Supplementary Material:

The discovery of integrated gene networks for autism and related disorders

Fereydoun Hormozdiari*, Osnat Penn*, Elhanan Borenstein and Evan E. Eichler

MAGI

Here, we provide the formal problem definition for the detection of modules using mutations in cases and the details of the method.

Formal problem definition

We define a disease module related to neurodevelopment as a set of genes that are highly co-expressed during brain development, highly connected in a PPI network, and have a low number of loci with severe mutations in control population (ESP) and a high number of *de novo* mutations in cases.

More formally, given a list of mutations in cases and two types of networks, namely a PPI network $G_{ppi} = (V_{ppi}, E_{ppi})$ and a co-expression network $G_{coexp} = (V_{coexp}, E_{coexp})$, and a lower size bound, k, we are interested in finding a module, i.e., a set of genes $M = \{g_1, g_2, \dots, g_z\}$, of size at least k (i.e., $|M| = z \geq k$) that maximizes an objective function and satisfies certain constraints, as described below.

Objective function definition

We first define a scoring function for each gene in $V = V_{ppi} \cap V_{coexp}$. For each gene g_i we assign a score $S(g_i)$ based on the number of *de novo* missense mutations (m) and LoF mutations (l) observed in that gene and a prior probability, p_{g_i}, of observing a mutation in that gene. For simplicity, we are currently taking the gene's length as the sole determinant of the prior value: $p_{g_i} = \frac{length(g_i)}{\sum_{g \in V} length(g)}$ where $length(g_i)$ is representing the coding sequence length of gene g_i. However, this function can easily be made more complex to account for other parameters such as the percent GC content of each gene. Thus, the binomial probability of observing by chance x total mutations in gene g_i out of T total mutations observed in all genes is $P_{x,T}(g_i) = \binom{T}{x} p_{g_i}^{x} (1 - p_{g_i})^{T-x}$. To separate the missense and LoF mutations, we denote T_L and T_M as the total number of LoF and missense mutations, respectively, seen in all the genes in V and define the score of each gene g_i as follow:

$$S(g_i) = \begin{cases} -\log(P_{g_i}) = -\log(P_{l,T_L}(g_i) \cdot P_{m,T_M}(g_i)) & if \ \ l + m > 0 \\ 0 \ \ if \ l + m = 0 \end{cases}$$

The advantage of this score is that we can easily define the score of a pathway as the summation of scores of each gene:

$$S(\{g_1, g_2, \dots, g_k\}) = -\log(P_{g_1} \cdot P_{g_2} \dots P_{g_k}) = -\sum_{i=1}^{k} log\ P_{g_i} = \sum_{i=1}^{k} S(g_i)).$$

Finally, for a module $M = \{g_1, g_2, \dots, g_z\}$ of size z, we define a z-score motivated objective scoring function: $Obj(M) = \frac{\sum_{g \in V}(S(g) - \mu)}{\sqrt{z}}$ similar to the one used in previous studies (Ideker et al., 2002). In this equation μ is the mean value of the score of all the genes considered (which is a constant for each given set of *de novo* mutations).

Module constraints definition

A set of genes *M* is defined as a module if:

1) The genes in *M* are highly connected in the PPI network:
    a) The genes in M form a connected induced subgraph of the PPI network. For this constraint to minimize the effect of edges found in E_{ppi} that may not hold in human brain tissue/cells, we only consider PPI edges between genes that are also highly co-expressed. In practice, we define highly co-expressed edges as ones in the top 5% of co-expression values in E_{coexp} ($r^2 > 0.37$ for the BrainSpan dataset used here).
    b) The average density of PPI edges in the selected module, *Avg_Density(M)*, should be significantly higher than what is expected for a connected component of size k (the lower bound on the size of the desired module – i.e., $|M| = z \geq k$). More formally, assuming we have a lower bound threshold on the average density of protein interactions for different sizes of modules ($\alpha(x): x \to R$), we require $Avg\_Density(M) = 2\frac{|E_{ppi}(M)|}{z(z-1)}$ (i.e., the total number of observed edges over the maximum possible edges in *M*) to be at least $\alpha(k)$.
2) The genes in *M* are highly co-expressed in the co-expression network:
    a) Any pair of genes in *M* should be at least moderately co-expressed. More formally, we demand that for any two genes $g_i, g_j \in M$, the co-expression between genes g_i and g_j would be higher than a threshold. In practice, the cutoff is set to be the median of all co-expression values in the co-expression network. This condition guarantees that no two genes are in the same module if their co-expression is less than the median of all co-expression values.
    b) The average co-expression between all the genes in the selected module *M* should be higher than what is expected for a set of size k (the lower bound on the size of

the desired module – i.e., $|M| = z \geq k$). More formally, assuming we have a lower bound threshold on the average co-expression for different sizes of modules ($\beta(x): x \to R$), we demand that $Coexp(M) = \frac{\sum_{g_i g_j \in M} CoExp(g_i, g_j)}{z(z-1)} > \beta(k)$.

3) The total number of LoF loci in control population (e.g., the ESP) for genes in modules should be less than a predefined threshold (the results provided are for a maximum of five LoF loci).

Algorithm description

To tackle the problem, MAGI performs three major steps: (1) detecting high-scoring seed pathways; (2) merging the seed pathways into modules; and (3) improving the modules by adding or removing single genes (local search).

The output of the method is: (1) *M_Best*, i.e., the module that maximizes the score; (2) *M_Extended*, i.e., a set of genes that appear in more than 5% of the suboptimal modules (modules with score within the top 1 percentile that also overlap *M_Best*); and (3) a "confidence score" calculated for every gene that represent the percentage of suboptimal modules this gene is found in.

Seed detection

The seed pathways are high-scoring simple paths in the PPI network (5-8 genes) that satisfy the constraints. Specifically, to satisfy constraint 1.a, the seeds are defined using only edges found in the PPI network and show high co-expression. We represent this graph as $G(V, E)$, where V is the list of genes and E is the list of the PPI edges showing high co-expression. However, there is no polynomial solution for discovering the full set of simple pathways in the graph. Therefore, we are using an approach based on the color-coding algorithm (Alon et al., 1995) to find the pathways with the maximum score. This approach also allows us to limit the total number of LoF mutation loci found in controls (thus satisfying constraint 3) and outputs a set of high-scoring paths in addition to the maximum-score pathway.

The color-coding approach is an efficient method for finding simple paths of size $h \leq \log(|V|)$ in polynomial time. A simple extension of this method allows for finding a path that maximizes a certain score function that is assigned to each node (e.g., Dost et al., 2008). This color-coding approach has gained popularity in finding pathways in PPI (e.g., QPath (Shlomi et al., 2006), Qnet (Dost et al., 2008)) or counting paths and trees (Alon et al., 2008). In short, this approach involves two steps: (a) randomly color the graph's nodes with h different colors and (b) a dynamic programming algorithm for finding the colorful path (i.e., a simple path that covers all the h colors exactly once) that maximizes the score. Iterations of these two steps are needed since the optimal path is not necessarily colorful at each iteration. It was shown that expected $O(e^h)$ iterations are required to find the optimal path (Alon et al., 1995).

We have slightly modified the dynamic programming step of the color-coding algorithm to allow for a limit on the total number of mutations found in controls, denoted by Δ. Each gene $g \in V$ is assigned with a score $S(g)$ (as explained above) and an integer penalty function $C(g): g \to I$ representing the number of mutation loci seen in controls (e.g., ESP). Given a colored graph, a pathway length h, and a fixed threshold Δ, our aim is to find a simple path with a maximum score, i.e., $S(path) = \sum_{g_i \in path} S(g_i)$, while keeping $C(path) = \sum_{g_i \in path} C(g_i) \leq \Delta$. This can be done easily by adding a new dimension of size Δ to the dynamic programing matrix, representing the penalty. Let $W(v, c, J)$ denote the maximum score of a simple path that ends at gene $v \in V$, has a total penalty c, and covers every color in the set J. The recursion formula to fill this matrix is: $W(v, c, J) = max_{u \in V} \begin{cases} W\big(u, c - C(v), J - j(v)\big) + S(v) & (u, v) \in E \\ -\infty \end{cases}$ $W\big(v, C(v), j(v)\big) = S(v)$ as the base case.

In the above recursion formula, $j(v)$ represents the color assigned to gene v. Once the matrix is filled, the optimal path can be easily traced back from the cell with $\max(W(v, c, J))$ where J is the set of all colors.

To make sure that the seeds are independent from each other, we consider only the optimal colorful path from the dynamic programming step of each coloring iteration as a potential seed. In practice, we run 1000 iterations for each threshold (Δ = 0, 1, 2 and 3) and possible path length (h = 5, 6, 7, or 8) to produce a total of 16,000 potential pathways seeds (there are potential repetitions of seeds based on the random coloring). We then define "high-scoring seeds" as having a score higher than half the score of the optimal seed of the same category (i.e., a pathway with the same Δ and h). Only these highly scored paths are used in the next step to create the modules.

Creating modules

We model the clustering of the seeds into modules as a random walk in a graph created from the seed pathways. Each node in this graph represents a seed and two nodes (i.e., two different seeds) are connected if their union satisfies the constraints (as defined above). The random walk starts from a random node and tries to merge that seed with one of its neighbors to generate a module that satisfy the constraints, as long as the score does not decrease. This procedure continues until no more seeds can be added. Finally a series of local search moves are applied to find a local optimal solution.

More formally, assuming $U = \emptyset$, starting from a random node r (i.e., a seed pathway), the genes in this node (seed) are added to the module, $U \leftarrow U \cup r$. We then randomly traverse to one of the neighbors of node r, denoted as r', as long as $U \cup r'$ satisfies the constraints and $Obj(U \cup r') \geq Obj(U)$. We then assign $U \leftarrow U \cup r'$and continue the random walk from node r' till no more iterations can be made.

Finally, we apply a local search procedure that adds, removes and swaps genes from the set *U* as long as the total score improves and the conditions are satisfied, thus reaching a local maximum. In practice, since the minimum size of a module is unknown, we run the method for $k=\{15, 20, 25, 30, 35\}$ to create 100,000 clusters per *k* and chose the module with the highest score among them.

Discovery of additional modules (iterative running of MAGI)

To investigate the existence of additional distinct modules, it is possible to remove the genes found in *M_Best* module from the PPI and co-expression networks and then iteratively run MAGI. An additional filter is then applied to exclude non-significant modules (in terms of the score or the number of LoF mutations covered) or modules that overlap with modules already found in previous iterations.

We investigated the existence of a third module for the ASD+ID cohort by re-running MAGI after removing the genes found in *M1_Best* and *M2_Best*, and we found a significant module M3 (*M3_Extended* genes are provided in Supplementary Table 1). However *M3_Extended* was not considered for further analysis, since it highly overlaps with *M1_Extended* (over 25% of the genes in *M3_Extended*). In addition, *M1_Extended* was found to be highly co-expressed with *M3_Extended* and there are also many PPI interactions connecting them.

We also applied the same process to the epilepsy and schizophrenia cohorts. The second epilepsy module and the second schizophrenia module did not cover a significant number of *de novo* LoF mutations and thus were not considered for further analysis. Furthermore, the second schizophrenia module overlapped with the schizophrenia M1 module. Repeating the same process on the schizophrenia cohort for the third time produced an additional module which is significant in terms of both its score and the number of LoF mutations it covers and does not overlap with any of the two previous modules (this new module was denoted as M2 for schizophrenia).

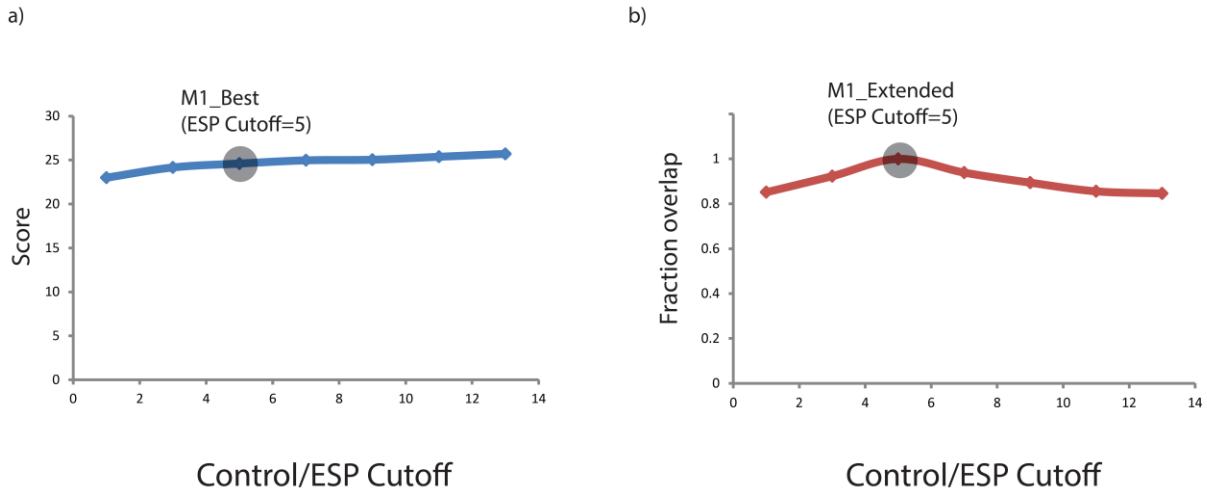
Robustness of MAGI to changes in its parameters and thresholds

We have thoroughly tested the robustness of the modules reported in this manuscript to various changes in MAGI's parameters. Throughout the robustness analysis we have calculated the fraction of overlap among two modules as the size of the intersection of the two gene sets divided by the average size of the two sets.

1) The identified modules are robust to the ESP/control threshold

To test whether the results are robust to the control threshold, we have applied MAGI to the ASD+ID data with different control thresholds ranging from 1 to 13 (Supplementary Figure 1). We found that the score of the *M_Best* does not change dramatically (a maximum increase of

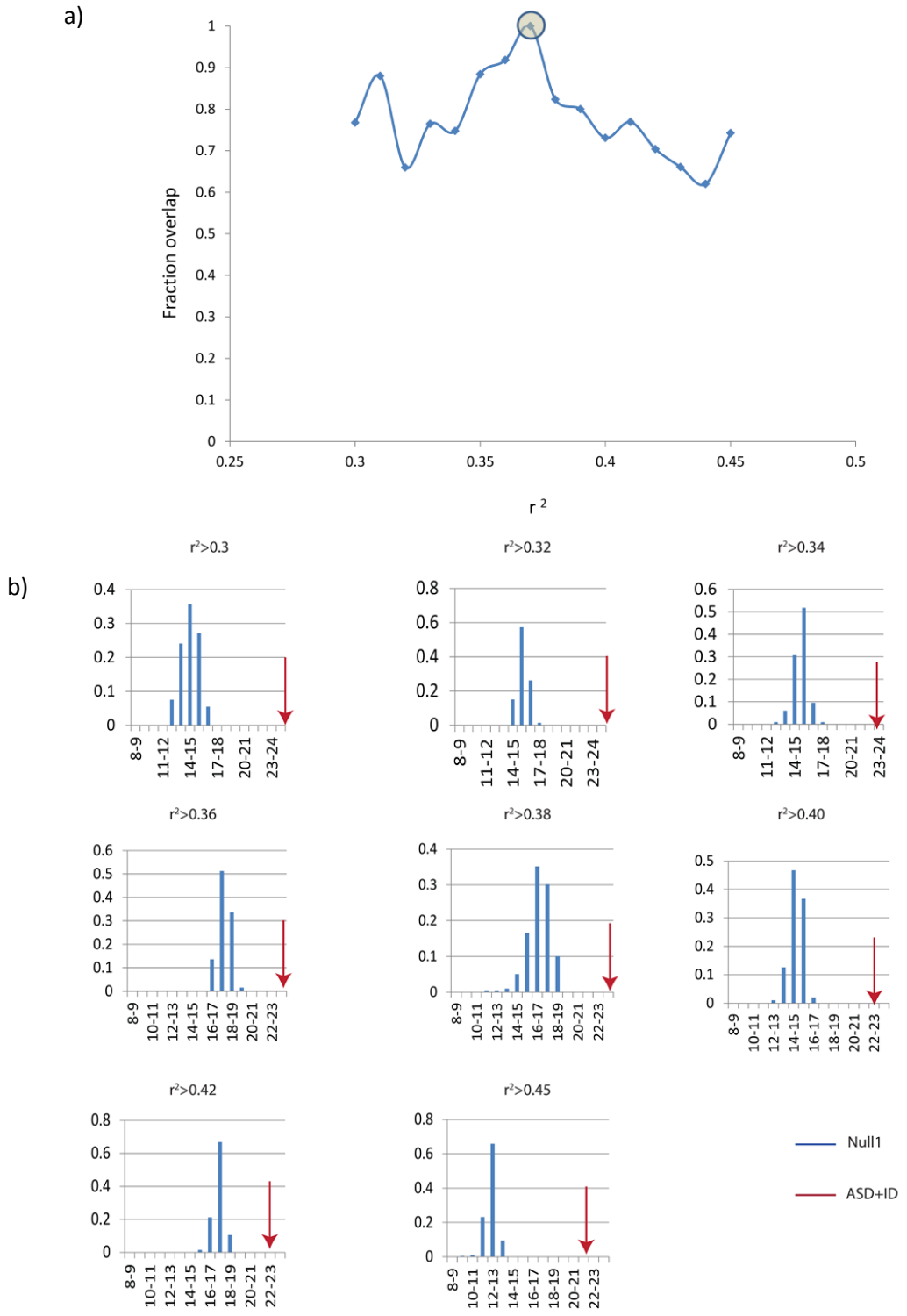
10%) when the control threshold increases to almost three times the threshold value used in the paper (Supplementary Figure 1a.). In addition, the genes covered by the modules found for different thresholds highly overlap. Comparing the resulted *M1_Extended* modules found using these thresholds to the *M1_Extended* reported in the main paper (i.e., upper bound of 5 LoF mutations in control), we found that there is over 85% overlap between them (Supplementary Figure 1b).



Supplementary Figure 1. a) The total score of *M1_Best* for different control (ESP) thresholds. b) The fraction of overlap of the extended module found for different control (ESP) thresholds with the *M1_Extended* module reported in the paper.

2) Robustness to changes in the co-expression threshold for defining "highly co-expressed" genes.

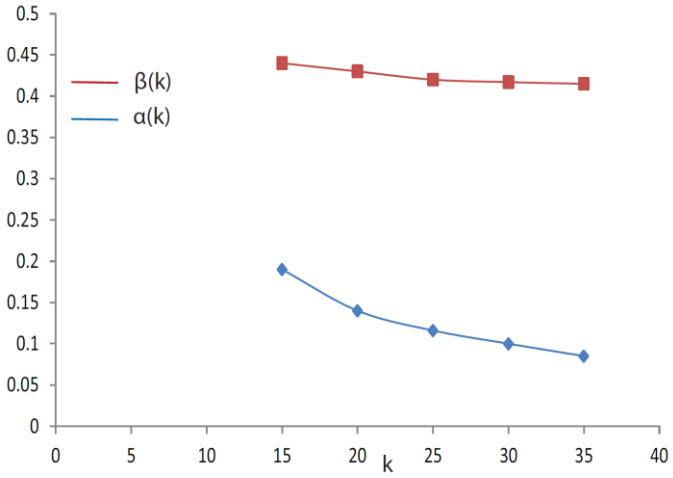
As mentioned in constraint 1a above, we only consider PPI edges between genes that are also highly co-expressed. In the modules reported in the paper we used the top 5% value of the co-expression values (i.e., $r^2>0.37$) as the threshold for highly co-expressed genes. Here, we tested how the modules would differ if we change this threshold. We changed this parameter from $r^2>0.3$ to $r^2>0.45$ (i.e., $|r|>0.5477$ to $|r|>0.6708$) and compared the *M_Best* modules found versus the *M1_Best* module reported in the paper. Interestingly, the overlap was quite high (>70%) supporting that MAGI is robust to changes to this parameter. The fraction of overlap of best modules found versus *M1_Best* is shown in the Supplementary Figure 2.



Supplementary Figure 2. a) Overlap of the *M_Best* modules found using different co-expression cutoffs and *M1_Best* (found using the default cutoff r2>0.37), marked with a circle. b) *M_Best* using different co-expression cutoffs was found to have a significantly higher score in comparison to the top-scored module (200 simulations) using the null model Null-1 (p<0.005).

