

SUPPLEMENTAL INFORMATION

SUPPLEMENTAL DATA

Table S1 The Characteristics of Brain Modules and Corresponding Gene Expression Traits. Detailed information regarding module- and gene-centered findings of the present study (modules with 50 or more unique gene symbols), including module MDC, enrichment of functional categories, intra-module connectivity, expression values for the top hubs in each module, external signatures (CNS cell types, the postsynaptic density proteome, the MEMN), gene expression correlations to Braak stage and frontal atrophy, results from the RNA sequencing of microglia cells over-expressing intact or truncated TYROBP, and lists of all eSNPs and corresponding transcripts detected in the present study at FDR of 10%.

EXTENDED EXPERIMENTAL PROCEDURES

Study Populations and Clinical Data

1647 frozen tissue samples from cerebellum (CB), visual cortex (VC) BA17 and dorsolateral prefrontal cortex (PFC) BA9 were provided by the Harvard Brain Tissue Resource Center (HBTRC) at McLean Hospital (Belmont, MA). These regions were selected because PFC is commonly affected in LOAD while VC and CB remain less affected throughout most of the disease progression (Braak and Braak, 1991). All autopsied brains, were collected from subjects with LOAD diagnosis or from normal non-demented subjects, for whom both the donor and the next of kin had completed the HBTRC informed consent (<http://www.brainbank>). Tissue collection and the research were conducted according to the HBTRC guidelines (<http://www.brainbank>). The HBTRC samples were primarily of Caucasian ancestry, as only eight non-Caucasian outliers were identified and therefore excluded for further analysis. Post-mortem interval (PMI) was 17.8 ± 8.3 hours, sample pH was 6.4 ± 0.3 and RNA integrity number (RIN) was 6.8 ± 0.8 for the average sample in the overall cohort. These were composed of 376 LOAD patients and 173 non-demented subjects. Braak stage, general and regional atrophy, gray and white matter atrophy and ventricular enlargement were assessed and catalogued by pathologists at McLean Hospital (Belmont, MA). The clustering of LOAD related pathology traits and age is shown in **Figure 2**. It is of note that age formed a separate cluster from most of the LOAD neuropathology, however hippocampal atrophy was positively correlated with age ($r = 0.28$) while Braak staging of LOAD pathology was negatively correlated with age ($r = -0.15$). The negative correlation between age and Braak stage in the HBTRC sample most likely reflects the higher degree of Tau-related neurofibrillary tangle load associated with the earlier onset and more severe LOAD cases that skew the correlation structure, as has been previously noted in combined familial and sporadic LOAD cases (Nochlin et al., 1993). Finally, for the immune/microglia comparative study we used a brain tissue sampling (PFC, CB and VC) of 194 Huntington disease (HD) patients from the HBTRC, where each of the 582 HD patient sample went through identical procedure as described for the LOAD and non-demented brain samples. In addition, the severity of pathology in the HD autopsy brains was determined from the Vonsattel grading system (Vonsattel et al., 1985) and the *HTT* CAG repeat allele size was determined using a modification of

previously described polymerase chain reaction amplification assay (Warner et al., 1993).

Brain samples for testing the replication of the *TYROBP* differential expression were obtained from 21 National Alzheimer's Coordinating Center (NACC) brain banks and from the Miami Brain Bank as previously described (Myers et al., 2007; Webster et al., 2009). The Rush Alzheimer's Disease Center series consists of two cohorts: the Religious Orders Study (ROS) and the Memory and Aging Project (MAP) (Bennett et al., 2006a; Bennett et al., 2006b; Bennett et al., 2005). cRNA was hybridized as before to Illumina HT-12 Expression Bead Chip (48,803 transcripts) via standard protocols using an Illumina Bead Station 500GX (Webster et al 2009). Disease status included 377 with LOAD, 119 classified with mild cognitive impairment (MCI) and 359 healthy non-demented controls. Brain regions sampled in these patients were 726 prefrontal and 129 temporal cortex samples.

Array Design, Gene Expression Normalization, Covariate Analysis and Genotyping

RNA preparation and array hybridizations applied custom microarrays manufactured by Agilent Technologies consisting of 4,720 control probes and 39,579 probes targeting transcripts representing 25,242 known and 14,337 predicted genes. One μg of total RNA from each of the 2229 (LOAD, non-demented, HD) brain tissues was reverse transcribed and labeled with either Cy3 or Cy5 fluorochrome. Purified Cy3 or Cy5 complementary RNA was hybridized to at least two single microarrays with fluor reversal for 24 hours in a hybridization chamber, washed, and scanned using a laser confocal scanner. Arrays were quantified on the basis of spot intensity relative to background, adjusted for experimental variation between arrays using average intensity over multiple channels, and fitted to an error model to determine significance (type I error) as previously described (Emilsson et al., 2008). This microarray dataset is MIAME compliant and the raw gene expression data for all hybridizations together with information related to demographics, disease state (LOAD or non-demented) and technical covariates (pH, RIN, PMI, batch, preservation method) are available via the GEO database (GEO accession number GSE44772 including the regional-specific subseries GSE44768, GSE44770, GSE44771).

Gene expression was reported as the mean-log ratio of individual microarray intensities relative to average intensities of all samples. Gene expression data were generated using Rosetta Resolver gene expression analysis software (version 7.0, Rosetta Biosoftware) and MATLAB (The MathWorks). To remove bias in expression profiles related to potential latent variables unrelated to underlying biological processes, we implemented a normalization method based on control probes present on the microarrays. We separated the control probes into two classes: (1) specialty probes, such as spike-in probes or other probes designed to monitor the quality of the microarrays; and (2) border probes used to describe the geometry of the microarray. We then identified the Principal Components (PC) explaining the variability of each control probe class and then identified the same components from a randomly permuted data set. We performed 10,000 permutations for each set of control probes and selected Principal Components with P-values defined as $\{\text{number of } [\text{var}(\text{random-PC}_i) > \text{var}(\text{PC}_i)]\} / (\text{number of permutations}) < 10^{-4}$. The expression data in each tissue for each probe are thus the residuals from a linear model fitting incorporating the significant PCs.

Next we analyzed the contribution of each of the experimental covariate's age, gender, sex and technical covariates. The Kolmogorov-Smirnov (KS) test was used to capture the difference of the correlation P-value distributions for each covariate per brain region and condition. **Figure S1** shows the correlation P-values between a specific covariate and all the probes based on the expression data before (red curve, raw) and after (black curve, corrected) the adjustment in LOAD or non-demented brains. Particular attention was paid to established covariates of mRNA quality that have been demonstrated in brain tissues, which included: (1) Extreme confounding of demographic variables in disease vs. controls; it was not possible to match samples based on demographic variables including age, sex, and ethnicity as many of these factors differ in their distributions substantially across the disease groups. Ethnicity was determined using STRUCTURE resulting in the exclusion of eight outliers (non-Caucasians), which were removed from all studies. (2) Pre-/post-mortem quantitative and quality factors including PMI, RIN, tissue pH status, manner of death and/or agonal state. Seven samples with RIN<5 were removed from the analysis. (3) Batch effects; day-to-day variability in the amplification process can influence the observed gene expression, even when profiled against a common-pool. For this particular data set, the batch effects are further exacerbated by the two additional facts: (i) the demographic compositions of the two batches differ, and (ii) for technical reasons, as it is preferable to have your pool sample be a representation of the greater population (mean-log ratio of individual microarray intensities was relative to average intensities of all samples in the final set). In summary, the distribution of P-values obtained from modeling the covariates listed above strongly suggest that adjusting for them would improve the downstream analysis steps. Therefore we chose a robust linear regression model for covariate corrections as $\text{rlm}(\text{expression} \sim \text{RIN} + \text{pH} + \text{PMI} + \text{age} + \text{batch} + \text{preservation} + \text{gender})$ and then the residuals were used for further testing including the genetics of gene expression, network construction, differential gene expression and gene expression clinical trait correlations carried out in the present study.

DNA isolation was performed as previously described (Emilsson et al., 2008). Each sample was genotyped on two different platforms, the IlluminaHumanHap650Y array and a custom Perlegen 300K array (a focused panel for detection of singleton SNPs). Counting only the union of markers from both genotyping platforms (114,925 SNPs were in the intersection), a total of 838,958 unique SNPs were used for the analysis. Finally, *APOE* genotypes were obtained through restriction fragment length polymorphism (RFLP) analysis (Lai et al., 1998).

Constructing the Multi-tissue Co-expression Networks

We constructed a multi-tissue co-expression network that simultaneously captured the intra- and inter-tissue gene-gene interactions between the LOAD and normal states. These networks can be characterized only when multiple tissues are monitored in a population of individuals donating all tissues (PFC, CB and VC). For construction of the multi-tissue brain network, we used the top one-third (T1) or 13,193 of the most variable genes (inter-individual variability) in a given tissue and physiological state (LOAD or non-demented). The most variable gene expression traits in the cohort were defined based on a previously described and validated error model testing the intensity between the experimental and reference channel, to obtain P-values for each individual expression trait and then compute the standard deviation (SD) of $-\log_{10}(\text{P-value})$ for each trait over all samples profiled for a given tissue. Then we rank ordered all of the genes profiled in each tissue based on this

SD value (rank ordered in descending order). Genes that fall at the top of this rank ordered list can be considered as the most differentially expressed or variable genes in the study. We have previously shown that this type of ranking approach well captures the most active genes in a given set of tissue samples (Emilsson et al., 2008; He et al., 2003).

