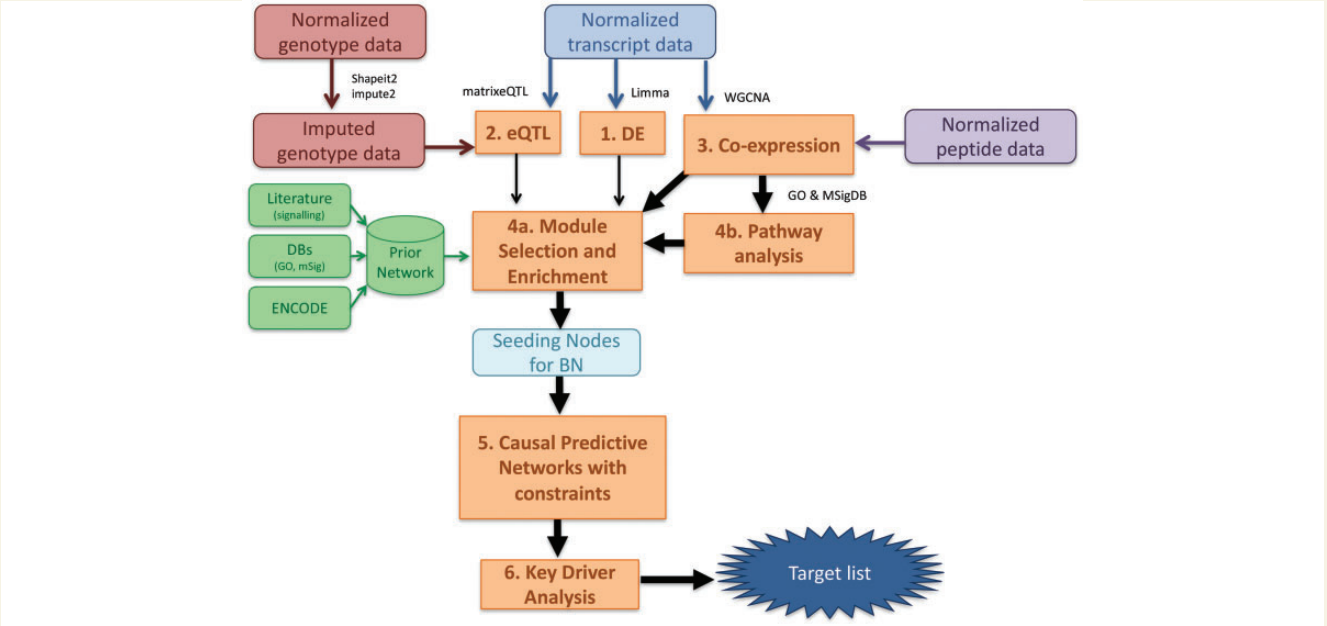

Figure 1 Analysis pipeline. A summary of the steps that were taken on the processed data is shown. Round rectangles indicate input data, green round rectangles indicate input data from external sources, and orange squares indicate processes and outputs from those processes. Steps are numbered on the figure. See main text for further detail. BN = Bayesaian network; DBs = databases; DE = differential expression; eQTL = expression quantitative trait loci; GO = Gene Ontology database; mSig = Molecular Signatures Database; WGCNA = weighted correlation network analysis.

Network analyses were used to capture more complex relationships between groups of data. We performed both analysis mapping how expression profiles were related between multiple transcript and peptide targets (co-expression network analysis) as well as mapping the causal structures within the data (causal network analysis). Causal network analysis is an expansion of co-expression analysis, in that the relationships between transcript and or peptide targets are given an order and direction in these predictions. For example, in co-expression analysis, relationships are mapped such that target A contacts targets B and C. In causal analysis, target A could be mapped upstream of targets B and C and could contact B before C. For the causal network analyses, we used a novel expansion of standard Bayesian network approaches, which allows for the analysis of opposite causality (i.e. both negative and positive correlations). This causal predictive network type of analysis is more comprehensive than testing single transcripts or proteins against single nucleotide polymorphisms (SNPs), which is the common procedure in expression quantitative trait loci studies and is the procedure carried out in many prior reports including the human proteome (Garge *et al.*, 2010; Stark *et al.*, 2014). Additionally, the multiple layers of regulation that can occur between DNA and protein make the protein-SNP relationships more complex and multivariate network approaches are capable of capturing relationships among all targets. Our final step is to narrow down our list of targets from the causal predictions by performing a key driver analysis. Key driver analysis involves looking for targets (formally, nodes within the causal structure) that have a higher number of connections than would be expected given a background.

While it was important to identify targets in brain, the causal consequences of changing levels of targets are statistically inferred; therefore, validating the predictions is critical. We used several different cell lines to validate targets outside the context of human brain tissue, measuring levels of amyloid-β_{40}, amyloid-β_{42}, total tau and phosphorylated (p)-tau to examine the downstream consequences of changing predicted target transcript and protein expression. Our hypothesis is that novel findings will be acting on a background of pathological expression of both amyloid-β and tau, i.e. our effects would act as modifiers of known pathology.

We present an integrated, multi-level analysis of how the analysis of DNA, RNA, and protein data can facilitate the study of the relationships among genes and proteins and their impact on the human brain in the context of late-onset Alzheimer's disease. These targets are vetted through a multi-pass validation procedure including multiple types of analysis, replication across multiple datasets, *in silico* predictions and *in vitro* validations.

Materials and methods

All procedures are extensively detailed in the Supplementary material.

Samples

KRONOSII is a subset of data already presented (Corneveaux *et al.*, 2010) and contains samples from Alzheimer's Disease Research Center-funded US brain banks as well as six European and British brain banks. KRONOSII is a convenience cohort with low secondary pathology (i.e. Lewy body disease) and high pathology load in the late-onset Alzheimer's disease affected samples and low pathology load for controls. The second set (RUSH) includes subjects from two large, prospectively followed cohorts maintained by investigators at Rush University Medical Center in Chicago, IL: The Religious Orders Study and the Memory and Aging Project. The RUSH set is an epidemiologically based cohort with a greater mix of pathologies and pathological staging. There are 168 late onset Alzheimer's disease-affected samples and 177 unaffected samples with all datasets collected for the KRONOSII cohort. From the RUSH cohort 141 late-onset Alzheimer's disease affected samples and 153 unaffected samples with all datasets were collected. The average age for the KRONOSII cohort is 81, with 59% female subjects. The average age of the RUSH cohort is 88 and 63% of the subjects are female. Tissue sections were taken from frontal (82% of the sample) and temporal (18% of the sample) cortical regions.

Data collection

Genomic DNA samples were analysed on the Genome-Wide Human SNP 6.0 Array (Affymetrix) according to the manufacturer's protocols. Birdsuite (Korn *et al.*, 2008) was used to call SNP genotypes from CEL files. The DNA quality control pipeline was similar to that described in Anderson *et al.* (2010). cRNA was hybridized to Illumina HumanRefseq-HT-12 v2 Expression BeadChip. Expression profiles were extracted, background was subtracted and missing bead types imputed using the BeadStudio software. Normalization for the RNA profiles was performed using lumi (Du *et al.*, 2008) and limma (Ritchie *et al.*, 2015). Sample data were adjusted for several biological covariates (gender, age at death and cortical region) and several methodological covariates (institute source of sample, post-mortem interval, detection and hybridization date). Tandem mass spectrometry analysis was performed using an Exactive Orbitrap mass spectrometer (Thermo Scientific) outfitted with a custom electrospray ionization (ESI) interface. Identification and quantification of peptides was performed using the accurate mass and time tag approach (Zimmer *et al.*, 2006). Decon2LS was used for peak-picking and for determining isotopic distributions and charge states (Jaitly *et al.*, 2009). De-isotoped spectral information was loaded into VIPER to find and match features to the peptide identifications in the accurate mass and time tag database (Monroe *et al.*, 2007). Relative peptide quantitation was based on ratios between intensities of natural ^{16}O isotope containing peptides and reference peptides labeled with stable ^{18}O isotope at the carbonyl group at the C-terminus of the peptide.

Data analysis

Our data analysis pipeline is shown in Fig. 1. This was a multi-pass selection procedure to both uncover late-onset Alzheimer's disease risk targets and place them in the context of upstream regulation (allelic information) and downstream outputs (transcripts and peptides). Our goal was to identify a minimal set of high-confidence targets for validation. Our pipeline was performed in KRONOSII and RUSH separately after normalization to ensure independent replication.

Differential expression

Differential expression analysis was performed using limma (Ritchie *et al.*, 2015) comparing late-onset Alzheimer's disease and pathologically confirmed controls. Each dataset (KRONOSII, RUSH) was run independently. Multiple testing adjustment was performed using Benjamini-Hochberg correction [5% false discovery rate (FDR)]. Results were used to define seeding sets for downstream analysis.

Expression quantitative trait loci

MatrixeQTL (Shabalin, 2012) was used to predict allele-transcript relationships. Each dataset (KRONOSII, RUSH) was run independently. Permutations were used to correct for both the dependence between individual tests and for multiple testing (Supplementary material).

Network analysis

We carried out network analyses that took as input genomic, transcriptomic and proteomic profiles from the two datasets (KRONOSII and RUSH), in addition to external data derived from the literature, pathway databases (Molecular Signatures Database, Gene Ontology Database), and the Roadmap initiatives (Roadmap Epigenomics Consortium *et al.*, 2015). Our goal was to produce an output list of the main biological processes that are dysregulated in late-onset Alzheimer's disease, as well as a small list of the top key drivers impacting late-onset Alzheimer's disease associated processes. KRONOSII and RUSH were treated as independent datasets and the effects were compared across sets to determine replicated targets. Our pipeline included the following procedures (Fig. 1, Steps 3–6, dark orange squares): (i) constructing co-expression networks to identify sets of co-regulated genes associated with late-onset Alzheimer's disease pathology (Step 3) and determining pathways enriched in each network module (Step 4b); (ii) determining seeding gene sets associated with late-onset Alzheimer's disease pathology (Step 4a, Module Selection and Module Enrichment); (iii) building multiscale causal predictive networks (Step 5); and (iv) determining the key drivers that modulate states of the causal predictive network subnetworks (Step 6).

Co-expression networks

We constructed co-expression networks separately in controls and late-onset Alzheimer's disease samples. Additionally, co-expression networks were constructed separately in KRONOSII and RUSH. Single-scale networks consisted of transcripts only or proteins only. Multi-scale networks included transcripts plus proteins, with reduction of the transcript set to modules that were most enriched for differentially expressed genes. Our transcript co-expression networks consisted of all 15 297 transcripts (Supplementary material: Ancillary Dataset 4, KRONOSII Transcript co-expression networks; and Ancillary Dataset 5, RUSH Transcript co-expression networks), the protein-only co-expression networks consisted of 1931 peptides (Supplementary material: Ancillary Dataset 6, KRONOSII Peptide co-expression networks; and Ancillary Dataset 7, RUSH Peptide co-expression networks), and the multiscale co-expression networks consisted of 15 297 transcripts and 1931 peptides (Supplementary material:

Ancillary Dataset 8, KRONOSII multiscale co-expression networks; and Ancillary Dataset 9, RUSH multiscale co-expression networks). Prior to building our networks, we hypothesized that using peptide information may be more informative than using protein aggregate information in the context of networks (i.e. collapsing all peptides mapping to a single gene to one target). This hypothesis was based on data indicating that the four peptides mapping to *MAPT* showed quite different signals, with two peptides differentially expressed and two peptides not-significantly different. As network analysis is based on correlative structures and not genomic locations, we hypothesized that peptides that had functions implicated in late-onset Alzheimer's disease pathogenesis (i.e. the two differentially expressed *MAPT* peptides), might cluster separately from peptides that were unchanged in late-onset Alzheimer's disease (i.e. the two non-significant *MAPT* peptides). To test this, we constructed a co-expression network with only peptide data. For both KRONOSII and RUSH, peptides of the same protein did not always cluster in the same modules (Fig. 3A–D), supporting that using individual peptide level data may capture more diverse biology, since individual peptides may track completely different transcript modifications that may not have correlated levels of expression due to alterations in function. Examining the specific test case of *MAPT* peptides, in both KRONOSII and RUSH [which were quality controlled together, but predictions were run independently (Supplementary material)], the two differentially expressed peptides clustered together, and one of the non-significant peptides was consistently in another module. The last peptide (MAPT_HLSNVSSTGSIDMVDSPQLATLADEVSASLAK) was somewhat noisy, mapping outside of the differential expression peptide module in RUSH, but within that module in KRONOSII.

