*Human Molecular Genetics, 2007, Vol. 16, No. 14* ***1682–1698***
*doi:10.1093/hmg/ddm116*
*Advance Access published on May 21, 2007*

# Genome-wide expression profiling of lymphoblastoid cell lines distinguishes different forms of autism and reveals shared pathways†

**Yuhei Nishimura[1,2,3], Christa L. Martin[7], Araceli Vazquez Lopez[2], Sarah J. Spence[1,4,5], Ana Isabel Alvarez-Retuerto[4], Marian Sigman[4,6], Corinna Steindler[8], Sandra Pellegrini[8], N. Carolyn Schanen[9], Stephen T. Warren[7] and Daniel H. Geschwind[1,2,3,4,*]**

[1]Center for Autism Research and Treatment, [2]Program in Neurogenetics, Department of Neurology, [3]Center for Neurobehavioral Genetics, [4]Department of Psychiatry, [5]Department of Pediatrics, [6]Department of Psychology, Semel Institute for Neuroscience and Human Behavior, David Geffen School of Medicine, University of California Los Angeles, Los Angeles, CA 90095, USA, [7]Department of Human Genetics, Emory University, Atlanta, GA 30322, USA, [8]Unite de Signalisation des Cytokines, Institut Pasteur, Paris 75724, France and [9]Center for Pediatric Research, Nemours Biomedical Research, Alfred I. duPont Hospital for Children, Wilmington, DE 19803, USA

Received February 5, 2007; Revised and Accepted March 20, 2007

**Autism is a heterogeneous condition that is likely to result from the combined effects of multiple genetic factors interacting with environmental factors. Given its complexity, the study of autism associated with Mendelian single gene disorders or known chromosomal etiologies provides an important perspective. We used microarray analysis to compare the mRNA expression profile in lymphoblastoid cells from males with autism due to a fragile X mutation (FMR1-FM), or a 15q11–q13 duplication (dup(15q)), and non-autistic controls. Gene expression profiles clearly distinguished autism from controls and separated individuals with autism based on their genetic etiology. We identified 68 genes that were dysregulated in common between autism with FMR1-FM and dup(15q). We also identified a potential molecular link between FMR1-FM and dup(15q), the cytoplasmic *FMR1* interacting protein 1 (*CYFIP1*), which was up-regulated in dup(15q) patients. We were able to confirm this link *in vitro* by showing common regulation of two other dysregulated genes, *JAKMIP1* and *GPR155*, downstream of *FMR1* or *CYFIP1*. We also confirmed the reduction of the *Jakmip1* protein in *Fmr1* knock-out mice, demonstrating *in vivo* relevance. Finally, we showed independent confirmation of roles for *JAKMIP1* and *GPR155* in autism spectrum disorders (ASDs) by showing their differential expression in male sib pairs discordant for idiopathic ASD. These results provide evidence that blood derived lymphoblastoid cells gene expression is likely to be useful for identifying etiological subsets of autism and exploring its pathophysiology.**

## INTRODUCTION

Genetic factors are significant determinants of autism spectrum disorders (ASDs) pathophysiology (1–4). Yet, identification of causal genes has been hampered by genetic and phenotypic heterogeneity (1–18). Thus, it seems reasonable to accelerate the gene discovery process by using combinations of experimental approaches, such as the study of 'single gene' or more simple causes, such as chromosomal copy number imbalances whose phenotypes include ASD (1–4). One such disorder is fragile X syndrome (FXS) (1–4,19), which is caused by an expansion of the trinucleotide repetitive sequence (CGG)n in the promoter region of the fragile X mental retardation 1 (*FMR1*) gene located at Xq27.3 (20). This mutation causes a significant deficit of the *FMR1* protein (FMRP) and a phenotype including cognitive

*To whom correspondence should be addressed at: 710 Westwood Plaza, Los Angeles, CA 90095-1769, USA. Tel: +1 3102066814; Fax: +1 3102672401; Email: dhg@ucla.edu
†The microarray expression data from this study have been submitted to GEO under accession number GSE7329.

impairment and other behavioral abnormalities that overlap with ASD. The prevalence of ASD among FXS cases has been estimated at 15–33% (21,22) and ~1–3% of those with autism and no obvious physical features of FXS are found to have FMR1-FM (19,23).