Specifically, for each of the three T1 data sets we first assigned each gene a unique identifier by combining probe ID and tissue name, then aligned the samples, and finally merged the three sets into a multi-tissue expression data set which included 39,579 genes and 129 AD samples. The combined 39,579 genes were then mapped to the three tissues in normal non-demented data set in a supervised manner, which led to another data set of 39,579 genes and 99 normal samples. The LOAD and normal multi-tissue data sets were independently processed through the weighted gene co-expression network analysis (Zhang and Horvath, 2005). The weighted network analysis begins with a matrix of the Pearson correlations between all gene pairs, then converts the correlation matrix into an adjacency matrix using a power function $f(x)=x^\beta$. The parameter β of the power function is determined in such a way that the resulting adjacency matrix, i.e., the weighted co-expression network, is approximately scale-free. To measure how well a network satisfies a scale-free topology, we use the fitting index proposed by Zhang et al. (Zhang and Horvath, 2005), i.e., the model fitting index R^2 of the linear model that regresses $\log(p(k))$ on $\log(k)$ where k is connectivity and $p(k)$ is the frequency distribution of connectivity. The fitting index of a perfect scale-free network is 1. The connectivity between genes or k_{ij} is a transformed correlation between the expression profiles of two genes, $|r(i,j)|^\beta$, with r as the Pearson correlation coefficient. The parameter β (>0) of the power function is determined in such a way that the global probability distribution of the resulted connectivity values for all the gene pairs is scale free. Thus, k_{ij} is a continuous value ranging from 0 to 1.

To explore the modular structures of the co-expression network, the adjacency matrix is further transformed into a topological overlap matrix (TOM) (Ravasz et al., 2002). As the topological overlap between two genes reflects not only their direct interaction but also their indirect interactions through all the other genes in the network, previous studies (Ravasz et al., 2002; Zhang and Horvath, 2005) have shown that topological overlap leads to more cohesive and biologically more meaningful modules. To identify modules of highly co-regulated genes, we used average linkage hierarchical clustering to group genes based on the topological overlap of their connectivity, followed by a dynamic cut-tree algorithm to dynamically cut clustering dendrogram branches into gene modules (Langfelder et al., 2008). To distinguish between modules, each module was assigned a unique color identifier, with the remaining, less well connected genes colored grey. To compare and contrast two multi-tissue networks, we combined their TOM heat-maps into a single large network. In the combined heat-map, the upper panel shows the hierarchical clustering on the TOM of the LOAD network while the color bar below represents the gene modules. Similarly, the lower panel represents the TOM from normal multi-tissue network. The color intensity in the map represents the interaction strength between genes. This connectivity map highlights how genes in the multi-tissue transcriptional networks fall into distinct network modules, where genes within a given module are more highly interconnected with each other (blocks along the diagonal of the matrix) than with genes in other modules.

Defining the Modular Differential Connectivity

We developed two measures to quantify the difference between the connectivity among a set of genes (or module, denoted Ω) in LOAD vs. normal non-demented networks. Given a set of N genes and two networks, x and y , the first MDC measurement is the ratio of the average connectivity among the N genes in the network x to that among the same gene set in network y , specified by the formula below:

$$\delta_{\Omega}(x, y) = \frac{\sum_{i=1}^{N-1} \sum_{j=i+1}^N k_{ij}^x}{\sum_{i=1}^{N-1} \sum_{j=i+1}^N k_{ij}^y}, \quad (1)$$

where, k_{ij} is the connectivity between two genes i and j in a given network.

Obviously, $\delta_{\Omega}(x, y)$ and $\delta_{\Omega}(y, x)$ are reciprocal, i.e. $\delta_{\Omega}(x, y) = \frac{1}{\delta_{\Omega}(y, x)}$.

The significance or false discovery rate (FDR) of the statistic MDC can be accessed by permuting the data underlying the two networks. We differentiate two scenarios, gain of connectivity ($\delta_{\Omega}(x, y) > 1$) and loss of connectivity ($\delta_{\Omega}(x, y) < 1$). Given M permutations, FDR of MDC is computed as follows:

$$FDR(\delta_{\Omega}(x, y) > 1) = \frac{1}{M} \sum_{p=1}^M \delta_{\Omega}(x, y) > \delta_{\Omega}(x^p, y^p) \sum_p \delta_{\Omega}(x, y) > \delta_{\Omega}(x^p, y^p),$$

or

$$FDR(\delta_{\Omega}(x, y) < 1) = \frac{1}{M} \sum_{p=1}^M \delta_{\Omega}(x, y) > \delta_{\Omega}(x^p, y^p) \sum_p \delta_{\Omega}(x, y) < \delta_{\Omega}(x^p, y^p),$$

where x^p and y^p are the networks derived from the permuted data. To rigorously access significance of MDC, we estimate two types of FDR estimates, one based on shuffled samples, i.e. networks with non-random nodes but random connections, and the other based on shuffled gene labels, i.e. networks with random nodes but non-random connections, and then we select the larger value as the final FDR estimate.

Differential Gene Expression and Module Relevance to LOAD Pathology

To identify differential expression of individual genes between LOAD and non-demented controls, Agilent gene expression data generated for each brain region was first adjusted using robust linear regressions for age, gender, PMI, tissue preservation method, tissue pH, RIN, and sample preparation batch. The residuals following regression were analyzed by ANOVA to identify reporters which showed differential expression between LOAD and control samples in each tissue, and a Monte Carlo analysis was performed by permuting sample order to estimate the false-discovery rate (FDR). Differentially expressed reporters were defined as those with a Bonferroni corrected P-value < 0.05 ; this corresponds to a nominal $P < 2.46 \times 10^{-7}$ and is well below the 0% FDR threshold for all three brain regions tested.

To examine how each gene module was related to LOAD neuropathology traits, we first performed principal component analysis (PCA) for each module and then computed module-trait relevance using two complementary approaches: (i) the correlation between the first principal component (Module Eigengene) and each trait and (ii) the correlation (the square root of R-square) between the top principal components and each trait through multivariate regression model. The significance (P-value) and FDR of each correlation was also calculated. FDR was estimated through random permutation of sample names of the trait data. A module is

associated with a trait if both correlation P-value and FDR are below 0.05. The total number of traits associated with a module is used to quantify the association of a module with LOAD.

Detection of eSNPs in the Different Brain Regions

Each of the 838,958 SNPs was tested for association to all 39,579 expression traits using Kruskal Wallis test based on the residuals, left after correction for covariates using robust linear regression as described in the section '*Array Design, Gene Expression Normalization, Covariate Analysis and Genotyping*'. Expression quantitative trait loci (e)QTLs for gene expression traits were determined by identifying the SNP most strongly associated with each expression trait profiled on the array over all the 838,958 genotyped SNPs. *Cis* analysis was limited to SNPs located within 1 Mb of either side of the transcription start or end within the gene body, while *trans* effects were defined as the associated SNPs located on a different chromosome to the physical location of the corresponding probe (Emilsson et al., 2008). The association P-value was adjusted to control for testing of multiple SNPs and expression traits using two different methods: (1) a highly conservative Bonferroni correction method to constrain the study-wide significance level, and (2) an empirical FDR method (Storey and Tibshirani, 2003) that constrains the overall rate of false positive events. For *cis* eQTL, to achieve a study-wide significance level of 0.05, the Bonferroni adjusted P-value threshold was computed as $0.05/(39,579 \times N_i)$, where N_i denotes the number of SNPs tested for trait i within the two Mb window, over all 39,579 expression traits tested. The nominal P-value to meet this significance threshold is 3.0×10^{-8} . The Bonferroni adjustment method can be conservative when there is dependence among the expression traits and among the SNP genotypes. Given that strong correlation structures exist among expression traits and among SNP genotypes in a given linkage disequilibrium (LD) block, the Bonferroni adjustment may be overly conservative. Therefore, we used an empirical FDR method based on permutations that accounts for the correlation structures among the expression traits and among the SNP genotypes. We constrained the empirically determined FDR to be less than 10%. Here, the FDR was estimated as the ratio of the average number of eQTLs found in datasets with randomized sample labels to the number of eQTLs identified in the original data set. Since the number of tests was large, we found the empirical null distribution was very stable and ten permutation runs were sufficient for convergence to estimate FDR. FDR computation was performed separately for *cis* and *trans* associations resulting in nominal P-value cutoffs of 5.0×10^{-5} and 1.0×10^{-8} for *cis* and *trans* eQTLs, respectively. The significance of the *trans* eQTL was also assessed by the Bonferroni method and by constraining the empirically determined FDR to be less than 10%. In the case of *trans* eQTL, all 838,958 SNPs were tested for association to each of the 39,280 expression traits. Therefore, the Bonferroni adjusted P-value threshold was computed as $0.05/(838,958 \times 39,579) = 1.5 \times 10^{-12}$. **Table S1** lists all *cis* and *trans* acting eSNPs detected in each brain region and condition at FDR of 10% and provides nominal P-values as well.

For the present work we pay particular attention to *cis* acting eSNPs for the following reasons: (1) for construction of causal probabilistic Bayesian networks as described in the section '*Reconstruction of the Bayesian Networks and Identification of Key Causal Regulators*' and (2) for studying the enrichment of eSNPs in LOAD associated modules. Brain eSNPs were compared to the 111 gene modules identified from the multi-tissue gene co-expression network analysis and the

enrichment of *cis* eSNPs tested using the Fisher's exact test (FET) to assess the significance of the overlap between each module and *cis* eSNPs, correcting for number of modules as eSNPs are regarded as a single category.

Reconstruction of the Bayesian Networks and Identification of Key Causal Regulators

Reconstruction of the Bayesian networks: Bayesian networks are directed acyclic graphs in which the edges of the graph are defined by conditional probabilities that characterize the distribution of states of each node given the state of its parents. The network topology defines a partitioned joint probability distribution over all nodes in a network, such that the probability distribution of states of a node depends only on the states of its parent nodes: formally, a joint probability distribution $p(X)$ on a set of nodes X can be decomposed as $p(X) = \prod_i p(X^i | \text{Pa}(X^i))$, where $\text{Pa}(X^i)$ represents

the parent set of X^i . In our networks, each node represents transcription expression of a gene. These conditional probabilities reflect not only relationships between genes, but also the stochastic nature of these relationships, as well as noise in the data used to reconstruct the network.