Causal predictive networks

While co-expression networks allow for descriptive characterizations of gene-protein relationships, causal relationships prediction is necessary for ordering of the network data into a hierarchy of relationships that in turn enables key driver analyses. While co-expression networks reflect only associative relationships, Bayesian networks infer directed edges that represent the direction of information flow. Bayesian network analysis can capture non-linear and combinatorial interactions. One limitation to standard Bayesian network analysis is that sometimes substructures within a Bayesian network are contradictory, which results in many directed edges having low confidence. To address this inherent limitation, we developed a novel causal predictive network approach, integrating a top-down Bayesian network approach with bottom-up causal inference that takes into account known causal relationships, which breaks the symmetry among contradictory causal structures and thus leads to higher confidence in edge directions.

The complexity of network building is a function of the number of nodes considered and sample size. We used all peptides in the network constructions; however, given the large number of probes used to query gene expression levels, we reduced the number of transcript probes to use in the causal predictive network reconstruction without losing important late-onset Alzheimer's disease gene and pathway information. We built gene-only co-expression networks and identified those modules enriched for differentially expressed genes, and then restricted causal predictive network construction to this subset of coherent late-onset Alzheimer's disease focused gene sets.

We focused our search on the identification of key drivers of network states associated with late-onset Alzheimer's disease, and thus used only late-onset Alzheimer's disease datasets. The seeding gene sets for both the KRONOSII and RUSH late-onset Alzheimer's disease datasets included modules enriched for differentially expressed transcript targets; therefore, pathways of relevance for late-onset Alzheimer's disease pathology were selected. We expanded these sets to include more than just differentially-expressed transcripts by including priors from a literature-based brain-specific network. Given the modest number of peptides measured, all peptides were used in the network models. Transcript data were reduced to the most crucial targets (Module Selection) and then expanded by including additional targets from the same pathways in curated databases (Module Enrichment). To ensure robust replication, KRONOSII and RUSH were pipelined as separate sets.

Key driver analysis

After the causal predictive network analysis was performed, the resulting predictive network models were examined using a key driver analysis algorithm. Key drivers are targets that have a significant impact on the regulatory states of other targets. Key drivers were predicted separately for KRONOSII and RUSH and overlaps determined. The late-onset Alzheimer's disease-associated subnetworks to which key driver analysis was applied were generated by projecting multiple different datasets onto the networks. First, we projected the module enrichment set only including differentially expressed transcripts from KRONOSII or RUSH. Second, we projected each module in its entirety including transcripts from KRONOSII or RUSH. Next, we projected the full differentially expressed transcript set from either KRONOSII or RUSH. For the peptide data, we first projected sets including both transcripts and peptides, performing the analysis separately on KRONOSII and RUSH. Finally, the entire peptide set was projected onto the transcript-peptide causal network, performing the analysis separately on KRONOSII and RUSH.

Data validation

While the identification of key drivers using the above approach is completely data driven, the inferences are statistical in nature and our results need to be verified experimentally. We used two different human cell lines which model the canonical amyloid-β and tau late-onset Alzheimer's disease pathways to validate hits outside the context of human brain tissue. Amyloid-β accumulation in plaques is thought to be the primary event in Alzheimer's disease pathogenesis (Hardy and Selkoe, 2002) and tau accumulation is one of the hallmark early features of Alzheimer's disease (Serrano-Pozo *et al.*, 2011); therefore, our hypothesis was that targets of interest should affect amyloid-β and tau levels further in the context of ongoing pathology. It is crucial for these studies to prove specificity of our effects to the specific single targets involved in late-onset Alzheimer's disease. While we and others have found general pathways such as inflammation that are involved in the pathogenic late-onset Alzheimer's disease process (Zhang *et al.*, 2013); those pathways have also been found in several other diseases (Miller and Raison, 2016; Miller *et al.*, 2017); and thus, while interesting, are

inappropriate for validating that our targets are specific to late-onset Alzheimer's disease declines and not neurodegeneration in general.

We used two different lines. First, all targets were transduced into the APP_{swe} HEK293 line. This is a human embryonal kidney cell line expressing amyloid-β complementary DNA bearing a double mutation [K595N and M596L; HEK293sw; gift from D. Selkoe, Boston, MA (Citron *et al.*, 1992)]. These cells produce 89-fold more *APP* mRNA than cells without the mutations (*t*-test *P*-value = 0.006). Levels of total tau protein and hyperphosphorylated tau were assessed in an H4 neuroglioma cell line engineered to overexpress four repeat tau [H4–4R0N, gift from T. Dunkley, Phoenix, AZ (Azorsa *et al.*, 2010)]. The H4–4R0N line produces ~5-fold more tau protein than cells without the construct. Using these lines demonstrated both specificity of effects to late-onset Alzheimer's disease, as amyloid-β processing in particular is a hallmark sign of disease, and replicated causality, since targets were modelled individually via overexpression or knockdown. Showing effects of the expression of single targets on amyloid-β and tau in external cell systems alleviates any concerns that computationally predicted brain tissue effects were merely the result of tissue degradation or age-related degeneration in general.

Data availability

All data are available through links at the Laboratory of Functional Genomics website (http://labs.med.miami.edu/myers/LFuN/LFuN.html). Further information and requests for resources and reagents should be directed to the corresponding author.

Results

APP and MAPT

It is possible that since both amyloid-β and tau proteins are deposited in late-onset Alzheimer's disease brains, they might be difficult to detect via tandem mass spectrometry techniques because of inefficient digestion; therefore, we first examined our data for known targets to make sure that corresponding peptides could be detected. To increase power, this analysis was performed across both series at once (n = 320 late-onset Alzheimer's disease, 338 pathology free controls, 115 mild cognitive impairment). The Consortium to Establish a Registry for Alzheimer's Disease (CERAD) (Mirra *et al.*, 1991) and Braak (Braak and Braak, 1995) staging were performed in each series. One peptide mapping to amino acids 17–28 of the amyloid-β peptide (LVFFAEDVGSNK, detected in 47% of the series) and four tau peptides (HLSNVSSTGSIDMVDSPQLATLADEVSASLAK, detected in 76% of the series; HVPGGGSVQIVYKPVDLSK, detected in 96% of the series; SGYSSPGSPGTPGSR, detected in 60% of the series; and IGSLDNITHVPGGGNK, detected in 98% of the series) were detected, thus demonstrating that our protocol was able to detect known late-onset Alzheimer's disease peptides. These peptides were examined to see if they were at increased levels in late-onset Alzheimer's disease as would be expected based on the amyloid hypothesis and our own work on microtubule associated protein tau (Myers *et al.*, 2005, 2007*b*). There were significant increases with the APP peptide and with two out of the four tau peptides detected (APP_LVFFAEDVGSNK: F = 33.23, P-value = 5.783×10^{-14}; MAPT_HVPGGGSVQIVYKPVDLSK: F = 44.77, P-value < 2.2×10^{-16}; MAPT_IGSLDNITHVPGGGNK: F = 90.137, P-value < 2.2×10^{-16}; MAPT_HLSNVSSTGSIDMVDSPQLATLADEVSASLAK: F = 0.4054, P-value = 0.667; MAPT_SGYSSPGSPGTPGSR: F = 1.4254, P-value = 0.242).

Peptide profiles were also examined to determine consistency with respect to late-onset Alzheimer's disease pathology. CERAD scores are a measure of neuritic plaque density corrected for age [see Table 1 in Mirra *et al.* (1991)]. The series mostly contains the oldest CERAD age group (age >75) where CERAD scores reflect neuritic plaque density and thus, this staging reflects amyloid-β levels. A consistent elevation in levels of APP_LVFFAEDVGSNK peptide was seen as CERAD scores progressed from 0 (no plaques) to C (moderate to frequent plaques depending on the age bracket). Braak staging quantifies the amount and cortical distribution of neurofibrillary tangles. Neurofibrillary tangles are composed of paired helical filaments of hyper-phosphorylated tau, and therefore, we hypothesized that as neurofibrillary tangle pathology was more widespread, tau peptides would show increased expression. This occurred for two of the tau peptides (MAPT_HVPGGGSVQIVYKPVDLSK and MAPT_IGSLDNITHVPGGGNK), which were the same peptides that were differentially expressed. Examining the alignment of these two peptides within the *MAPT* gene, the two peptides that were differentially expressed and correlated with Braak staging (MAPT_HVPGGGSVQIVYKPVDLSK and MAPT_IGSLDNITHVPGGGNK) aligned to the microtubule binding repeat regions of *MAPT*, whereas the two peptides that were not differentially expressed or correlated with Braak score aligned just outside the binding repeat regions.

Thus, our peptide data captured the known key targets involved in late-onset Alzheimer's disease pathogenesis. Peptides aligning to these targets were differentially distributed both with late-onset Alzheimer's disease diagnosis and post-mortem pathological measures. This finding gave us further confidence to proceed with an analysis of all datasets in both cohorts.

All targets

Differential expression

There were 8044 significantly differentially expressed transcripts in the KRONOSII set (Supplementary material: Ancillary Dataset 1, Differentially Expressed Transcripts) and 347 transcripts in the RUSH series (Supplementary material: Ancillary Dataset 1, Differentially Expressed Transcripts). These transcripts were used to seed the construction of the network models. For the peptide data, there

were 176 significant differentially expressed peptides comparing late-onset Alzheimer's disease and controls in the KRONOSII data (Supplementary material: Ancillary Dataset 2, Differentially Expressed Peptides) and 29 in the RUSH series (Supplementary material: Ancillary Dataset 2, Differentially Expressed Peptides). In comparing mild cognitive impairment tissue profiles to pathologically confirmed controls, no significant differences were observed; therefore, mild cognitive impairment data were only used to confirm levels from key targets and not for network construction.

Expression quantitative trait loci

We analysed allelic-transcript correlations and allelic-peptide correlations; however, only allelic-transcript relationships had significant expression quantitative trait loci. This is not surprising given that there is a direct relationship between DNA alleles and downstream RNA expression and furthermore, the path between DNA alleles to peptide profiles is considerably more convoluted. Additionally, peptide datasets are sparse, which can complicate analysis; therefore, the analysis of peptide data in the context of allelic drivers requires a more complex approach than standard protein quantitative trait loci single target metrics. For our data, peptides were incorporated at the multiscale network analysis level, which will allow for RNA to act as an intermediary signal (see the following section).

In a prior analysis (Webster *et al.*, 2009), we found that ~9% of transcripts we tested showed a genome-wide significant correlation with SNP genotype using a bootstrapping procedure for correction. In our current datasets, for KRONOSII ~12% of transcript probes tested were significantly correlated with allele dosage in *cis* (5% FDR), and for RUSH ~18% of transcript probes tested were significantly correlated with allele dosage in *cis* (5% FDR). Of these *cis* expression quantitative trait loci detected in each cohort, 1975 expression quantitative trait loci were overlapping between the sets (Supplementary material: Ancillary Dataset 3, Cis Overlapping expression quantitative trait loci), a 1.6-fold enrichment over what would have been expected by chance (Fisher's exact $P < 2.23 \times 10^{-191}$). Additionally, we replicated ~50% of the *cis* expression quantitative trait loci detected in our first report (Webster *et al.*, 2009), even though these studies comprised different sample sets and profiled using different microarrays. We also detected 113 *trans* expression quantitative trait loci in the KRONOSII set and 246 *trans* expression quantitative trait loci in the RUSH set. Of the *trans* expression quantitative trait loci identified, 40 were overlapping between the KRONOSII and RUSH sets, a 21-fold enriched over what would be expected by chance (Fisher's exact $P < 10^{-16}$). All *cis* hits are mapped in Fig. 2A and B.

Networks

Multiscale co-expression network

Examining the multiscale aggregate transcript-peptide networks, most modules were either predominantly comprised transcripts or peptides, demonstrating that these datasets are quantitatively independent (Supplementary Fig. 2). This result was not unexpected given previous correlation coefficient estimates of 0.27 between gene and corresponding peptide expression traits (Ghazalpour *et al.*, 2011). There are technical differences in collecting oligonucleotide profiles versus amino acid profiles, which may result in low levels of correlation. There are also biological differences in the way that oligonucleotides and amino acids are handled within a cell. Cleavage, sorting, and the timing of degradation can all act to vary amino acid profiles away from oligonucleotide profiles in biologically meaningful ways. *APP* represents a good example of this effect in late-onset Alzheimer's disease; transcript profiles are unchanged, but through differential cleavage, peptide profiles are altered. This further emphasizes the importance of analysing both RNA and peptide. Since multiscale co-expression network analysis is rooted in correlations, this method was not appropriate for discovery of connections between transcripts and peptides. Thus, we developed causal predictive networks, conditionally testing relationships and facilitating the identification of non-linear relationships between transcripts and peptides.

