Co-expression Profiling of Autism Genes in the Mouse Brain

Idan Menashe1,2,*, Pascal Grange3,#, Eric C. Larsen1, Sharmila Banerjee-Basu1, Partha P. Mitra3

1 MindSpec, McLean, Virginia, United States of America, 2 Department of Public Health, Faculty of Health Sciences, Ben-Gurion University of the Negev, Beer Sheva, Israel, 3 Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, United States of America

Abstract

Autism spectrum disorder (ASD) is one of the most prevalent and highly heritable neurodevelopmental disorders in humans. There is significant evidence that the onset and severity of ASD is governed in part by complex genetic mechanisms affecting the normal development of the brain. To date, a number of genes have been associated with ASD. However, the temporal and spatial co-expression of these genes in the brain remain unclear. To address this issue, we examined the co-expression network of 26 autism genes from AutDB (http://mindspec.org/autdb.html), in the framework of 3,041 genes whose expression energies have the highest correlation between the coronal and sagittal images from the Allen Mouse Brain Atlas database (http://mouse.brain-map.org). These data were derived from in situ hybridization experiments conducted on male, 56-day old C57BL/6J mice co-registered to the Allen Reference Atlas, and were used to generate a normalized co-expression matrix indicating the cosine similarity between expression vectors of genes in this database. The network formed by the autism-associated genes showed a higher degree of co-expression connectivity than seen for the other genes in this dataset (Kolmogorov–Smirnov $P=5 \times 10^{-28}$). Using Monte Carlo simulations, we identified two cliques of co-expressed genes that were significantly enriched with autism genes (A Bonferroni corrected $P < 0.05$). Genes in both these cliques were significantly over-expressed in the cerebellar cortex ($P=1 \times 10^{-5}$) suggesting possible implication of this brain region in autism. In conclusion, our study provides a detailed profiling of co-expression patterns of autism genes in the mouse brain, and suggests specific brain regions and new candidate genes that could be involved in autism etiology.

Introduction

Autism spectrum disorder (ASD) is one of the most prevalent and highly heritable neurodevelopmental disorders in humans [1–3]. There is strong evidence that the onset and severity of ASD is governed in part by complex molecular mechanisms affecting the normal development of the brain [1,4]. While no major anatomical pathology have been observed in brains of ASD cases [5], various molecular and neuroimaging studies have linked several brain regions to ASD. For example, Voineagu et al. have found differences in gene expression patterns in the cortex of ASD brain [6]. Cortical regions has also been highlighted in neuroimaging studies of autistic brains along the cerebellum and other brain areas [7,8]. In addition, other studies have pointed to various molecular mechanisms that might be altered in the autistic brain [9–11]. In this realm, genes involved in synapse formation and brain circuitry are consistently found to be dysregulated in people with ASD [12–14].

Recent genomic advances have led to the discovery of diverse genetic loci linked to ASD, including chromosomal aberrations [15,16], copy number variations [11,17,18] and both common and rare single nucleotide variations (SNVs) [19–24]. Consequently, to date, more than 330 candidate genes have been associated with ASD susceptibility [25] and many more are projected to be found. However, despite the plethora of genetic variations associated with ASD, the molecular mechanisms and neuroanatomical structures underlying ASD traits remain largely unclear.

The mouse model system provides a convenient and safe approach to experimentally study neuroanatomical mechanisms and candidate genes for autism susceptibility [26–28]. At present, dozens of single gene knockout and transgenic mice models have been used to elucidate neuropathology that might underlie the autism-like behaviors [29]. Despite the obvious genetic and neuroanatomical differences between mouse and human, mouse models are extremely valuable and effectively used in dissecting out the role of specific gene, pathway, neuron subtype, or brain region in a particular abnormal behavior shared by both these mammals. In this realm, the Allen Brain Atlas of the mouse [30–33] provides a comprehensive source of genome-wide high-resolution atlas of gene expression throughout the adult mouse brain. In this study, we utilized this database to examine the spatial co-expression characteristics of genes associated with autism susceptibility. Consequently, we identified several co-expression characteristics of autism genes in the mouse brain.
gene networks that are enriched with autism genes highlighting potential candidate genes and brain regions implicated in autism.