Bayes formula allows us to determine the likelihood of a network model M given observed data D as a function of our prior belief that the model is correct and the probability of the observed data given the model: $P(M | D) \sim P(D | M) * P(M)$. The number of possible network structures grows super-exponentially with the number of nodes, so an exhaustive search of all possible structures to find the one best supported by the data is not feasible, even for a relatively small number of nodes. We employed Monte Carlo Markov Chain (MCMC) (Madigan, 1995) simulation to identify potentially thousands of different plausible networks, which are then combined to obtain a consensus network (see below). Each reconstruction begins with a null network. Small random changes are then made to the network by flipping, adding, or deleting individual edges, ultimately accepting those changes that lead to an overall improvement in the fit of the network to the data. We assess whether a change improves the network model using the Bayesian Information Criterion (BIC)(Schwarz, 1978), which avoids over-fitting by imposing a cost on the addition of new parameters. This is equivalent to imposing a lower prior probability $P(M)$ on models with larger numbers of parameters.

Although edges in Bayesian networks are directed, we can't infer causal relationships from the structure directly in general. For example, in a network with two nodes, X^1 and X^2 , the two models $X^1 \rightarrow X^2$ and $X^2 \rightarrow X^1$ have equal probability distributions as $p(X^1, X^2) = p(X^2 | X^1) p(X^1) = p(X^1 | X^2) p(X^2)$. Thus, by the data itself, we can't infer whether X^1 is causal to X^2 , or vice versa. In a more general case, a network with three nodes, X^1 , X^2 , and X^3 , there are multiple groups of structures that are mathematically equivalent. For example, the following three different models, M1: $X^1 \rightarrow X^2, X^2 \rightarrow X^3$, M2: $X^2 \rightarrow X^1, X^2 \rightarrow X^3$, and M3: $X^2 \rightarrow X^1, X^3 \rightarrow X^2$, are Markov equivalent (which means that they all encode for the same conditional independent relationships). In the above case, all three structures encode the same conditional independent relationship, $X^1 \not\perp X^3 | X^2$, X^1 and X^3 are independent conditioning on X^2 , and they are mathematically equal

$$\begin{aligned}
p(X) &= p(M1 | D) = p(X^2 | X^1)p(X^1)p(X^3 | X^2) \\
&= p(M2 | D) = p(X^1 | X^2)p(X^2)p(X^3 | X^2) \\
&= p(M3 | D) = p(X^2 | X^3)p(X^3)p(X^1 | X^2)
\end{aligned}$$

Thus, we cannot infer whether X^1 is causal to X^2 or vice versa from these types of structures. However, there is a class of structures, V-shape structure (e.g. $Mv : X^1 \rightarrow X^2, X^3 \rightarrow X^2$), which has no Markov equivalent structure. In this case, we can infer causal relationships. There are more parameters to estimate in the Mv model than M1, M2, or M3, which means a large penalty in BIC score for the Mv model. In practice, a large sample size is needed to differentiate the Mv model from the M1, M2, or M3 models.

Incorporating genetic data as a structure prior in the Bayesian network reconstruction process: In general, Bayesian networks can only be solved to Markov equivalent structures, so that it is often not possible to determine the causal direction of a link between two nodes even though Bayesian networks are directed graphs. However, the Bayesian network reconstruction algorithm can take advantage of the experimental design by incorporating genetic data to break the symmetry among nodes in the network that lead to Markov equivalent structures, thereby providing a way to infer causal directions in the network in an unambiguous fashion (Zhu et al., 2004). We modified the reconstruction algorithm to incorporate eSNP data as priors (see schematic **Figure S3**), in the following way: genes with *cis* eSNP (Schadt et al., 2008) are allowed to be parent nodes of genes without *cis* eSNPs, but genes without *cis* eSNPs are not allowed to be parents of genes with *cis* eSNPs, $p(trans \rightarrow cis) = 0$. We have shown that integrating genetic data such as *cis* acting eSNPs or eQTLs (excluding edges into certain nodes) improves the quality of the network reconstruction by simulations (Zhu et al., 2007) and by experimental validations (Zhu et al., 2004; Zhu et al., 2008). We note that in applying this particular version of the Bayesian network reconstruction algorithm (i.e. incorporating genetic information as a prior) and if genetic information is not available or is ignored, the population is simply treated as a population with random genetic perturbations.

Averaging network models: Searching optimal Bayesian network structures in a given dataset is a NP-hard problem. We employed an MCMC method to do local search of optimal structures as described above. As the method is stochastic, the resulting structure will be different for each run. In our process, 1,000 Bayesian networks were reconstructed using different random seeds to start the stochastic reconstruction process. From the resulting set of 1,000 networks generated by this process, edges that appeared in greater than 30% of the networks were used to define a consensus network. A 30% cutoff threshold for edge inclusion was based on our simulation study (Zhu et al., 2007), where a 30% cutoff yields the best tradeoff between recall rate and precision. The consensus network resulting from the averaging process may not be a Bayesian network (a directed acyclic graph). To ensure the consensus network structure is a directed acyclic graph, edges in this consensus network were removed if and only if (1) the edge was involved in a loop, and (2) the edge was the most weakly supported of all edges making up the loop.

Bayesian network for individual co-expression module: The computational complexity of our MCMC method described above is expressed as $O(N^4)$, where the number of nodes included in the network reconstruction process is N . It is practically impossible to construct a global Bayesian network including all 39,000

genes from three different brain regions. Thus, we constructed a Bayesian network for each individual co-expression module. Following the procedure described above, 1,000 Bayesian networks were reconstructed using different random seeds to start the reconstruction process. From the resulting set of 1,000 networks generated by this process, edges that appeared in greater than 30% of the networks were used to define a consensus network. Our previous simulation study shows that the 30% inclusion threshold results in a stable structure and achieves the best tradeoff between precision and recall (Zhu et al., 2007).

Identification of key causal regulators: For each Bayesian network of individual modules, we further identified the regulators by examining the number of N-hop downstream nodes (NHDN) for each gene in the directed network (Wang et al., 2012; Yang et al., 2010; Zhu et al., 2007). For a given network, let μ be the numbers of N-hop downstream nodes and d be the out degrees for all the genes. Genes with a number of NHDN greater than $\bar{\mu} + \sigma(\mu)$, were nominated as causal regulators. The regulators with degree above $\bar{d} + 2\sigma(d)$, where d denotes the number of downstream genes, become key causal regulators of a corresponding network module associated with LOAD differential connectivity. These criteria identified genes with number of downstream nodes and number of out links significantly above the corresponding average value.

Mouse Microglia Cultivation, Cell Transduction and Flow Cytometry Analysis

Embryonic stem cell-derived microglia cells (ESdM) were generated from C57BL/6 ES cells (ATCC number SCRC-1002) and were shown to be very similar to primary cultured microglia both phenotypically and functionally as previously described (Beutner et al., 2010). ESdM were cultured under serum-free conditions in medium consisting of DMEM/F12 (Gibco) supplemented with 1% N2 supplement (Invitrogen), 0.5 mM L-Glutamine (Gibco), 15 μ g/ml D-glucose (Sigma), and 100 μ g/ml penicillin/streptomycin (Gibco). After reaching a confluency of more than 80%, cells were split using cell scrapers (Sarstedt), suspended in medium, centrifuged and plated in fresh medium.

Lentiviral vectors of the third generation (pLenti6/V5, Invitrogen) were used for transduction of microglia (ESdM). The CMV promoter in the pLenti6/V5 vector was replaced by the elongation factor-1 α (EF1 α) promoter. Full length mouse *Tyrobp* (cds from ntd 115 to 393, accession NM_011662) and truncated mouse *Tyrobp* (with the ITAM motifs removed, cds from ntd 115 to 324, accession NM_011662) were tagged with the FLAG sequence at the 5', HA sequence at the 3' and the IgK leader sequence and cloned independently downstream of the EF1 α promoter. An IRES-enhanced green fluorescent protein (eGFP) gene was inserted directly downstream of the *Tyrobp* sequences. The correct nature of all cloned sequences was confirmed by automated sequencing (Seqlab, Germany). For over-expression studies, microglia (ESdM) were transduced using the corresponding IRES-eGFP constructs containing either full length mouse *Tyrobp* or truncated *Tyrobp* (ITAM motifs removed). For a control, microglia (ESdM) were transduced with the modified pLenti vector expressing only IRES-eGFP. Lentiviral transduction was performed using the ViraPower transduction kit (Invitrogen). Supernatant was removed 24 hours after infection and replaced with medium. Cells expressing eGFP were isolated by flow cytometry sorting (BD, DiVacellsorter) to obtain a pure population. We confirmed successful over-expression of the transduced genes through the flow cytometry analysis that detects the expression levels in single cells. Further, we detected

comparable protein expression levels in the flow cytometry between the full-length *Tyrobp* and truncated *Tyrobp* transduced microglial cell lines as shown in **Figure S6**.