In these multiscale co-expression networks (transcript plus protein), there were 26 modules in KRONOSII late-onset Alzheimer's disease and 32 modules in RUSH late-onset Alzheimer's disease (Supplementary material: Ancillary Dataset 8, KRONOSII multiscale co-expression networks; and Ancillary Dataset 9, RUSH multiscale co-expression networks). For the control data, there were 25 and 32 modules for KRONOSII and RUSH, respectively (Supplementary material: Ancillary Dataset 8, KRONOSII multiscale co-expression networks; and Ancillary Dataset 9, RUSH multiscale co-expression networks). Figure 3E–H gives the functional enrichments for each module for the aggregate multivariate co-expression network predictions. The five most significantly enriched biological processes that replicated across the late-onset Alzheimer's disease KRONOSII and RUSH datasets were: (i) generation of precursor metabolites and energy; (ii) tissue development; (iii) response to unfolded protein; (iv) defence response; and (v) hydrogen peroxide catabolic process. In the control modules, the four most enriched biological processes that replicated across the KRONOSII and RUSH datasets were: (i) response to virus; (ii) response to unfolded protein; (iii) regulation of action potential in neuron; and (iv) RNA metabolic process.

While comparing the molecular signatures database and gene ontology processes gives a general idea of the overlap between KRONOSII and RUSH, this is not a direct comparison of module membership. Membership of all KRONOSII late-onset Alzheimer's disease modules was compared to RUSH late-onset Alzheimer's disease modules and KRONOSII control modules to RUSH control modules (Fig. 3I and J). In no case was there perfect overlap between KRONOSII and RUSH structure; however, there were several modules that had some degree of overlap, showing similarities between the two datasets.

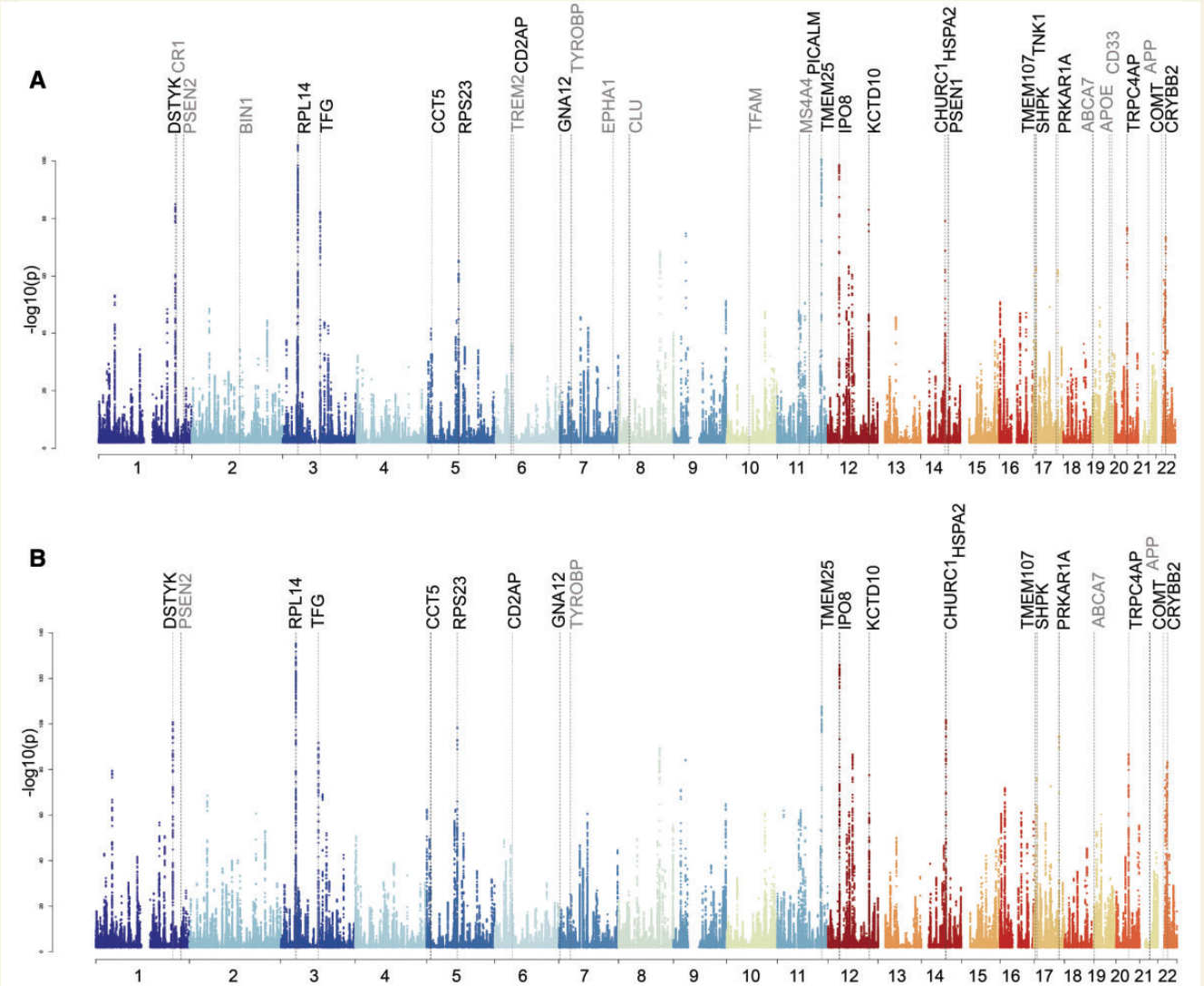

Figure 2 Cis expression quantitative trait loci. All *cis* expression quantitative trait loci hits are plotted for (**A**) KRONOSII and (**B**) RUSH. Results from each chromosome (*x*-axis) are highlighted in a different colour. Each point denotes one *cis* SNP–probe relationship. Prior genome-wide association studies and key driver hits are marked by dashed grey lines.

Multi-scale causal predictive network

There were six modules that were enriched for differentially expressed genes in the KRONOSII set and five modules in the RUSH set. The final directed networks consisted of 1931 peptides and 2465 interactions for KRONOSII peptides only network, and 1931 peptides and 2524 interactions for RUSH peptides only network; 8153 genes and 10 848 interactions for KRONOSII transcript only network, and 7357 genes and 9962 interactions for RUSH transcript only network; 10 160 genes, 1931 peptides, and 14 103 interactions for KRONOSII multiscale network, and 9338 genes, 1931 peptides, and 13 478 interactions for RUSH multiscale network.

Key driver analysis

The key driver results are shown in Fig. 4A and B. In total, there were 100 transcripts appearing in at least two networks and 105 peptides identified as key drivers. Of these key drivers, 80 transcript key drivers and 53 peptide key drivers were replicated between KRONOSII and RUSH. We selected three transcript hits and four peptide hits for experimental validation. Targets were prioritized that were differentially expressed and/or if they were an expression quantitative trait loci in the transcript dataset.

Data validation

Of the seven targets, one target (*ST18*) was not followed due to construct size and cost. The other six constructs were tested in the HEK293 and H4 lines. Of the other targets, three (*HSPA2*, *GNA12*, *COMT*) were overexpressed in at least one late-onset Alzheimer's disease cohort, and two were under expressed in late-onset Alzheimer's disease (*PDHB* and *RGS4*) (Table 1). For these constructs we replicated the late-onset Alzheimer's

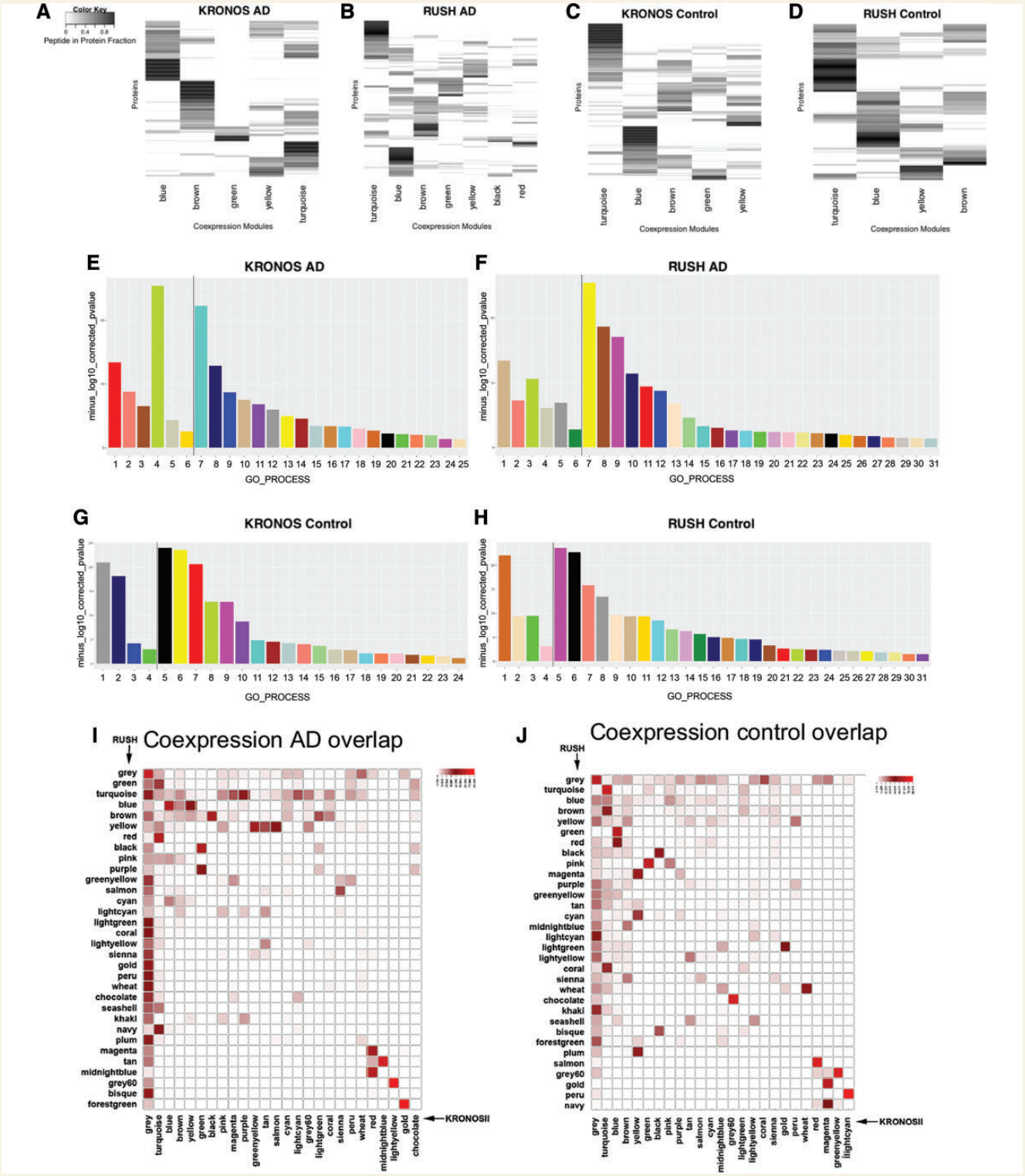

Figure 3 Networks. Shown are the fraction of peptides mapping to their corresponding gene target in each module used in the analysis for the (**A**) KRONOSII late-onset Alzheimer's disease set, (**B**) Rush late-onset Alzheimer's disease set, (**C**) KRONOSII Control set, and (**D**) RUSH Control Set. Darker colours indicate all peptides for a given target mapped to both the same module as well as to the same gene target. As can be seen on the figure, there is an imperfect correlation between module membership, gene mapping and peptide identity. Testing for whether counts of peptides for a particular protein mapped to the same or different modules was significant in both the KRONOSII (Fisher's exact *P*-value = 0.0002, alpha = 0.05), and RUSH sets (Fisher's exact *P*-value = 0.05, alpha = 0.05). In **E–H,** Gene ontology pathways are shown for modules from multiscale co-expression predictions that are enriched for differentially expressed targets from the (**E**) KRONOSII late-onset Alzheimer's disease dataset, (**F**) RUSH late-onset Alzheimer's disease dataset, (**G**) KRONOSII pathology-free dataset and (**H**) RUSH pathology-free dataset. The *x*-axis plots each module and *y*-axis is the $-\log_{10}$ *P*-value of the enrichment analysis. Modules and processes to the left of the line

(continued)

disease state, overexpressing *HSPA2*, *GNA12*, *COMT* and knocking down *PDHB* and *RGS4*. *CCT5* was not differentially expressed, but was followed as a key driver peptide. *CCT5* was overexpressed as a first pass of replication. Our goal was to obtain consistent measures of changes of amyloid-β and/or tau at multiple time points (48, 72 and 96 h) after transduction.