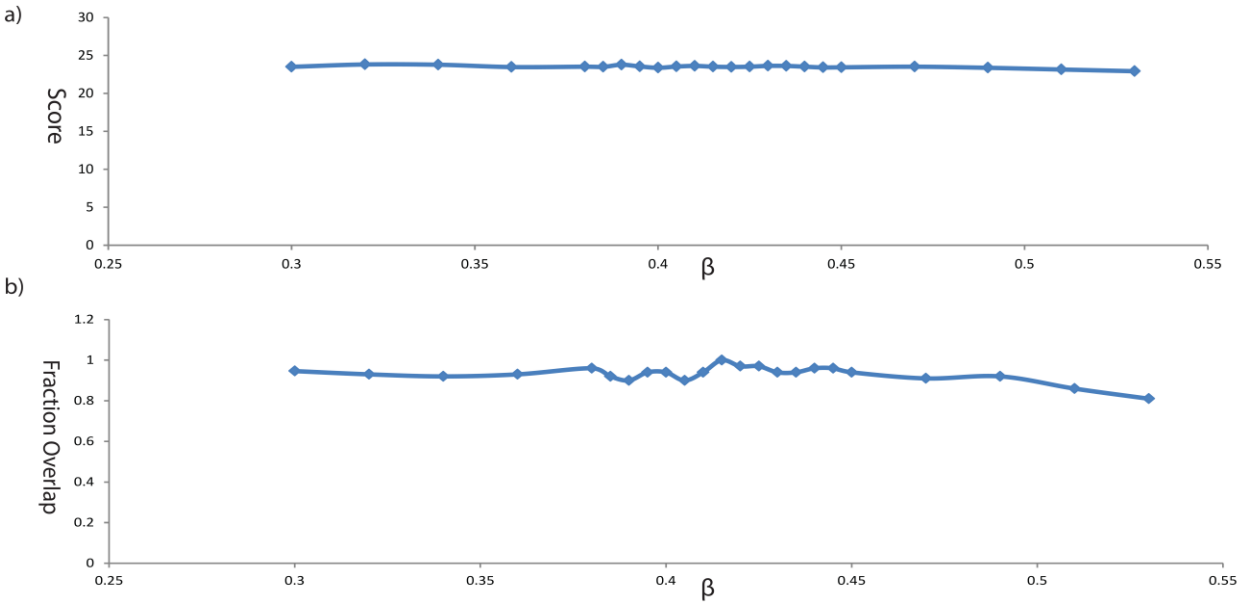
3) Robustness to average density ($\alpha(k)$) and average co-expression ($\beta(k)$) thresholds

To set the thresholds for $\alpha(k)$ and $\beta(k)$, we have created 1,000,000 connected components of different sizes (k = 15, 20, 25, 30 and 35) using a random walk strategy in the PPI network. We then set $\alpha(k)$ to be the 95th percentile of the average density of these components. To set the threshold $\beta(k)$, we used a more conservative formula and take the 95th percentile of $(\frac{\sum_i x_i^3}{n})^{1/3}$ where x_i are the co-expression values (r^2) between every two genes in the component (Supplementary Figure 3).

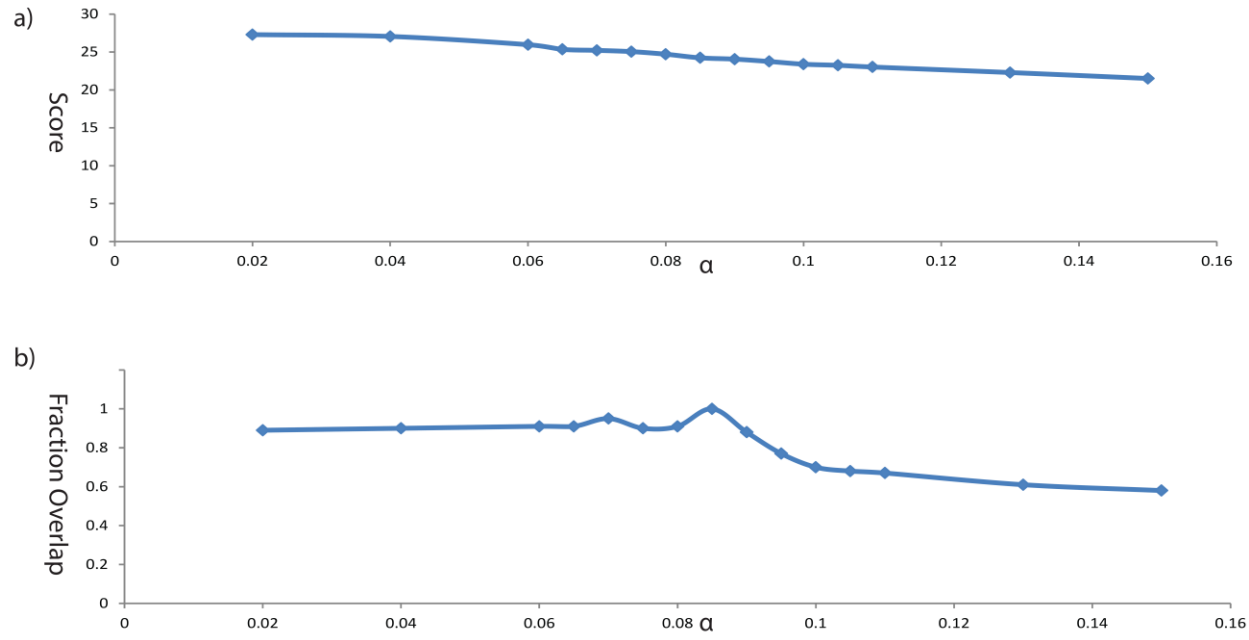


Supplementary Figure 3. The thresholds of average co-expression and edge density used for different *k* values.

We investigated the robustness of MAGI to each of these parameters using the ASD+ID dataset. We found that the resulted modules are highly robust to moderate changes to both parameters in terms of the score of the identified modules as well as the genes included in them. The change in the score of the modules found for different values of β (ranging from 0.3 to 0.53) is less than 3% (Supplementary Figure 4a). The total overlap between the extended modules for different thresholds of β (ranging from 0.3 to 0.53) and the module found using the default β=0.415 (for k=35) is higher than 93% on average (Supplementary Figure 4b). This shows that MAGI in general and the ASD+ID modules specifically are very robust to changes in β paramter. Similar robustness results were obtained for the α parameter (Supplementary Figure 5).



Supplementary Figure 4. Robustness of the ASD+ID modules to changes in the $\boldsymbol{\beta}$ parameter. The seed pathways detected for the ASD+ID mutations were used to generate 5000 clusters, resulting *M1_Best* and *M1_Extended*_for different $\boldsymbol{\beta}$ values ranging from 0.3 to 0.53. In all these analyses *k* was set to be 35. a) The score of *M1_Best* for different values of $\boldsymbol{\beta}$. b) The fraction of overlap between *M1_Extended* modules found for different values of $\boldsymbol{\beta}$ in comparison to *M1_Extended* module found using the default $\boldsymbol{\beta}$=0.415 for *k*=35.



Supplementary Figure 5. Robustness of the ASD+ID modules to changes in the $\boldsymbol{\alpha}$ parameter. The seed pathways detected for the ASD+ID mutations were used to generate 5000 clusters, resulting *M1_Best* and *M1_Extended*_for different $\boldsymbol{\alpha}$ values ranging from 0.02 to 0.15. In all these analyses *k* was set to be 35. a) The score of *M1_Best* for different values of $\boldsymbol{\alpha}$. b) the fraction of overlap between *M1_Extended* found for different values of $\boldsymbol{\alpha}$ in comparison to *M1_Extended* found using the default $\boldsymbol{\alpha}$=0.085 for *k*=35.

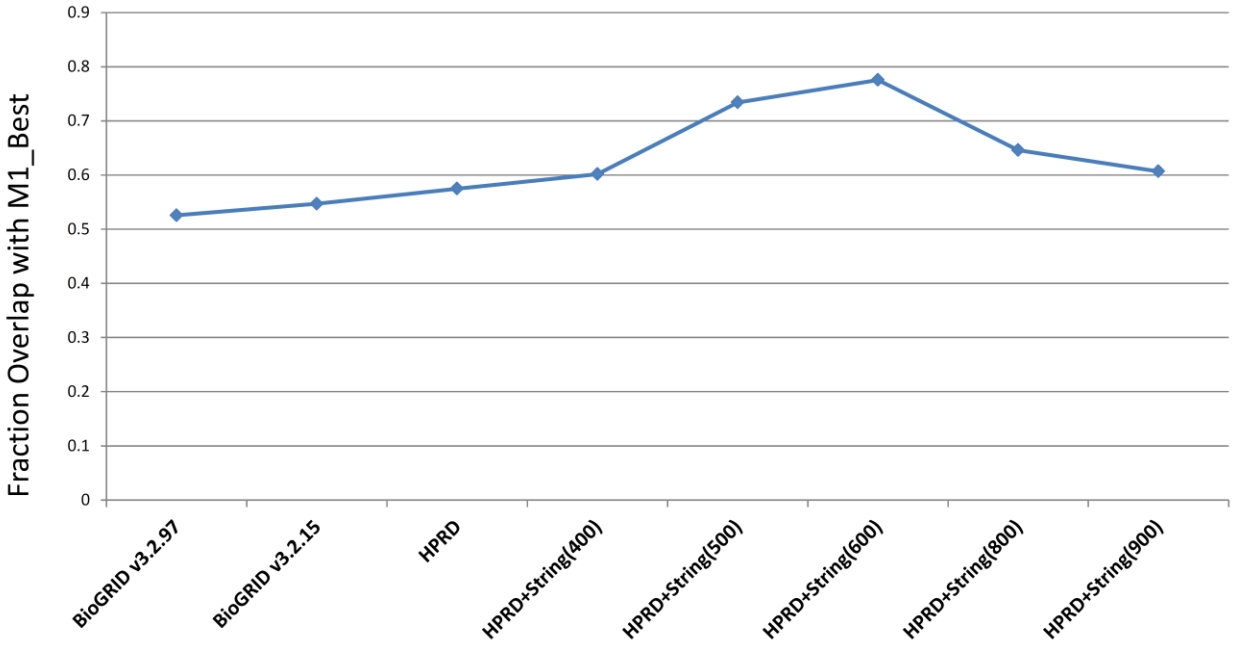
4) Robustness to PPI networks

The reported M1 and M2 modules were found using the union of two PPI networks, namely, the HPRD dataset and the STRING dataset (with confidence score higher than 700 and experimental score higher than 400). To test the robustness of these results to changes in the input PPI network, we run MAGI on various PPI networks, given in Supplementary Table 1.

Supplementary Table 1. Properties of the different PPI networks used to test the robustness of MAGI.

PPI network	**#Nodes**	**#Edges**
HPRD	9607	39042
HPRD+STRING (Confidence score>400)	12023	77233
HPRD+ STRING (Confidence score>500)	11612	72162
HPRD+ STRING (Confidence score>600)	11375	68349
HPRD+ STRING (Confidence score>700)	**10581**	**52801 (Used in the paper)**
HPRD+ STRING (Confidence score>800)	10377	49363
HPRD+ STRING (Confidence score>900)	10191	46305
BioGrid – v 3.2.97	14882	112859
BioGrid – v3.2.15	18359	147715

We compared the *M1_Best* modules found using all these PPI networks to the *M1_Best* module reported for the HPRD+STRING (confidence score>700). The overlap of the modules is shown in Supplementary Figure 6.



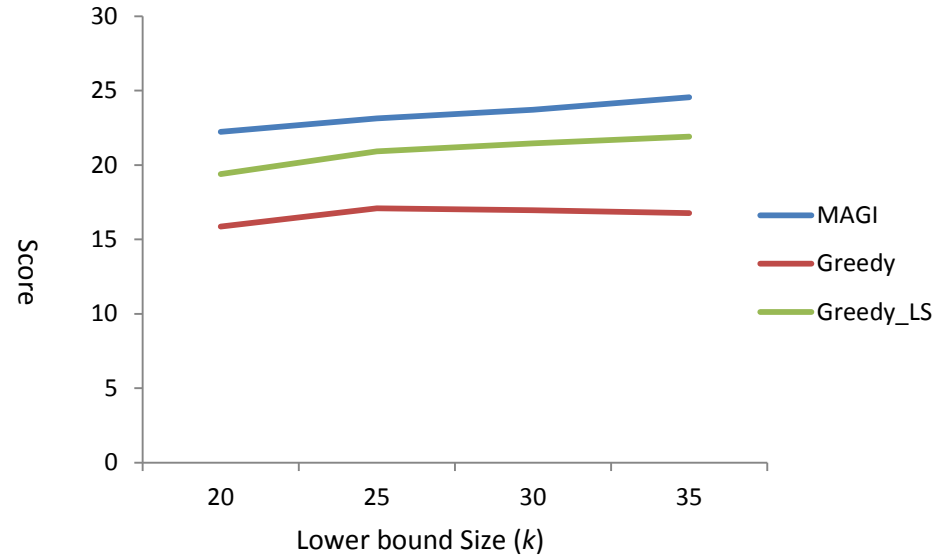
Supplementary Figure 6. Fraction of overlap between *M1_Best* and *M_Best* modules produced by MAGI using different PPI networks as input. Parameters were fixed to the same values used to discover *M1_Best*.

MAGI's performance

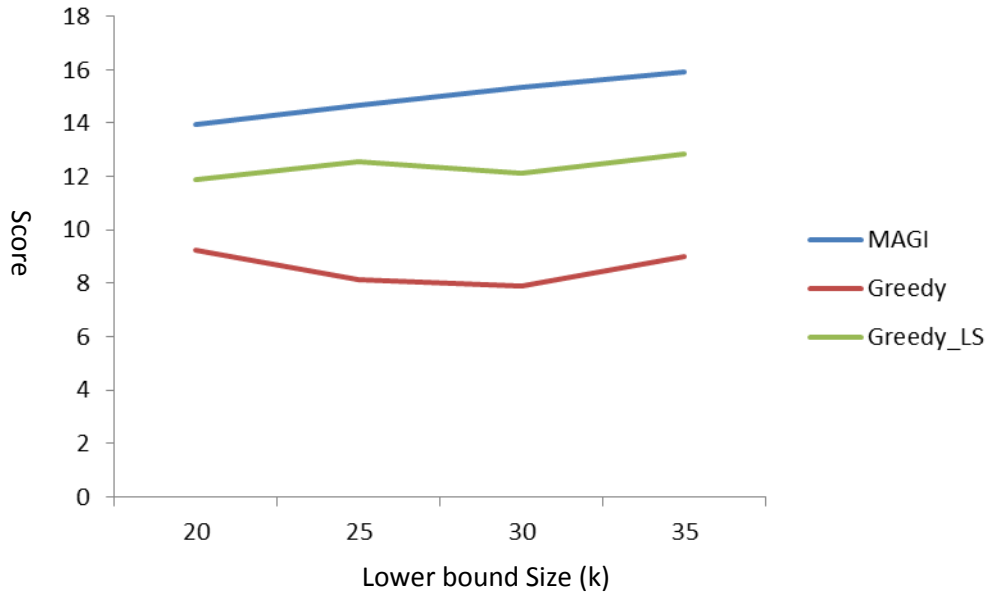
1) MAGI outperforms both the naïve greedy and greedy with local search methods

We compared the results of MAGI for ASD+ID probands and siblings run for different *k* values against two standard approaches of greedy and greedy combined with local search. As the set of constraints and objective functions defined for MAGI are not covered by any available tool, we cannot compare our results with a direct implementation of the published tools for module discovery. However, we can implement and test the main approaches suggested for finding modules by imposing the same constraints as in MAGI. For this, we have implemented a standard greedy approach (similar to what was done in NETBAG (Gilman et al., 2011) and an approach that combines greedy with local search (Greedy_LS) and compared their results against MAGI. The results show that for all different sizes MAGI outperforms the modules discovered using the greedy approaches (Supplementary Figures

Supplementary Figure 7 andSupplementary Figure 9). Moreover, the greedy with local search approach, which similar to MAGI guarantees to find a local maximum, also underperforms in comparison to MAGI.



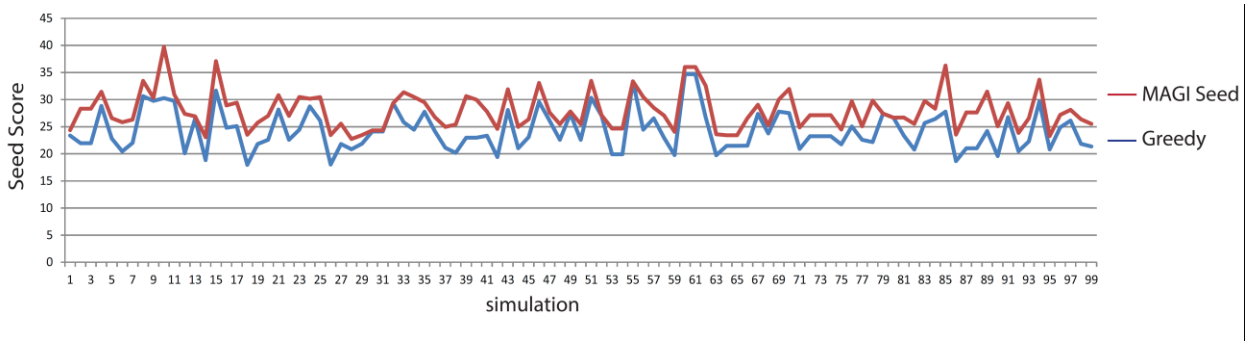
Supplementary Figure 7. Comparison of the scores of the best modules found using MAGI (blue line), a greedy approach (red line) and a greedy approach with additional local search (green line). The three methods were applied on the ASD+ID dataset using different *k* values.



Supplementary Figure 8. Comparison of the scores of the best modules found using MAGI (blue line), a greedy approach (red line), and a greedy with additional local search (green line) using the sibling data.

2) MAGI seeds selection outperforms greedy seeds selection

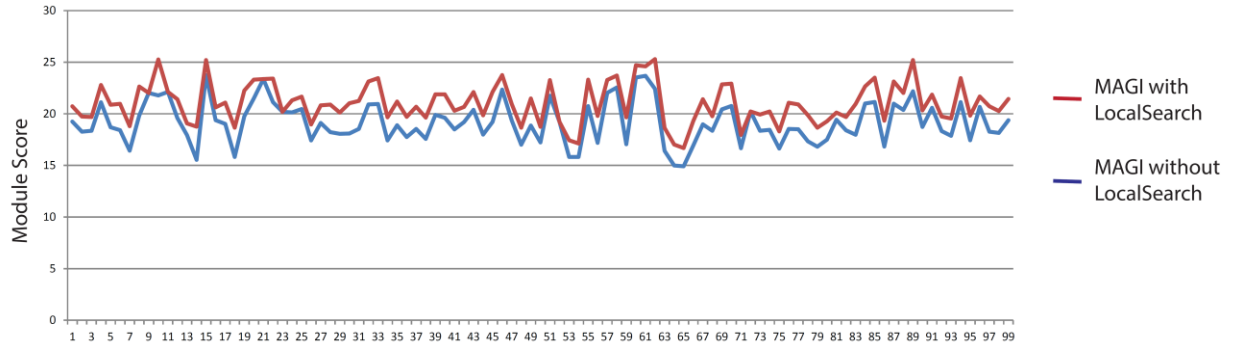
We compared the score of the seed pathways (of size 8) found using MAGI's color-coding approach to the score of the seed pathways found using a greedy traversal of the networks. On the real data, the highest scoring seed using MAGI's color coding approach was found to have a higher score (42.37) compared to the score of the greedy graph traversal starting from 1000 random genes (for which the score was 40.1). Similarly, applying 100 simulations of randomly distributing the ASD+ID mutations (accounting for gene length), Supplementary Figure 9 shows that as expected the best seed found using the color-coding algorithm consistently outperformed the one found using the greedy traversal approach.



Supplementary Figure 9. The maximal score of the seed pathways of length 8 found using MAGI discovery (red) compared to that of the seeds found using a greedy traversal (blue).

3) Local search improvement

We also compared the score of the *M_Best* module found using MAGI default clustering approach that includes a local search step after merging the seed pathways (using the random-walk) to the *M_Best* module found without using the local search step. Using 100 simulations of random permutations of the mutations, we found that in most cases (99 out of 100) the local search step improves the total score of the resulted module (Supplementary Figure 10).



Supplementary Figure 10. The score of best module found using MAGI with and without the final local search step.

Comparison of *M1_Extended* and *M2_Extended* to previously published networks

We have compared *M1_Extended* and *M2_Extended* to recently published co-expression based modules for ASD (Parikshak et al., 2013; Willsey et al., 2013), PPI-based approaches (AXAS(Cristino et al., 2014), SteinerNet (Tuncbag et al., 2012)) as well as other previously published modules associated with autism (NETBAG (Gilman et al., 2011), DAWN (Liu et al., 2014)). Some of these modules overlap with *M1_Extended* and/or *M2_Extended* (Supplementary Table 2).