RNA Sequencing, Sample Preparation and Data Analysis

To access the genome-wide gene expression changes in response to the perturbation of *Tyrobp*, we prepared twelve samples from mouse microglia cell lines, four from each of the three previously described groups: full length *Tyrobp* over-expressing cells, truncated *Tyrobp* over-expressing cells and GFP expressing control cells. For each sample, about one microgram of total RNA was used for the preparation of the sequence library using RNA TruSeq Kit supplied by Illumina (Cat # 1004814). Briefly, rRNA was depleted from total RNA using ribozero kit (Invitrogen) to enrich polyadenylated coding RNA and non-coding RNA. The ribominus RNA was then fragmented in the presence of divalent cations at 94°C. The fragmented RNA was converted into double stranded cDNA. After polishing the ends of the cDNA, adenine base was added at the 3' ends following which Illumina supplied specific adaptors were ligated. The adaptor ligated DNA was size selected to get an average of 200 bp insert size using AmpPure beads, and amplified by 15 cycle PCR. The PCR DNA was then purified using AmpPure beads to get the final sequence library ready for sequencing. The insert size and DNA concentration of the sequence library was determined on Agilent Bioanalyzer. Each RNA sequence library was layered on one of the eight lanes of the Illumina flow cell at appropriate concentration and bridge amplified to get around 35-40 million raw reads. The DNA reads on the flow cell were then sequenced on HiSeq 2000 using 100 bp single end recipe. The rate of sequencing was around 1.2 hours per base sequenced. Read mapping was done using the TopHat (Trapnell et al., 2009) RNA-seq aligner. We allow up to 2 mismatches in each aligned segment and keep only reads mapping less than 5 times in the genome using the default quality cutoffs. The resulting alignments to genomic intervals and predicted splice sites are stored as a binary BAM file (Li et al., 2009).

Differential gene expression with a statistical significance $P < 0.05$ (12 tests), based on a t-test, was used to identify the differentially expressed gene sets. This yielded 2638 and 3415 differentially expressed genes for the full length *Tyrobp* and the truncated *Tyrobp* expressions, respectively. The mean FDR for the two signatures was 2.4% and 1.8%, respectively. To compute the FDR, we first constructed a pair of random case and control groups, each of which was comprised of two samples from a true case group and two from a true control group, then selected the most differentially expressed genes based on the same criteria. The FDR was the percentage of the original differentially expressed genes which were found in the gene pool based on the permutation. Such a procedure was applied to all 36 possible permutations.

Acknowledgments for Additional Brain Samples

Additional tissues include samples from the following sites: 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 University of Michigan (NIH grant P50-AG08671); Dr. Roger Albin, Lisa Bain, Eszter Gombosi, 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): Joseph Rogers, PhD, Thomas G. Beach, MD, PhD, Lucia I. Sue University of Miami/NPF Brain Endowment Bank: Deborah C. Mash, MD, Margaret J Basile, Mitsuko Tanaka Oregon Health & Science University: Randy Wotljer, PhD 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, South West Dementia Brain Bank (funding via numerous sources including the Higher Education Funding Council for England (HEFCE), Alzheimer's Research Trust (ART), BRACE as well as North Bristol NHS Trust Research and Innovation Department and DeNDROn): Seth Love, MD, Patrick Kehoe, PhD, Laura Palmer, The Netherlands Brain Bank (funding via numerous sources including Stichting MS Research, Brain Net Europe, Hersenstichting Nederland Breinbrekend Werk, International Parkinson Fonds, Internationale Stichting Alzheimer Onderzoek): Inge Huitinga, MD, Marleen Rademaker, Michiel Kooreman, Institut de Neuropatologia, Servei Anatomia Patologica, Universitat de Barcelona: Isidre Ferrer Abizanda, MD, PhD, Susana Casas Boluda.

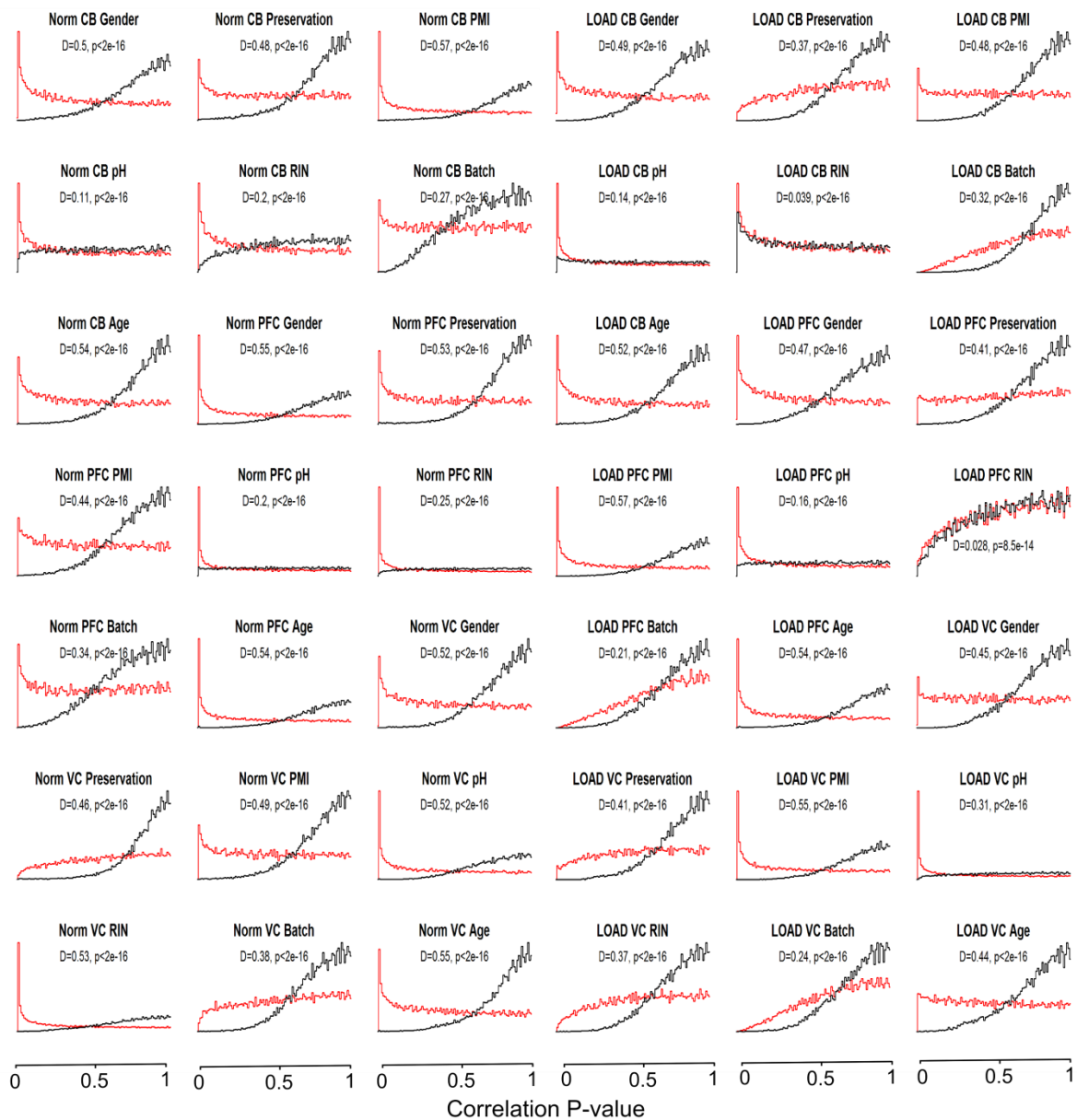


Figure S1 Correcting for Covariate Effects on the Expression Data in the Non-demented and LOAD Brains, Related to Figure 1. Each plot shows the distribution of the correlation P-values (x-axis) between a specific covariate and all the probes based on the expression data before (red curve, raw) and after the adjustment (black curve, corrected). The Kolmogorov-Smirnov (KS) test was used to capture the difference of the correlation P-value distributions. The D-statistics and P-values of the KS test are shown here. Y-axis denotes the frequency, or the fraction of genes showing correlation to a given covariate

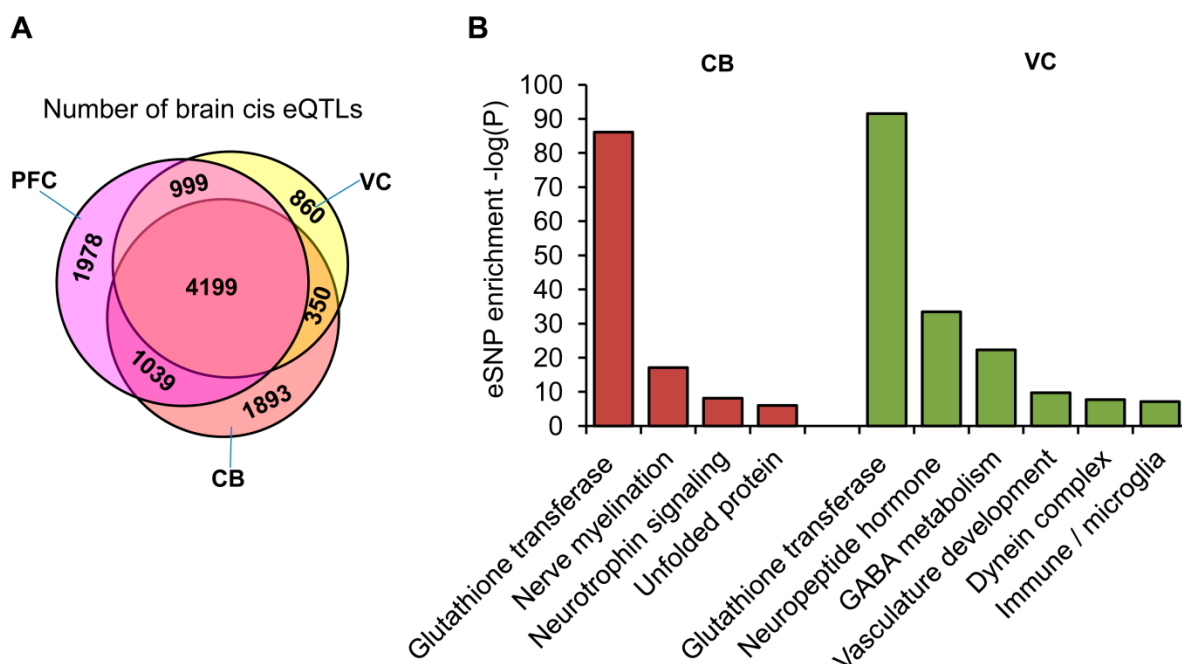


Figure S2 Brain Cis eSNPs and Their Enrichment in Modules, Related to Figure 4. (A) The number and overlap of all *cis* eSNPs detected in the different brain regions in the present study at FDR of 10%. **Table S1** lists all *cis* and *trans* eSNPs detected in the present study. Detailed statistical procedure related to the eSNP detection is provided in the **Extended Experimental Procedures**. (B) We tested the enrichment of brain eSNPs in the differentially connected modules of the multi-tissue co-expression network in LOAD as per brain region. Here we show the VC and CB modules with at least 100 gene members and showing significant enrichment of brain eSNPs (**Figure 4C** highlights eSNP enrichment in PFC).