Methods

Construction of the autism genes dataset (AutRef84)

The autism gene database, AutDB [34], was used to construct a set of 84 genes, AutRef84, strongly implicated in autism pathogenesis [25,29]. Genes contained in AutDB were classified into four genetic categories: (1) Rare: rare submicroscopic copy number variants and single gene disruptions or mutations directly linked to ASD; (2) Syndromic: genes implicated in syndromes in which a subset of affected individuals also develop autistic symptoms; (3) Association: small risk-conferring candidate genes with common polymorphisms that have been identified from genome-wide association studies in idiopathic ASD cases; and (4) Functional: functional candidate genes relevant to ASD that are not covered by any of the other genetic categories. AutRef84 was generated by filtering out functional candidate genes that lack any experimentally derived genetic link to ASD as well as genes that solely belong to the Association classification. The resulting dataset consisted of 64 genes classified within the rare classification and 20 within the syndromic classification (Supplementary Table S1).

Gene expression dataset

We utilized gene expression data available from the Allen Brain Atlas (ABA) of the mouse brain (http://mouse.brain-map.org) which contains voxelized expression profiles for ~20,000 genes derived from in situ hybridization (ISH) experiments conducted on male, 56-day old C57BL/6J mice) [30–33]. Gene expression profiles in the ABA were generated for the mouse brains from processed image data for the sagittal sections of a single hemisphere and, for 4,104 genes of high neurobiological interest, coronal and sagittal sections across the whole brain. We focused on genes for which brain-wide data are available. The ISH data were co-registered to the Allen Reference Atlas, which is partitioned into 49,742 cubic voxels of size 200 microns. A gene labeled g has a profile of expression energy presented as a function $E(v,g)$ over the brain, where $v$ is a voxel label. The Allen Mouse Brain Atlas Addiction Database (http:// addiction.brainarchitecture.org/) used this expression dataset to create a high quality brain expression dataset, called $A_{high}$ which consists of 3,041 genes with the most highly correlated expression energy profiles between the coronal and sagittal sections [35]. This gene-expression dataset was used for the co-expression analysis discussed below.

Gene co-expression analysis

For the set of 3,041 genes defined above, a gene-by-gene co-expression matrix was computed as follows:

$$C(g, g') = \frac{\sum_{v=1}^{V} E(v,g)E(v,g')}{\sqrt{\sum_{v=1}^{V} E(v,g)^2 \sum_{v=1}^{V} E(v,g')^2}}$$  \hspace{1cm} (1)

where $V = 49,742$ is the total number of voxels in the brain. The quantity $C(g,g')$ is the cosine similarity between the gene-expression profiles of genes $g$ and $g'$ with values ranging between zero and one by construction. The motivation for choosing the cosine similarity as the co-expression measure in this study, is its connection to the difference between the energies of two brain-wide functions. The proposed co-expression measure (Equation 1) can also be described as a simple decreasing function of the squared energy of the difference between two brain-wide functions. The co-expression matrix $C$ is symmetric, with ones on the diagonal. Its size is 3,041, the number of genes in the full $A_{high}$ dataset. Given a subset of this dataset, the co-expression matrix of this subset can be obtained by extracting the sub-matrix of $C$ corresponding to the position of the genes in the full dataset.

To determine whether the autism genes are more co-expressed than other genes expected by chance, the cumulative distribution function (CDF) of the entries of the autism-related and $A_{high}$ co-expression matrices were compared, and a two-sample Kolmogorov–Smirnov test was used to determine whether they were drawn from the same probability distribution. In addition, we compared the connectivity of the co-expression network of the autism genes to co-expression networks of the same size randomly drawn from $A_{high}$. For that, Monte Carlo simulations [36] were conducted to generate 100,000 random gene sets sampled from $A_{high}$ and the co-expression matrices of these sets of genes were extracted from the co-expression matrix $C$.