For *RGS4*, there was a significant downregulation of amyloid-β_{40} at all time points measured, but no effect on amyloid-β_{42} (Supplementary Fig. 3E). This drop in amyloid-β levels is counter to the effects seen in brain tissue. In the brain tissue having less *RGS4* was toxic and therefore, amyloid-β should be increased with less *RGS4*. Alternatively, the lower levels of *RGS4* could be reflecting end-stage protective compensatory mechanisms, and in that context the results make sense. The amyloid-β results for *PDHB* were for the most part non-significant, with only one time-point showing a difference in amyloid-β_{40} (Supplementary Fig. 4E). For tau, there were no significant results with *RGS4* nor was there any trend in the data (Supplementary Fig. 4E). For *PDHB*, there was no change in total tau and p-tau was significant at two out of three time points measured (Supplementary Fig. 4E). These effects are consistent with the brain tissue data, since there was less expression of *PDHB* in late-onset Alzheimer's disease brains therefore, tau should be increased with knockdown.

Several of the key driver overexpressed targets (*CCT5*, *COMT* and *GNA12*) significantly changed either levels of tau and p-tau (*CCT5* and *COMT*) or amyloid-β_{40} and amyloid-β_{42} (*GNA12*), but not both consistently (Supplementary Figs 5–7). Most of these results matched to what would be expected from the profiles in brain tissue, i.e. increases in the canonical pathological proteins with target overexpression. *CCT5* was the exception and showed consistent decreases in amyloid-β_{42}, counter to the expected overexpression. It is notable that there is less total RNA present with *CCT5* transduction (Supplementary Fig. 5); thus, this effect may be secondary to HEK cell death. Additionally, as with *RGS4*, *CCT5* overproduction could be protective and compensatory to pathogenic processes. Finally, *CCT5* is not significantly changed in terms of differential expression (Supplementary Fig. 5A); therefore, more complex modelling rather than just overexpression may be required.

Our best validated target from the key driver prediction using the transcript dataset was *HSPA2*. This target significantly elevated amyloid-β_{40} and amyloid-β_{42} at all time

Figure 3 Continued

are replicated across sets. Colours are kept consistent to arbitrary assignments by Weighted Gene Co-expression Network Analysis. Processes are as follows. (**E**) (KRONOSII AD): 1, Generation of precursor metabolites and energy; 2, Tissue development; 3, Response to unfolded protein; 4, Defence response; 5, Hydrogen peroxide catabolic process_1; 6, Hydrogen peroxide catabolic process_2; 7, Translational elongation; 8, Gluconeogenesis; 9, Neurological system process; 10, Response to stress; 11, Glial cell differentiation; 12, Blood vessel development; 13, Cellular catabolic process; 14, Respiratory electron transport chain_1; 15, Respiratory electron transport chain_2; 16, Axon guidance; 17, Response to electrical stimulus; 18, Response to chemical stimulus; 19, Negative regulation of cellular biosynthesis; 20, DNA recombination; 21, Regulation of microtubule-based process; 22, Ether metabolic process; 23, Negative regulation of transcription from RNA polymerase II promoter; 24, Intracellular protein transmembrane import; 25, Telomere maintenance. (**F**) (RUSH AD): 1, Generation of precursor metabolites and energy; 2, Tissue development; 3, Response to unfolded protein; 4, Defence response; 5, Hydrogen peroxide catabolic process_1; 6, Hydrogen peroxide catabolic process_2; 7, Nuclear-transcribed mRNA catabolic process; 8,Type I interferon signalling pathway; 9, Cellular respiration_1; 10, Ion transport; 11, Ensheathment of neurons; 12, Cellular respiration_2; 13, Protein polymerization; 14, Regulation of protein complex disassembly; 15, Negative regulation of gene expression; 16, Cellular macromolecule metabolic process; 17, Neuron development; 18, Positive regulation of MAPK cascade; 19, Microtubule bundle formation; 20, RNA Methylation; 21, Single-organism behaviour; 22, Carboxylic acid metabolic process; 23, Glycosphingolipid metabolic process; 24, Regulation of mRNA catabolic process; 25, Cell volume homeostasis; 26, Oxidation-reduction process; 27, Mitotic Spindle assembly checkpoint 28, G2 DNA damage checkpoint; 29, Regulation of defence response to virus by virus; 30, RNA processing; 31, Positive regulation of signalling. (**G**) (KRONOSII control): 1, Response to virus; 2, Response to unfolded protein; 3, Regulation of action potential in neuron; 4, RNA metabolic process; 5, Organic substance catabolic process; 6, Regulation of immune system process; 7, Cellular respiration; 8, Monosaccharide biosynthetic process; 9, Single-organism transport; 10, Extracellular matrix organization; 11, Generation of precursor metabolites and energy; 12, Hydrogen transport; 13, Hydrogen peroxide catabolic process; 14, Single-multicellular organism process; 15, RNA splicing via transesterification reactions with bulged adenosine as nucleophile; 16, Protein dephosphorylation; 17, Synapse organization; 18, Retrograde vesicle-mediated transport Golgi to ER; 19, Negative regulation of cellular carbohydrate metabolic process; 20, Synaptic transmission; 21, Amyloid precursor protein metabolic process; 22, Centrosome duplication; 23, Cellular response to growth factor stimulus; 24, Cilium assembly. (**H**) (RUSH control): 1, Response to virus; 2, Response to unfolded protein; 3, Regulation of action potential in neuron; 4, RNA metabolic process; 5, Defence response_2; 6, Cellular membrane organization; 7, Generation of precursor metabolites and energy_1; 8, Generation of precursor metabolites and energy_3; 9, Translational termination; 10, Cellular response to zinc ion; 11, Ion transport; 12, Defence response_1; 13, Protein folding; 14, Angiogenesis; 15, Skeletal muscle cell differentiation; 16, Proton transport; 17, Hydrogen peroxide metabolic process; 18, Generation of precursor metabolites and energy_2; 19, Regulation of synaptic plasticity_2; 20, Glutamate receptor signalling pathway; 21, Peptidyl-glutamic acid modification; 22, Regulation of RNA metabolic process; 23, Androgen receptor signalling pathway; 24, Nuclear-transcribed mRNA catabolic process exonucleoytic; 25, Synapse maturation,; 26, Regulation of synaptic plasticity_1; 27, Cerebral cortex development; 28, Dephosphorylation; 29, RNA stabilization; 30, Phagocytosis; 31, Response to copper ion. (**I** and **J**) Heatmaps of the overlap between KRONOSII and RUSH for the (**I**) late onset Alzheimer's disease datasets and (**J**) control datasets. Again, seeding datasets and module prediction were performed completely independently for each dataset; therefore, this is a true replication. As can be seen, many modules had low overlap between sets; however, there were several modules where membership was highly overlapping (dark red) indicating that module prediction can replicate from series to series.

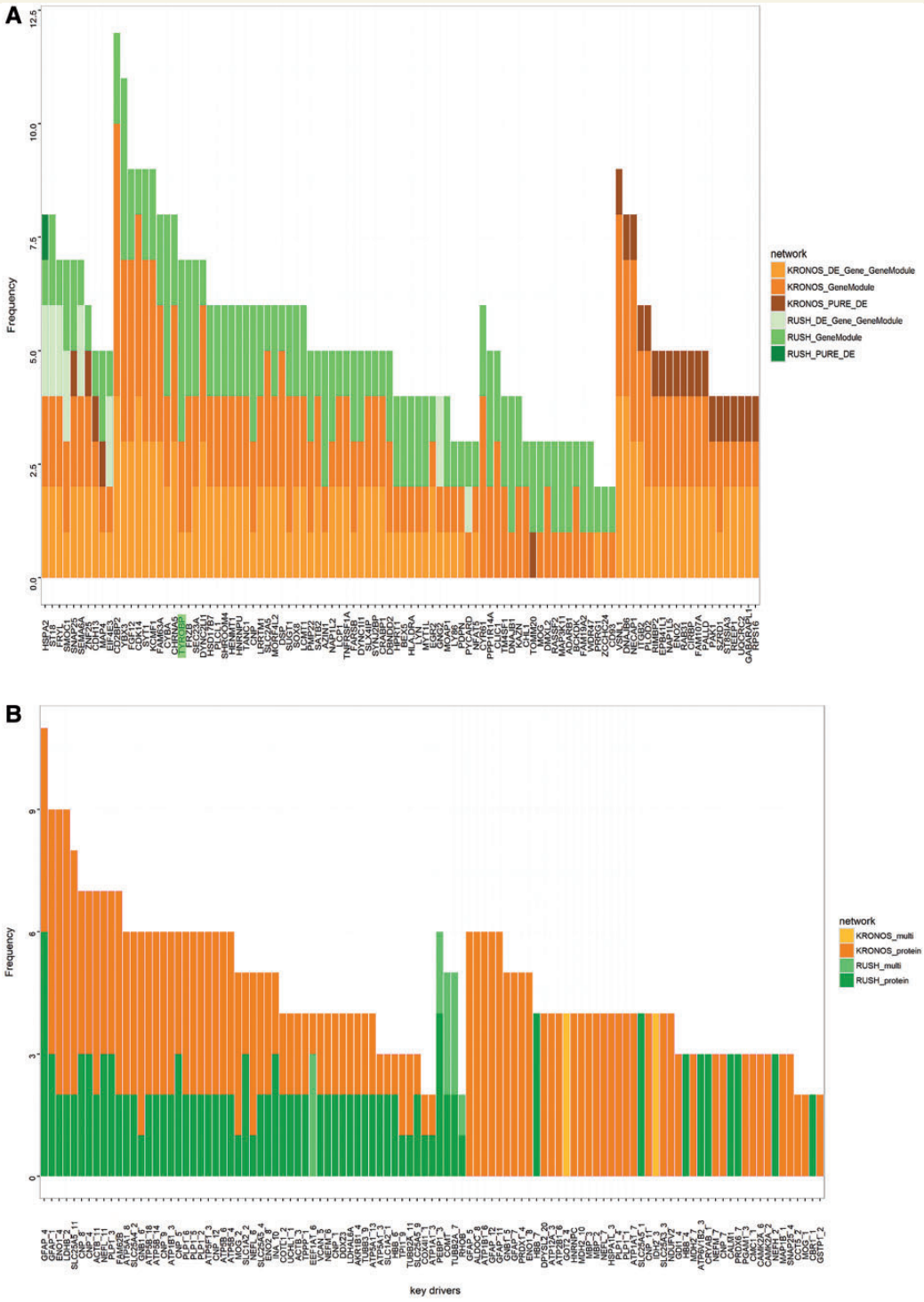

Figure 4 Key driver analysis. (**A**) Transcripts. Shown is the graph counting the significant over-representation of particular key drivers in the networks using the transcript dataset as the projection series. Four separate networks were examined: (i) KRONOSII causal predictive transcript network; (ii) RUSH causal predictive transcript network; (iii) KRONOSII causal predictive transcript and peptide network; and (iv) RUSH causal predictive transcript and peptide network. The colour of the boxes represents which dataset the key driver originates from, and the shade represents which seeding gene list it belongs to. There were six seeding gene lists used: (i) the intersection of each module transcripts with differentially expressed transcripts from KRONOSII (KRONOS_DE_Gene_GenModule); (ii) the module transcripts from KRONOSII (KRONOS_GeneModule); (iii) the full differentially expressed transcript set from KRONOSII (KRONOS_PURE_DE); (iv) the intersection of each module transcripts with differentially expressed transcripts from RUSH (RUSH_DE_Gene_GenModule); (v) the module transcripts from RUSH (RUSH_GeneModule); and (vi) the full differentially expressed transcript set from RUSH (RUSH_PURE_DE). The *x*-axis includes the top key drivers, the *y*-axis counts the number of times the target is a key driver in any of the modules. Targets can be counted greater than four times if they appear in multiple replicated modules. Green highlights *TYROBP*. (**B**) Peptides. Shown is the graph counting the significant over-representation

(continued)

points measured (Fig. 5B and E). Tau and p-tau were also elevated, with every time point significant except for the first collection of p-tau (Fig. 6C and D). Correcting for cell densities, *HSPA2* gave a 1.8-fold increase in pathologically processed amyloid-β_{40}, a 1.6-fold increase in amyloid-β_{42}, a 2.2-fold increase in total tau and a 3.4-fold increase in p-tau (Table 2). While these changes are modest, given that in our cells amyloid-β and tau protein are already overexpressed ~20-fold and ~5-fold, it is encouraging that further consistent increases can be obtained.