Another disorder associated with susceptibility for ASD is a maternally inherited duplication of 15q11–q13 (dup(15q)) (1–4,24). Multiple repeat elements within the region mediate a variety of rearrangements, including interstitial duplications, interstitial triplications and supernumerary isodicentric marker chromosomes (25). Dup(15q) occurs with an estimated frequency of 1:600 children with developmental delay (26) and is the most common copy number variation causing ASD (3,24). Several lines of evidence (24,27) suggest that dysregulation of non-imprinted genes in the duplicated region may contribute to the autistic phenotype observed in dup(15q).

Therefore, we reasoned that the identification of genes whose expression is dysregulated by both FMR1-FM and dup(15q) may be relevant to ASD, since the two genetic abnormalities represent cases where single mutations, either a trinucleotide repeat or copy number variation (28), cause ASD. We also wanted to examine, as a proof of principle, whether lymphoblast gene expression profiles identified by microarrays could differentiate these single mutation 'simple' causes of autism from each other and controls. This would provide a basis for further application of these methods in idiopathic autism, where more multigenic inheritance and environmental influences may be involved (1–4).

Recently, several studies have suggested that lymphoblastoid cells can be used to detect biologically plausible correlations between candidate genes and neuropsychiatric diseases, including Rett syndrome (29), non-specific X-linked mental retardation (30), bipolar disorder (31), FXS (32) and dup(15q) (27). In the present study, we investigated whether gene expression profiles of lymphoblastoid cells could be used to: (i) differentiate autistic subjects who were ascertained and diagnosed as having ASD in the Autism Genetic Resource Exchange (AGRE) (33) repository into etiological categories (FMR1-FM and dup(15q)) and (ii) identify common genes and pathways shared by FMR1-FM and dup(15q) that might be relevant to autism pathophysiology. Here, we demonstrate that gene expression profiles were able to clearly distinguish individuals based on their etiology. We also identified 68 genes dysregulated in both autism with FMR1-FM and dup(15q). Interestingly, we identified a molecular connection between FMR1-FM and dup(15q), *CYFIP1*, which was significantly induced in dup(15q) and is known to antagonize certain aspects of FMRP function (34). We further demonstrated that the expression of Janus kinase and microtubule interacting protein 1 (*JAKMIP1*) and G protein-coupled receptor 155 (*GPR155*) were commonly dysregulated by the reduction of *FMR1* or induction of *CYFIP1 in vitro*. The expression of *Jakmip1* was also dysregulated in the brain of the *Fmr1* knock-out mouse. Finally, we were able to show that *JAKMIP1* and *GPR155* were dysregulated in males with ASDs relative to their non-affected siblings, providing independent confirmation to suggest these genes are associated with ASD.

## RESULTS

### Hierarchical clustering and principal component analysis distinguished individuals based on genetic etiology

We analyzed the whole-genome mRNA expression profile in lymphoblastoid cells from 15 autistic males (8 autistic males with FMR1-FM and 7 autistic males with dup(15q)) and 15 non-autistic control males from AGRE (Supplementary Material, Table S1) using Agilent Whole Genome Human Microarrays. Overall, of 41 000 probes analyzed, 31 044 probes, representing 23 822 genes, were expressed in the lymphoblastoid cells. To find genes that were differentially expressed across the three subject groups, the expression profile of the lymphoblastoid cells was subjected to Analysis of Variance (ANOVA) (35). The ANOVA identified 293 probes (277 genes) below a defined false discovery rate (FDR) threshold of 5% (Supplementary Material, Table S2). It has been shown that the expression of *FMR1* is decreased in lymphoblastoid cells with FMR1-FM (36) and that the expression of *UBE3A* is increased in lymphoblastoid cells with dup(15q) (37). Concordant with these reports, *FMR1* and *UBE3A* were among the 293 differentially expressed probes, providing independent controls for the microarray analysis.