Given the co-expression matrix of $G$ genes of interest, one can consider the underlying weighted graph with nodes corresponding to genes, and the weight of links equals to the co-expression of the nodes [37]. The matrix can be cut at any value $\rho$ between zero and one, resulting in links with weights lower than the threshold $\rho$. At any value of the threshold, the connected components (sets of connected genes) can be computed using Tarjan’s algorithm. In particular, the maximum size $M(\rho)$ and average size $A(\rho)$ of connected components can be calculated for all gene sets over different co-expression values. If $N(\rho)(k)$ is the number of connected components in the co-expression matrix of $G$ genes defined by $\rho$ threshold and contain exactly k genes out of $G$ genes in the gene set, the maximum and average sizes are expressed as follows:

$$A(\rho) = \frac{\sum_{k=1}^{G} kN(\rho)(k)}{\sum_{k=1}^{G} N(\rho)(k)}$$ \hspace{1cm} (2)

$$M(\rho) = \max\{k\in[1...G], N(\rho)(k) > 0\}$$ \hspace{1cm} (3)
Figure 1. Co-expression characteristics of gene networks. Distributions of co-expression characteristics are depicted for the autism gene network (Red) and for the 3041 genes in the mouse Allen Brain Atlas database (Blue). (A) Cumulative distribution functions (CDFs). Black arrow indicates the maximal difference between the two CDFs. (B) Average and maximum sizes of connected components for different co-expression thresholds are plotted for the autism gene network (Red) and 1000 randomly generated gene networks of size G = 26 (Cyan).

doi:10.1371/journal.pcbi.1003128.g001
## Co-expression of Autism Genes in the Mouse Brain

| Threshold | # of autism genes | # of all genes | p-value |
|-----------|-------------------|----------------|---------|
| 70.5% | 4 | 33 | 0.0002 |
| 79.5% | 2 | 6 | 0.0008 |
| 59.3% | 12 | 578 | 0.0014 |
| 62.4% | 3 | 30 | 0.0015 |
| 61.2% | 10 | 410 | 0.0023 |
| 75.2% | 3 | 33 | 0.0024 |
| 55.2% | 10 | 468 | 0.0026 |
| 57.5% | 13 | 760 | 0.0053 |
| 49.2% | 18 | 1355 | 0.0076 |
| 71.8% | 4 | 95 | 0.0088 |
| 68.2% | 6 | 225 | 0.0100 |
| 53.3% | 11 | 649 | 0.0112 |
| 47.6% | 19 | 1496 | 0.0117 |
| 59.5% | 4 | 109 | 0.0125 |
| 58.5% | 5 | 174 | 0.0140 |
| 61.3% | 11 | 678 | 0.0185 |
| 54.2% | 14 | 1005 | 0.0224 |
| 66.4% | 7 | 340 | 0.0232 |
| 73.2% | 10 | 644 | 0.0323 |
| 86.6% | 4 | 144 | 0.0342 |
| 70.1% | 12 | 850 | 0.0356 |
| 82.7% | 6 | 299 | 0.0361 |
| 71.8% | 11 | 750 | 0.0363 |
| 52.6% | 17 | 1411 | 0.0400 |
| 62.8% | 15 | 1179 | 0.0408 |
| 56.9% | 6 | 310 | 0.0427 |
| 48.1% | 18 | 1549 | 0.0450 |
| 47.6% | 14 | 1095 | 0.0462 |
| 76.2% | 3 | 95 | 0.0472 |
| 84.3% | 4 | 169 | 0.0515 |
| 88.7% | 2 | 44 | 0.0540 |
| 51.2% | 15 | 1226 | 0.0550 |
| 73.0% | 5 | 250 | 0.0587 |
| 85.1% | 3 | 109 | 0.0620 |
| 74.6% | 9 | 623 | 0.0642 |
| 48.3% | 13 | 1033 | 0.0676 |
| 74.0% | 4 | 186 | 0.0710 |
| 65.7% | 4 | 189 | 0.0751 |
| 74.6% | 9 | 632 | 0.0752 |
| 63.1% | 14 | 1178 | 0.0877 |
| 62.8% | 8 | 560 | 0.0886 |
| 49.1% | 12 | 977 | 0.0909 |
| 53.7% | 16 | 1419 | 0.0919 |
| 72.5% | 10 | 766 | 0.0964 |
| 76.6% | 7 | 482 | 0.1080 |
| 64.6% | 13 | 1137 | 0.1334 |
| 77.6% | 6 | 412 | 0.1351 |
| 67.0% | 7 | 541 | 0.1563 |
| 78.5% | 5 | 345 | 0.1630 |
| 88.0% | 4 | 259 | 0.1749 |
| 77.6% | 7 | 563 | 0.1830 |
| 75.6% | 8 | 663 | 0.1868 |
| 84.3% | 5 | 366 | 0.1883 |
| 83.3% | 5 | 376 | 0.2167 |
| 62.2% | 5 | 384 | 0.2270 |
| 91.0% | 2 | 143 | 0.3420 |
| 80.3% | 6 | 557 | 0.3435 |
| 89.4% | 2 | 178 | 0.4485 |
| 77.4% | 6 | 706 | 0.5724 |
Further, we applied the classical definition of “clique” from graph theory [38,39] to our co-expression matrix to characterize networks of co-expressed genes such that every gene in the network is connected to all other genes at a co-expression threshold $\rho$.