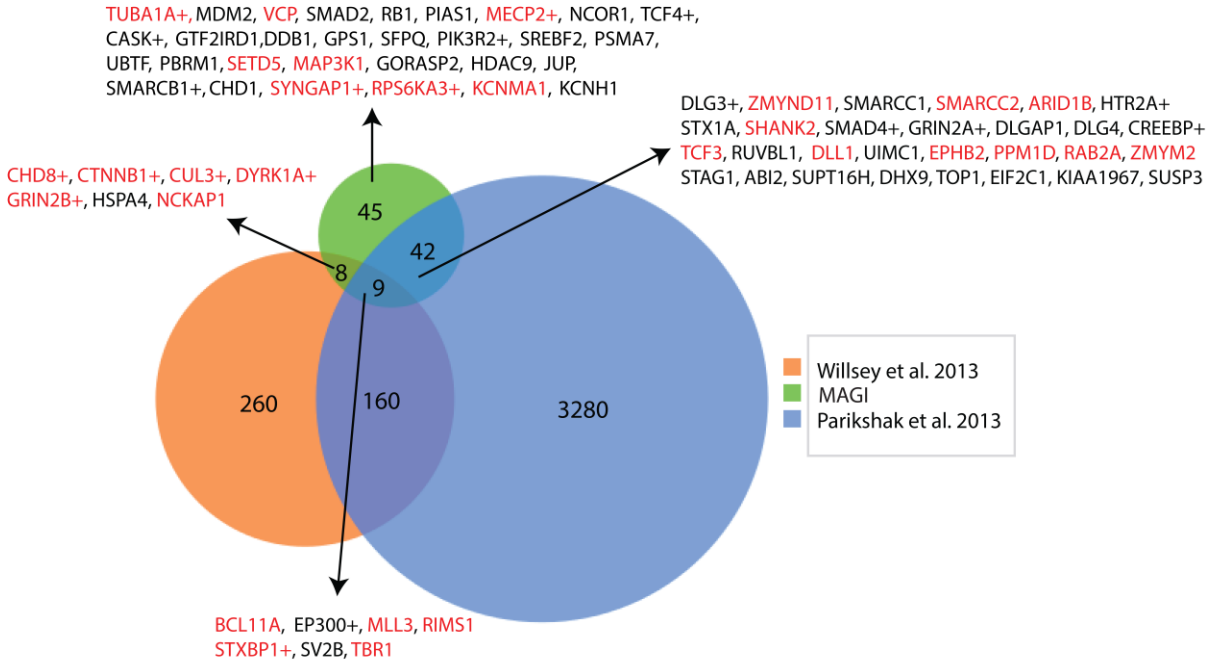
Supplementary Table 2. Overlap of *M1_Extended* and M2_Extended with modules previously published (Cristino et al., 2014; Gilman et al., 2011; Parikshak et al., 2013; Willsey et al., 2013). The table is sorted based on the overlap among the modules. Three values are presented for every two modules compared: (i) |module1-module2|, i.e., the number of genes that are found in module1 but are not found in module2; (ii) |module1 & module2 |, i.e., the number of genes found in the intersection of module1 and mofule2; (iii) |module2-module1|, i.e., the number of genes that are found in module2 but are not found in module1. Only modules for which the intersection with *M1_Extended* or *M2_Extended* is not empty are shown. The table is sorted by the fraction of overlap from high to low.

module1 (MAGI)	module2	|module1 - module2|	|module1 & module2|	|module2 - module1|
M2_Extended	DAWN_C6	19	5	18
M1_Extended	SteinerNet	35	45	403
M2_Extended	DAWN_C1	21	3	29
M1_Extended	DAWN_C1	74	6	26
M2_Extended	AXAS_M24	22	2	19
M1_Extended	DAWN_C4	76	4	12
M1_Extended	Willsey_Per3_5	71	9	154
M1_Extended	AXAS_M1	53	27	673
M1_Extended	AXAS_M3	65	15	380
M1_Extended	DAWN_C2	77	3	16
M2_Extended	AXAS_M11	20	4	109
M1_Extended	Willsey_Per4_6	73	7	160
M1_Extended	Parikshak_M3	55	25	825
M1_Extended	Willsey_Per8_10	73	7	181
M2_Extended	SteinerNet	13	11	437
M1_Extended	DAWN_C7	78	2	11
M1_Extended	DAWN_C5	78	2	13
M2_Extended	NETBAG	22	2	70
M2_Extended	Parikshak_M16	15	9	449
M2_Extended	AXAS_M2	11	13	667
M1_Extended	AXAS_M13	77	3	95

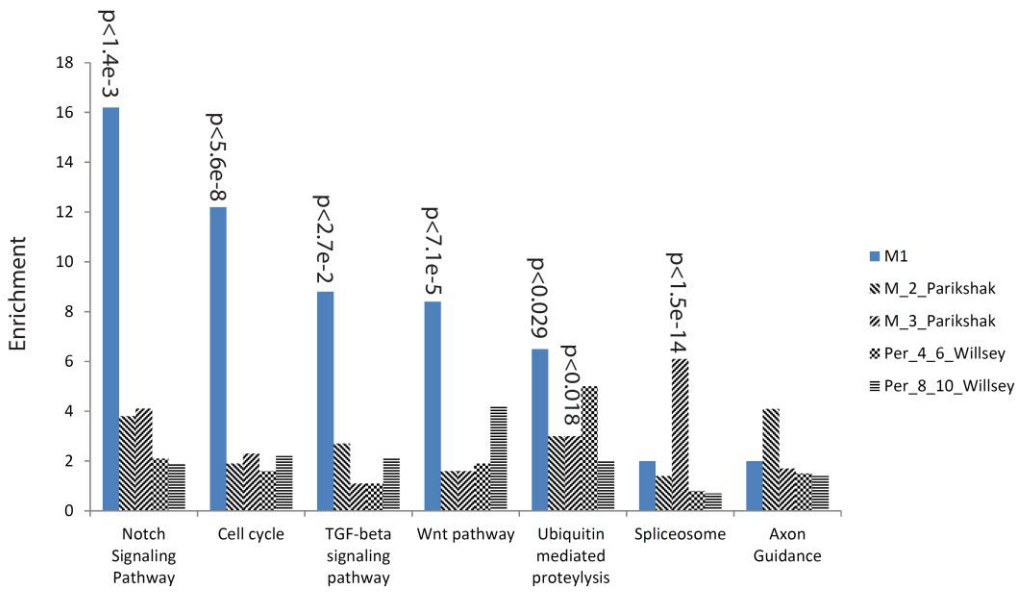
M2_Extended	Willsey_Per3_5	21	3	160
M2_Extended	Willsey_Per4_6	21	3	164
M1_Extended	AXAS_M4	75	5	265
M1_Extended	AXAS_M16	78	2	60
M1_Extended	AXAS_M8	77	3	151
M1_Extended	AXAS_M15	78	2	80
M1_Extended	DAWN_C6	79	1	22
M2_Extended	Parikshak_M13	17	7	697
M2_Extended	Willsey_Per8_10	22	2	186
M1_Extended	AXAS_M2	73	7	673
M1_Extended	Parikshak_M2	73	7	681
M2_Extended	AXAS_M12	23	1	110
M2_Extended	AXAS_M10	23	1	113
M1_Extended	AXAS_M19	79	1	57
M1_Extended	NETBAG	79	1	71
M1_Extended	AXAS_M14	79	1	81
M1_Extended	AXAS_M9	79	1	118
M1_Extended	AXAS_M7	79	1	154
M1_Extended	Parikshak_M16	78	2	456
M1_Extended	AXAS_M6	79	1	197
M1_Extended	AXAS_M5	79	1	198
M2_Extended	AXAS_M1	23	1	699
M2_Extended	Parikshak_M17	23	1	801
M1_Extended	Parikshak_M17	79	1	801

1) Comparison of *M1_Extended* and *M2_Extended* to recently published co-expression based networks for ASD (Parikshak et al., 2013; Willsey et al., 2013)

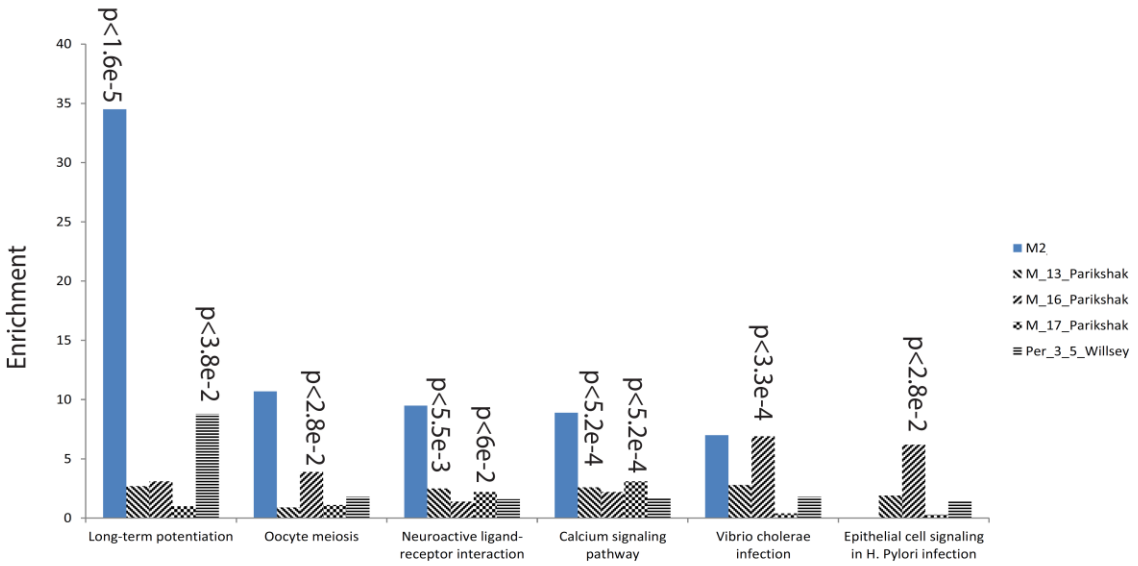
Two recent papers by Parikshak et al. (2013) and Willsey et al. (2013) provided modules using co-expression networks from the BrainSpan Atlas and the same set of *de novo* mutations from ASD studies. Since each method satisfies a different set of conditions, comparing them directly is quite challenging. Here we compare the significant modules reported in these publications (modules M2, M3, M13, M16 and M17 for Parikshak et al. (2013), and modules Per_4_6, Per_8_10 and Per_3_5 for Willsey et al. (2013)) with our *M1_Extended* and *M2_Extended* based on three different categories: 1) overlap with the other approaches in terms of genes included in the modules (Supplementary Figure 11); 2) enrichment in known KEGG pathways using the DAVID tool (Supplementary Figures Supplementary Figure 12and **Error! Reference source not found.**); and 3) enrichment in genes associated with neurodevelopmental diseases (Supplementary Figures Supplementary Figure 14 and Supplementary Figure 15), based on OMIM (www.omim.org) or AISS (Krumm et al., 2013).



Supplementary Figure 11. Venn diagram of the total number of genes reported as part MAGI's modules (green) compared to the ones found in the significant modules for ASD reported by Willsey et al. (2013) (orange), and Parikshak et al. (2013) (blue). Gene names are reported for genes that carry *de novo* mutations in probands or genes that are associated with neurodevelopmental diseases (based on OMIM). Genes with *de novo* truncating mutations are colored in red. Genes that are associated with neurodevelopmental diseases are marked with "+".

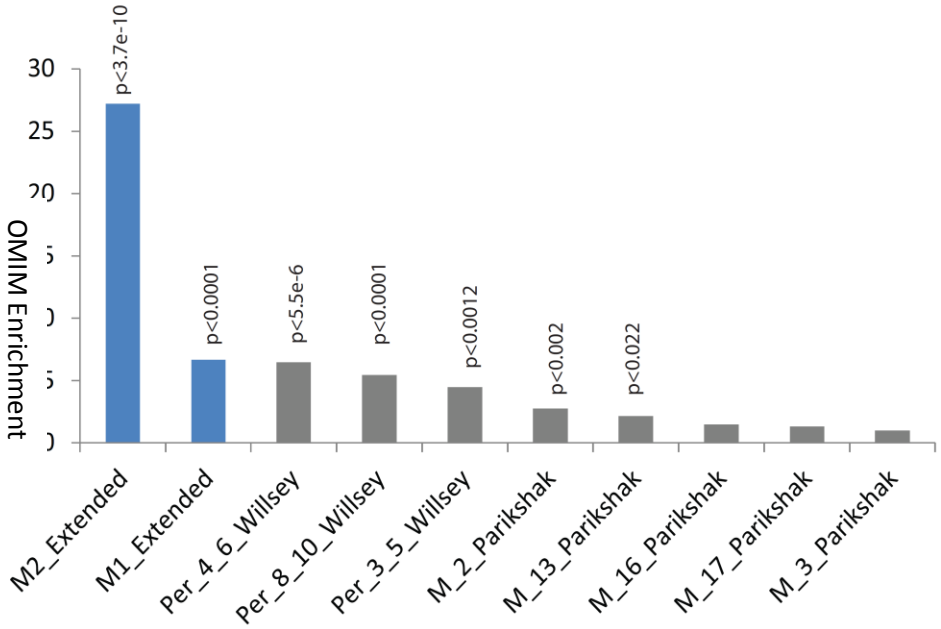


Supplementary Figure 12. The enrichment of KEGG pathways (using the DAVID tool) for *M1_Extended* and modules previously reported (Parikshak et al., 2013; Willsey et al., 2013) that show expression patterns during brain development similar to *M1_Extended*. Only pathways significant (*p<0.05)* in at least one of the modules predicted in any of the three studies are shown. Cancer-related pathways are not shown. P-values are presented for *p<0.05* (after Bonferroni correction).

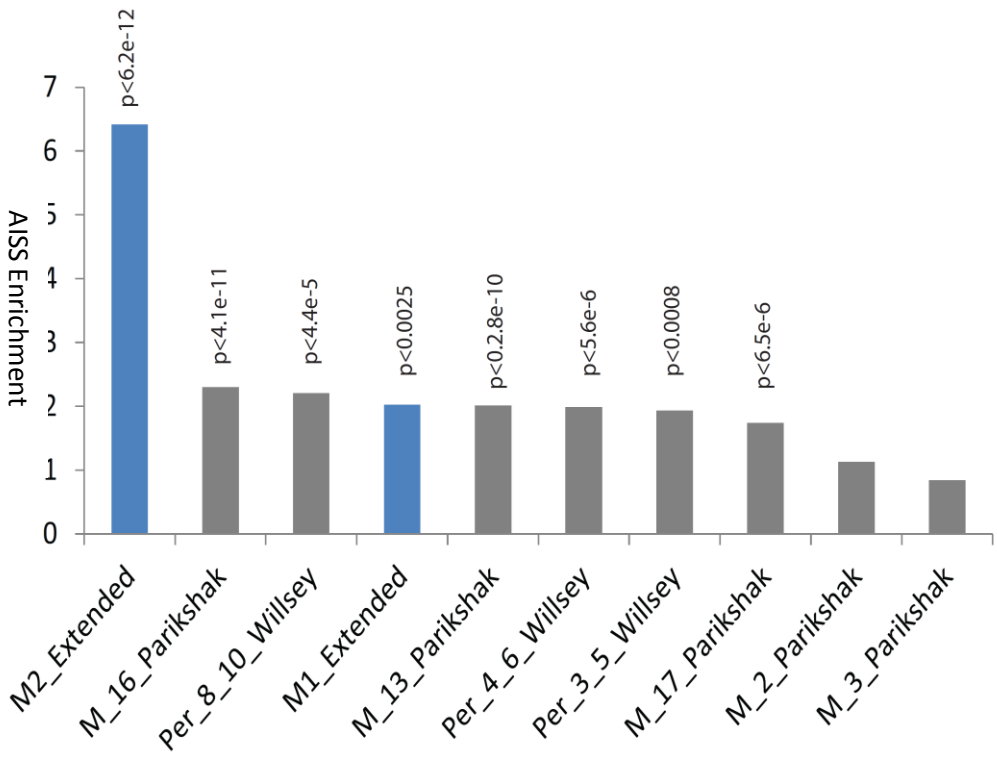


Supplementary Figure 13. The enrichment of KEGG pathways (using the DAVID tool) for *M2_Extended* and modules previously reported (Parikshak et al., 2013; Willsey et al., 2013) that show expression patterns during brain development similar to *M2_Extended*. Only pathways significant (*p<0.05)* in at least one of the modules predicted in any of the three studies are shown. Cancer-related pathways are not shown. P-values are presented for *p<0.05* (after Bonferroni correction). For *M2_Extended* the most significant pathway is the long-term potentiation pathway (known to be related to passing signaling between neurons).

We compared the enrichment of *M1_Extended* and *M2_Extended* in genes associated with neurodevelopment diseases (OMIM and AISS) to that of the significant modules reported in the other two studies (Parikshak et al., 2013; Willsey et al., 2013). For OMIM, we used the set of genes that are found in the PPI network and are annotated in the OMIM dataset with the explicit terms: "autism", "mental retardation", "development delay", "macrocephaly" or "microcephaly". For AISS, we used the union of the genes with known neurodevelopment phenotype (Krumm et al., 2013) and the OMIM set. The p-value provided is using Fisher's exact test. Only *p*-values lower than 0.05 are reported (after Bonferroni correction).



Supplementary Figure 14. Enrichment of different modules in genes in the OMIM set. P-values are presented for $p<0.05$.



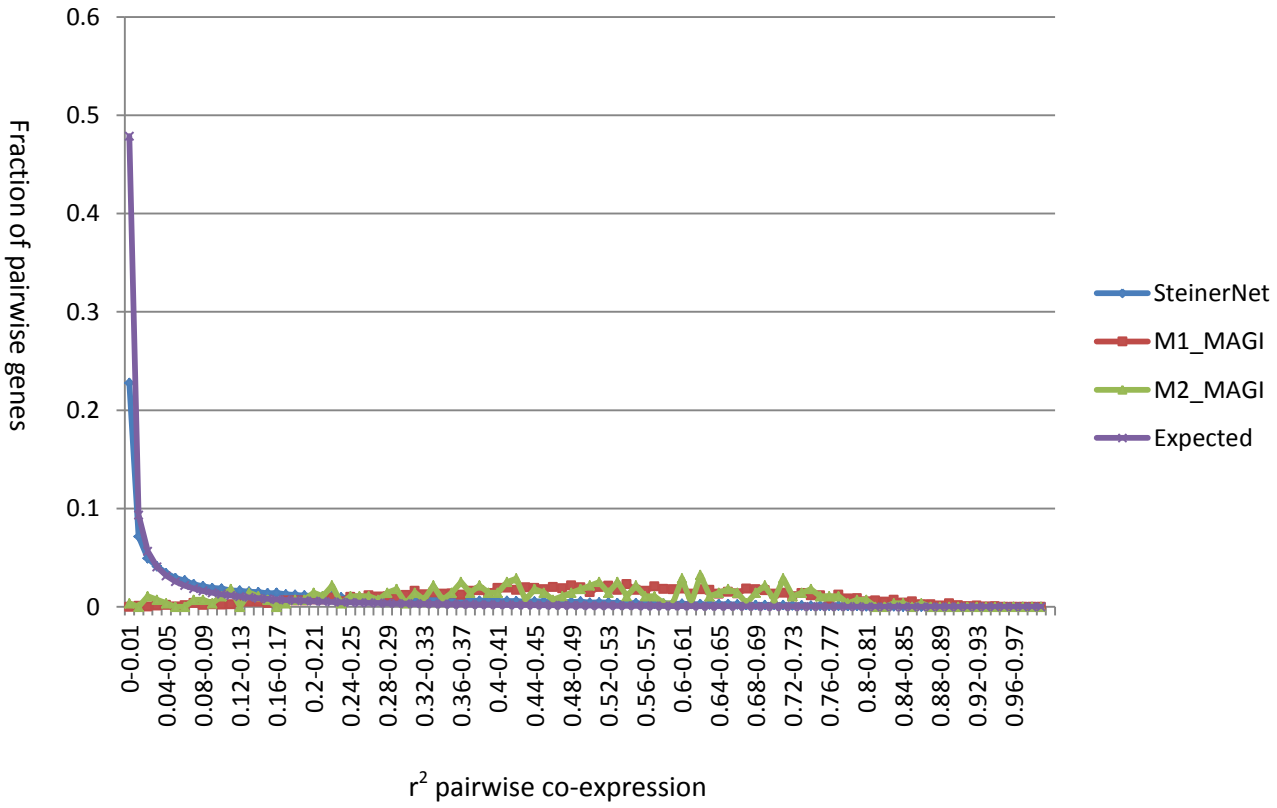
Supplementary Figure 15. Enrichment of different modules in genes in the AISS set. P-values are presented for *p<0.05*.