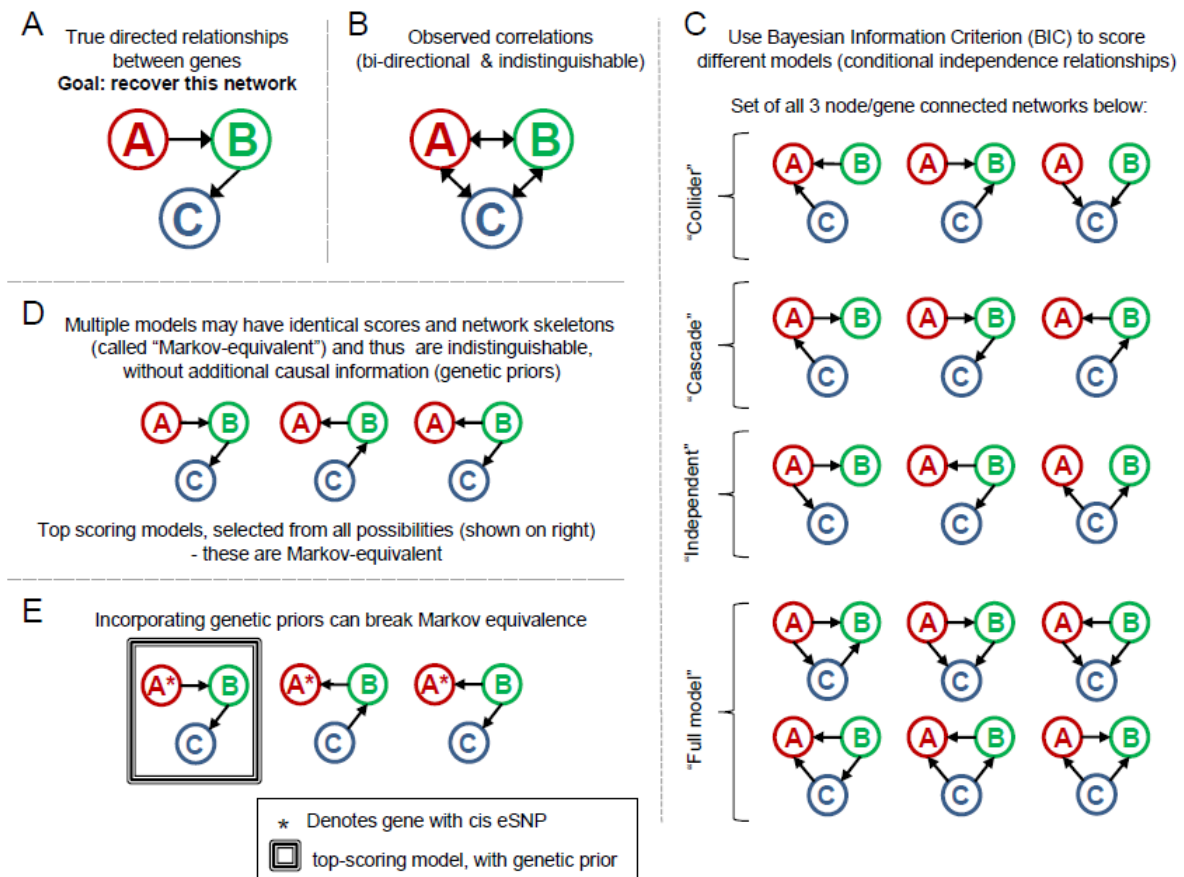


Figure S3 Schematic Drawing for Deriving Causal Gene-Gene Relationships from Gene Expression and eSNP Data, Related to Experimental Procedures.

Correlations between gene expression values alone do not indicate the true causal structure of gene networks. Our goal is to identify the true causal structure between transcripts (A) given observed correlation between gene expression (B), through a combination of model evaluation and incorporation of genetic priors as "causal anchors" (C-E). There are many possible causal relationships between even small groups of genes (C) and we use the Bayesian Information Criterion (BIC) to evaluate the posterior probability of each structure (shown in C), given the expression data. Several of these model structures may be statistically indistinguishable (D). Therefore, additional information is required to determine which of these models most accurately represents the data. We use a hard prior to put genes with strong *cis* eSNPs as head nodes (causal anchors) to distinguish these models (E). This is justified because genetic perturbations (*cis* eSNPs in the case of this dataset) always precede changes in expression, and conversely expression changes do not lead to DNA alterations.

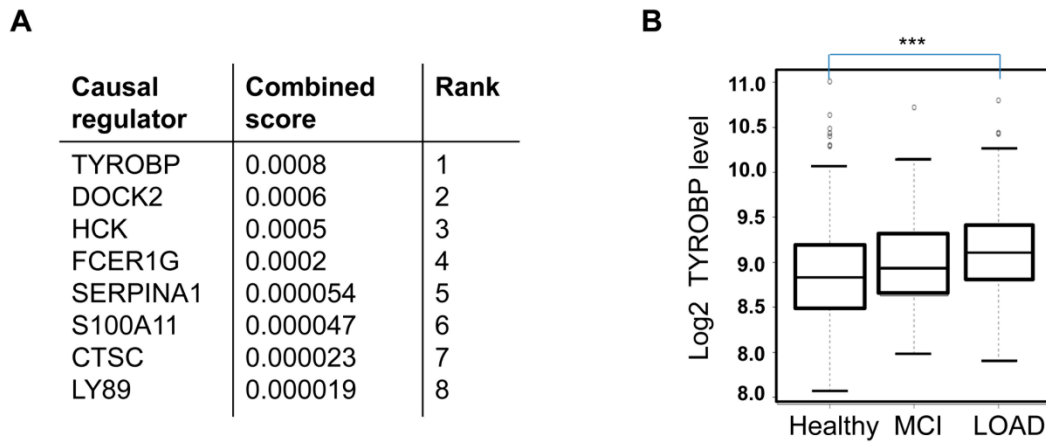


Figure S4 Ranking of Causal Regulators, Differential Expression and Over-expression of *TYROBP*, Related to Table 1. (A) There were eight common causal regulators for all immune networks. We used a combined ranking score based on degree regulated nodes in the network related to each of the common causal regulators and the significance of the differential expression in LOAD brains, where *TYROBP* ranked the highest (see **Results**). (B) *TYROBP* levels in PFC were compared between 377 LOAD patients, 119 with the diagnosis of mild cognitive impairment (MCI) and 359 non-demented controls. The raw t-test P-value was 5.1e-05 while the Benjamini-Hochberg adjusted P-value=5e-04. See **Extended Experimental Procedures** for details on study groups.

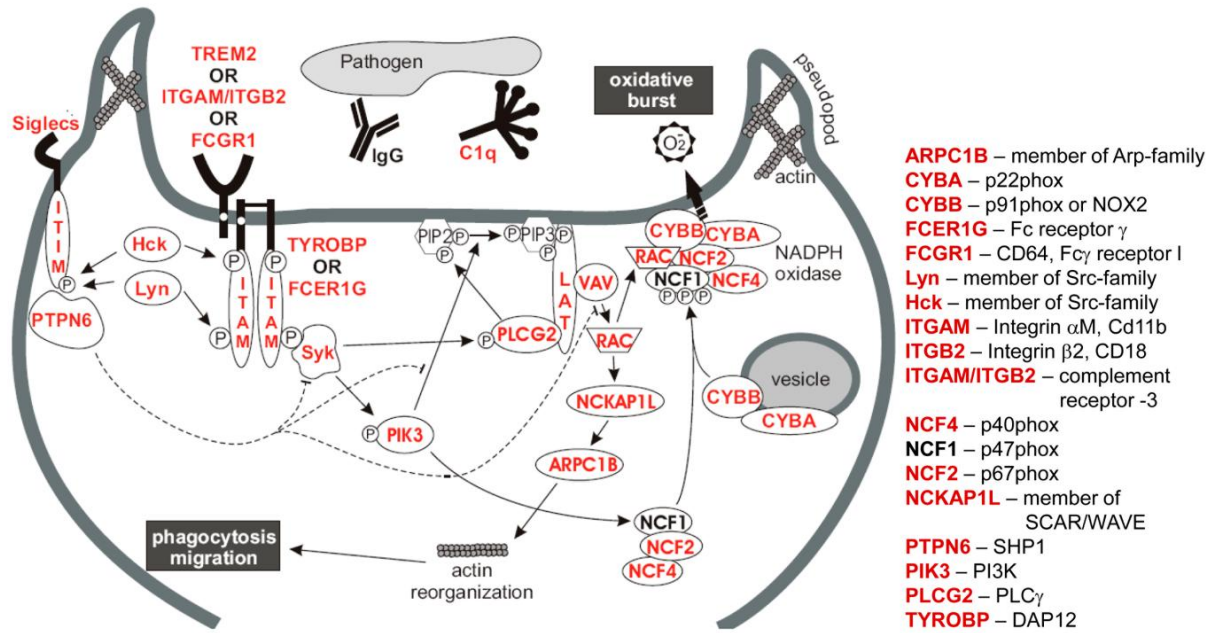


Figure S5 The Microglia Pathogen Phagocytosis Pathway, Related to Figure 5. Genes derived from the LOAD immune and microglia enriched module are marked in red. Pathogens are recognized by complement C1q or immunoglobulin (IgG) that bind to microglia complement receptors (e.g. ITGAM/ITGB2) or Fc-receptors (e.g. FCGR1) that signal via the immunoreceptor tyrosine-based activation motif (ITAM)-containing adaptor molecules TYROBP or FCER1G, respectively. Alternatively, pathogens are directly recognized by classical innate immune receptors (e.g. TREM2) that require the interaction with TYROBP for further signalling.