**Anatomical and functional characterization of co-expression cliques**

Next, we aimed at evaluating the unique anatomical properties of co-expression cliques that are significantly enriched with autism genes. For that, we first identified virtually all cliques in the dataset that contained $n_g \geq 2$ autism genes. Then, we used Monte Carlo simulations using 100,000 randomly generated gene sets of size $G = 26$ to compute the likelihood of each clique to contain at least $n_g$ autism genes. Finally, for cliques that were significantly enriched with autism genes, we calculated the sum of the normalized expression profiles of the genes in the clique as follow:

$$S_{\text{clique}}(v) = \sum_{g \in \text{clique}} E_{\text{norm}}(v,g),$$

where

$$E_{\text{norm}}(v,g) = \frac{E(v,g)}{\sqrt{\sum_{w=1}^{G} E(w,g)^2}}$$

We then used $S_{\text{clique}}$ to examine the neuroanatomical properties of the genes included in the clique as follows [40]. For a given brain region $\phi$ in the Allen Reference Atlas [41], the fitting score between the expression profile and the region $\phi$ is defined as the cosine similarity between the expression profile and the characteristic function $\chi_{\phi}$ of the brain region.

**Figure 2. Cliques of co-expression genes delineated by co-expressed autism genes.** The 59 cliques of co-expression genes containing $\geq 2$ autism genes identified in our data are listed in the rows together with the minimal co-expression value in the clique, and numbers of autism genes and number of total genes in the clique. A p-value indicates the likelihood of finding this number of autism genes in the clique (based on 100,000 Monte Carlo simulations). The 26 autism genes included in the study are depicted in columns with black and white fillings indicating their presence and absence from a clique respectively.

doi:10.1371/journal.pcbi.1003128.g002

**Figure 3. Cliques of co-expressed genes.** Venn diagram for the ten cliques of co-expressed genes that are highly enriched with autism genes. For each clique, the following parameters are given: the number of autism genes/number of all genes in the clique, the co-expression threshold (in %) of the clique, and the p-value of the autism gene enrichment (using 100,000 Monte Carlo resampling procedure).

doi:10.1371/journal.pcbi.1003128.g003
where $x_v(v)$ equals one if voxel labeled v belongs to region $v$, and zero otherwise. This fitting score would equal one if the sum of the gene expression in the clique were proportional to the characteristic function of the region $v$, and zero if it were entirely supported outside the region. The fitting score defined in the above equation equals the co-expression between a (hypothetical) gene whose expression profile would be $S_{clique}$, and a (hypothetical) gene whose expression would coincide exactly with region $v$. Fitting scores of a given clique can be computed in all brain regions, and the distribution of these fitting scores can be simulated by repeatedly drawing random cliques of genes (with the same number of genes) from $A_{best}$. These analyses were performed using a commercial software package (MATLAB R2011b, The MathWorks Inc., Natick, MA, 2000).

Finally, we used the Bioconductor GOstats package in R software [42] to assess whether genes belonging to a co-expression clique, also share other functional or molecular properties. The absolute list of GO terms were obtained using both a (a) cut-off $= 2 \times \text{ratio (fg/fc)}$ [where fg = frequency of occurrence of a GO term in the given gene set, fc = frequency of its occurrence in the complete list of human genes] and (b) cut-off $= \text{median value of ratio (fg/fc)}$. Only significant terms ($P < 0.01$) with an associated gene count $> 5$ were considered.