2) Comparison of *M1_Extended* and *M2_Extended* to other networks for ASD (AXAS, SteinerNet, NETBAG, and DAWN)

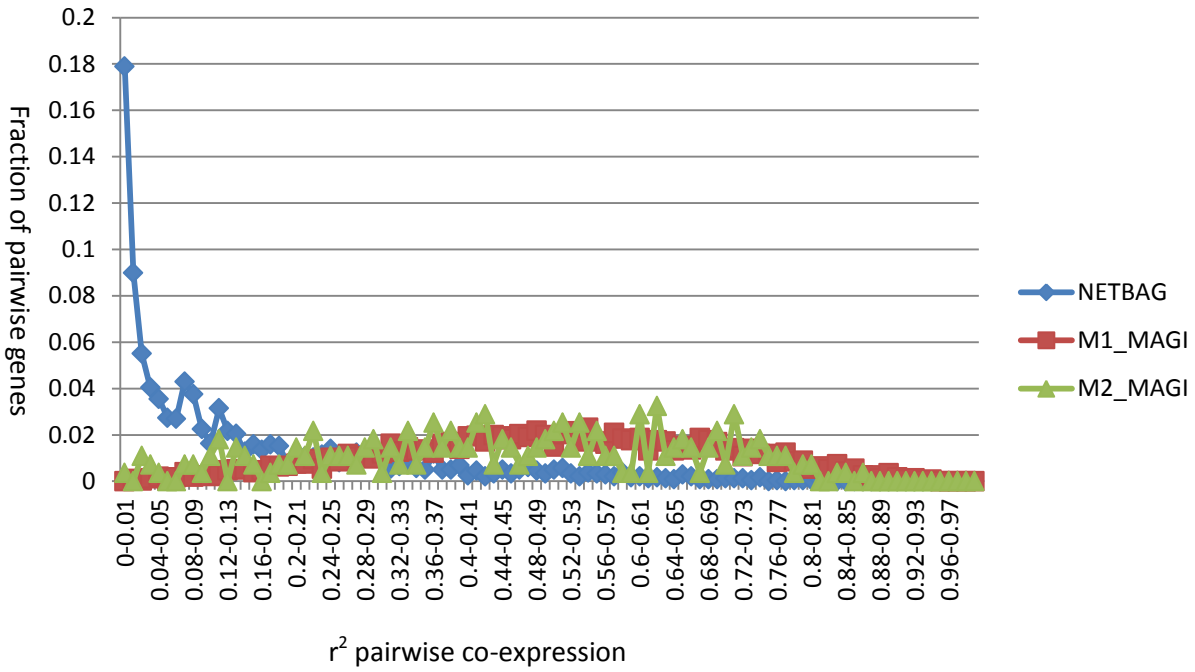
For AXAS comparisons, we obtained 30 modules from the author of AXAS, from which we considered 26 modules that include at least 10 genes. For the Steiner tree comparisons, we have applied the tool SteinerNet (http://fraenkel.csbi.mit.edu/steinernet/) to construct the Steiner tree using MAGI's gene scores for each node and PPI edges that were all given the same score (1.0). The resulted connected component contained 448 nodes (genes). Interestingly, *M1_Extended* and *M2_Extended* overlap more significantly with the SteinerNet module compared to random modules (connected components in PPI) of the same size ($p<0.001$ and $p<0.01$, respectively). For NETBAG comparisons, we used the autism module reported in the NETBAG manuscript (Gilman et al., 2011). This module includes 72 genes. For DAWN comparisons, we used the seven modules reported in Figure 4 of the DAWN manuscript (Liu et al., 2014).

The SteinerNet tool was run on the PPI network only and therefore the resulting module was not required to show high gene co-expression. Similarly, NETBAG and the AXAS modules did not take into account co-expression information. Here, we show that these modules have a higher fraction of genes that show no co-expression during brain developed compared to MAGI's modules.

We compared the pairwise distribution of gene co-expression values for different modules. A substantial number of gene pairs (>10%) in the AXAS, SteinerNet and NETBAG modules were found to have a very low co-expression value (r^2<0.01). In contrast, almost none of the gene pairs found in the reported MAGI's module have such a low co-expression. Finally most modules by DAWN (except 2) have over 10% of their gene pairs with very low co-expression value.



Supplementary Figure 16. A histogram-like distribution of pairwise co-expression values (r^2) between every two genes in the SteinerNet module (blue), *M1_Extended* (red), and *M2_Extended* (green), respectively. Modules detected by MAGI do not have any pair of genes with r^2<0.01, while the SteinerNet module has around 23% of such pairs. For comparison, the distribution of the pairwise co-expression values of all the genes is shown (purple).



Supplementary Figure 17. A histogram-like distribution of pairwise co-expression values (r^2) between every two genes in the NETBAG module (blue), *M1_Extended* (red) and *M2_Extended* (green) modules, respectively. Modules detected by MAGI do not have any pair of genes with $r^2<0.01$, while the NETBAG module has around 18% of such pairs.

Finally for every AXAS module, over 15% of gene pairs were found to have a co-expression value lower than 0.01. The exact percentage of gene pairs with such a low co-expression value is provided in supplementary table 3.

We also compared the average number of truncating mutations per gene seen in normal population for each module, using the Exome Sequencing Project (ESP). We expect that disease modules would be depleted for genes for which truncating mutations are common in the general population. On average, genes in the union of *M1_Extended* and *M2_Extended* have 0.17 truncating mutations in the ESP. In contrast, from the genes in the NETBAG module and the SteinerNet module have on average 1.87, and 1.96 truncating mutations, respectively. The average number of truncating mutations in ESP for each of the AXAS modules is reported in supplementary table 3.

Supplementary Table 3. Comparison to the AXAS networks in terms of co-expression and truncated mutations found in the ESP dataset.

	Fraction of gene pairs with co-expression r^2<0.01	Avg # truncating mutations per gene in ESP
M1_AXAS	0.28	1.152857143
M2_AXAS	0.3	1.482352941
M3_AXAS	0.17	1.420253165
M4_AXAS	0.23	1.637037037
M5_AXAS	0.3	1.59798995
M6_AXAS	0.47	2.181818182
M7_AXAS	0.27	1.425806452
M8_AXAS	0.26	1.824675325
M9_AXAS	0.29	1.394957983
M10_AXAS	0.24	1.631578947
M11_AXAS	0.26	1.876106195
M12_AXAS	0.36	1.189189189
M13_AXAS	0.29	1.510204082
M14_AXAS	0.15	2.329268293
M15_AXAS	0.27	1.707317073
M16_AXAS	0.22	1.85483871
M17_AXAS	0.34	2.803278689
M18_AXAS	0.18	2.254237288
M19_AXAS	0.12	1.120689655
M20_AXAS	0.29	0.98
M21_AXAS	0.22	1.931818182
M22_AXAS	0.27	1.666666667
M23_AXAS	0.18	1.647058824
M24_AXAS	0.3	1.80952381
M25_AXAS	0.64	1.3
M26_AXAS	0.37	0.384615385
M1_MAGI	0	0.1625
M2_MAGI	0	0.208333333

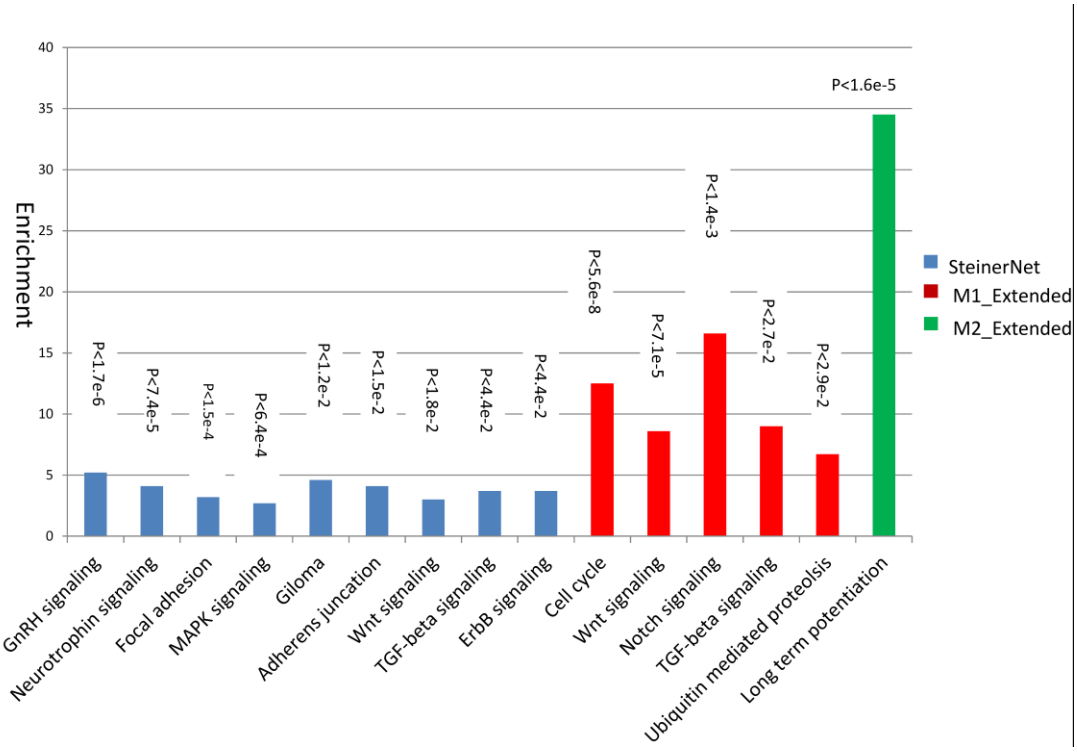
The fraction of genes pairs with average co-expression lower than 0.01 and the total number of truncating mutations found in the ESP for the DAWN modules are given in supplementary table 4.

.

Supplementary Table 4. The average LoF mutations in ESP per gene for modules reported in DAWN and fractionation of pairwise genes with co-expression r^2<0.01.

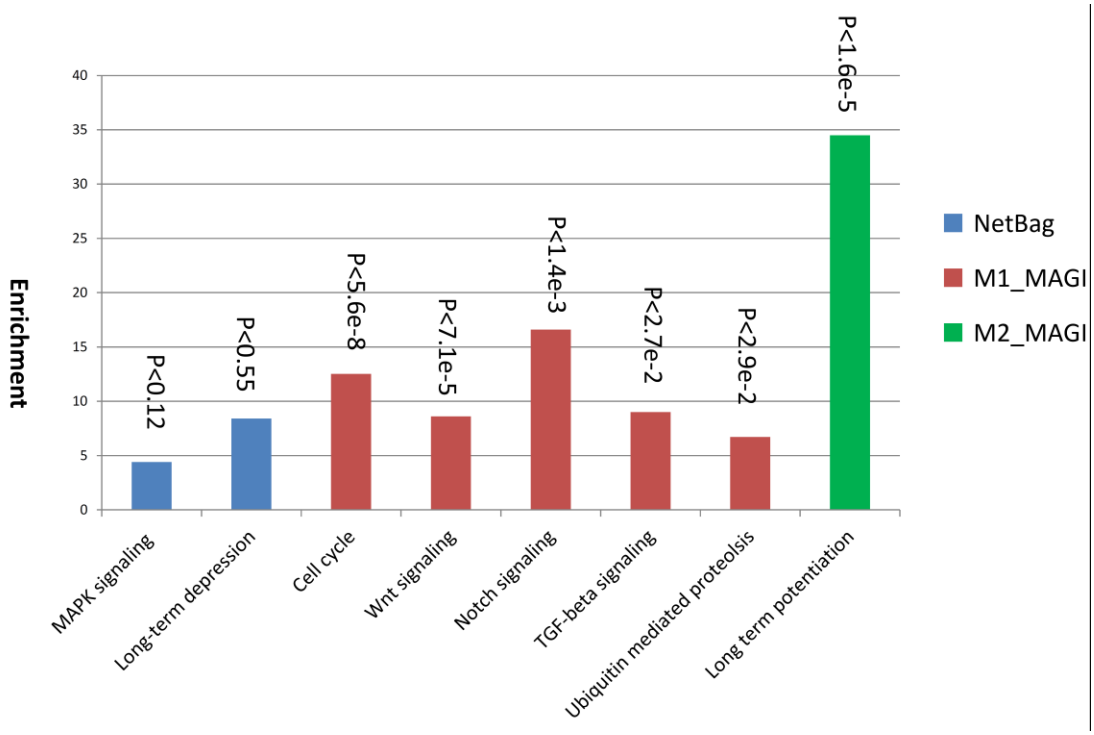
	Fraction of gene pairs with co-expression r^2<0.01	Avg # truncating mutations per gene in ESP
C1	0.185	0.9
C2	0.09	2
C3	0.22	1.4
C4	0.12	1
C5	0.35	0.6
C6	0.2	0.6
C7	0.025	1
M1_MAGI	0	0.1625
M2_MAGI	0	0.208333333

Comparison of KEGG pathways enrichment for *M1_Extended*, *M2_Extended*, NETBAG, SteinerNet, and the AXAS modules with the highest overlap with MAGI's modules (M1_AXAS, and M2_AXAS) is shown in Supplementary Figures Supplementary Figure 18, Supplementary Figure 19, Supplementary Figure 20, and Supplementary Figure 21.

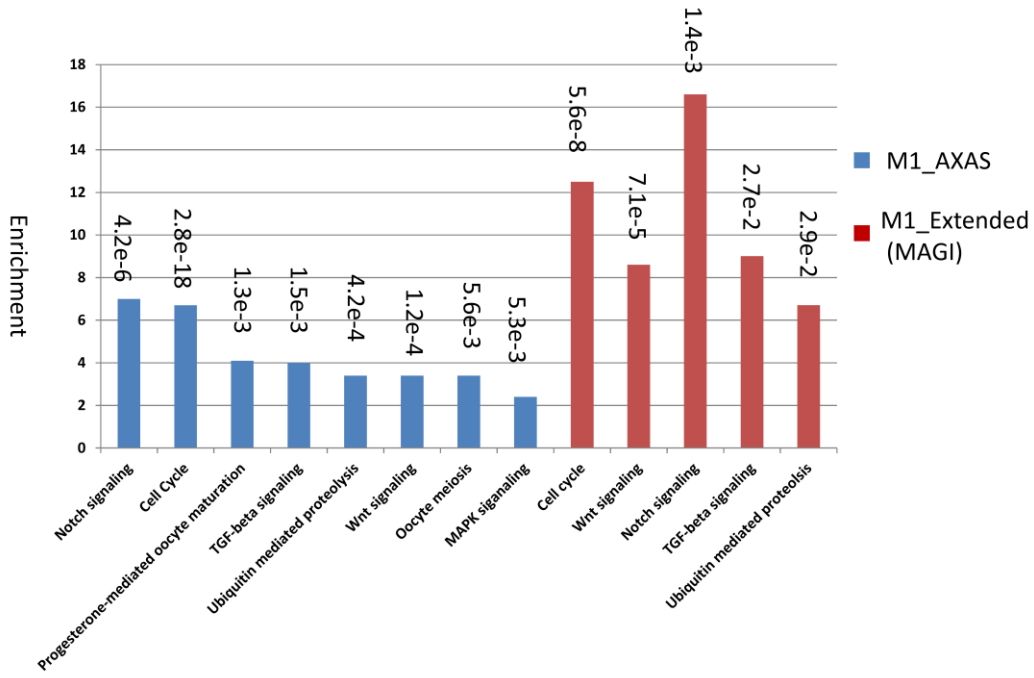


Supplementary Figure 18. The enrichment for significant KEGG pathways after Bonferroni correction (p<0.05) for MAGI's modules and the SteinerNet module. Modules reported by MAGI are found to have higher enrichment to known pathways. The two common pathways between *M1_Extended* and SteinerNet are Wnt and TGF-beta

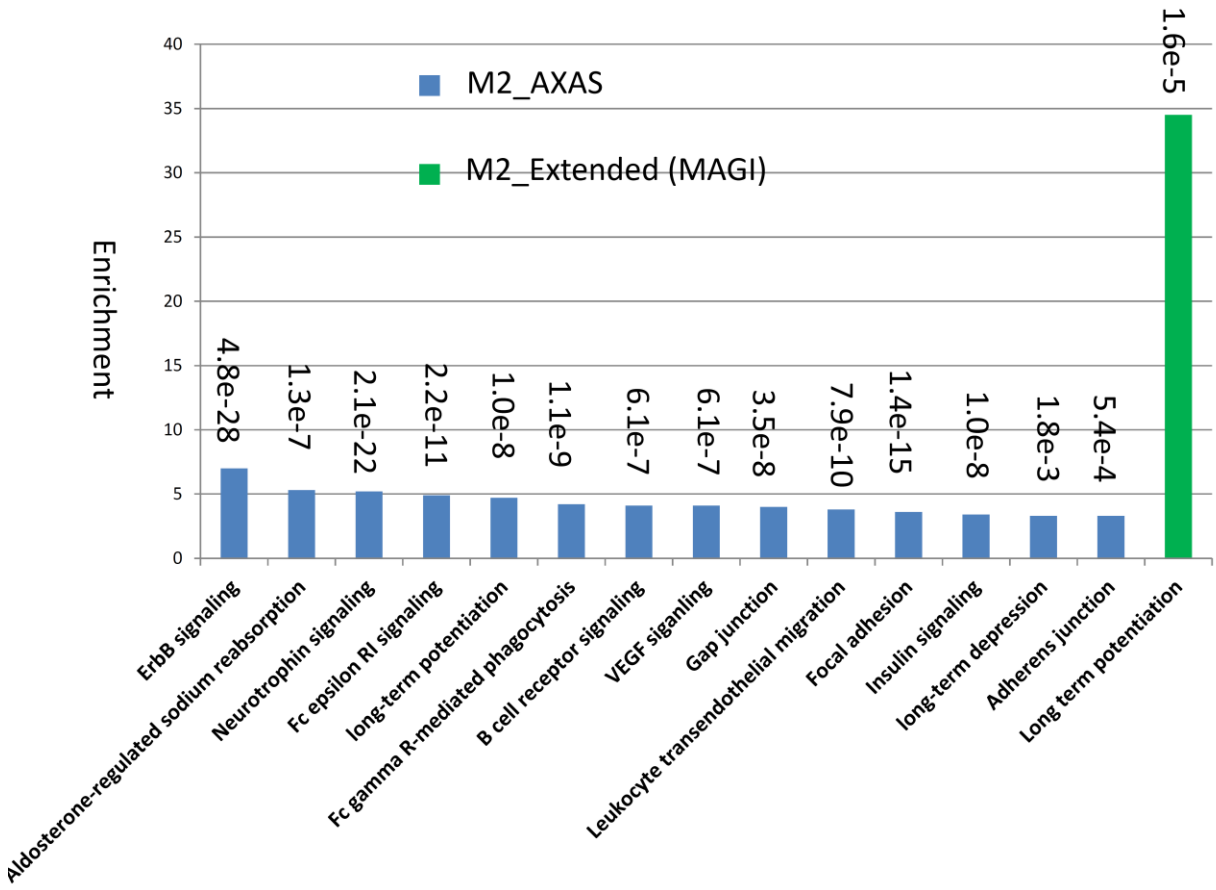
signaling, for which the MAGI module *M1_Extended* is found to have higher enrichment with lower p-value compared to the SteinerNet module. Cancer-related pathways are not shown.



Supplementary Figure 19. The enrichment for significant pathways after Bonferroni correction ($p<0.05$) from KEGG database for MAGI's modules and the NETBAG module. The NETBAG module was not found to be significantly enriched in any of the known KEGG pathways. The two KEGG pathways that are presented for NETBAG are the ones with the lowest P-values. Cancer-related pathways are not shown.

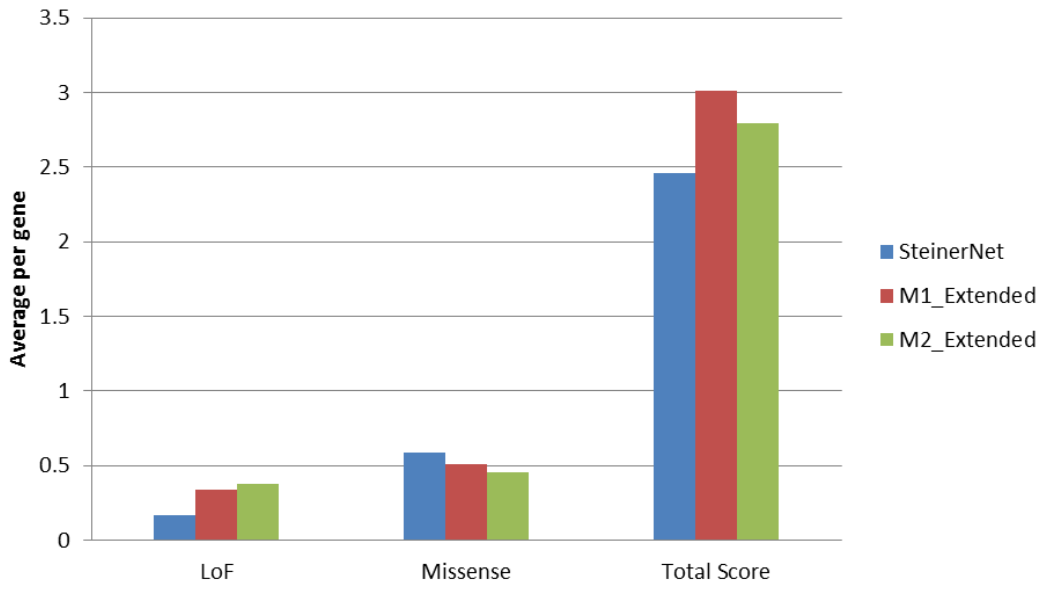


Supplementary Figure 20. Enrichment of significant KEGG pathways ($p<0.05$ after Bonferroni correction) for *M1_Extended* (red) and the first AXAS module (blue), M1_AXAS, which has the highest number of shared genes with *M1_Extended*. Cancer-related pathways are not shown.