Discussion

Through our analysis of pathologically confirmed brain tissues, we have shown the following: (i) DNA–RNA–protein networks are robust and replicable; (ii) protein profiling uncovered novel key drivers and was crucial to understanding data outputs; (iii) inclusion of mild cognitive impairment subjects added modest value to the screen, since there were no significant differentially expressed genes; (iv) having two distinct datasets was crucial, since not all mapped processes replicated; (v) the defence response is a major driver of late-onset Alzheimer's disease differences, which replicates prior findings (Zhang *et al.*, 2013); (vi) there are other replicated major processes beyond defence response, indicating the potential for further hits; (vii) replicated key drivers have downstream effects on the amyloid-β and tau canonical pathways; and (viii) by examining each target in isolation in systems that mimic late-onset Alzheimer's disease pathology, we have firmly demonstrated that the effects we have reported are both specific to late-onset Alzheimer's disease and are not the result of secondary declines due to technical artefacts, agonal state or neurodegeneration in general. There are some limitations to the work. First, while our sample sizes are appropriate for 'omics work given the hypothesis-free nature of network analysis, our sample is smaller than most genome-wide association studies screens. Second, given the nature of brain tissue and the extensive data collection involved in this work, there is the potential for noise additions at each step. We accounted for this by performing replications with two independent series and cell culture work on each single target to validate effects, but it is still a possible factor. Finally, much more extensive phenotyping needs to be performed to determine the exact nature of the relationships between targets and APP or tau response. The work presented is an initial step of many to dissect the true nature of this pathology.

Single target effects

Single targets were assessed using differential expression and expression quantitative trait loci analysis. It is notable that there was a considerable difference in differential expression outcomes whereby the RUSH cohort had fewer differences. Since data were normalized at the same time, using the same procedures, this is likely to not be quality control variability. It is possible that this is because of the differences in the nature of the collections. KRONOSII is an extremely selected cohort, with little secondary pathology. This is not the case with RUSH where there is a greater mixture of pathologies. We had significant overlap between sets for expression quantitative trait loci predictions, both in the *cis* sets and in the *trans* set as well as our original report.

Major network effects

There were five main processes that were significantly enriched in the late-onset Alzheimer's disease modules and replicated across the KRONOSII and RUSH datasets: generation of precursor metabolites and energy, tissue development, response to unfolded protein, defence response, and hydrogen peroxide catabolic process. Generation of precursor metabolites and energy are processes involved in the mitochondrial electron transport chain or glycolysis. This is consistent with existing data in that mitochondrial dysfunction in late-onset Alzheimer's disease has been mapped in many studies (Hong *et al.*, 2008; Moreira *et al.*, 2010). Additionally, changing glycolytic pathways might change susceptibility to amyloid-β late-onset Alzheimer's disease pathology (Fu *et al.*, 2015) and cell mis-metabolism is likely a general process in neurodegeneration (Ngo and Steyn, 2015). Tissue development is a broad term with ~30 subprocesses, including tissue regeneration as a subclass. This module likely reflects processes involved in tissue repair and maintenance. The unfolded protein response (UPR) is a mechanism for cells to compensate for accumulation of unfolded proteins within the endoplasmic reticulum. This response involves an upregulation of resident

Figure 4 Continued
of particular key drivers in the networks using the peptide dataset as the projection series. Four separate networks were examined: (i) KRONOSII causal predictive peptide network; (ii) RUSH causal predictive peptide network; (iii) KRONOSII causal predictive transcript and peptide network; and (iv) RUSH causal predictive transcript and peptide network. The colour of the boxes represents which dataset the key driver originates from and the shade represents which seeding gene list it belongs to. There were four seeding gene lists used: (i) the full set of transcripts and peptides from KRONOSII (KRONOS_multi); (ii) the entire peptide set from KRONOSII (KRONOS_protein); (iii) the full set of transcripts and peptides from RUSH (RUSH_multi); and (iv) the entire peptide set from RUSH (RUSH_protein). The *x*-axis includes the top key drivers, the *y*-axis counts the number of times the target is a key driver in any of the modules from the module enrichment set. Targets can be counted greater than two times if they appear in multiple replicated modules.

Table 1 Summary of pipelined targets

Name	Gene, bp	Construct, bp	Dataset	Load expression	Transduction	HEK293SW	H4–4R0N
HSPA2	7332	8990	KD transcript KRONOSII and RUSH; DE KRONOSII; no expression quantitative trait loci	KRONOSII UP RUSH NS	Worked	Increased amyloid-β 40 and 42	Increased tau and p-tau
ST18	299 047	10214	KD transcript KRONOSII and RUSH; DE KRONOSII; no expression quantitative trait loci	KRONOSII UP RUSH NS	Not ordered	Not measured	Not measured
RGS4	8196	872	DE KRONOSII; no expression quantitative trait loci	KRONOSII DOWN RUSH NS	Worked	Decreased amyloid-β 40	No significant change
CCT5	16460	8696	KD protein KRONOSII; expression quantitative trait loci KRONOSII and RUSH	KRONOSII NS RUSH NS	Worked	Decreased amyloid-β 42	Increased tau and p-tau
COMT	28235	7866	KD protein RUSH; DE KRONOSII; no expression quantitative trait loci	KRONOSII UP RUSH NS	Worked	No significant change	Increased tau and p-tau
GNA12	116224	8216	DE KRONOSII; expression quantitative trait loci KRONOSII and RUSH	KRONOSII UP RUSH UP	Worked	Increased amyloid-β 40 and 42	No significant change
PDHB	6222	8720	DE KRONOSII; expression quantitative trait loci KRONOSII and RUSH	KRONOSII DOWN RUSH NS	Worked	No consistent significant change	Increased p-tau some time points

The table lists the summary data for all targets pipelined in the validation screen testing how targets affect levels of pathologically processed amyloid-β and tau. The first column lists each target and the second and third columns list gene and construct sizes. Datasets examined are listed in the fourth column and include: (i) whether the target was a key driver in the differentially expressed transcript or peptide projected series and replicated in both datasets; (ii) whether the target was differentially expressed or for peptides, whether the matching transcript was differentially expressed; and (iii) whether the target was an expression quantitative trait loci. The fifth column lists the differential expression results and the direction of effect in late-onset Alzheimer's disease samples. The sixth column lists the transduction status. *ST18* was not ordered due to size and cost. The last two columns list the summary of results for the HEK293sw and the H4–4R0N line. DE = differential expression; DOWN = decreased expression in Alzheimer's disease; KD = key driver; NS = non-significant; UP = increased expression in Alzheimer's disease.

chaperone genes to facilitate folding. Specific to late-onset Alzheimer's disease, the UPR has been implicated in familial Alzheimer's presenilin 1 (PSEN1) toxicity (Katayama *et al.*, 1999) as well as more generally in the earliest stages of late-onset Alzheimer's disease neuropathology (Hoozemans *et al.*, 2009). Our finding that defence response is a replicated process is in line with our prior report (Zhang *et al.*, 2013), where *TYROBP* was mapped as a major effect and generally immune and inflammatory processes are overrepresented. *TYROBP* was a key driver of effects in both series. Finally, both KRONOSII and RUSH were enriched for targets that are involved in pathways resulting in the breakdown of hydrogen peroxide. Due to high oxygen consumption, brain tissue generates hydrogen peroxide along with other reactive oxygen species. Amyloid-β can bind catalase and specifically inhibit the breakdown of hydrogen peroxide (Milton, 1999) and this may contribute to neuropathology.

Specific network targets

Of the selected targets, we obtained results for six of them: one target from the transcript set, two from the peptide set and three targets based on differential expression or eQTL significance. In the transcript key driver dataset, *HSPA2* gave the most robust results through our pipelines. The transcript was a key driver in both datasets, appearing four times in the KRONOSII predictions and four times in the RUSH predictions. It was also differentially expressed in the KRONOSII dataset. Examining *HSPA2* in the context of overexpressed amyloid-β and tau gave significant results in all cases; *HSPA2* overexpression further drove production including the more toxic amyloid-β_{42} and p-tau subspecies.

HSPA2 is a member of the larger Hsp70 group of heat shock protein genes. Heat shock proteins were first identified for their role in protein folding and the chaperone system; however, further data now indicate a wider involvement in a vast array of cell processes such as synaptic transmission, autophagy, endoplasmic reticulum stress response, protein kinase and cell death signalling (Stetler *et al.*, 2010). Hsp70 has been extensively studied in Alzheimer's disease and Down syndrome. Counter-intuitively, elevations in Hsp70 levels are thought to be neuroprotective (Muchowski and Wacker, 2005; Leak, 2014); however, to date, most of the studies have been focused on the stress induced forms (Hsp70-A1 and Hsp70-A2), which are encoded by the *HSPA1A* and *HSPA1B* genes and not *HSPA2* (Leak, 2014). In mammalian systems there are 13 separate Hsp70 genes and *HSPA2* encodes the minor form of the constitutively active species of Hsc70, with the major Hsc70 form being encoded by *HSPA8*. Besides differences in activity and response, there are differences in expression between Hsp70 and Hsc70 with Hsc70 being the major form in brain (Daugaard *et al.*, 2007). Thus, our finding that *HSPA2* elevations act to proliferate late-onset Alzheimer's disease pathology