Results

Overall, 26 genes were found in the intersection of the autism-genes dataset and the dataset of high-quality expression genes from the Allen Brain Atlas (ABA) of the mouse brain ($AutRef8 \cap Abest = 26$). These autism-related genes showed a higher degree of co-expression connectivity than all other genes in this dataset (Kolmogorov–Smirnov $P = 5 \times 10^{-28}$). Comparing the empirical distributions of co-expression values of the autism genes to the other genes in the Allen dataset revealed that the largest deviation between these distributions was at co-expression value of 47.53%
Furthermore, we evaluated the connectivity of genes across different co-expression values. Here too, the average size of connected components (See methods) among autism gene was consistently larger than seen in 1000 randomly generated gene sets from the Allen database.

Next, we examined cliques (see methods) of co-expressed genes delineated by autism genes inter-connected at co-expression values of $\geq 47.53\%$. A total of 59 overlapping cliques were characterized containing on average 563.5 genes and 8.3 autism genes (Figure 2). Finally, using Monte Carlo simulations, we identified ten cliques that were significantly enriched with autism genes at $P < 0.01$ (Figure 3). Of note, the two top ranked cliques remained significant ($P < 0.05$) even after accounting for multiple testing using the conservative Bonferroni correction.

The top-ranked clique in our analysis (hereafter will be referred as Clique I), was delineated by the autism genes: *Ptchd1*, *Galnt13*, *Dpp6* and *Astn2*, and included another 29 genes, inter-connected with a co-expression values of $\geq 70\%$ (Supplementary Table S2). The second top-ranked clique (hereafter will be referred as Clique II) included the autism genes: *Rims3*, and *Astn2*, and another four genes, all inter-connected at $\geq 79\%$ co-expression level (Supplementary Table S3). Examining the neuroanatomical expression properties across a set of 134 brain regions of the left hemisphere (41 of which are cortical, and 93 subcortical) grouped by the 12 main brain regions according to the Allen Reference Atlas, revealed a significant over-expression of genes belonging to Cliques I in the cerebellar cortex (Figures 4A, 4B, Supplementary Figure S1). Genes belonging to Clique II also showed a slight over-expression in the cerebellar cortex (Figures 5A, 5B), as well as in several cortical regions, however these signals were much weaker than the one of clique I. Next, we asked if the over-expression in the cerebellum is a unique property of these two cliques. For that, we examined the neuroanatomical expression of all cliques in Figure 2 and found only seven other cliques showing a similar over-expression in the cerebellar cortex. Interestingly, these cliques had substantial gene overlap with cliques I&II, and were ranked high in their autism gene enrichment scores (Supplementary Table S4), thus supporting the illumination of the cerebellar cortex in this analysis.

Finally, we asked if these two co-expressed cliques are associated with particular cell type or any other functional or cellular property. Examining cell-type-specific microarray data which we

---

**Figure 1A.** Furthermore, we evaluated the connectivity of genes across different co-expression values. Here too, the average size of connected components (See methods) among autism gene was consistently larger than seen in 1000 randomly generated gene sets from the Allen database. **Figure 1B.** The fitting score in the cerebellar cortex is highest among all other brain regions and higher than expected by chance ($P = 0.018$, based on 100,000 random permutations). Some other regions in the cerebral cortex also show a slight deviation from expression values expected by chance ($P < 0.05$). The brain regions of the ABA at 200 micron resolutions (one dot per region on the figure) are grouped into the following main regions: COR (cerebral cortex), OLF (olfactory areas), Hi (hippocampal region), RHi (Retrohippocampal region), STR (striatum), PAL (pallidum), THA (thalamus), HYP (hypothyamus), MID (midbrain), PON (pons), MED (medulla), CER (cerebellum).

doi:10.1371/journal.pcbi.1003128.g005
have for 64 cell types [43] revealed that 4 of them (stellate basket cells, granule cells, oligodendrocytes and Purkinje cells) are considerably populating the cerebellum (Supplementary Figure S2). Further, we looked through different coronal sections through the cerebellum from the Allen Reference Atlas [41] and visually compared them to sections of the sum of expressions of genes in Cliques I & II. Figure 6 shows the normalized volumetric expression quantities of both cliques along with the closest coronal section of the mouse brain in the Allen Reference Atlas. One can see that the voxels with the most intense expression in both cliques tend to follow the granular layer. Hence, the results of these analyses suggest that genes in both clique I & II tend to be over-expressed in granule cells in the cerebellar cortex. Using the Bioconductor GOstats package in R, we two biological processes: “Transmission of nerve impulse” ($P=0.001842$), and “Ion transport” ($P=0.000733$) and one cellular component “Vesicles” ($P=0.001134$) that were enriched with genes from Clique I (Supplementary Table S5). Unfortunately, the number of genes in Clique II was too small for this analysis.