Supplementary Figure 21. Enrichment of significant KEGG pathways ($p<0.05$ after Bonferroni correction) for *M2_Extended* (green) and the first AXAS module (blue), M2_AXAS, which has the highest number of shared genes with *M2_Extended*. Cancer-related pathways are not shown.

We also compared the average number of LoF and missense mutations per gene, as well as the average gene score, between the SteinerNet module and the *M1_Extended* and *M2_Extended* modules. *M1_Extended* and *M2_Extended* were found to be more enriched in LoF mutations and have a higher average gene score compared to the SteinerNet module (Supplementary Figure 22).



Supplementary Figure 22. A comparison of the average number of LoF mutations and missense mutations per gene, as well as the average gene score (i.e., summation of score of all genes in each module divided by size of the module), is shown for the SteinerNet module and the *M1_Extended* and *M2_Extended* modules.

Enrichment for KEGG pathways of *M1_Extended* and *M2_Extended*

To calculate the confidence intervals for pathway enrichment for both *M1_Extended* and *M2_Extended* we used Fisher's exact test to calculate the *p*-value, confidence intervals and odds ratio (OR), as reported below:

Supplementary Table 5. OR and confidence intervals for the significant KEGG pathways found for *M1_Extended*.

M1_Extended	p-value	OR	Confidence interval
Cell cycle	6.50E-011	19.3	[8.6, 40.6]
Wnt signaling pathway	1.20E-007	11.9	[5.09, 25.8]
Notch signaling pathway	1.10E-006	22.1	[7.1, 57.4]
TGF-beta signaling pathway	4.40E-005	11.1	[3.7, 27.8]
Ubiquitin mediated proteolysis	6.60E+005	8.2	[3.01, 19.4]
Adherens junction	0.00027	10.1	[3, 27.1]
Jak-STAT signaling pathway	0.0061	4.7	[1.4, 12.5]

For *M2_Extended* the highest enrichment that was found was for the long-term potentiation pathway with $p<7.3e-9$, OR=69 and confidence interval between 18 and 246.

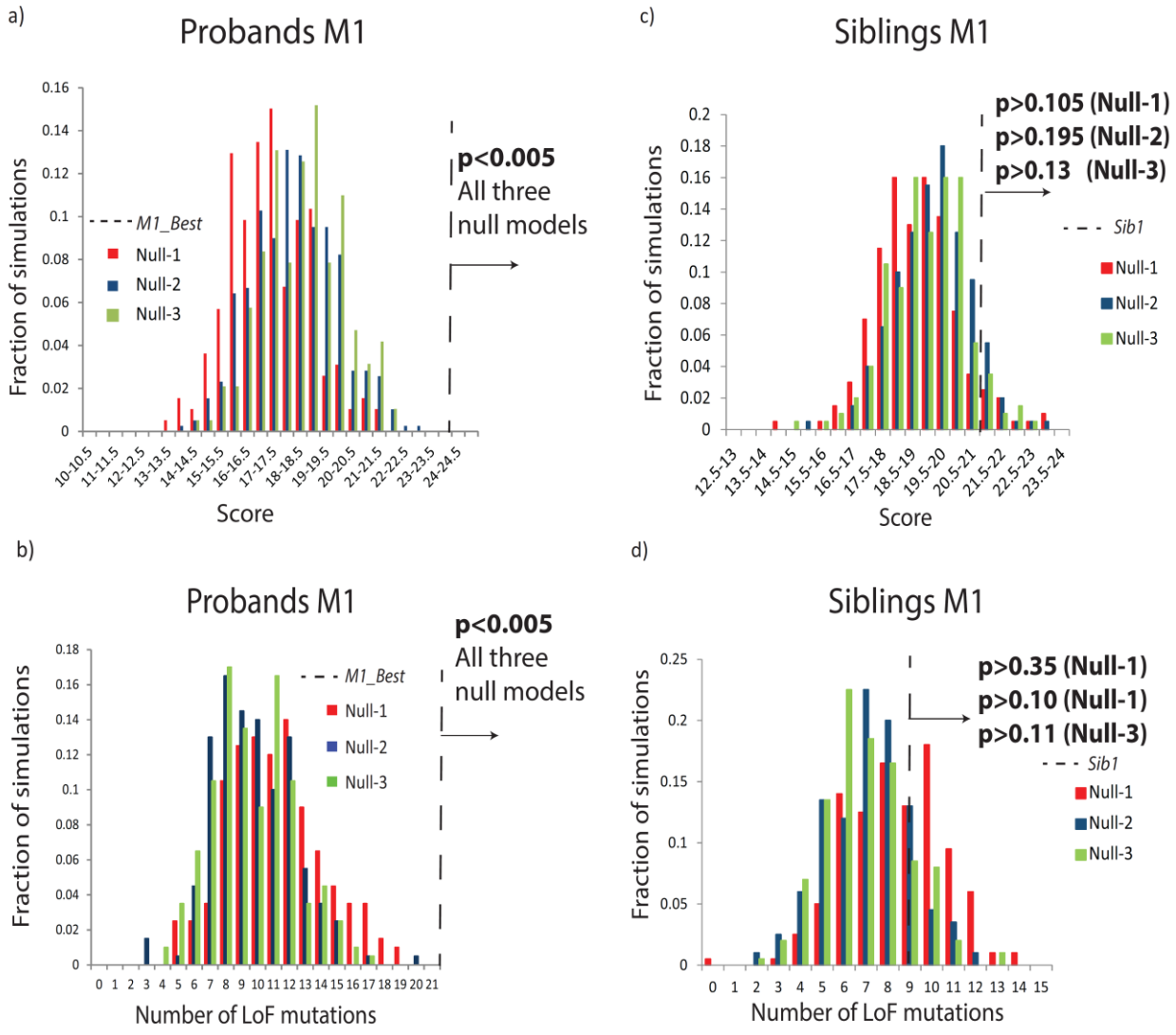
Significance of *M1_Best* and *M2_Best* vs. *Sib1_Best* and *Sib2_Best*

For calculating the significance of modules we used simulations in which we permuted the mutations based on three different null models (Null-1, Null-2 and Null-3). For Null-1 we distributed the mutations based on gene length (i.e., the probability of a mutation being assigned to a gene was directly correlated to its coding sequence length). For Null-2 and Null-3 the probability of assigning a mutation to a gene was based on observation of mutations in ESP dataset (Tennessen et al., 2012) or the whole genome as reported in (Kong et al., 2012). For Null-2 and Null-3 these probabilities were calculated as described in (O'Roak et al., 2012).

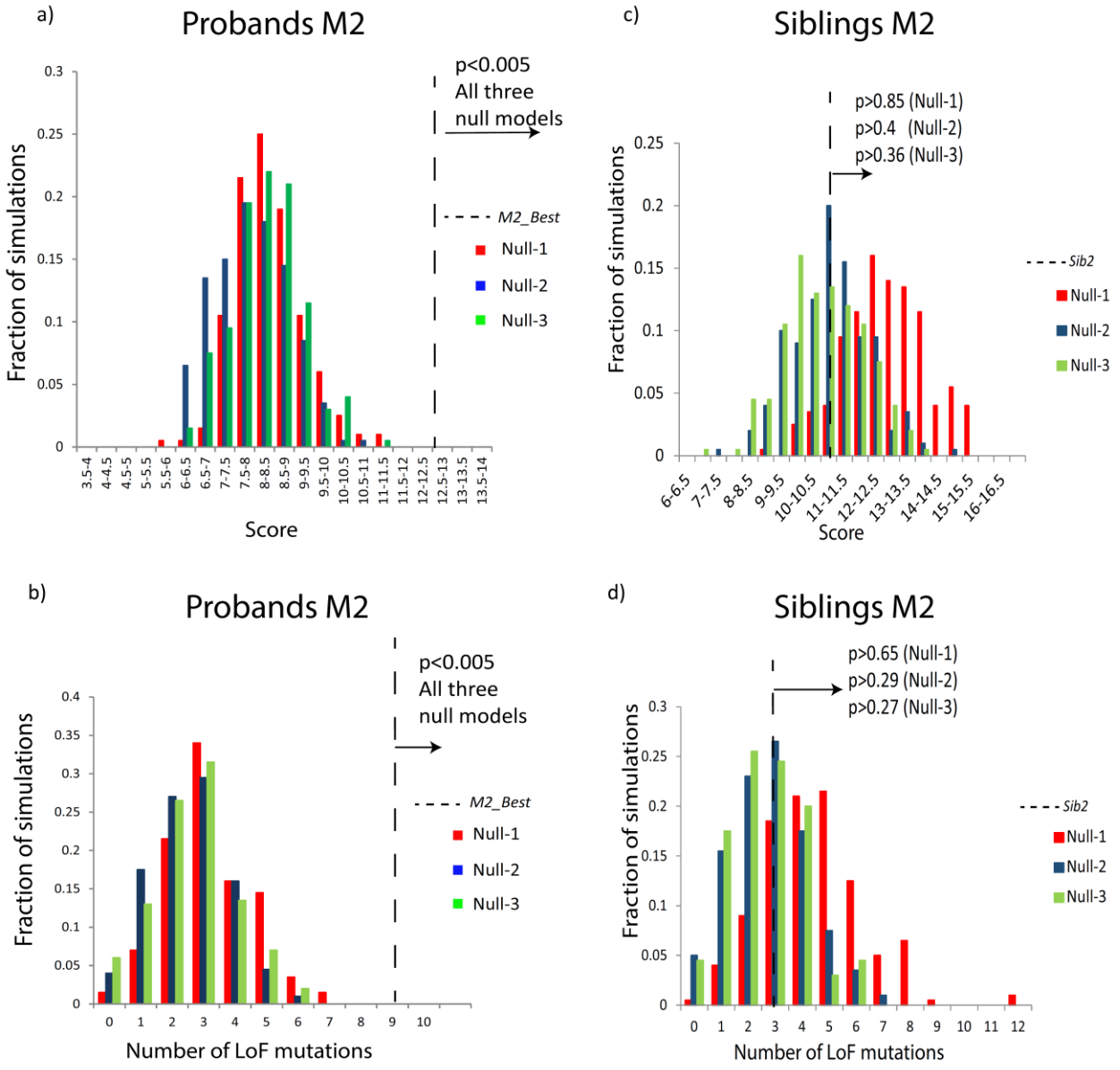
For Null-2 and Null-3 the distribution of the mutations was also aware of the degree distribution of genes with mutation in PPI network. To address the simulation, we adopted an approach similar to that reported in Jia et al. (2012). In each simulation, we shuffle the mutations from the real data based on the probabilities calculated for each null model, such that every mutation from the real data is assigned to a new gene. In addition, to account for the PPI degree, mutations are assigned only to genes for which the degree in the PPI is similar to that of the original gene carrying the real mutation. If the degree of the original gene is lower than 2^5, the new gene is required to have the exact same degree. For degrees higher or equal to 2^5, we consider three category of [2^5-2^6, 2^6-2^7, >2^7] and assign mutations only to genes that are in the same category as the original gene.

The reason for using these categories is that the higher the degree, the fewer the genes sharing that same degree in the PPI. This approach guarantees that the genes carrying the permutated mutations based on Null-2 and Null-3 have a similar degree distribution to that of the original data.

The *p*-values for probands and siblings using the three Null modules for M1 and M2 are shown below.

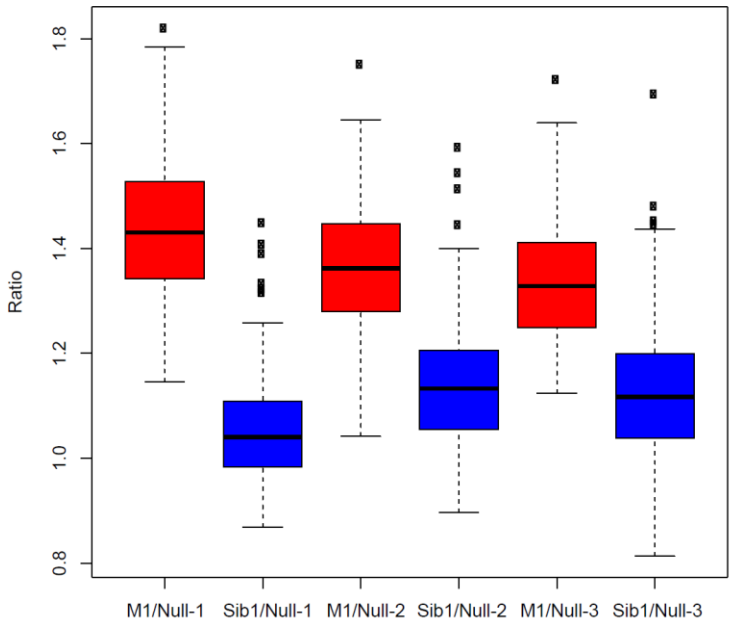


Supplementary Figure 23. Significance of M1. a) The *M1_Best* score (dashed black line) shown in comparison to the top-scored module of 200 simulations using three different null models (Null-1, Null-2, and Null-3). b) The number of LoF mutations covered by the top-scoring module (*M1_Best*) found using proband mutations versus the number of simulated LoF mutations covered by the top-scoring modules found under the same simulations. c) The score of the top module found using sibling and control mutations (dashed black line) in comparison to the top-scored module of 200 simulations with the same number of mutations using the three null models. The sibling simulations were performed without using the ESP constraint (although similar results were obtained when an ESP constraint was applied). d) The number of LoF mutations covered by the top-scoring module found using sibling mutations versus the number of simulated LoF mutations covered by the top-scoring modules found using the three null models. Sibling simulations were performed without filtering based on ESP controls.

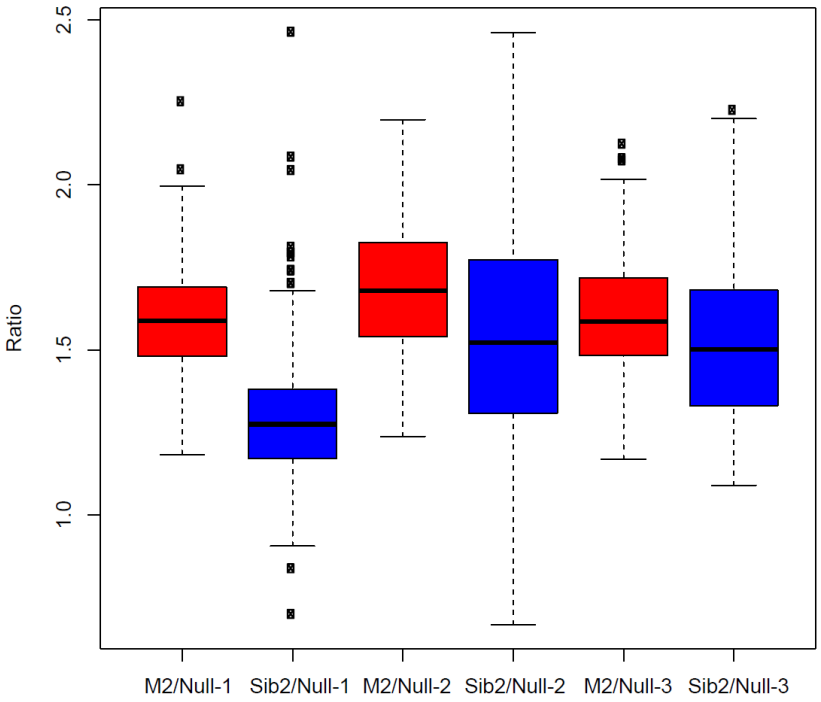


Supplementary Figure 24. Significance of M2. See legend of Supplementary Figure 23.

To compare the scores of the best modules found using probands data to the score of the best modules found using the three null models, we run 200 simulations for each null model. We then compared the ratio of the scores of *M1_Best* (or *M2_Best*) to the score of the best modules found on the simulation datasets models (Supplementary FiguresSupplementary Figure 25Supplementary Figure 26). We found that the *M1_Best* module has a score that is approximately 30% higher than null models. We repeated the analysis for the best modules found using sibling datasets. Clearly, the modules found using the probands mutations have significantly higher scores than null models, while the score of the modules found for siblings are quite similar to that of the null models. In all these analyses we used the exact same control threshold.



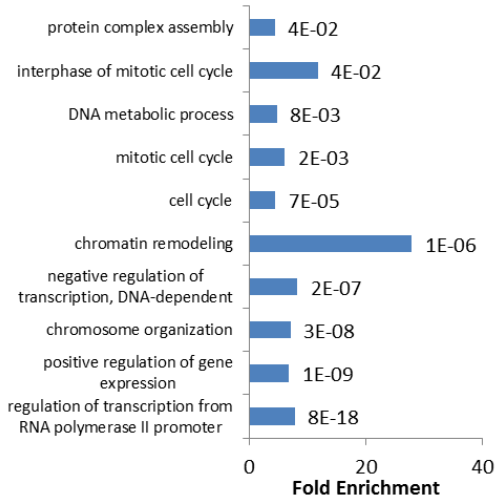
Supplementary Figure 25. The ratio of the score of the highest scoring module found using ASD+ID dataset (*M1_Best*) vs. the highest scoring modules found using Null1, Null2 and Null3 for 200 different simulations is shown as a red bar chart. The ratio of the highest scoring modules found using sibling data (Sib1) against Null1, Null2 and Null3 is depicted as the blue bars.



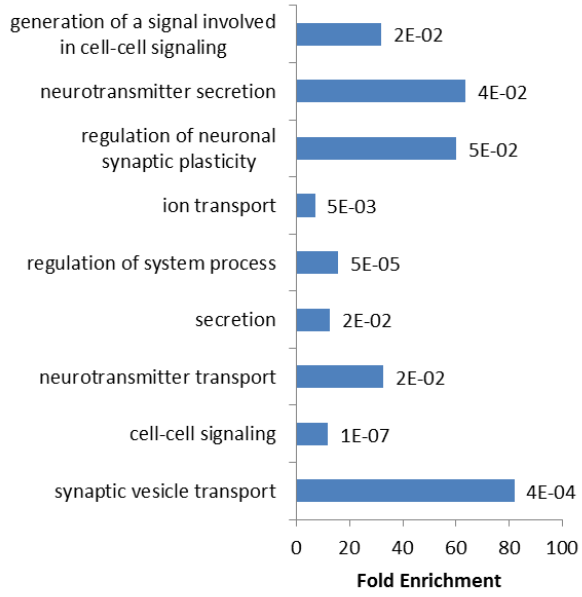
Supplementary Figure 26. The ratio of the score of the highest scoring second module found using ASD+ID dataset (*M2_Best*) vs. the highest scoring modules found using Null1, Null2 and Null3 for 200 different simulations is shown as a red bar chart. The ratio of the highest scoring second modules found using sibling data (Sibs) against Null1, Null2 and Null3 is depicted as blue bars.

Gene Ontology (GO)

GO terms with *p*-value<0.05 (with FDR correction) according to DAVID (Huang da et al., 2009) were given together with the enrichment fold as input to REVIGO (Supek et al., 2011). The summarized list is presented in Supplementary Figures Supplementary Figure 27 and Supplementary Figure 28.

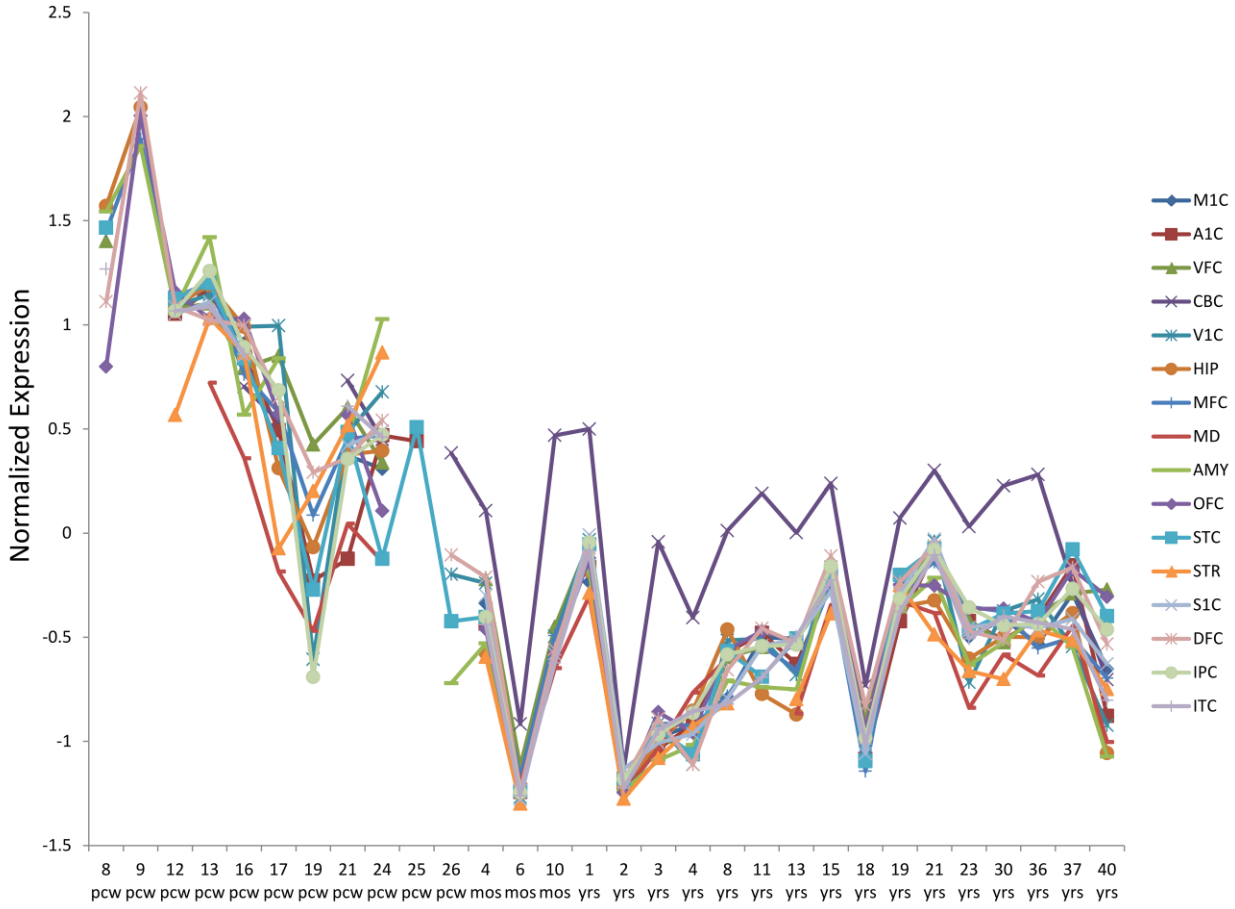


Supplementary Figure 27. GO terms enrichment in *M1_Extended*. Column labels show the p-values after FDR correction.