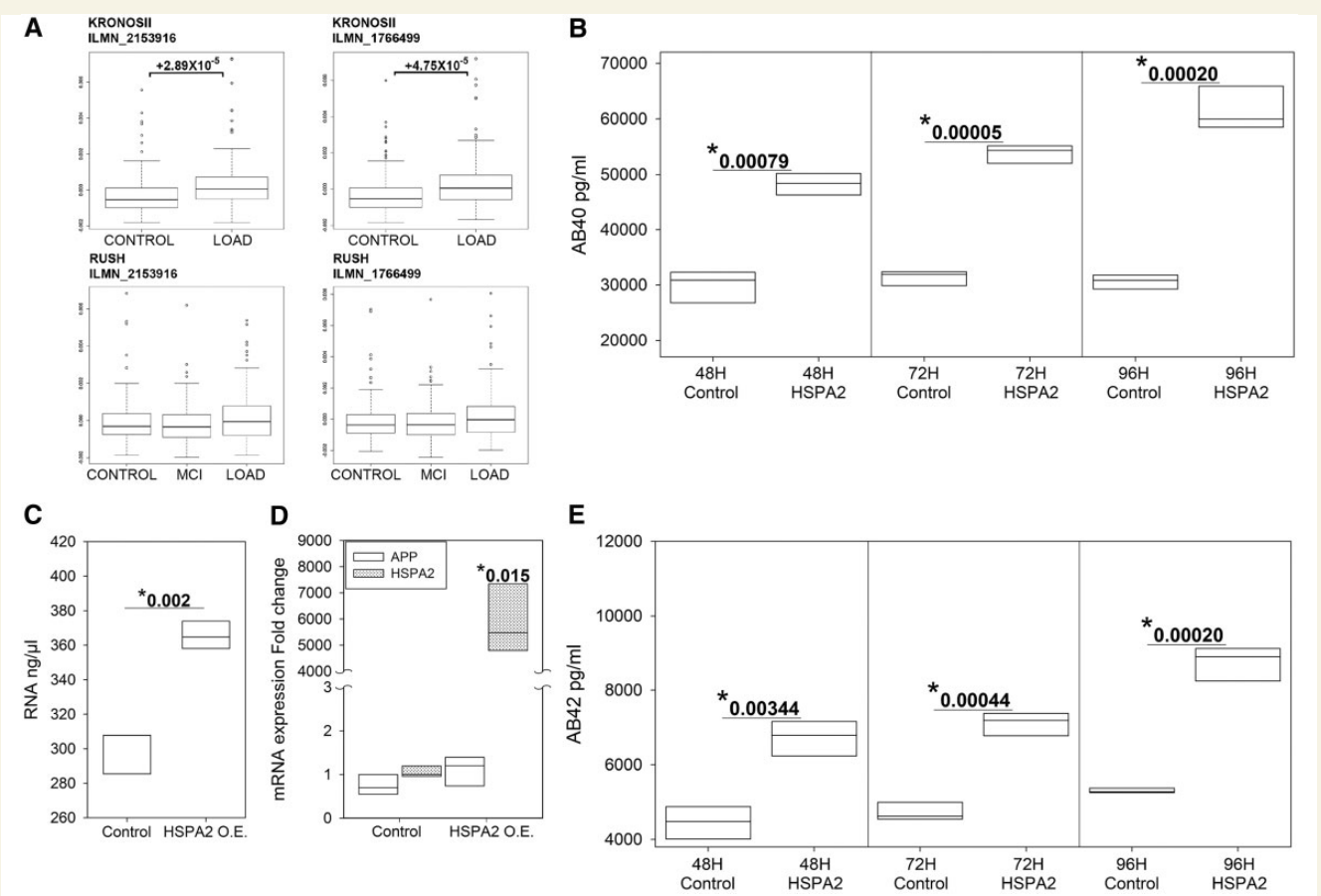

Figure 5 HSPA2 APP measures. (**A**) Shows the level of transcript expression for HSPA2 for the two detected probes in both KRONOSII (*top*) and RUSH (*bottom*). Only KRONOSII showed significant differential expression. The levels of total RNA (**C**, 96 h shown, measured as a surrogate of the level of cell death), transcript overexpression both for target and *APP* (**D**, 96 h shown, boxplot of three replicates), amyloid-β_{40} peptide levels (**B**) and amyloid-β_{42} peptide levels (**E**) for three repeat measures of conditioned media at three different time points of the top key driver target in the HEK293sw cell line are plotted. Measurements are taken at 48, 72 and 96 h post transduction. Control = measurements from cells transduced with an empty vector; HSPA2 = measurements from cells transduced with target. $^{+}$limma *P*-value; **t*-test *P*-value.

might not be counter to the previous results, but instead represent a separate related pathway.

Our finding that *HSPA2* acts to specifically increase levels of amyloid-β_{40}, amyloid-β_{42}, total tau and p-tau suggests that changes in *HSPA2* are not merely due to secondary or technical effects such as agonal state, but are specific to late onset Alzheimer's disease processes. Additionally, performing an unbiased hypothesis-free screen was crucial to detect *HSPA2* since it is neither the most studied inducible Hsp70 form, nor even the major component of the constitutive form.

Interestingly, *HSPA2* maps within a non-significant linkage peak found on chromosome 14 using a series of sibling pairs collected from late-onset Alzheimer's disease families (Myers *et al.*, 2002). *HSPA2* was originally suggested as an early-onset Alzheimer's disease gene due to its location close to markers thought to map near the AD3 locus (Cruts *et al.*, 1995) as well as its known significant association with Alzheimer's disease pathology (Hamos *et al.*, 1991). After a reassessment of the linkage maps, the AD3 locus was eventually mapped as PSEN1 (Sherrington *et al.*, 1995). The late-onset Alzheimer's disease chromosome 14 locus has yet to be mapped. It is notable that this is not a significant linkage peak, but it is intriguing that *HSPA2* maps within the region.

RGS4 was under-expressed in late-onset Alzheimer's disease and decreased amyloid-β_{40} at all time points, with no effects on amyloid-β_{42}, tau and p-tau; thus, it was one of the weakest of all of our validations. RGS proteins activate GTP hydrolysis by the alpha subunit of heterotrimeric G proteins and by this means inhibit G-protein coupled receptor (GPCR) signalling. *RGS4* has been associated with the development of schizophrenia (Chowdari *et al.*, 2002). GPCR signalling is generally implicated in amyloid-β processing (Thathiah *et al.*, 2009) and targeting GPCRs has been suggested as a possible therapeutic pathway; however, *Rgs4* knockout mice did not have any deficits in associative learning or working memory (Grillet *et al.*, 2005), so it may be that knockdown of *RGS4* is not the best GPCR pathway target and *GNA12* may be a more viable alternative.

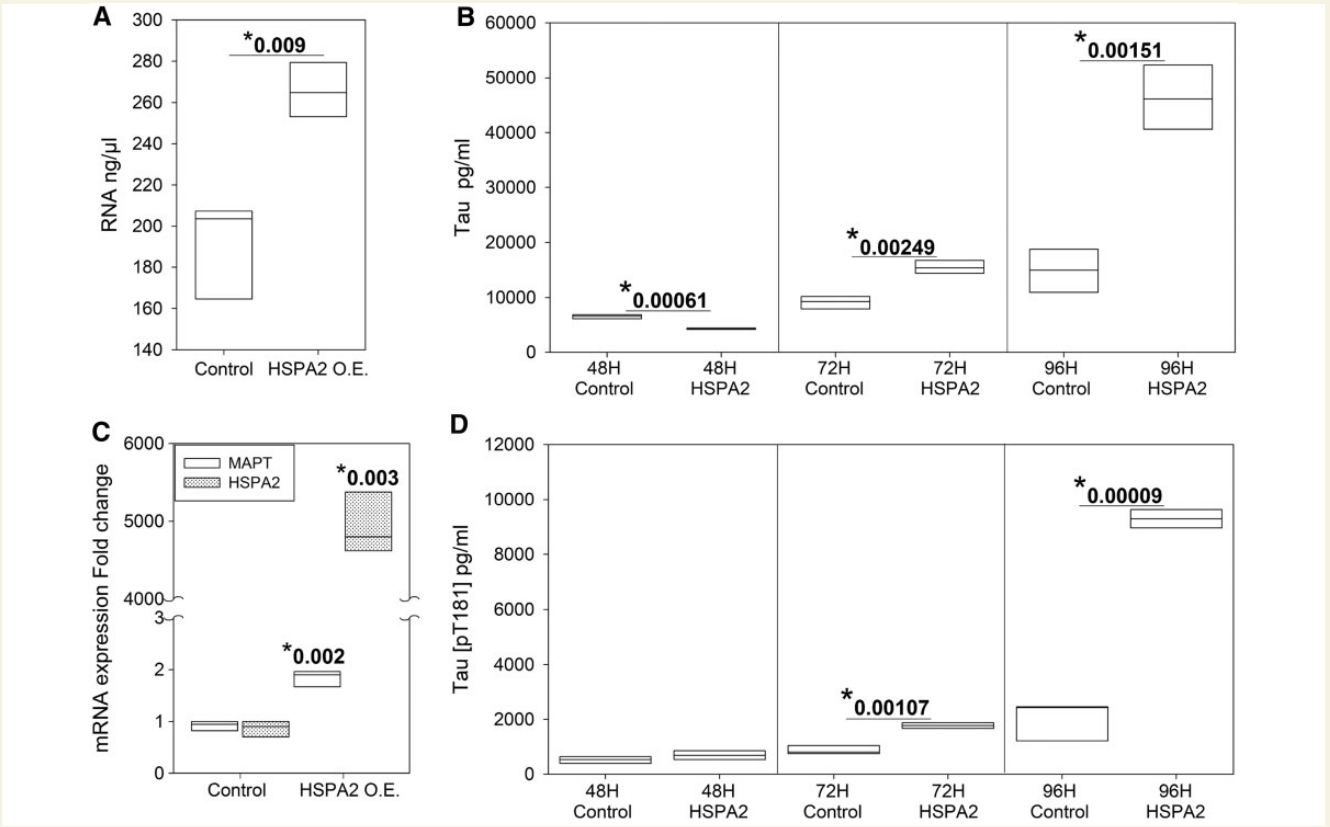

Figure 6 HSPA2 tau measures. Shown in the figure are the levels of total RNA (**A**, 96 h shown, measured as a surrogate of the level of cell death), transcript overexpression both for target and *MAPT* (**C**, 96 h shown, boxplot of three replicates), total tau peptide levels (**B**) and p-tau peptide levels (**D**) for three repeat measures of conditioned media at three different time points of the top key driver target in the H4–4R0N cell line. Measurements are taken at 48, 72 and 96 h post transduction. Control = measurements from cells transduced with an empty vector; *HSPA2* = measurements from cells transduced with target. *t-test *P*-value.

Table 2 HSPA2 fold-change

Cell line	Peptide	Empty vector	HSPA2 OE	Fold change
HEK293sw	Amyloid-β_{40}, pg/ml	30 656	54 523	1.8
	Amyloid-β_{42}, pg/ml	4826	7533	1.6
H4–4R0N	Tau[Total], pg/ml	10 190	22 055	2.2
	Tau[pT181], pg/ml	1141	3918	3.4

Fold-change calculations for HSPA2 in the HEK293sw line and H4–4R0N lines. OE = overexpression.

In the peptide dataset, *CCT5* was counted in two KRONOSII modules and *COMT* was counted in five RUSH modules. Both of these targets significantly elevated tau and p-tau at most time points; however, for *COMT* there was no effect on amyloid-β_{40} or amyloid-β_{42} levels. *CCT5* only had a significant effect on amyloid-β_{42} in that there were decreased levels of amyloid-β_{42} released at multiple time points, but it is possible this change is a secondary effect due to cell death, since total RNA levels were decreased over the course of the experiment.

CCT5 is a member of the same chaperonin complex as *HSPA2* (Neef *et al.*, 2014). Specifically, *CCT5* is a member of the TCP1 complex, also known as TRiC. This complex folds various proteins including actin (Gao *et al.*, 1992) and tubulin (Yaffe *et al.*, 1992). Hsp70 co-purifies with *CCT* and it's possible that they directly interact (Lewis *et al.*, 1992; Kubota *et al.*, 1994); however, it is unclear whether those results are specific to the forms we have mapped.

COMT is the major catecholamine degrading enzyme, acting in neurons and microglia after uptake from the synaptic cleft. It was originally mapped as a gene of interest for schizophrenia (Mier *et al.*, 2010). There have been some studies of *COMT* variation and Alzheimer's disease risk;

however, none of the large genome-wide association studies meta-analyses (Lambert *et al.*, 2013) have replicated *COMT* as a gene for late-onset Alzheimer's disease. Our own work did not show a direct relationship between DNA alleles and *COMT* mRNA or peptide expression.

In our cohorts, *COMT* was seen to be overexpressed in late-onset Alzheimer's disease brains, which would be consistent with findings of reduced dopamine in Alzheimer's brain tissues (Storga *et al.*, 1996). *COMT* overexpression increased levels of tau and p-tau, but not amyloid-β. This result appears to be at odds with the known synergies between tau and *COMT*. Activation of dopamine D1 receptors causes protein kinase A (PKA) activation and results in a downstream chain of events leading to tau hyperphosphorylation (Lebel *et al.*, 2009); therefore, higher levels of synaptic dopamine and lower *COMT* activity in late-onset Alzheimer's disease would be expected. However, imaging data have shown that low dopamine tone is associated with an upregulation of binding at D1 receptors (Guo *et al.*, 2003), and the higher activity Val allele (similar to an increase in *COMT* expression) shows increases in D1 receptor binding (Slifstein *et al.*, 2008). Thus, our results would indicate a model whereby there is mis-regulation of the dopamine-*COMT* systems and insufficient brain compensatory mechanisms result in tau changes.

Two targets were followed because they were expression quantitative trait loci in both sets and differentially expressed in KRONOSII. *GNA12* was upregulated in late-onset Alzheimer's disease and *PDHB* was downregulated in late-onset Alzheimer's disease in the KRONOSII data.