Discussion

In this study, we explored the co-expression network 26 autism genes within the framework of 3,041 genes exhibiting the highest-quality expression data in the Allen Mouse Brain Atlas database [30–33]. The significantly tighter co-expression connectivity among the 26 autism genes than other genes, implies common functional properties for these genes in the mouse brain. Further investigation into the co-expression patterns of these genes revealed two cliques of co-expressed genes that were significantly dominated by autism genes. Genes in both these cliques shown significant over-expression in the cerebellar cortex, and particularly in sections that are predominantly populated by granular cells. Some regions of the cerebral cortex are also highlighted by the second clique (Figure 5), but to a lesser extent than the cerebellar cortex. Another recent study of our group examining the expression of the same autism gene set (AutRef84) in different human tissues, found a statistically significant enrichment in the frontal cortex [44]. The cerebral cortex was highlighted in other neuroanatomical studies of autism in both human [45,46] and mouse [47] and is known to play a central role in cognitive and emotional processing [48], which are key deficits in autism and other neuropsychiatric disorders. In addition, a recent neuroimaging study [49] highlighted functional sub-regions in the cerebellum as playing a role in both motor and cognitive tasks. Other genes associated with autism have been shown to play a role in the development of this region [50–53]. Our results, provide additional support in the potential involvement of the cerebellum in
Co-expression of Autism Genes in the Mouse Brain

Supporting Information

Figure S1  ISH images of the autism genes of Clique I. Imaged sections of ISH-treated brains (close to bregma), for (A) Astn2, (B) Galnt13, and (C) Pchd1 with the cerebellar cortex clearly visible. (TIF)

Figure S2  Estimated brain-wide density for (A) Stellate Basket Cells, (B) Granule Cells, and (C) Mature Oligodendrocytes. These data correspond to microarray data from [56], using the data sets estimated in [43] to have the highest purity. Interestingly, the estimated brain-wide densities are almost zero in the cerebral cortex, suggesting that cell types characterized by their transcriptomes are indeed specific to the cerebellar cortex. (TIF)

Table S1 AutRef84: A Reference set of Rare and Syndromic ASD-linked genes. (XLSX)

Table S2 Co-expression of genes in Clique I. The pairwise co-expression values between all 33 genes in Clique I are depicted and color-coded. The four genes already associated with autism are highlighted in yellow. (XLSX)

Table S3 Co-expression of genes in Clique II. The pairwise co-expression values between all six genes in Clique II are depicted and color-coded. The two genes already associated with autism are highlighted in yellow. (XLSX)

Table S4 Overlap of cliques displaying overexpression in the cerebellum with cliques 1 & II. (DOC)

Table S5 Results of the Bioconductor GOstat analysis for genes belonging to Clique I. (DOC)

Author Contributions

Conceived and designed the experiments: IM PG ECL SBB PPM. Performed the experiments: PG ECL. Analyzed the data: IM PG. Contributed reagents/materials/analysis tools: SBB PPM. Wrote the paper: IM PG.