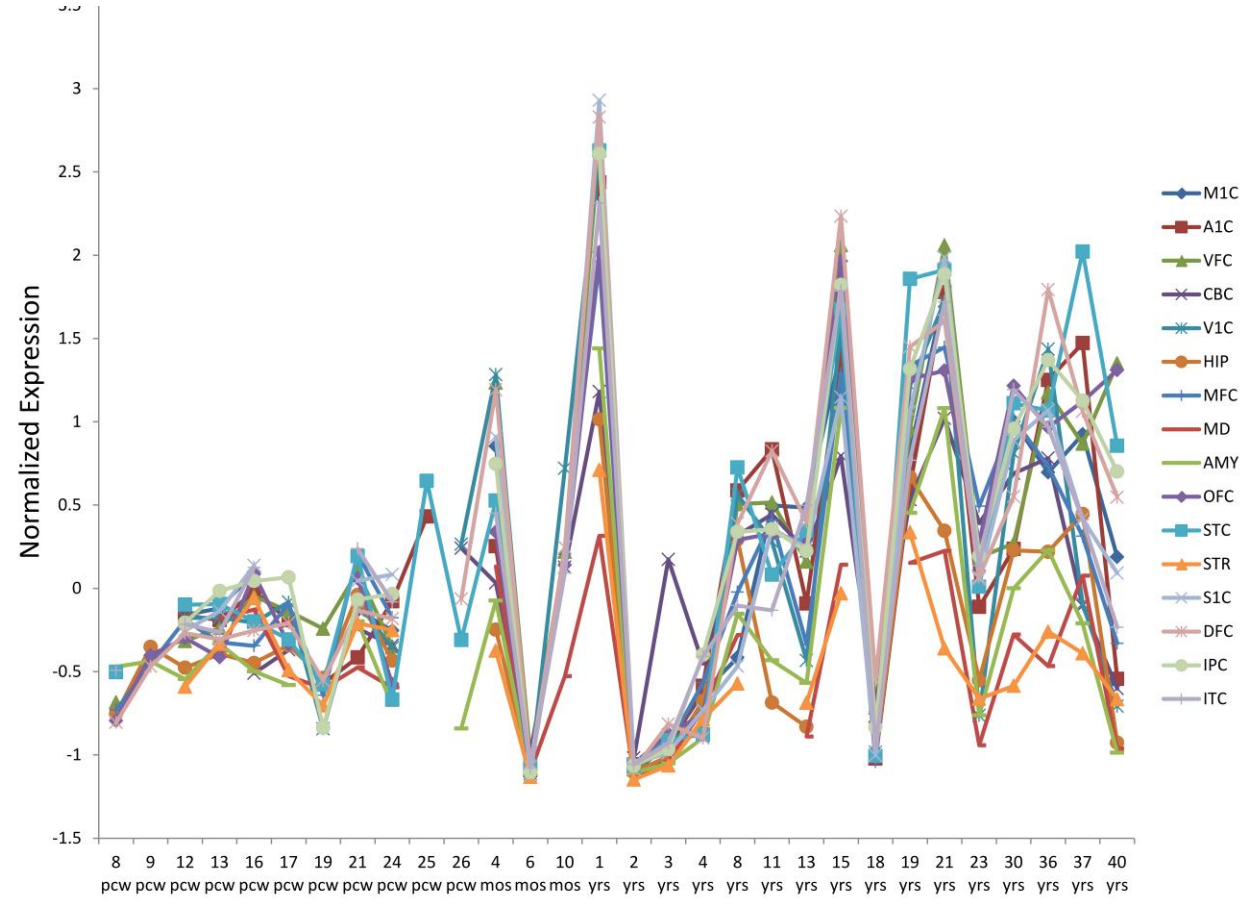


Supplementary Figure 28. GO terms enrichment in *M2_Extended*. Column labels show the *p*-values after FDR correction.

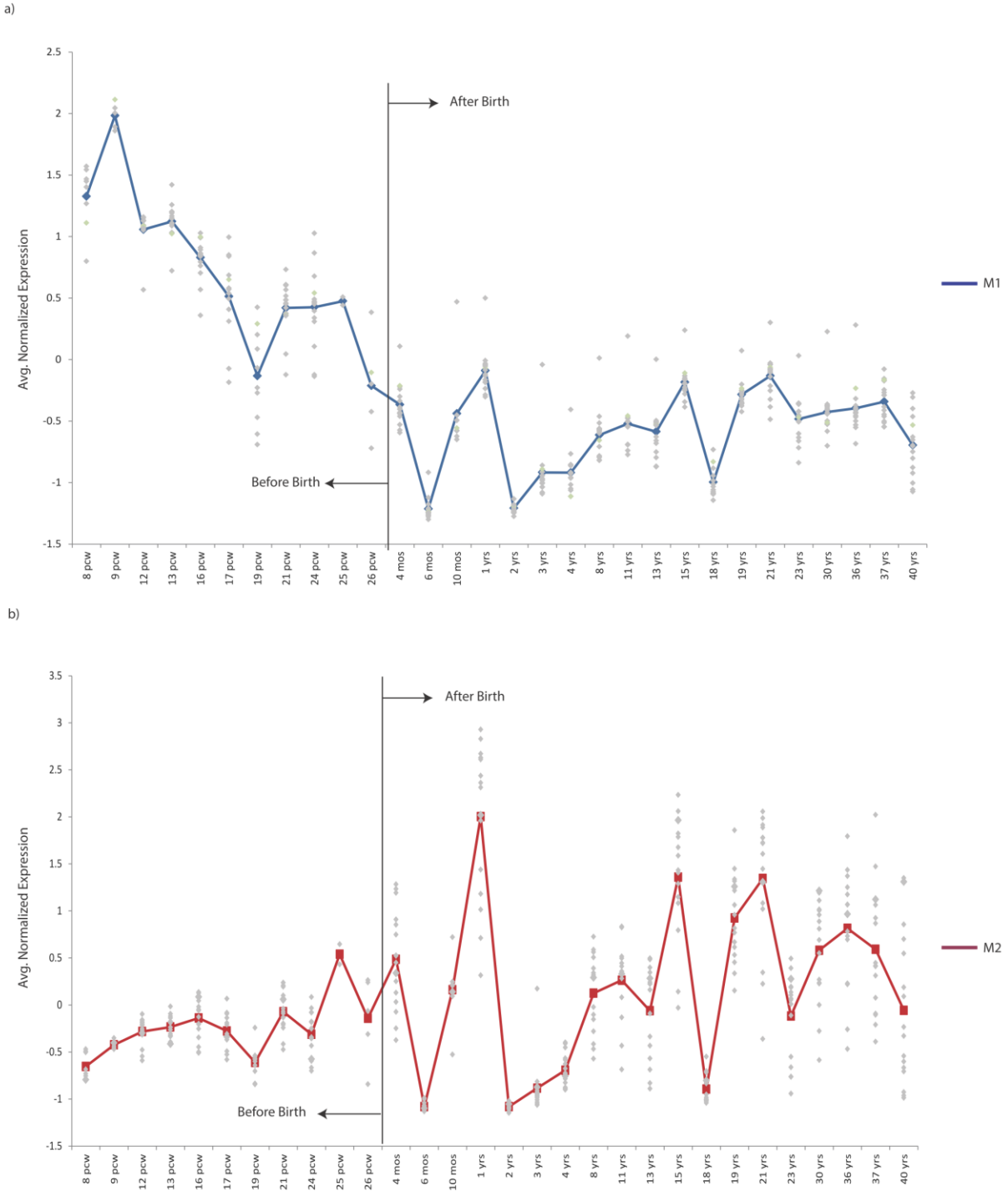
Gene expression of *M1_Extended* and *M2_Extended*



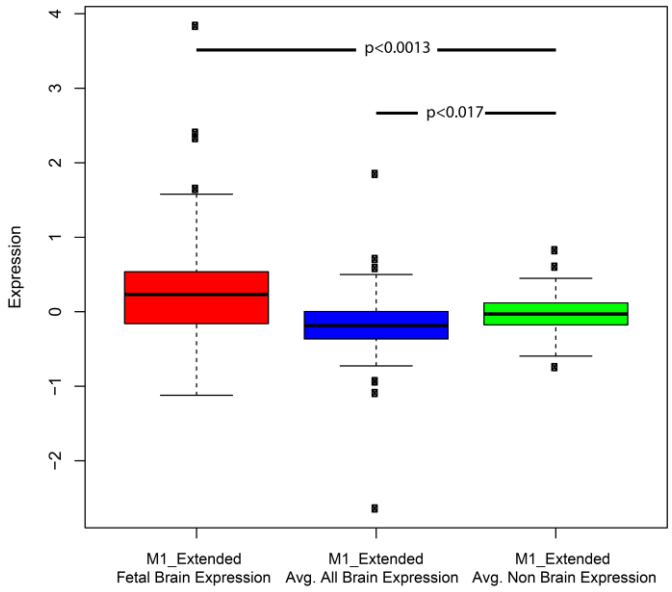
Supplementary Figure 29. Normalized expression of different brain subtissues for *M1_Extended* during brain development.



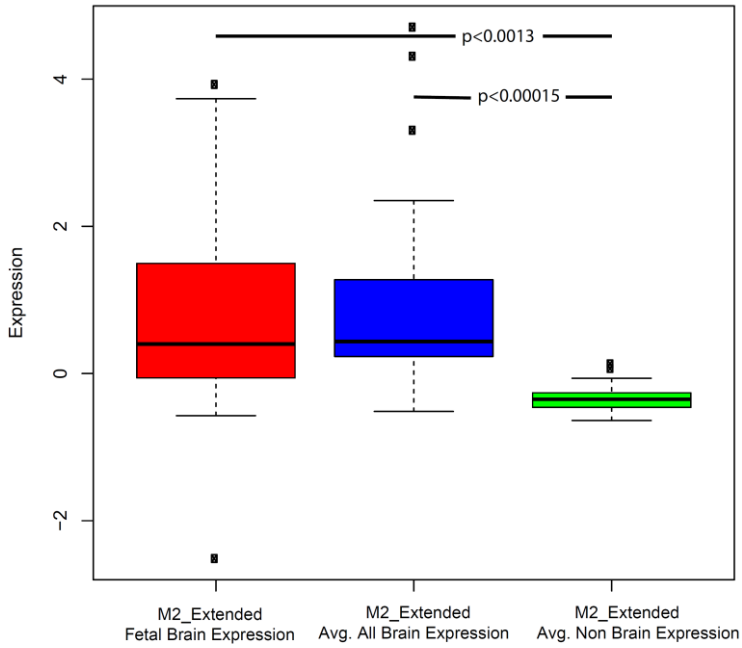
Supplementary Figure 30. Normalized expression of different brain subtissues for *M2_Extended* during brain development.



Supplementary Figure 31. Average normalized expression of all brain subtissues for the modules (a) *M1_Extended* and (b) *M2_Extended* during brain development. Gray dots represent the average normalized expression of all the genes in the module for each time point and subtissue.



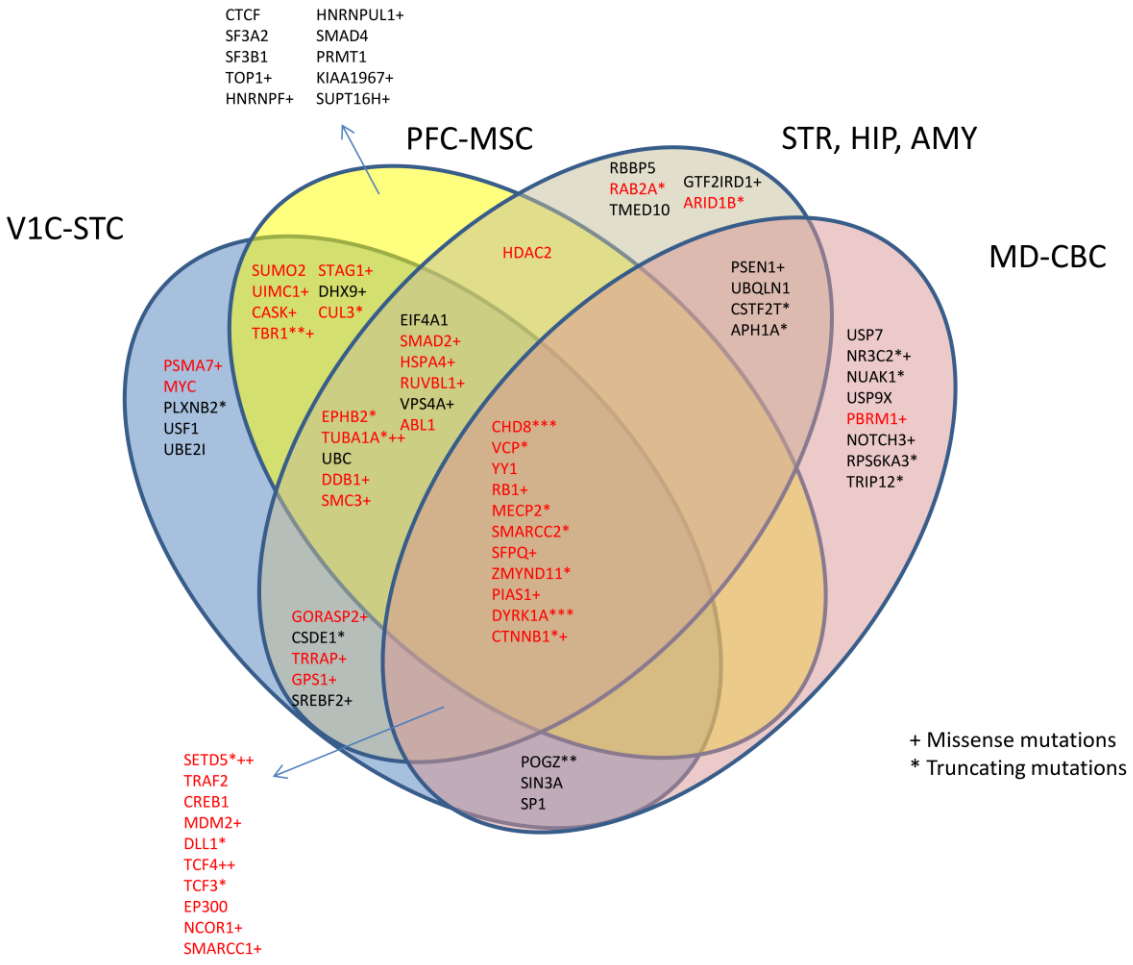
Supplementary Figure 32. Expression of *M1_Extended* in different tissues, using the Gene Atlas (Gene Expression Omnibus accession number GSE1133).



Supplementary Figure 33. Expression of *M2_Extended* in different tissues, using the Gene Atlas (Gene Expression Omnibus accession number GSE1133).

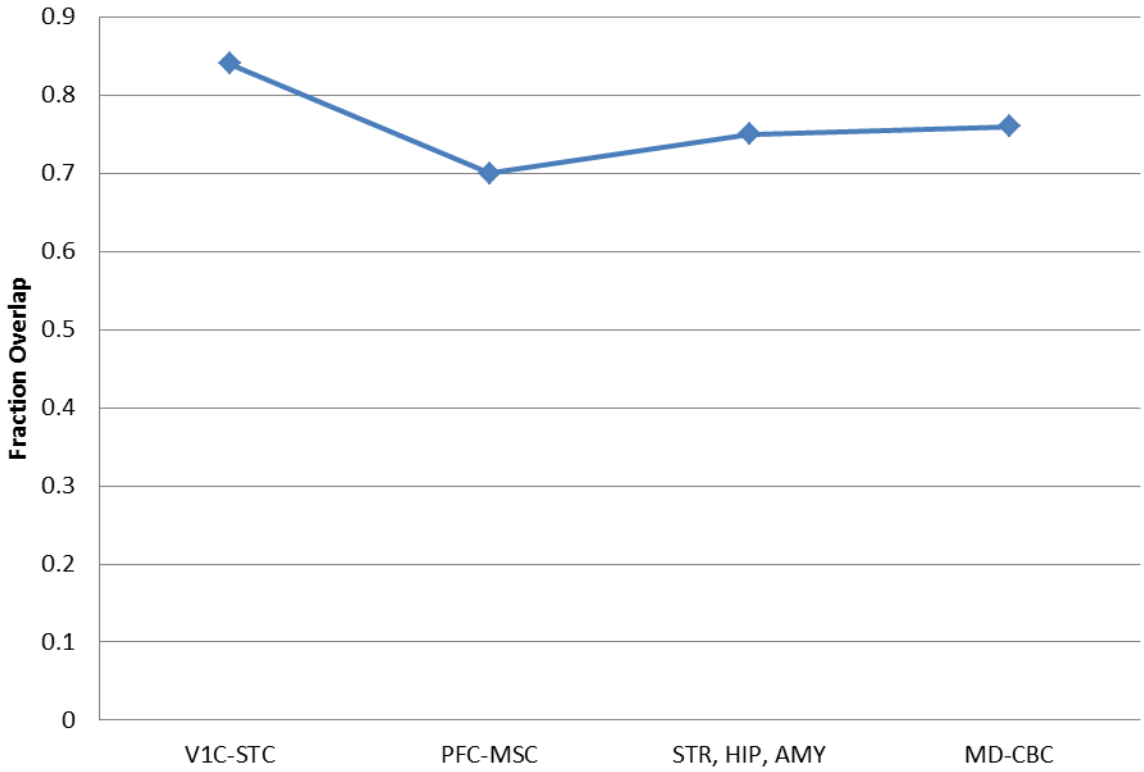
Detecting modules for specific brain regions

To detect modules for specific brain regions, we have calculated the co-expression between every pair of genes (using Pearson correlation coefficient) during four main clusters of brain regions, considering the times points from 8PCW to 1 year after birth. The clusters that were considered are the ones previously analyzed by (Willsey et al., 2013) namely: (1) Primary visual cortex - superior temporal cortex, or V1C-STC cluster; (2) Prefrontal and primary motor-somatosensory cortex or PFC-MSC cluster; (3) Striatum (STR), hippocampus (HIP), and amygdaloid complex (AMY); (4) Mediodorsal nucleus of thalamus - cerebellar cortex or MD-CBC cluster. We then applied MAGI using these co-expression values and detect the *M1_Best* and *M2_Best* modules for each of the brain regions clusters. The resulted *M1_Best* modules are different to some extent from the full-brain *M1_Best* but still show a remarkable overlap with the modules found using the full expression dataset. Specifically, the *M1_Best* and *M2_Best* modules found using the MD-CBC cluster show the least overlap with the full-brain modules. The majority of the genes found in *M1_Extended* and *M2_Extended* are not shared among the best detected modules found for the MD-CBC brain region. For comparison, we also used the neocortex (including both the V1C-STC and the PFC-MSC cluster) as an additional cluster, similar to Kang et al. (2011).