As shown in Figure 1A and B, hierarchical clustering using the 293 probes clearly classified individuals based on their genotype. The 293 probes were also subjected to principal component analysis (PCA). As shown in Figure 1C, three dominant PCA components that contained 70% of the variance in the data matrix clearly separated individuals based on genetic etiology. In this plot, the first principal component axis accounted for 56% of the variance in the data set and clearly separated autism with FMR1-FM and dup(15q) from controls, whereas the second principal component (PC2) accounted for 10% of the variance and segregated autism with FMR1-FM from autism with dup(15q). The top 10 genes contributing to PC2 include *FMR1*, *UBE3A*, *CYFIP1*, non-imprinted in Prader-Willi/Angelman syndrome 2 (*NIPA2*) and hect domain and RLD 2 (*HERC2*). The latter four genes are all located in 15q11–q13. These results suggest that the selective reduction of *FMR1* and the selective induction of the four genes located in 15q11–q13 differentiated autism with FMR1-FM from autism with dup(15q). These data provide a critical proof of principle that the gene expression profile of lymphoblastoid cells is useful in the study of autism.

### Microarray analyses revealed the significant overlap of FMR1-FM and dup(15q)

To identify the set of the most robustly differentially expressed genes in each group, we applied three different statistical methods, ANOVA, Significant Analysis of Microarray (SAM) (38) and Rank Product Analysis (RankProd) (39). SAM is a modified *t*-test statistic, whereas RankProd is a non-parametric statistic that detects items that are consistently highly ranked in a number of lists. SAM identified 5139 probes and 1281 probes as significant (FDR $< 5\%$) in autism with FMR1-FM and dup(15q), respectively (Fig. 2A and B, Supplementary Material, Tables S3 and S4). RankProd

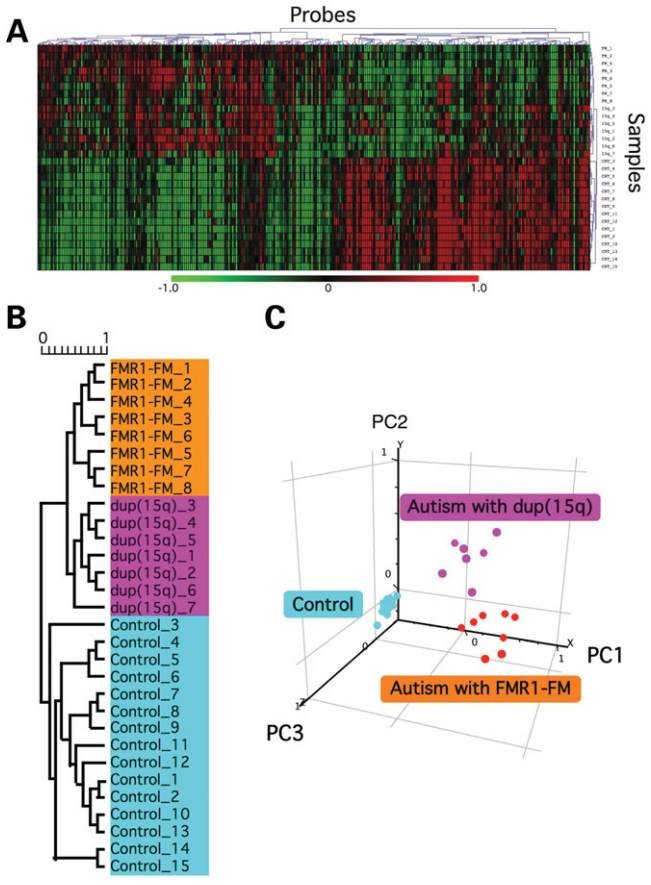


**Figure 1.** Hierarchical clustering and PCA differentiate individuals based on their etiology. ANOVA identified 293 probes with significantly different expression between autism with FMR1-FM ($n = 8$), autism with dup(15q) ($n = 7$) and control ($n = 15$). The probes were subjected to hierarchical clustering and PCA. (**A**) Hierarchical clustering of the 30 individuals and genes. Each row represents an individual and each column represents one of the 293 probes. A pseudo-colored representation of the relative intensity is shown, such that a red color indicates high expression and green color low expression, with the scale shown below. Relative distance of each probe (horizontal axis) and individuals (vertical axis) are also demonstrated. (**B**) Enlargement of the hierarchical clustering dendrogram of the sample in (A). All eight autism with FMR1-FM, seven autism with dup(15q) and 15 controls correctly clustered within their etiological categories. The scale shows the Spearman rank correlation coefficient used to construct the dendrogram. (**C**) PCA of the expression profile of the 293 probes from 30 individuals. Shown here are three principal components. Autism with FMR1-FM are depicted as orange, autism with dup(15q) as magenta and control as cyan. The individuals are clustered according to their genetic etiologies.