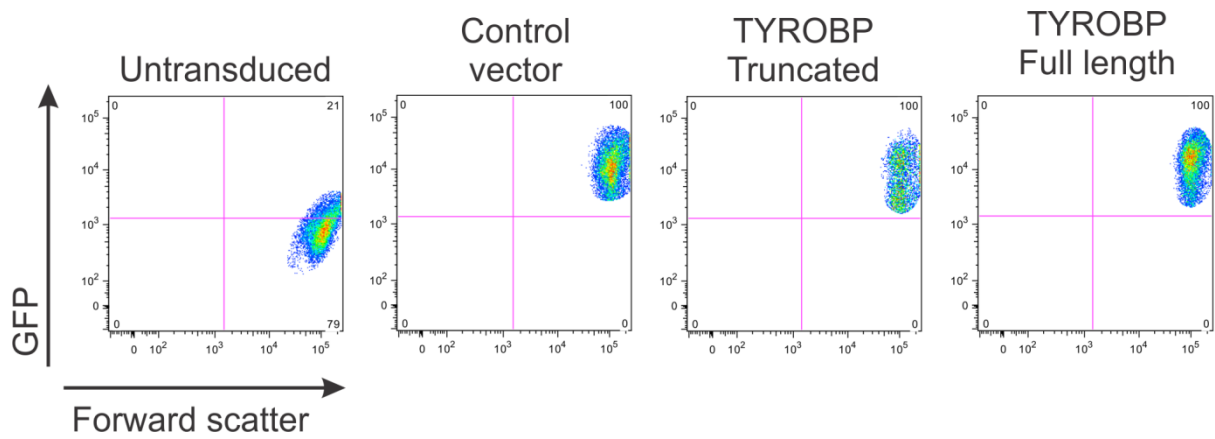
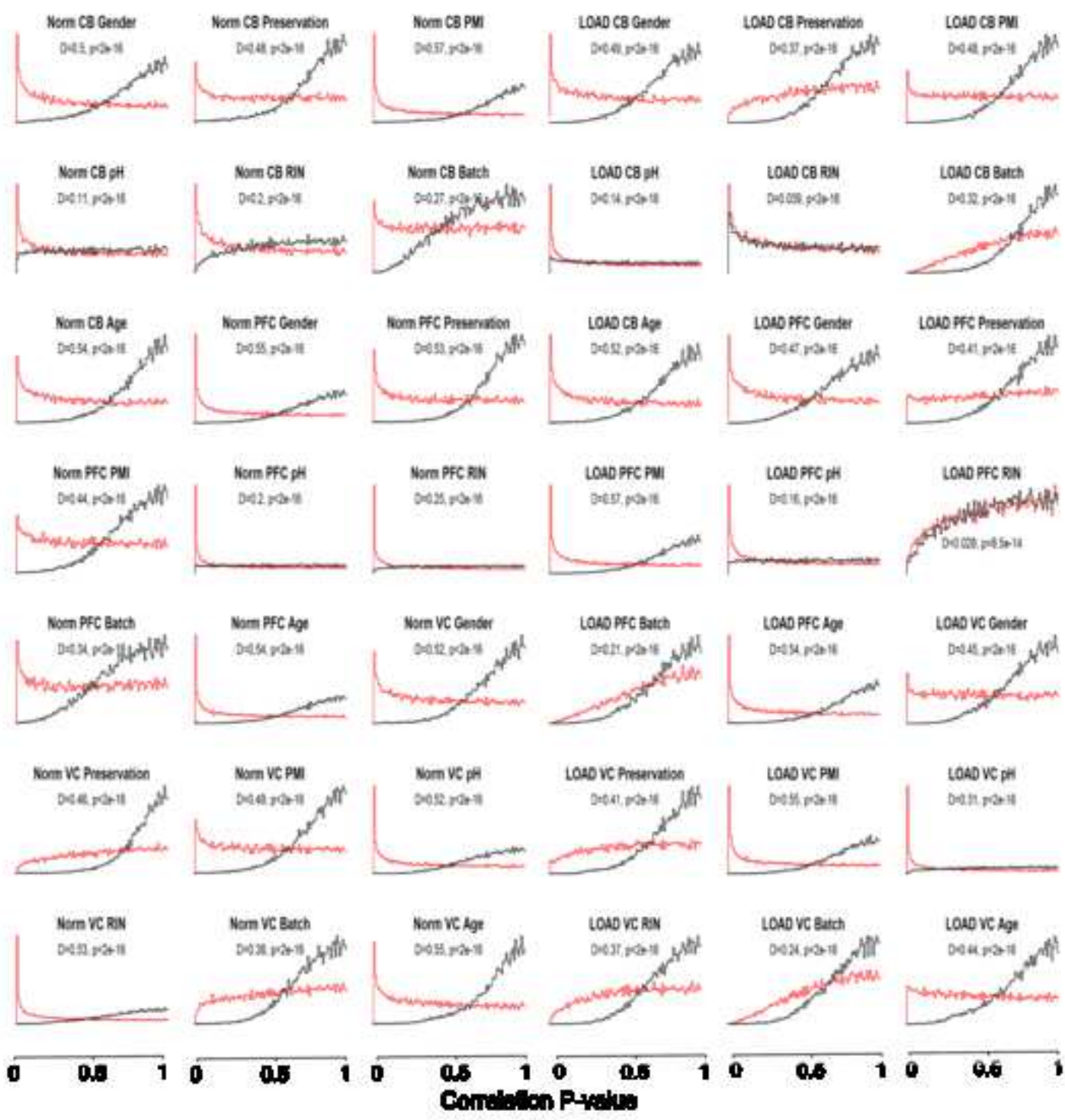


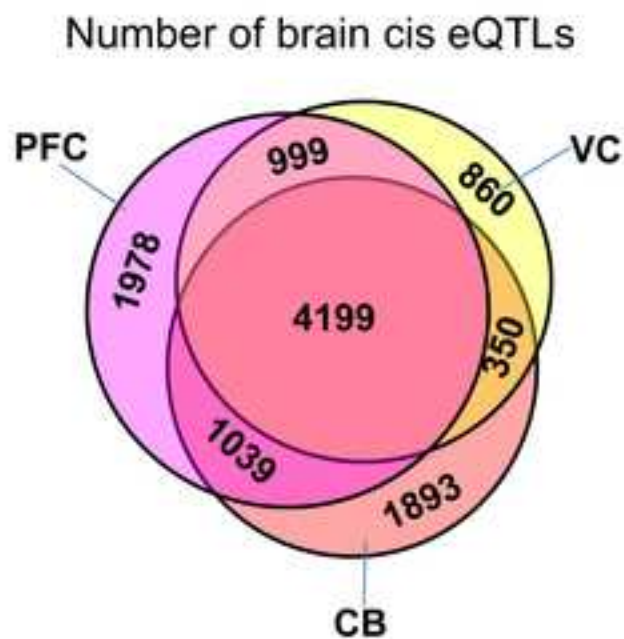
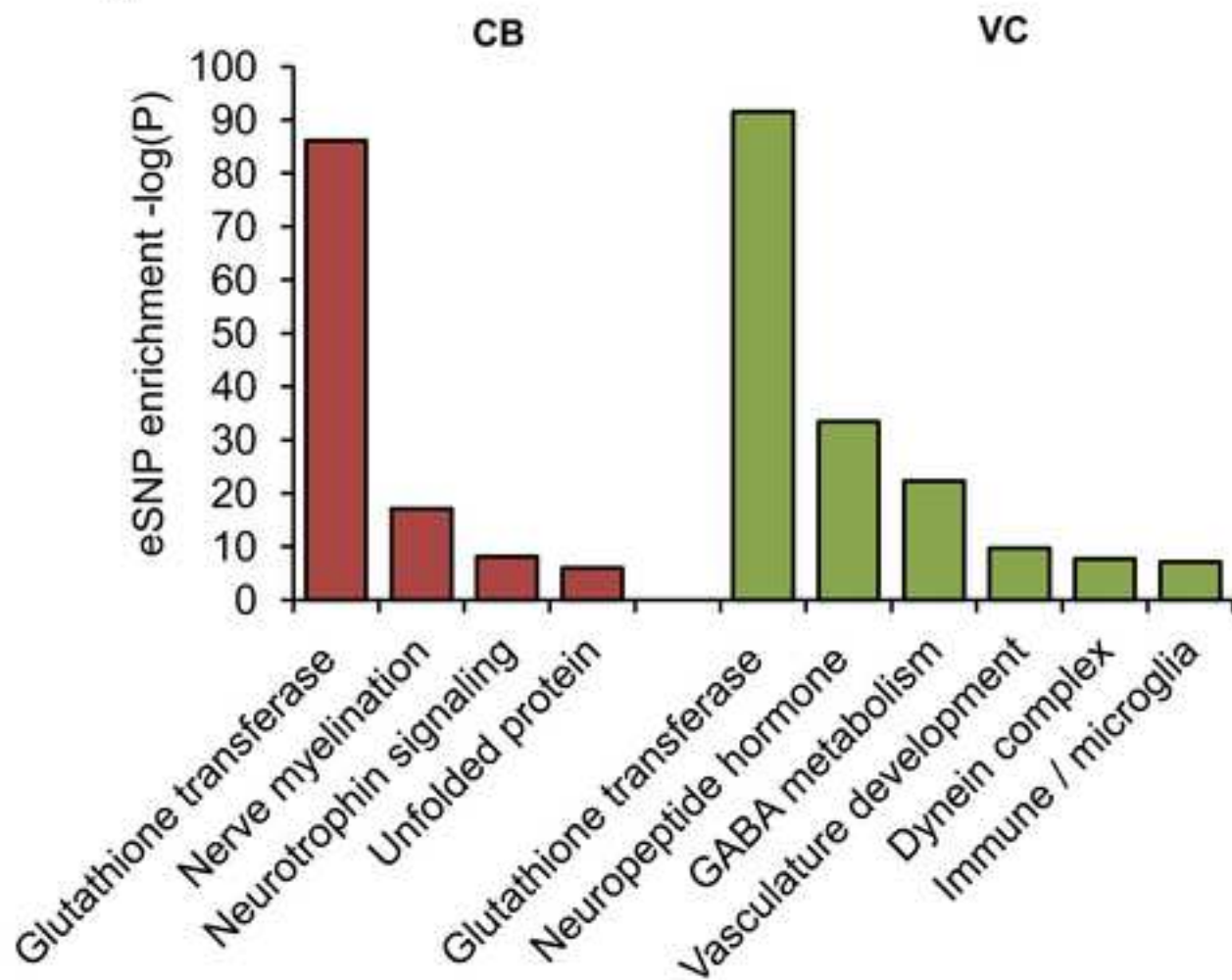
Figure S6 Over-expression of Intact or Truncated TYROBP in Mouse Microglia Cells, Related to Figure 6. Flow cytometry analysis of microglia transduced with full length mouse TYROBP-IRES-GFP (TYROBP full length), truncated mouse TYROBP-IRES-GFP (TYROBP truncated), green fluorescent protein (GFP) control vector or untransduced microglia. Transduced GFP proteins were detected at similar levels in the microglial cells over-expressing either full length TYROBP, truncated TYROBP or the control vector. Untransduced microglia cells were used as control.