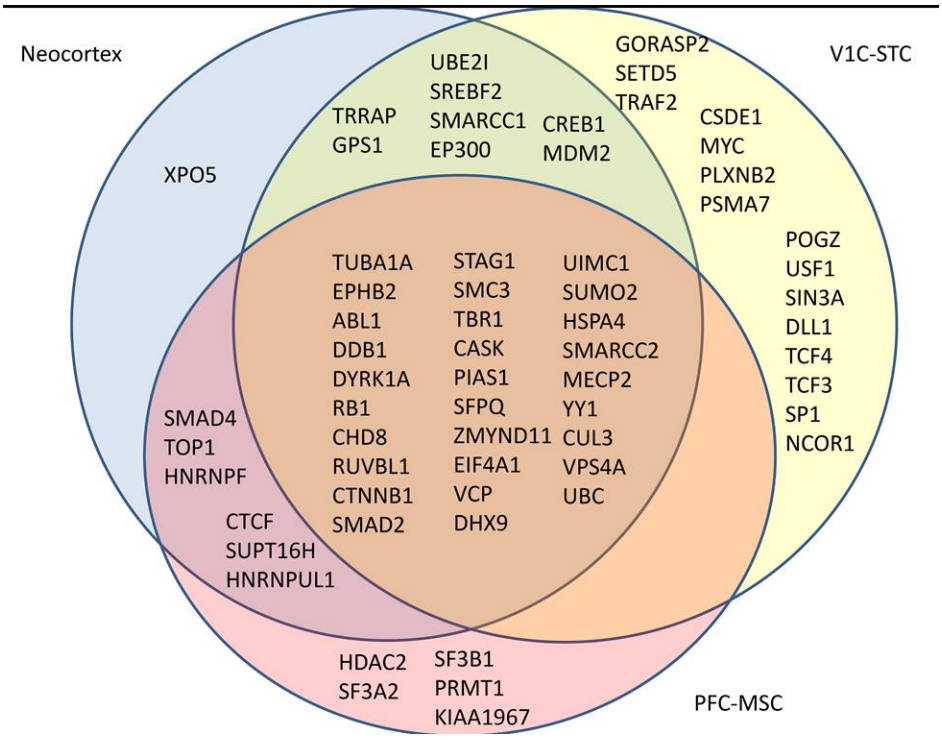
Supplementary Figure 34. Overlap among the genes found in *M1_Best* for specific brain regions. Marked in red are genes that are also part of the *M1_Best* calculated using the full-brain expression data.

Similar results were found when using Spearman's rho instead of Pearson correlation coefficient to calculate co-expression. The overlap among the modules found using the two different co-expression values (Pearson vs. Spearman) was found to be high (>0.7) for each of the brain regions.



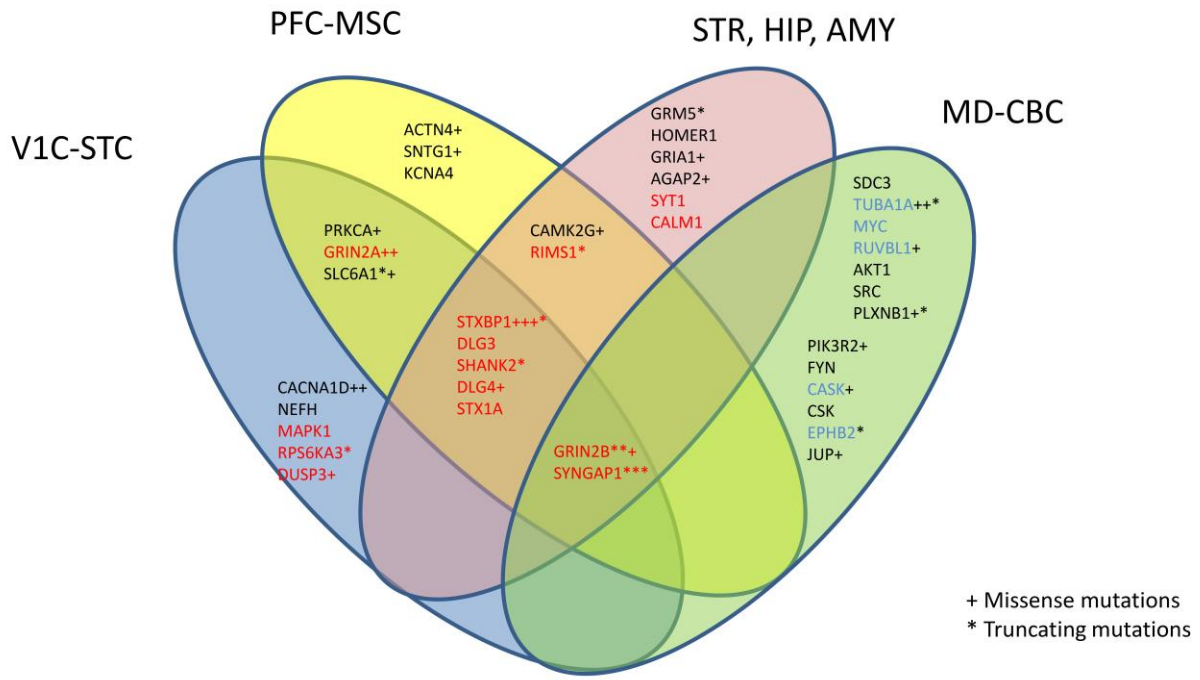
Supplementary Figure 35. The overlap between the extended modules found using co-expression values calculated using the Pearson correlation coefficient and Spearman's rho for four different brain regions.

Comparing the *M1_Best* module found for the neocortex region to the modules found for the V1C-STC and the PFC-MSC regions shows high overlap among the three modules.



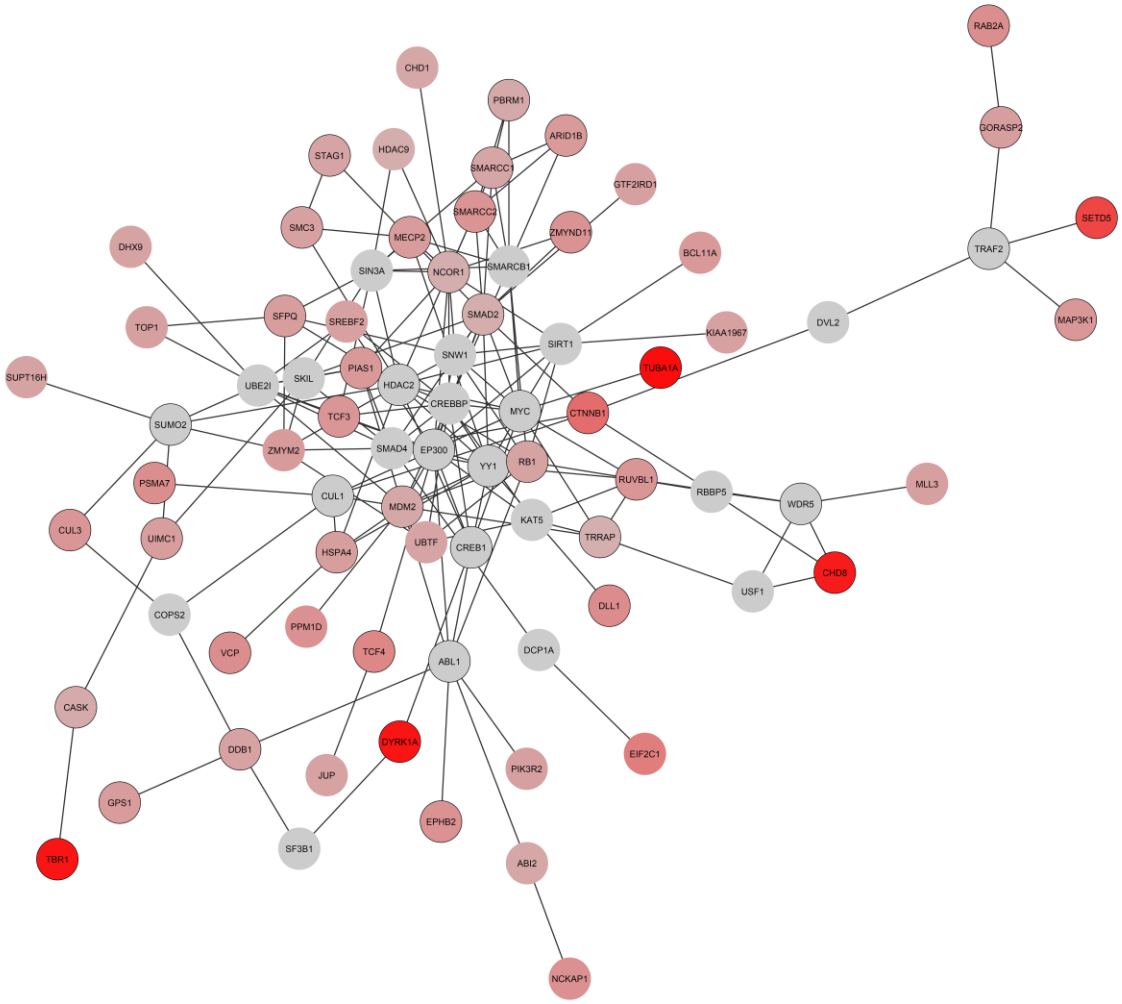
Supplementary Figure 36. Overlap among the genes found in *M1_Best* for the neocortex region and the V1C-STC and PFC-MSC brain regions.

Similarly, for *M2_Best* there is also a clear overlap among the modules found for three regions (V1C-STC; PFC-MSC; STR, HIP, and AMY) and the *M2_Best* module found for the full-brain expression data (using Pearson correlation coefficient for co-expression values). However, for the region MD-CBC, *M2_Best* is quite different.

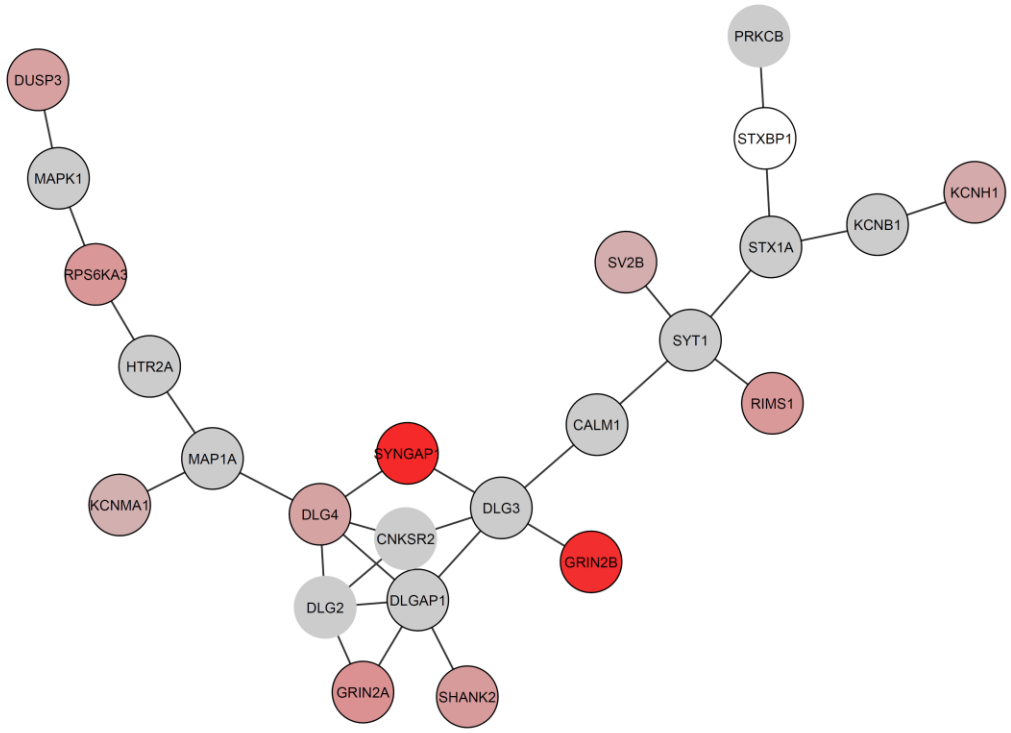


Supplementary Figure 37: Overlap among the genes found in *M2_Best* for specific brain regions. Marked in red are genes that are also part of the *M2_Best* calculated using the full-brain expression data.

Force-directed representation of MAGI's modules



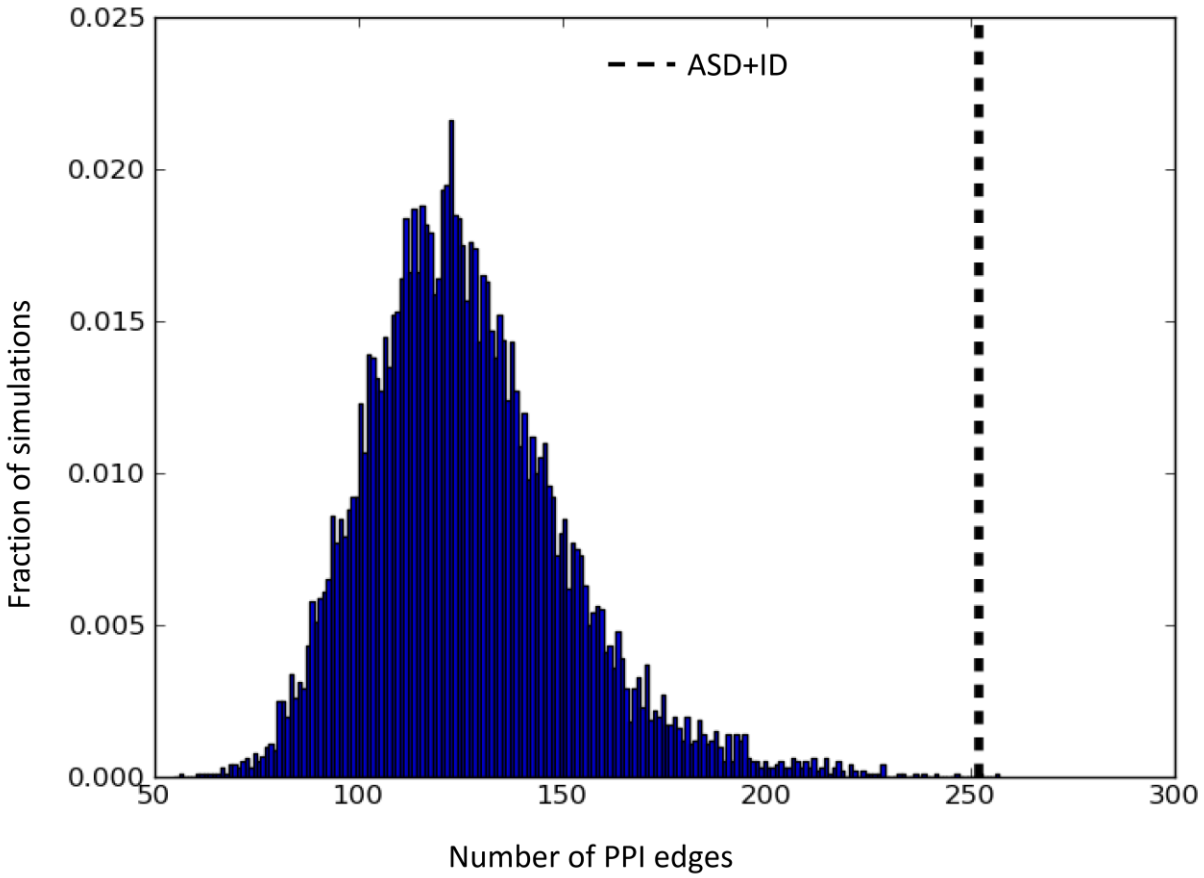
Supplementary Figure 38: Force-directed layout of *M1_Extended*. Genes detected as part of module *M1_Extended* are displayed as graph nodes. Node colors reflect the score of each gene based on the number and type of *de novo* mutations: the more intense red color indicates a higher score while gray indicates a score of zero (no *de novo* mutations observed). Edges (black lines) between two nodes represent genes that interact with each other according to the PPI network and are also highly co-expressed (Pearson correlation coefficient $r^2>0.37$, i.e., the genes are included in the top 5% of gene-pair co-expression during brain development). The network is displayed with the force-directed layout function of Cytoscape (Shannon et al., 2003) using the co-expression (r^2) as the edge weight.



Supplementary Figure 39. Force-directed layout of *M2_Extended*. Genes detected as part of module M2_Extended are displayed as graph nodes. Node colors reflect the score of each gene based on the number and type of *de novo* mutations: the more intense red color indicates a higher score while gray indicates a score of zero (no *de novo* mutations observed). Edges (black lines) between two nodes represent genes that interact with each other according to the PPI network and are also highly co-expressed (Pearson correlation coefficient r^2>0.37, i.e., the genes are included in the top 5% of gene pair co-expression during brain development). The network is displayed with the force-directed layout function of Cytoscape (Shannon et al., 2003) using the co-expression (r^2) as the edge weight.

Genes carrying *de novo* mutations are more connected

Looking at the entire PPI network, there are 517 genes carrying either LoF or missense *de novo* mutations, with 252 edges connecting them. We randomly sampled 10,000 sets of 517 genes from the PPI and found that a random set are connected with an average of 125.72 edges. Genes with either LoF or missense *de novo* mutations were found to be significantly connected compared to these random sets ($p<0.00029$).



Supplementary Figure 40. Genes carrying *de novo* LoF or missense mutations in probands that are included in the PPI network (n=517) are connected by a total of 252 protein interactions (black dashed line). For comparison, 10,000 random sets of genes of the same size (n=517) were sampled. The histogram shows the number of protein interactions connecting each such random set. Genes carrying *de novo* LoF or missense mutations are found to have a significantly higher number of protein interactions connecting them (p<0.00029).

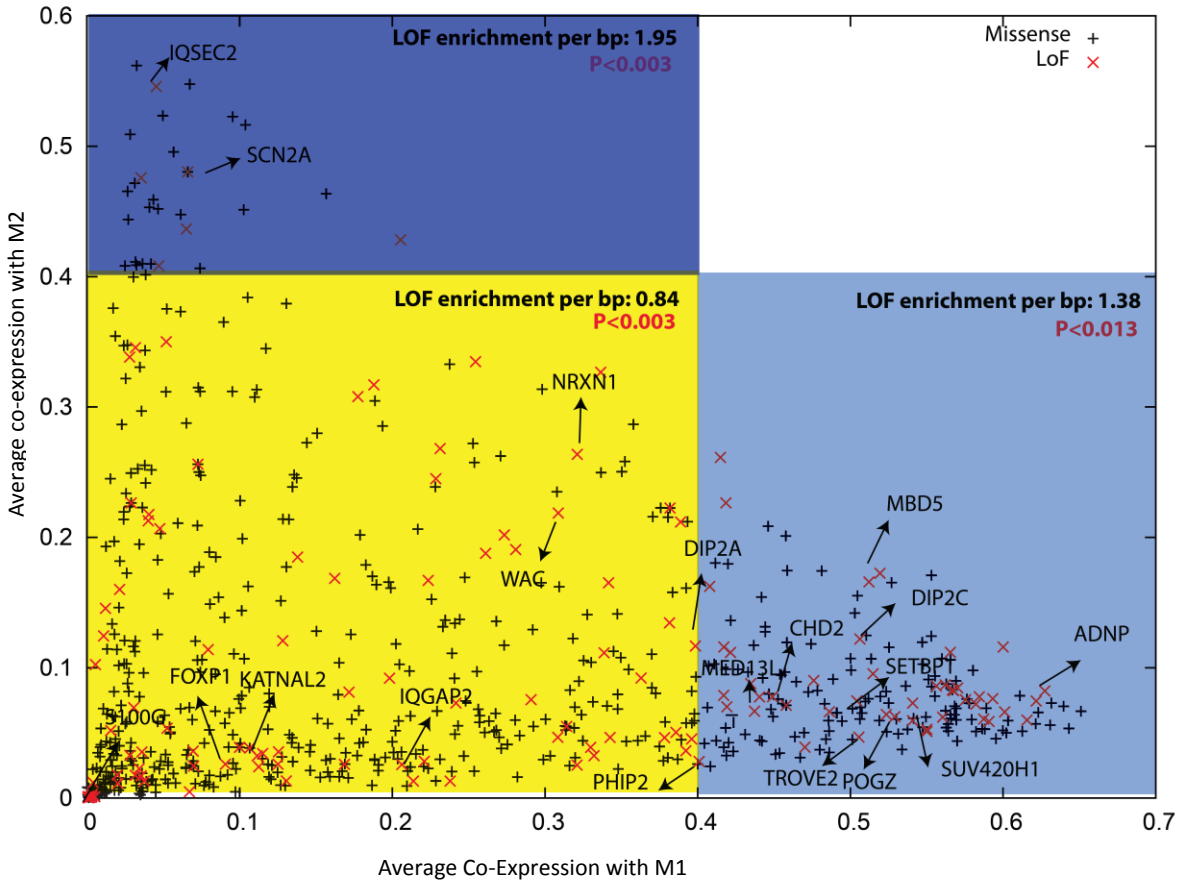
Similarly, considering only *de novo* missense mutations, there are 421 genes with missense mutations in the PPI network, connected by 180 edges. Using similar random sampling of sets of size 421, we found that the average number of edges in these samples is 83.0485 and that the genes carrying *de novo* missense mutations are significantly connected (p<0.00019998).