Like *RGS4*, *GNA12* is a part of the GPCR system. It is a member of the alpha class of heterotrimeric G proteins, which typically upon exchange of GDP to GTP, activates downstream intercellular signalling pathways. *GNA12* along with *GNA13* represents a fourth class of alpha subunits (Strathmann and Simon, 1991). *GNA12* can regulate actin cytoskeleton remodelling and along with its GPCR partners is a known activator of Rho (Riobo and Manning, 2005). Upstream coupling can occur through direct interactions with Rho guanine nucleotide exchange factors (RhoGEFs) (Hart *et al.*, 1998), *ARHGEF11* (Jackson *et al.*, 2001), cell adhesion molecules (Meigs *et al.*, 2001) and other effectors. HSP90 was shown to directly bind GNA12 (Vaiskunaite *et al.*, 2001), thus implicating *GNA12* in heat shock protein signalling. Hsc70 and Hsp90 have shown to demonstrate cooperative binding in other systems (Rajapandi *et al.*, 2000; Iwasaki *et al.*, 2010) and are linked together through Hsp70-Hsp90 organizing protein (Hop) (Johnson *et al.*, 1998). It has also been shown that the N-terminus of GNA12 contains a mitochondrial targeting sequence and is involved in the regulation of mitochondrial motility, morphology and membrane permeability (Andreeva *et al.*, 2008).

GPCRs are known to affect APP processing via actions on cleavage enzymes. Through direct coupling to beta and gamma secretase, GPR3 has been shown to potentiate gamma secretase APP cleavage in a screen for modulators of amyloid-β production (Thathiah *et al.*, 2009). GPR3 associates with the G(s) G protein subunit; therefore, it is unlikely that there is a direct interaction between GNA12 and GPR3. It remains to be seen whether there is a direct interaction between APP processing enzymes and GNA12, as with GPR3, but our data suggest there is some modification of APP processing by GNA12.

PDHB is a part of the PDH complex, which is a nuclear-encoded mitochondrial multi-enzyme complex that catalyses the overall conversion of pyruvate to acetyl-CoA and carbon dioxide and provides the primary link between glycolysis and the tricarboxylic acid (TCA) cycle. Mitochondrial dysfunction has been repeatedly mapped to late-onset Alzheimer's disease, and some studies have reported oxidative damage preceding plaque formation, indicating putative causation (Nunomura *et al.*, 2001).

Knockdown of *PDHB* increased levels of p-tau in our validation cell lines, but had no effect on total tau. PDH has been shown to directly bind GSK3B, which results in PDH phosphorylation and lowers PH activity. GSK3B also phosphorylates tau, thus the link between *PDHB* and tau effects in our data is probably via GSK3B (Hoshi *et al.*, 1996). In primary rat hippocampal culture, amyloid-β exposure inactivated PDH and resulted in mitochondrial dysfunction, lowered acetylcholine levels in cholinergic neurons and neuronal cell death (Hoshi *et al.*, 1996); therefore, amyloid-β could be upstream of *PDHB*, which would explain why we see no changes in amyloid-β. Our human data fit in nicely with effects mapped in animal models and through our analysis we can specifically target the toxic sub-component of PDH.

In summary, this study has mapped and validated six novel targets via a hypothesis-free approach to uncovering misregulation in DNA, RNA and protein relationships in late-onset Alzheimer's disease. Many of these targets replicated between two distinct neuropathological datasets that were collected in different manners (convenience and epidemiological) from different sources. Nominated targets had some effect on amyloid-β_{40}, amyloid-β_{42}, tau or p-tau in separate experiments using two different cell lines and individual transduction of each target, demonstrating both specificity to known late-onset Alzheimer's disease pathways as well as causality. It is notable that our selection pipeline was hypothesis-free; therefore, we did not enrich for hits that would yield positive results, and indeed some targets like *RGS4* were not as consistent in our assays. While our main targets mapped to pathways already implicated in late-onset Alzheimer's disease including the chaperonin complex, mitochondrial changes or the GPCR signalling pathway, many of our targets have not been studied in the context of late-onset Alzheimer's disease. Of all of our effects, *HSPA2* gave the most consistent results in that this target was a key driver in both datasets, differentially expressed and had the largest effects on levels of amyloid-β and tau in already overexpressed model systems.

Acknowledgements

This manuscript is dedicated to the memory of our dear friends and colleagues Christopher B. Heward and Jason J. Corneveaux. Requiescat in pace.

We thank the patients and their families for their selfless donations. Many data and biomaterials were collected from several National Institute on Aging (NIA) and National Alzheimer's Coordinating Center (NACC, grant #U01 AG016976) funded sites. Amanda J. Myers, PhD (University of Miami, Department of Psychiatry) and John A. Hardy, PhD (Reta Lila Weston Institute, University College London) collected and prepared the series. Marcelle Morrison-Bogorad, PhD., Tony Phelps, PhD and Walter Kukull PhD are thanked for helping to co-ordinate this collection. The directors, pathologist and technicians involved 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, Director, John F. Ervin, Stanford University: Dikran Horoupian, MD, Ahmad Salehi, MD, PhD, Massachusetts Alzheimer's Disease Research Center (P50 AG005134): E. Tessa Hedley-Whyte, MD, MP Frosch, MD, Karlotta Fitch, University of Michigan (NIH grant P50-AG053760): Dr. Roger Albin, Lisa Bain, Eszter Gombosi, University of Kentucky (NIH #AG05144): William Markesbery, MD, Sonya Anderson, Mayo Clinic, Jacksonville: Dennis W. Dickson, MD, Natalie Thomas, Washington University, St Louis Alzheimer's Disease Research Center (NIH #P50AG05681): Dan McKeel, MD, John C. Morris, MD, Eugene Johnson, Jr., PhD, Virginia Buckles, PhD, Deborah Carter, University of Washington, Seattle (NIH #P50 AG05136):Thomas Montine, MD, PhD, Aimee Schantz, MEd., Boston University Alzheimer's Disease Research Center (NIH grant P30-AG13846): Ann C. McKee, Carol Kubilus Banner Sun Health Research Institute Brain Donation Program of Sun City, Arizona (NIA #P30 AG19610; Arizona Alzheimer's Disease Core Center, Arizona Department of Health Services, contract 211002, Arizona Alzheimer's Research Center; Arizona Biomedical Research Commission, contracts 4001, 0011, 05_901 and 1001 to the Arizona Parkinson's Disease Consortium; Michael J. Fox Foundation for Parkinson's Research): Thomas G. Beach, MD, PhD, Lucia I. Sue, Geidy E. Serrano; 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, Rush University Medical Center, Rush Alzheimer's Disease Center (NIH #AG10161): David A. Bennett, M.D. Julie A. Schneider, MD, MS, Karen Skish, MS, PA (ASCP)MT, Wayne T Longman, University of Miami Brain Endowment Bank (supported in part by HHSN-271-2013-00030C and the McGowan Endowment): Deborah C. Mash, MD, Margaret J Basile, Mitsuko Tanaka, Oregon Health and Science University: Randy Wotljer, PhD.

Additional tissues include samples from the following sites: Newcastle Brain Tissue Resource (funding via the Medical Research Council, local NHS trusts and Newcastle University): C.M. Morris, MD, Ian G McKeith, Robert H Perry MRC London Brain Bank for Neurodegenerative Diseases (funding via the Medical Research Council): Simon Lovestone, Md PhD, Safa Al-Sarraj. MD, Claire Troakes, The Netherlands Brain Bank (funding via numerous sources including Stichting MS Research, Brain Net Europe, Hersenstichting Nederland Breinbrekend Werk, International Parkinson Fonds, Internationale Stiching Alzheimer Onderzoek): Inge Huitinga, MD, Marleen Rademaker, Michiel Kooreman, Institut de Neuropatologia, Servei Anatomia Patologica, Universitat de Barcelona: Isidre Ferrer, MD, PhD, Susana Casas Boluda.

Funding

This project was supported by grants from the National Center for Research Resources (P41RR018522) and the National Institute of General Medical Sciences (P41GM103493) from the National Institutes of Health as well as NIH EUREKA grant AG034504 to A.J.M. and NIA AG041232 to A.J.M and M.H. Tissue resources from the University of Michigan were funded by NIH grant P50-AG08671. Portions of this work were supported by NIH P41GM103493 (to R.D.S).The proteomics work was performed in the Environmental Molecular Science Laboratory, a U.S. Department of Energy (DOE) national scientific user facility at Pacific Northwest National Laboratory (PNNL) in Richland, WA. Battelle operates PNNL for the DOE under contract DE-AC05-76RLO 1830.

Competing interests

T.G.B. receives consulting fees from Prothena, Genetech and GlaxoSmithKline; R.L.A. sits on the following data safety and monitoring boards: IONIS-HTTRX (IONIS), LEGATO-HD (ICON/Teva), Biogen 251PP301/ BMS CN002012 (KlICON/BMS-Biogen); J.A.H. is on the Scientific Advisory Board for Ceracuity.

Supplementary material

Supplementary material is available at *Brain* online.

References

Anderson CA, Pettersson FH, Clarke GM, Cardon LR, Morris AP, Zondervan KT. Data quality control in genetic case-control association studies. Nat Protoc 2010; 5: 1564–73.

Andreeva AV, Kutuzov MA, Voyno-Yasenetskaya TA. G alpha12 is targeted to the mitochondria and affects mitochondrial morphology and motility. FASEB J 2008; 22: 2821–31.

Azorsa DO, Robeson RH, Frost D, Meec hoovet B, Brautigam GR, Dickey C, et al. High-content siRNA screening of the kinome identifies kinases involved in Alzheimer's disease-related tau hyperphosphorylation. BMC Genomics 2010; 11: 25.

Braak H, Braak E. Staging of Alzheimer's disease-related neurofibrillary changes. Neurobiol Aging 1995; 16: 271–8; discussion 8–84.

Chowdari KV, Mirnics K, Semwal P, Wood J, Lawrence E, Bhatia T, et al. Association and linkage analyses of RGS4 polymorphisms in schizophrenia. Hum Mol Genet 2002; 11: 1373–80.

Citron M, Oltersdorf T, Haass C, McConlogue L, Hung AY, Seubert P, et al. Mutation of the beta-amyloid precursor protein in familial Alzheimer's disease increases beta-protein production. Nature 1992; 360: 672–4.

Corneveaux JJ, Myers AJ, Allen AN, Pruzin JJ, Ramirez M, Engel A, et al. Association of CR1, CLU and PICALM with Alzheimer's disease in a cohort of clinically characterized and neuropathologically verified individuals. Hum Mol Genet 2010; 19: 3295–301.

Cruts M, Backhovens H, Theuns J, Clark RF, Le Paslier D, Weissenbach J, et al. Genetic and physical characterization of the early-onset Alzheimer's disease AD3 locus on chromosome 14q24.3. Hum Mol Genet 1995; 4: 1355–64.

Daugaard M, Rohde M, Jaattela M. The heat shock protein 70 family: highly homologous proteins with overlapping and distinct functions. FEBS Lett 2007; 581: 3702–10.

Du P, Kibbe WA, Lin SM. lumi: a pipeline for processing Illumina microarray. Bioinformatics 2008; 24: 1547–8.

Fei SS, Wilmarth PA, Hitzemann RJ, McWeeney SK, Belknap JK, David LL. Protein database and quantitative analysis considerations when integrating genetics and proteomics to compare mouse strains. J Proteome Res 2011; 10: 2905–12.

Foss EJ, Radulovic D, Shaffer SA, Ruderfer DM, Bedalov A, Goodlett DR, et al. Genetic basis of proteome variation in yeast. Nat Genet 2007; 39: 1369–75.

Fu W, Shi D, Westaway D, Jhamandas JH. Bioenergetic mechanisms in astrocytes may contribute to amyloid plaque deposition and toxicity. J Biol Chem 2015; 290: 12504–13.

Gao Y, Thomas JO, Chow RL, Lee GH, Cowan NJ. A cytoplasmic chaperonin that catalyzes beta-actin folding. Cell 1992; 69: 1043–50.

Garge N, Pan H, Rowland MD, Cargile BJ, Zhang X, Cooley PC, et al. Identification of quantitative trait loci underlying proteome variation in human lymphoblastoid cells. Mol Cell Proteomics 2010; 9: 1383–99.

Ghazalpour A, Bennett B, Petyuk VA, Orozco L, Hagopian R, Mungrue IN, et al. Comparative analysis of proteome and transcriptome variation in mouse. PLoS Genet 2011; 7: e1001393.

Grillet N, Pattyn A, Contet C, Kieffer BL, Goridis C, Brunet JF. Generation and characterization of Rgs4 mutant mice. Mol Cell Biol 2005; 25: 4221–8.