SUPPLEMENTAL REFERENCES

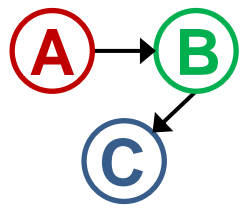
- Bennett, D.A., Schneider, J., Aggarwal, N., Arvanitakis, Z., Shah, R.C., Kelly, J., Fox, J.H., Cochran, E.J., Arends, D., Treinkman, A.D., *et al.* (2006a). Decision rules guiding the clinical diagnosis of Alzheimer's disease in two community-based cohort studies compared to standard practice in a clinic-based cohort study. *Neuroepidemiology* 27, 169-176.
- Bennett, D.A., Schneider, J.A., Arvanitakis, Z., Kelly, J.F., Aggarwal, N.T., Shah, R.C., and Wilson, R.S. (2006b). Neuropathology of older persons without cognitive impairment from two community-based studies. *Neurology* 66, 1837-1844.
- Bennett, D.A., Schneider, J.A., Buchman, A.S., Mendes de Leon, C., Bienias, J.L., and Wilson, R.S. (2005). The Rush Memory and Aging Project: study design and baseline characteristics of the study cohort. *Neuroepidemiology* 25, 163-175.
- Beutner, C., Roy, K., Linnartz, B., Napoli, I., and Neumann, H. (2010). Generation of microglial cells from mouse embryonic stem cells. *Nat Protoc* 9, 1481-1494.
- Braak, H., and Braak, E. (1991). Neuropathological staging of Alzheimer-related changes. *Acta Neuropathol* 82, 239-259.
- Emilsson, V., Thorleifsson, G., Zhang, B., Leonardson, A.S., Zink, F., Zhu, J., Carlson, S., Helgason, A., Walters, G.B., Gunnarsdottir, S., *et al.* (2008). Genetics of gene expression and its effect on disease. *Nature* 452, 423-428
- He, Y.D., Dai, H., Schadt, E., Cavet, G., Edwards, S.W., Stepaniants, S.B., Duenwald, S., Kleinhanz, R., Jones, A.R., Shoemaker, D.D., *et al.* (2003). Microarray standard data set and figures of merit for comparing data processing methods and experiment designs. *Bioinformatics* 19, 956-965.
- Lai, E., Riley, J., Purvis, I., and Roses, A. (1998). A 4-Mb high-density single nucleotide polymorphism-based map around human APOE. *Genomics* 54, 31-38.
- Langfelder, P., Zhang, B., and Horvath, S. (2008). Defining clusters from a hierarchical cluster tree: the Dynamic Tree Cut package for R. *Bioinformatics* 24, 719-720.
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25, 2078-2079.
- Madigan, D.A.Y., J. (1995). Bayesian graphical models for discrete data. *International Statistical Review* 63, 215-232.
- Myers, A.J., Gibbs, J.R., Webster, J.A., Rohrer, K., Zhao, A., Marlowe, L., Kaleem, M., Leung, D., Bryden, L., Nath, P., *et al.* (2007). A survey of genetic human cortical gene expression. *Nat Genet* 39, 1494-1499
- Nochlin, D., van Belle, G., Bird, T.D., and Sumi, S.M. (1993). Comparison of the severity of neuropathologic changes in familial and sporadic Alzheimer's disease. *Alzheimer Dis Assoc Disord* 7, 212-222.
- Ravasz, E., Somera, A.L., Mongru, D.A., Oltvai, Z.N., and Barabasi, A.L. (2002). Hierarchical organization of modularity in metabolic networks. *Science* 297, 1551-1555.
- Schadt, E.E., Molony, C., Chudin, E., Hao, K., Yang, X., Lum, P.Y., Kasarskis, A., Zhang, B., Wang, S., Suver, C., *et al.* (2008). Mapping the genetic architecture of gene expression in human liver. *PLoS Biol* 6, e107.
- Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics* 461-464 6, 461-464.
- Storey, J.D., and Tibshirani, R. (2003). Statistical methods for identifying differentially expressed genes in DNA microarrays. *Methods Mol Biol* 224, 149-157.

- Trapnell, C., Pachter, L., and Salzberg, S.L. (2009). TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* 25, 1105-1111.
- Vonsattel, J.P., Myers, R.H., Stevens, T.J., Ferrante, R.J., Bird, E.D., and Richardson, E.P., Jr. (1985). Neuropathological classification of Huntington's disease. *Journal of neuropathology and experimental neurology* 44, 559-577.
- Wang, I.M., Zhang, B., Yang, X., Zhu, J., Stepaniants, S.B., Zhang, C., Meng, Q., Peters, M., He, Y., Ni, C., et al. (2012). Systems analysis of eleven rodent disease models reveals an inflammatome signature and key drivers. *Mol Syst Biol* 8, 594. doi: 510.1038/msb.
- Warner, J.P., Barron, L.H., and Brock, D.J. (1993). A new polymerase chain reaction (PCR) assay for the trinucleotide repeat that is unstable and expanded on Huntington's disease chromosomes. *Molecular and cellular probes* 7, 235-239.
- Webster, J.A., Gibbs, J.R., Clarke, J., Ray, M., Zhang, W., Holmans, P., Rohrer, K., Zhao, A., Zhao A Fau - Marlowe, L., Marlowe L Fau - Kaleem, M., et al. (2009). Genetic control of human brain transcript expression in Alzheimer disease. *Am J Hum Genet* 84, 445-458.
- Yang, X., Zhang, B., Molony, C., Chudin, E., Hao, K., Zhu, J., Gaedigk, A., Suver, C., Zhong, H., Leeder, J.S., et al. (2010). Systematic genetic and genomic analysis of cytochrome P450 enzyme activities in human liver. *Genome Res* 20, 1020-1036.
- Zhang, B., and Horvath, S. (2005). A general framework for weighted gene co-expression network analysis. *Stat Appl Genet Mol Biol* 4, 12.
- Zhu, J., Lum, P.Y., Lamb, J., GuhaThakurta, D., Edwards, S.W., Thieringer, R., Berger, J.P., Wu, M.S., Thompson, J., Sachs, A.B., et al. (2004). An integrative genomics approach to the reconstruction of gene networks in segregating populations. *Cytogenet Genome Res* 105, 363-374.
- Zhu, J., Wiener, M.C., Zhang, C., Fridman, A., Minch, E., Lum, P.Y., Sachs, J.R., and Schadt, E.E. (2007). Increasing the power to detect causal associations by combining genotypic and expression data in segregating populations. *PLoS Comput Biol* 3, e69. .
- Zhu, J., Zhang, B., Smith, E.N., Drees, B., Brem, R.B., Kruglyak, L., Bumgarner, R.E., and Schadt, E.E. (2008). Integrating large-scale functional genomic data to dissect the complexity of yeast regulatory networks. *Nat Genet* 40, 854-861.