References

1. Levy SE, Mandell DS, Schultz RT (2009) Autism. Lancet 374: 1627–1638.
2. Lord C (2011) Epidemiology: How common is autism? Nature 474: 166–168.
3. Newschaffer CJ, Coven LA, Daniels J, Giarelli E, Grether JK, et al. (2007) The epidemiology of autism spectrum disorders. Annu Rev Public Health 28: 235–256.
4. Walsh CA, Morrow EM, Rubenstein JL (2008) Autism and brain development. Cell 135: 396–400.
5. Amaral DG, Schumann CM, Nordahl CW (2000) Neuroanatomy of autism. Trends Neurosci 23: 117–121.
6. Voineagu I, Wang X, Johnston P, Lowe JK, Tian Y, et al. (2011) Transcriptomic analysis of autistic brain reveals convergent molecular pathology. Nature 474: 380–384.
7. Carper RA, Courchesne E (2005) Localized enlargement of the frontal cortex in early autism. Biological Psychiatry 57: 126–133.
8. Courchesne E, Pierce K (2005) Why the frontal cortex in autism might be talking only to itself: local over-connectivity but long-distance disconnection. Current Opinion in Neurobiology 15: 225–230.
9. Elliotton M, Foldy C, Sharma M, Tabuchi K, Liu X, et al. (2011) Autism-linked neuroligin-3 R451C mutation differentially alters hippocampal and cortical synaptic function. Proc Natl Acad Sci U S A 108: 13764–13769.
10. Herbert MR (2011) SHANK3, the synapse, and autism. N Engl J Med 365: 175–177.
11. Gilman SR, Iossifov I, Levy D, Ronemus M, Wigler M, et al. (2011) Rare de novo variants associated with autism implicate a large functional network of genes involved in formation and function of synapses. Neuron 70: 890–907.
12. Garber K (2007) Neuroscience - Autism's cause may reside in abnormalities at the synapse. Science 317: 190–191.
13. Bourgeron T (2009) A synaptic trek to autism. Current Opinion in Neurobiology 19: 231–234.
14. Durand CM, Betancur C, Bockers TM, Boeckmann J, Chase P, et al. (2007) Mutations in the gene encoding the synaptic scaffolding protein SHANK1 are associated with autism spectrum disorders. Nat Genet 39: 25–27.
15. Sztatmari P, Paterson AD, Zwaigenbaum L, Roberts W, Brian J, et al. (2007) Mapping autism risk loci using genetic linkage and chromosomal rearrangements. Nat Genet 39: 319–320.
16. Jacquemont ML, Sanlaville D, Redon R, Raoul O, Cormier-Daire V, et al. (2006) Array-based comparative genomic hybridization identifies high frequency of cryptic chromosomal rearrangements in patients with syndromic autism spectrum disorders. J Med Genet 43: 843–849.
17. Cook EH, Jr., Scherer SW (2008) Copy-number variations associated with neuropsychiatric conditions. Nature 455: 919–923.
18. Levy D, Ronemus M, Yamrom B, Lee YH, Loetta A, et al. (2011) Rare de novo and transmitted copy-number variation in autistic spectrum disorders. Neuron 70: 896–897.

autism etiology, and suggest additional candidate genes that are also over-expressed in the cerebellar cortex.

Two recent transcriptomic analyses in human brains [6,34] revealed additional co-expression modules enriched with autism-associated genes. Some of these modules partially overlap with our findings in either gene content or brain regions, suggesting common functional and neuroanatomical properties of autism gene in both human and mouse brains. Together, these studies provide new insights into the specific gene networks and brain regions that could be involved in autism etiology.

A major strength of our study is the utilization of the Allen Mouse Brain Atlas [30–33] which comprises a high-resolution genome-wide exploration of gene expression in the adult mouse brain. This data allows one to explore gene expression properties up to a resolution of 200 microns, which provide a good distinction between different brain regions as well as potentially tell apart different sub-regions and cell types. Another advantage of this study is the focus on those genes exhibiting the highest expression correlation between the coronal and sagittal sections [35] as well as restricting the autism gene to a subset demonstrating, to the best of our knowledge, the most compelling associations to autism susceptibility. These strict criteria reduce the chances of erroneous results. Our study has also some pitfalls. First, the analyses were done on data from mouse brains. Since autism is a human condition, one may ask how well finding of this study apply to human brain. A recent study comparing postsynaptic protein composition between mouse and human suggest a high correlation between these two mammals in those matters [55]. Nevertheless, similar analyses in the human brain are still required to provide a finer validity to our findings. In addition, the strict criteria used here, restricted the number of studied genes to 3, 041 and 26 autism genes which are roughly represent 15% and 31% of the genes in the Allen Brain Atlas and AutDB datasets respectively. Such a small number of genes might results in false negatives and hence might miss other co-expression properties and brain regions associated with autism. Hence, larger studies are needed to complement the results of our analysis.

In conclusions, our study provides unique insights into the neuroanatomical co-expression properties of genes associated with autism in the mouse brain and suggest specific regions implicated in autism etiology. Complementing these findings with additional genomics and neuroimaging analyses from both mouse and human brains would help gaining a broader picture of the autistic brain.
Co-expression of Autism Genes in the Mouse Brain