Guo N, Hwang DR, Lo ES, Huang YY, Laruelle M, Abi-Dargham A. Dopamine depletion and *in vivo* binding of PET D1 receptor radioligands: implications for imaging studies in schizophrenia. Neuropsychopharmacology 2003; 28: 1703–11.

Hamos JE, Oblas B, Pulaski-Salo D, Welch WJ, Bole DG, Drachman DA. Expression of heat shock proteins in Alzheimer's disease. Neurology 1991; 41: 345–50.

Hardy J, Selkoe DJ. The amyloid hypothesis of Alzheimer's disease: progress and problems on the road to therapeutics. Science 2002; 297: 353–6.

Hart MJ, Jiang X, Kozasa T, Roscoe W, Singer WD, Gilman AG, et al. Direct stimulation of the guanine nucleotide exchange activity of p115 RhoGEF by Galpha13. Science 1998; 280: 2112–14.

Hong MG, Myers AJ, Magnusson PK, Prince JA. Transcriptome-wide assessment of human brain and lymphocyte senescence. PLoS One 2008; 3: e3024.

Hoozemans JJ, van Haastert ES, Nijholt DA, Rozemuller AJ, Eikelenboom P, Scheper W. The unfolded protein response is activated in pretangle neurons in Alzheimer's disease hippocampus. Am J Pathol 2009; 174: 1241–51.

Hoshi M, Takashima A, Noguchi K, Murayama M, Sato M, Kondo S, et al. Regulation of mitochondrial pyruvate dehydrogenase activity by tau protein kinase I/glycogen synthase kinase 3beta in brain. Proc Natl Acad Sci USA 1996; 93: 2719–23.

Iwasaki S, Kobayashi M, Yoda M, Sakaguchi Y, Katsuma S, Suzuki T, et al. Hsc70/Hsp90 chaperone machinery mediates ATP-dependent RISC loading of small RNA duplexes. Mol Cell 2010; 39: 292–9.

Jackson M, Song W, Liu MY, Jin L, Dykes-Hoberg M, Lin CI, et al. Modulation of the neuronal glutamate transporter EAAT4 by two interacting proteins. Nature 2001; 410: 89–93.

Jaitly N, Mayampurath A, Littlefield K, Adkins JN, Anderson GA, Smith RD. Decon2LS: an open-source software package for automated processing and visualization of high resolution mass spectrometry data. BMC Bioinformatics 2009; 10: 87.

Johnson BD, Schumacher RJ, Ross ED, Toft DO. Hop modulates Hsp70/Hsp90 interactions in protein folding. J Biol Chem 1998; 273: 3679–86.

Katayama T, Imaizumi K, Sato N, Miyoshi K, Kudo T, Hitomi J, et al. Presenilin-1 mutations downregulate the signalling pathway of the unfolded-protein response. Nat Cell Biol 1999; 1: 479–85.

Korn JM, Kuruvilla FG, McCarroll SA, Wysoker A, Nemesh J, Cawley S, et al. Integrated genotype calling and association analysis of SNPs, common copy number polymorphisms and rare CNVs. Nat Genet 2008; 40: 1253–60.

Kubota H, Hynes G, Carne A, Ashworth A, Willison K. Identification of six Tcp-1-related genes encoding divergent subunits of the TCP-1-containing chaperonin. Curr Biol 1994; 4: 89–99.

Lambert JC, Ibrahim-Verbaas CA, Harold D, Naj AC, Sims R, Bellenguez C, et al. Meta-analysis of 74 046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. Nat Genet 2013; 45: 1452–8.

Leak RK. Heat shock proteins in neurodegenerative disorders and aging. J Cell Commun Signal 2014; 8: 293–310.

Lebel M, Patenaude C, Allyson J, Massicotte G, Cyr M. Dopamine D1 receptor activation induces tau phosphorylation via cdk5 and GSK3 signaling pathways. Neuropharmacology 2009; 57: 392–402.

Lewis VA, Hynes GM, Zheng D, Saibil H, Willison K. T-complex polypeptide-1 is a subunit of a heteromeric particle in the eukaryotic cytosol. Nature 1992; 358: 249–52.

Meigs TE, Fields TA, McKee DD, Casey PJ. Interaction of Galpha 12 and Galpha 13 with the cytoplasmic domain of cadherin provides a mechanism for beta -catenin release. Proc Natl Acad Sci USA 2001; 98: 519–24.

Mier D, Kirsch P, Meyer-Lindenberg A. Neural substrates of pleiotropic action of genetic variation in COMT: a meta-analysis. Mol Psychiatry 2010; 15: 918–27.

Miller AH, Raison CL. The role of inflammation in depression: from evolutionary imperative to modern treatment target. Nat Rev Immunol 2016; 16: 22–34.

Miller K, Driscoll D, Smith LM, Ramaswamy S. The role of inflammation in late-life post-traumatic stress disorder. Mil Med 2017; 182: e1815–18.

Milton NG. Amyloid-beta binds catalase with high affinity and inhibits hydrogen peroxide breakdown. Biochem J 1999; 344 (Pt 2): 293–6.

Mirra SS, Heyman A, McKeel D, Sumi SM, Crain BJ, Brownlee LM, et al. The Consortium to Establish a Registry for Alzheimer's

Disease (CERAD). Part II. Standardization of the neuropathologic assessment of Alzheimer's disease. Neurology 1991; 41: 479–86.
Monroe ME, Tolic N, Jaitly N, Shaw JL, Adkins JN, Smith RD. VIPER: an advanced software package to support high-throughput LC-MS peptide identification. Bioinformatics 2007; 23: 2021–3.
Moreira PI, Carvalho C, Zhu X, Smith MA, Perry G. Mitochondrial dysfunction is a trigger of Alzheimer's disease pathophysiology. Biochim Biophys Acta 2010; 1802: 2–10.
Muchowski PJ, Wacker JL. Modulation of neurodegeneration by molecular chaperones. Nat Rev Neurosci 2005; 6: 11–22.
Myers A, Wavrant De-Vrieze F, Holmans P, Hamshere M, Crook R, Compton D, et al. Full genome screen for Alzheimer disease: stage II analysis. Am J Med Genet 2002; 114: 235–44.
Myers AJ. The age of the "ome": genome, transcriptome and proteome data set collection and analysis. Brain Res Bull 2012; 88: 294–301.
Myers AJ. AD gene 3-D: moving past single layer genetic information to map novel loci involved in Alzheimer's disease. J Alzheimers Dis 2013; 33 (Suppl 1): S15–22.
Myers AJ. The genetics of gene expression: multiple layers and multiple players. In: Coppola G, editor. The OMICs: applications in neuroscience. New York, NY: Oxford University Press; 2014. p. 132–52.
Myers AJ, Gibbs JR, Webster JA, Rohrer K, Zhao A, Marlowe L, et al. A survey of genetic human cortical gene expression. Nat Genet 2007a; 39: 1494–9.
Myers AJ, Kaleem M, Marlowe L, Pittman AM, Lees AJ, Fung HC, et al. The H1c haplotype at the MAPT locus is associated with Alzheimer's disease. Hum Mol Genet 2005; 14: 2399–404.
Myers AJ, Pittman AM, Zhao AS, Rohrer K, Kaleem M, Marlowe L, et al. The MAPT H1c risk haplotype is associated with increased expression of tau and especially of 4 repeat containing transcripts. Neurobiol Dis 2007b; 25: 561–70.
Neef DW, Jaeger AM, Gomez-Pastor R, Willmund F, Frydman J, Thiele DJ. A direct regulatory interaction between chaperonin TRiC and stress-responsive transcription factor HSF1. Cell Rep 2014; 9: 955–66.
Ngo ST, Steyn FJ. The interplay between metabolic homeostasis and neurodegeneration: insights into the neurometabolic nature of amyotrophic lateral sclerosis. Cell Regen 2015; 4: 5.
Nunomura A, Perry G, Aliev G, Hirai K, Takeda A, Balraj EK, et al. Oxidative damage is the earliest event in Alzheimer disease. J Neuropathol Exp Neurol 2001; 60: 759–67.
Piehowski PD, Petyuk VA, Orton DJ, Xie F, Moore RJ, Ramirez-Restrepo M, et al. Sources of technical variability in quantitative LC-MS proteomics: human brain tissue sample analysis. J Proteome Res 2013; 12: 2128–37.
Portelli MA, Siedlinski M, Stewart CE, Postma DS, Nieuwenhuis MA, Vonk JM, et al. Genome-wide protein QTL mapping identifies human plasma kallikrein as a post-translational regulator of serum uPAR levels. FASEB J 2014; 28: 923–34.
Rajapandi T, Greene LE, Eisenberg E. The molecular chaperones Hsp90 and Hsc70 are both necessary and sufficient to activate hormone binding by glucocorticoid receptor. J Biol Chem 2000; 275: 22597–604.
Riobo NA, Manning DR. Receptors coupled to heterotrimeric G proteins of the G12 family. Trends Pharmacol Sci 2005; 26: 146–54.
Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 2015; 43: e47.
Roadmap Epigenomics Consortium; Kundaje A, Meuleman W, Ernst J, Bilenky M, Yen A, et al. Integrative analysis of 111 reference human epigenomes. Nature 2015; 518: 317–30.
Serrano-Pozo A, Frosch MP, Masliah E, Hyman BT. Neuropathological alterations in Alzheimer disease. Cold Spring Harb Perspect Med 2011; 1: a006189.
Shabalin AA. Matrix eQTL: ultra fast eQTL analysis via large matrix operations. Bioinformatics 2012; 28: 1353–8.
Sherrington R, Rogaev EI, Liang Y, Rogaeva EA, Levesque G, Ikeda M, et al. Cloning of a gene bearing missense mutations in early-onset familial Alzheimer's disease. Nature 1995; 375: 754–60.
Slifstein M, Kolachana B, Simpson EH, Tabares P, Cheng B, Duvall M, et al. COMT genotype predicts cortical-limbic D1 receptor availability measured with [11C]NNC112 and PET. Mol Psychiatry 2008; 13: 821–7.
Stark AL, Hause RJ Jr, Gorsic LK, Antao NN, Wong SS, Chung SH, et al. Protein quantitative trait loci identify novel candidates modulating cellular response to chemotherapy. PLoS Genet 2014; 10: e1004192.
Stetler RA, Gan Y, Zhang W, Liou AK, Gao Y, Cao G, et al. Heat shock proteins: cellular and molecular mechanisms in the central nervous system. Prog Neurobiol 2010; 92: 184–211.
Storga D, Vrecko K, Birkmayer JG, Reibnegger G. Monoaminergic neurotransmitters, their precursors and metabolites in brains of Alzheimer patients. Neurosci Lett 1996; 203: 29–32.
Strathmann MP, Simon MI. G alpha 12 and G alpha 13 subunits define a fourth class of G protein alpha subunits. Proc Natl Acad Sci USA 1991; 88: 5582–6.
Thathiah A, Spittaels K, Hoffmann M, Staes M, Cohen A, Horre K, et al. The orphan G protein-coupled receptor 3 modulates amyloid-beta peptide generation in neurons. Science 2009; 323: 946–51.
Vaiskunaite R, Kozasa T, Voyno-Yasenetskaya TA. Interaction between the G alpha subunit of heterotrimeric G protein and Hsp90 is required for G alpha signaling. J Biol Chem 2001; 276: 46088–93.
Webster JA, Gibbs JR, Clarke J, Ray M, Zhang W, Holmans P, et al. Genetic control of human brain transcript expression in Alzheimer disease. Am J Hum Genet 2009; 84: 445–58.
Wu G, Nie L, Zhang W. Integrative analyses of posttranscriptional regulation in the yeast Saccharomyces cerevisiae using transcriptomic and proteomic data. Curr Microbiol 2008; 57: 18–22.
Yaffe MB, Farr GW, Miklos D, Horwich AL, Sternlicht ML, Sternlicht H. TCP1 complex is a molecular chaperone in tubulin biogenesis. Nature 1992; 358: 245–8.
Zhang B, Gaiteri C, Bodea LG, Wang Z, McElwee J, Podtelezhnikov AA, et al. Integrated systems approach identifies genetic nodes and networks in late-onset Alzheimer's disease. Cell 2013; 153: 707–20.
Zimmer JS, Monroe ME, Qian WJ, Smith RD. Advances in proteomics data analysis and display using an accurate mass and time tag approach. Mass Spectrom Rev 2006; 25: 450–82.