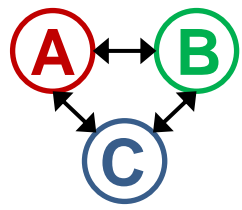


A**B**

A True directed relationships between genes
Goal: recover this network

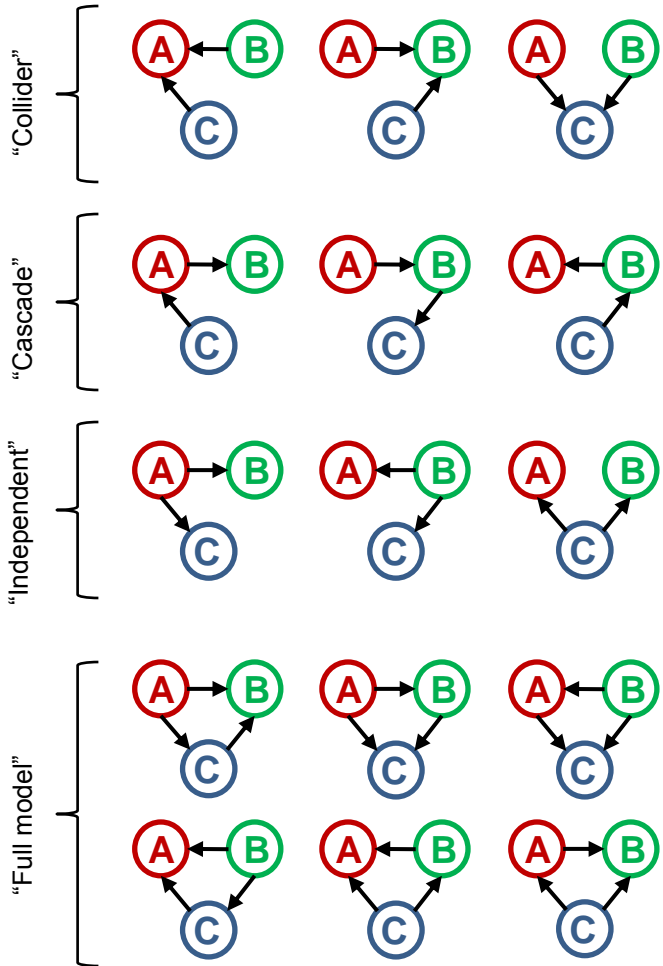


B Observed correlations (bi-directional & indistinguishable)

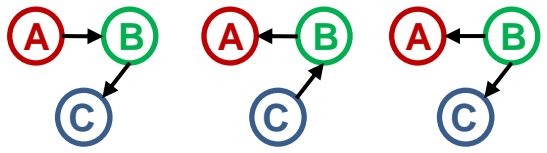


C Use Bayesian Information Criterion (BIC) to score different models (conditional independence relationships)

Set of all 3 node/gene connected networks below:

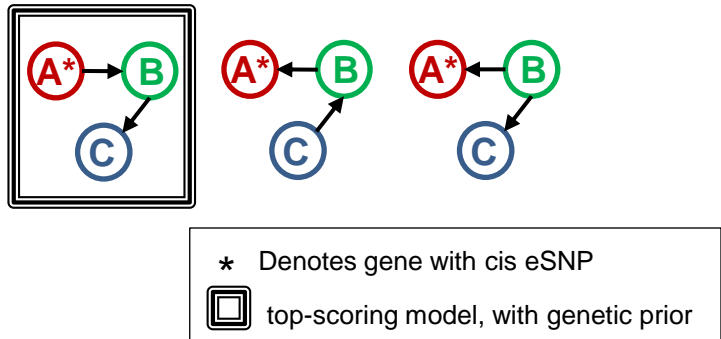


D Multiple models may have identical scores and network skeletons (called "Markov-equivalent") and thus are indistinguishable, without additional causal information (genetic priors)



Top scoring models, selected from all possibilities (shown on right) - these are Markov-equivalent

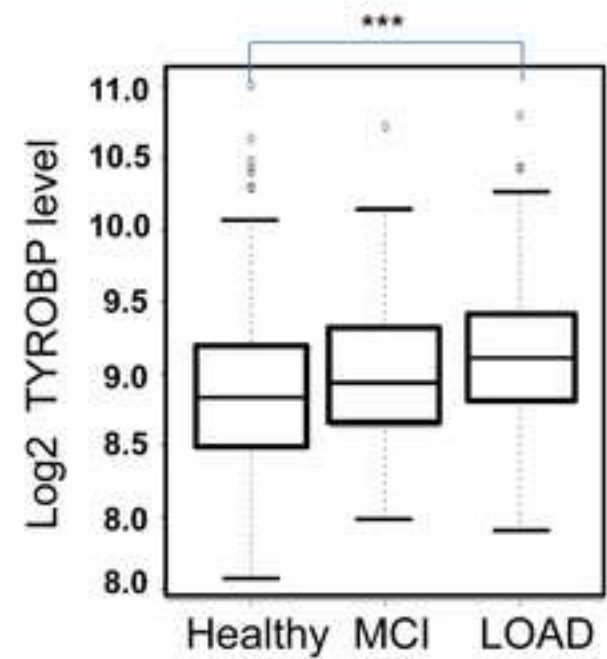
E Incorporating genetic priors can break Markov equivalence

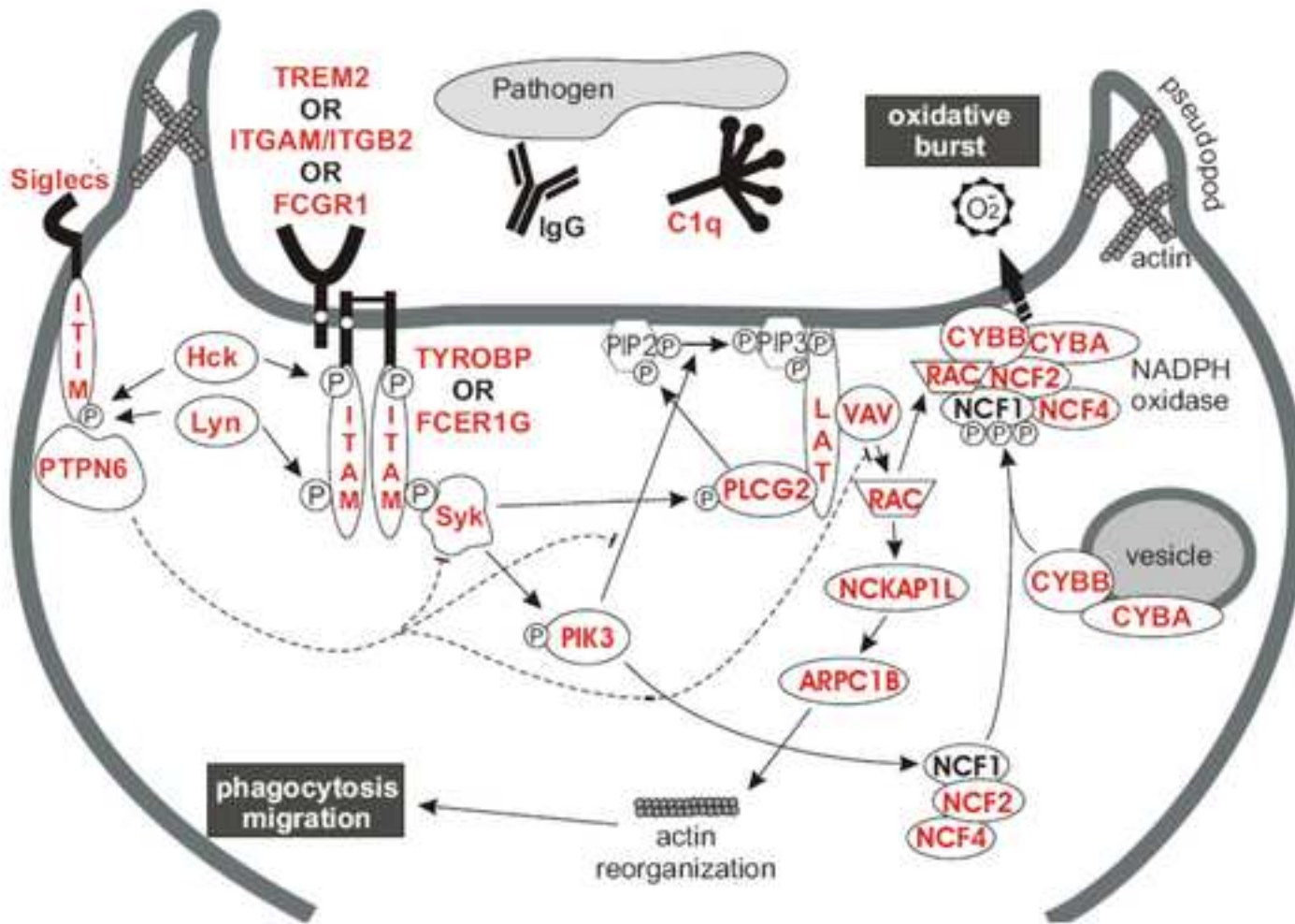


* Denotes gene with cis eSNP
 □ top-scoring model, with genetic prior

A

Causal regulator	Combined score	Rank
TYROBP	0.0008	1
DOCK2	0.0006	2
HCK	0.0005	3
FCER1G	0.0002	4
SERPINA1	0.000054	5
S100A11	0.000047	6
CTSC	0.000023	7
LY89	0.000019	8

B



- ARPC1B** – member of Arp-family
- CYBA** – p22phox
- CYBB** – p91phox or NOX2
- FCER1G** – Fc receptor γ
- FCGR1** – CD64, Fc γ receptor I
- Lyn** – member of Src-family
- Hck** – member of Src-family
- ITGAM** – Integrin α M, Cd11b
- ITGB2** – Integrin β 2, CD18
- ITGAM/ITGB2** – complement receptor -3
- NCF4** – p40phox
- NCF1** – p47phox
- NCF2** – p67phox
- NCKAP1L** – member of SCAR/WAVE
- PTPN6** – SHP1
- PIK3** – PI3K
- PLCG2** – PLC γ
- TYROBP** – DAP12