De novo LoF mutations are enriched in genes that co-express with MAGI's modules

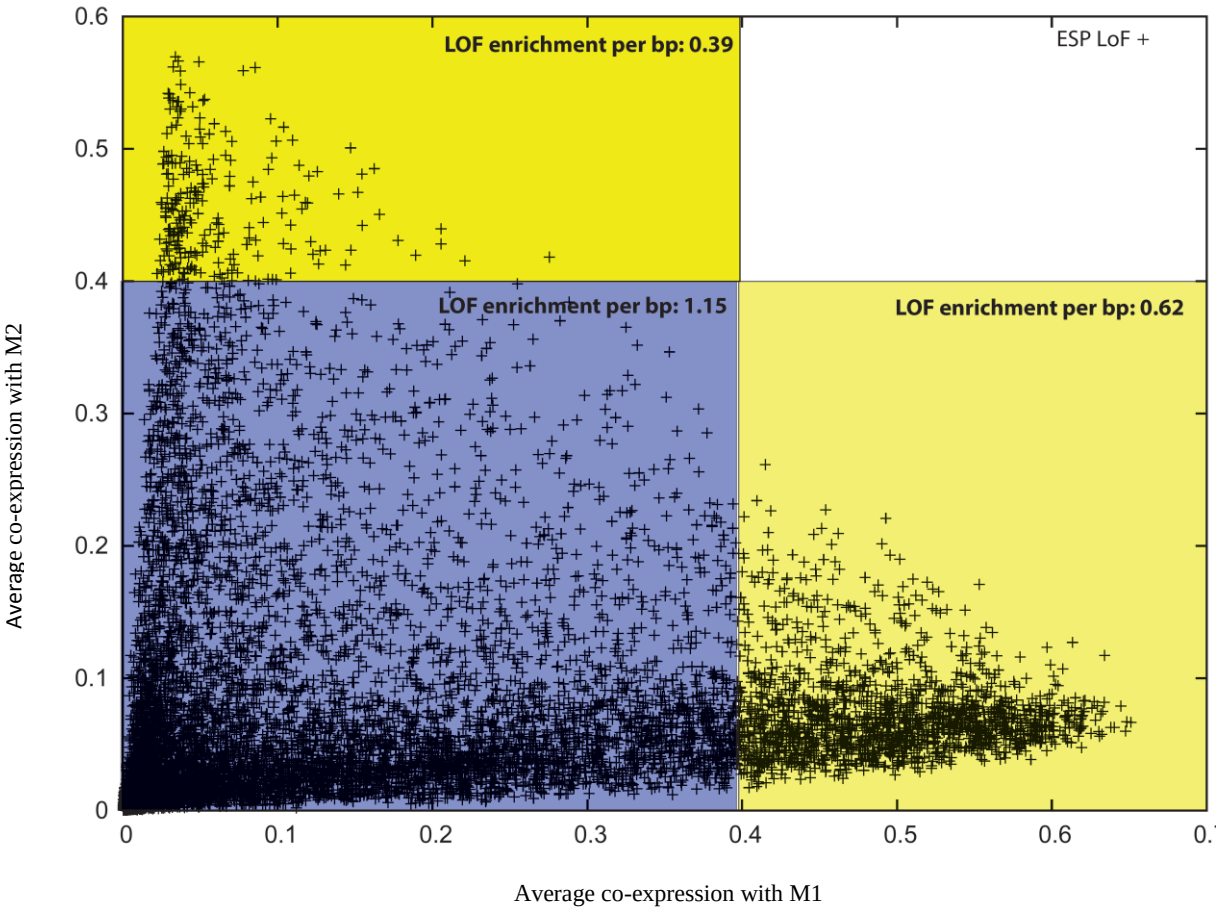
Many of the genes that carry *de novo* LoF or missense mutations in probands in the ASD+ID dataset were not identified as part of the two reported modules (*M1_Extended* and *M2_Extended*). We investigated the possibility that these genes are related to *M1_Extended* and

M2_Extended. For this, we calculated the average co-expression of each gene that is not covered by either *M1_Extended* or *M2_Extended* with the genes in *M1_Extended* and *M2_extended* modules most confident genes - only genes with confidence score >0.99 (in the inner most circle of Figures 2 and 3) were considered. Interestingly, genes showing high co-expression with *M1_Extended* or *M2_Extended* are significantly enriched (1.95x and 1.38x, respectively) in *de novo* LoF mutations, while genes that have low co-expression with both of them are significantly depleted (0.84X) in LoF mutation in probands (Supplementary Figure 41). The exact mirror image observed for LoF mutations in ESP/control data (

Supplementary Figure 42).



Supplementary Figure 41. Each point represents a gene with LoF mutations (green) or missense mutations (red) in ASD+ID probands. The x-axis represents the average co-expression with the genes in module *M1_Extended* (only genes with confidence score >0.99 were considered). The y-axis represents the average co-expression with the genes in module *M2_Extended*. The enrichment and *p*-values were calculated considering the length of the genes in each group. Enrichment and depletion are represented using blue and yellow background colors, respectively.

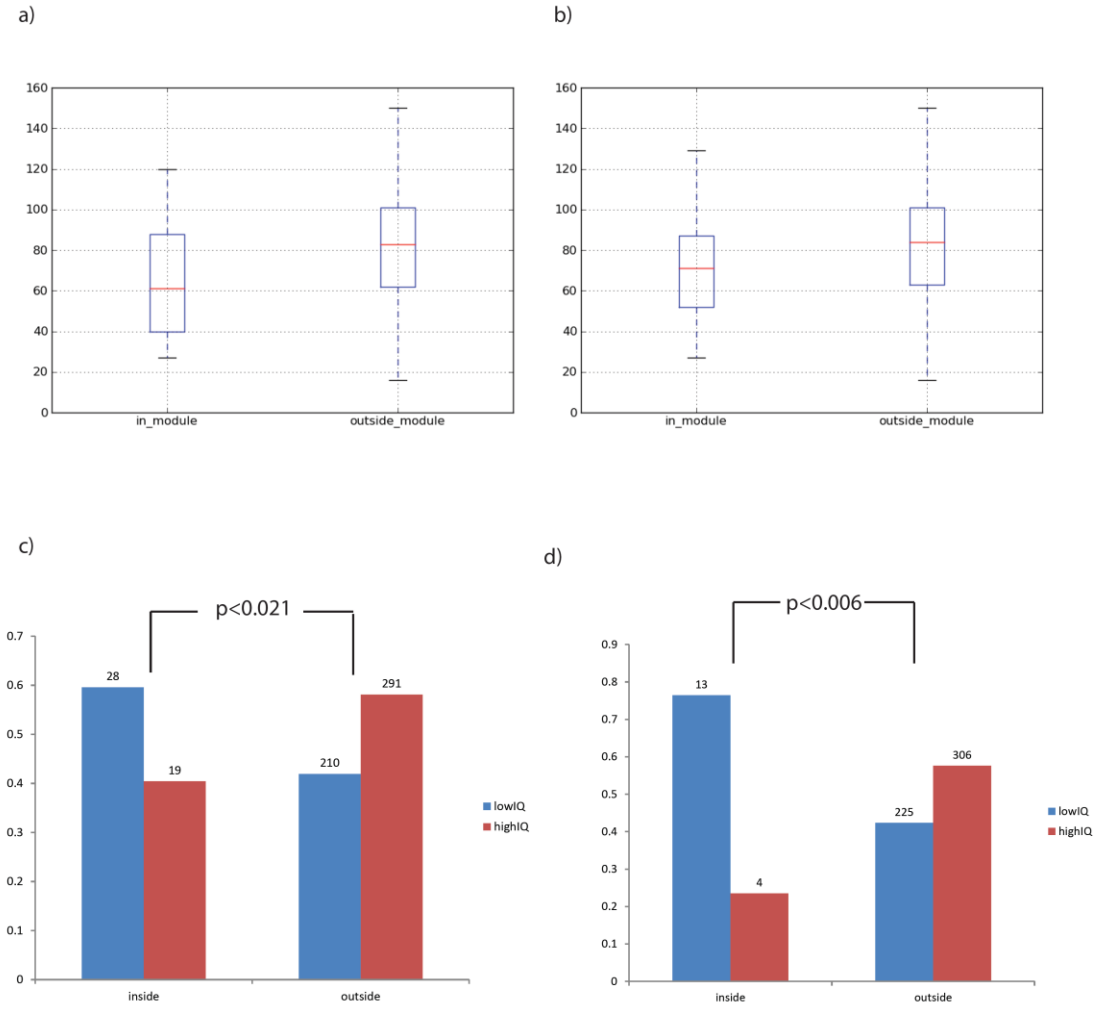


Supplementary Figure 42. Each point represents a gene with LoF mutations in the ESP. The x-axis is the average co-expression with the genes in module *M1_Extended* (the genes with confidence score >0.99) and y-axis is the average co-expression of the genes with the genes in module *M2_Extended*. Enrichment and depletion are represented using blue and yellow background colors, respectively.

IQ distinction of modules

We investigated the IQ of probands carrying *de novo* mutations in the detected modules. We found that probands with *de novo* LoF mutations in *M1_Best* have a significantly lower IQ compared to other probands that have LoF mutations outside of the module (n1 = 17, median1 = 61 vs. n2 = 108, median2 = 81.5) with $p<0.028$ (Mann-Whitney two-sided test). The significance increases when testing probands with *de novo* LoF or missense mutations in *M1_Best* (n1 = 37, median1 = 71) compared to probands with a *de novo* LoF or missense mutations only outside of *M1_Best* (n2 = 421, median2 = 84) with $p<0.008$. Similarly, probands with *de novo* mutations in *M1_Extended* were found to have significantly lower IQ than other probands when focusing on LoF mutations only ($p<0.018$) or LoF and missense mutations ($p<0.016$). Unfortunately, there is no available information on the IQ of individuals from the two ID studies, but most of the mutations covered by *M2_Extended* are coming from these studies. In order to account for these

two studies, we have therefore used a different strategy. We divided the probands from all six ASD+ID studies into two groups: (i) probands with ID, including all the probands from the two ID studies and probands from the four ASD studies that have IQ<70 and (ii) probands without ID, including all the probands from the four ASD studies that have IQ≥70. We found that the number of individuals with IQ<70 with LoF or missense mutations in *M2* (both *M2_Best* and *M2_Extended*) was highly enriched compared to probands with LoF or missense mutations outside the module (p<0.0062).

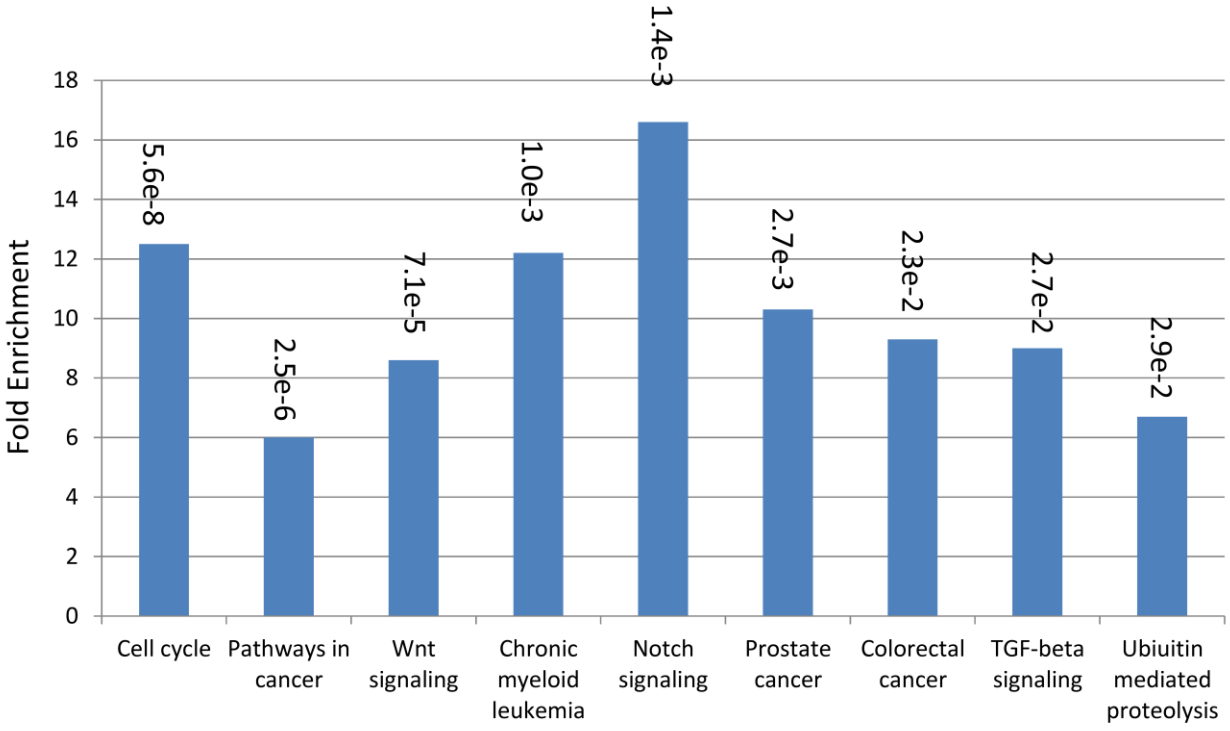


Supplementary Figure 43. (a) The IQ of the probands with *de novo* LoF mutations in *M1_Best* compared to the IQ of other probands with *de novo* LoF mutations. (b) The IQ of probands with *de novo* LoF or missense mutations inside *M1_Best* compared to the IQ of other probands with *de novo* LoF or missense mutations. (c) The number of probands with ID (blue bars) or without ID (red bars) that have *de novo* LoF or missense mutations in genes inside *M1_Best* or genes outside of *M1_Best*. (d) The number of probands with ID (blue bars) or without ID (red bars) that have *de novo* LoF or missense mutations in genes in *M2_Best* or outside *M2_Best*.

Similar results are found when looking at male and female samples separately. The average IQ of male probands with *de novo* LoF mutations inside *M1_Extended* (median IQ=76) is significantly

lower in comparison to other male probands with *de novo* LoF mutations (median IQ=83) with p<0.036. A similar trend is observed for female samples with median IQ of 52.5 for probands with LoF mutation inside of module versus other female probands with LoF mutations with median IQ of 75. However, due to the low number of female probands with *de novo* mutations in our cohort (n=30), the results are not significant (p<0.159).

KEGG pathway enrichment for *M1_Extended* with cancer pathways

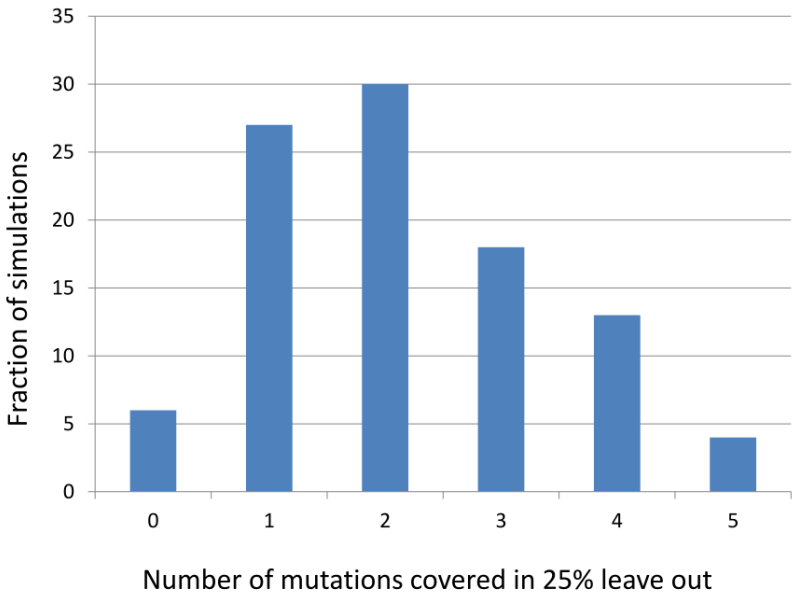


Supplementary Figure 44. The KEGG pathway enrichment for *M1_Extended* including the cancer-related pathways (which were excluded from previous figures).

Cross validation

We also performed a leave-25%-out cross-validation analysis. For this we have randomly partitioned the mutations to a training set of size 75% and a validation set of size 25%. We then run MAGI on the training set and calculate the number of mutations in the validation set that are found in the detected module. We ran the above experiment 100 times and the distribution on number of mutations covered from the smaller set is provided below. In 94% of these runs, at least one mutation from the validation set was covered in the resulted module, while the expected number of mutations found in a module of size ~50 genes is strictly lower than 1 (assuming total

number of genes is ~25,000). The distribution of number of mutations covered is given in Supplementary Figure 45.



Supplementary Figure 45. Histogram of the number of mutations covered in the leave-25%-out cross-validation (100 simulations).

References

Alon, N., Dao, P., Hajirasouliha, I., Hormozdiari, F., and Sahinalp, S.C. (2008). Biomolecular network motif counting and discovery by color coding. Bioinformatics *24*, i241-249.

Alon, N., Yuster, R., and Zwick, U. (1995). Color-Coding. J Assoc Comput Mach *42*, 844-856.

Cristino, A.S., Williams, S.M., Hawi, Z., An, J.Y., Bellgrove, M.A., Schwartz, C.E., Costa Lda, F., and Claudianos, C. (2014). Neurodevelopmental and neuropsychiatric disorders represent an interconnected molecular system. Molecular psychiatry *19*, 294-301.

Dost, B., Shlomi, T., Gupta, N., Ruppin, E., Bafna, V., and Sharan, R. (2008). QNet: a tool for querying protein interaction networks. Journal of computational biology : a journal of computational molecular cell biology *15*, 913-925.

Gilman, S.R., Iossifov, I., Levy, D., Ronemus, M., Wigler, M., and Vitkup, D. (2011). Rare de novo variants associated with autism implicate a large functional network of genes involved in formation and function of synapses. Neuron *70*, 898-907.

Huang da, W., Sherman, B.T., and Lempicki, R.A. (2009). Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic acids research *37*, 1-13.

Ideker, T., Ozier, O., Schwikowski, B., and Siegel, A.F. (2002). Discovering regulatory and signalling circuits in molecular interaction networks. Bioinformatics *18 Suppl 1*, S233-240.

Jia, P., Wang, L., Fanous, A.H., Pato, C.N., Edwards, T.L., International Schizophrenia, C., and Zhao, Z. (2012). Network-assisted investigation of combined causal signals from genome-wide association studies in schizophrenia. PLoS computational biology *8*, e1002587.
Kang, H.J., Kawasawa, Y.I., Cheng, F., Zhu, Y., Xu, X., Li, M., Sousa, A.M., Pletikos, M., Meyer, K.A., Sedmak, G., *et al.* (2011). Spatio-temporal transcriptome of the human brain. Nature *478*, 483-489.
Kong, A., Frigge, M.L., Masson, G., Besenbacher, S., Sulem, P., Magnusson, G., Gudjonsson, S.A., Sigurdsson, A., Jonasdottir, A., Jonasdottir, A., *et al.* (2012). Rate of de novo mutations and the importance of father's age to disease risk. Nature *488*, 471-475.
Krumm, N., O'Roak, B.J., Karakoc, E., Mohajeri, K., Nelson, B., Vives, L., Jacquemont, S., Munson, J., Bernier, R., and Eichler, E.E. (2013). Transmission disequilibrium of small CNVs in simplex autism. American journal of human genetics *93*, 595-606.
Liu, L., Lei, J., Sanders, S.J., Willsey, A.J., Kou, Y., Cicek, A.E., Klei, L., Lu, C., He, X., Li, M., *et al.* (2014). DAWN: a framework to identify autism genes and subnetworks using gene expression and genetics. Molecular autism *5*, 22.
O'Roak, B.J., Vives, L., Girirajan, S., Karakoc, E., Krumm, N., Coe, B.P., Levy, R., Ko, A., Lee, C., Smith, J.D., *et al.* (2012). Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations. Nature *485*, 246-250.
Parikshak, N.N., Luo, R., Zhang, A., Won, H., Lowe, J.K., Chandran, V., Horvath, S., and Geschwind, D.H. (2013). Integrative functional genomic analyses implicate specific molecular pathways and circuits in autism. Cell *155*, 1008-1021.
Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome research *13*, 2498-2504.
Shlomi, T., Segal, D., Ruppin, E., and Sharan, R. (2006). QPath: a method for querying pathways in a protein-protein interaction network. BMC bioinformatics *7*, 199.
Supek, F., Bosnjak, M., Skunca, N., and Smuc, T. (2011). REVIGO summarizes and visualizes long lists of gene ontology terms. PloS one *6*, e21800.
Tennessen, J.A., Bigham, A.W., O'Connor, T.D., Fu, W., Kenny, E.E., Gravel, S., McGee, S., Do, R., Liu, X., Jun, G., *et al.* (2012). Evolution and functional impact of rare coding variation from deep sequencing of human exomes. Science *337*, 64-69.
Tuncbag, N., McCallum, S., Huang, S.S., and Fraenkel, E. (2012). SteinerNet: a web server for integrating 'omic' data to discover hidden components of response pathways. Nucleic acids research *40*, W505-509.
Willsey, A.J., Sanders, S.J., Li, M., Dong, S., Tebbenkamp, A.T., Muhle, R.A., Reilly, S.K., Lin, L., Fertuzinhos, S., Miller, J.A., *et al.* (2013). Coexpression networks implicate human midfetal deep cortical projection neurons in the pathogenesis of autism. Cell *155*, 997-1007.