identified 2281 probes and 1444 probes as significant (FDR < 5%) in autism with FMR1-FM and dup(15q), respectively (Fig. 2A and B, Supplementary Material, Tables S5 and S6). The combination of ANOVA, SAM and RankProd identified 146 probes (120 genes) in autism with FMR1-FM and 97 probes (80 genes) in autism with dup(15q) (Fig. 2C). Eighty-three probes representing 68 genes were dysregulated in both autism with FMR1-FM and with dup(15q) (Table 1). This degree of overlap was highly significant (hypergeometric probability, $P = 1.2 \times 10^{-153}$). Fifty-two genes and 12 genes were selectively dysregulated in either autism with FMR1-FM and autism with dup(15q), respectively (Table 2), with fold changes ranging between 0.54 and 1.98 fold.

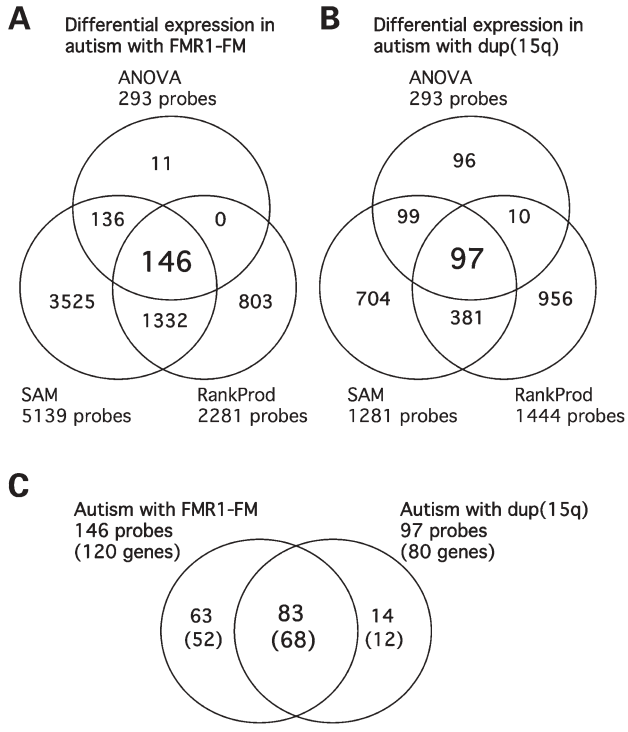


**Figure 2.** Differentially expressed probes identified by three different statistical methods, ANOVA, SAM and RankProd. Venn diagram showing the number of probes identified as differentially expressed between (**A**) autism with FMR1-FM ($n = 8$) and control ($n = 15$) and (**B**) autism with dup(15q) ($n = 7$) and control ($n = 15$). (**C**) Overlap of the differentially expressed probes (genes) in autism with FMR1-FM and dup(15q).

## Quantitative real-time PCR confirmed the differential expression identified by the microarray analysis

To validate the differential expression identified by microarray analysis using independent methods, we performed quantitative real-time PCR (qRTPCR) analysis of 19 genes chosen as a cross-section using the same samples used in the microarray analysis. The qRTPCR confirmed that 17 of the 19 genes were differentially expressed as expected by the microarray analysis (Fig. 3A–C). There was an overall highly significant correlation between microarray and qRTPCR results (Pearson correlation, $r = 0.57$, $P < 0.0001$). Fold changes observed with qRTPCR were typically higher that the smaller fold change observed on the arrays.

*CYFIP1* was one of the genes selectively induced in autism with dup(15q). Because *CYFIP1* is known to antagonize FMRP (34), we reasoned that the induction of *CYFIP1* in dup(15q) might explain some of the significant overlap between autism with FMR1-FM and with dup(15q). *JAKMIP1*, also known as *MARLIN-1*, was significantly induced in autism with FMR1-FM and had a positive trend in autism with dup(15q) ($P = 0.062$), suggesting that *JAKMIP1* could represent a commonly dysregulated pathway. In fact, RankProd identified *JAKMIP1* as a significantly up-regulated gene in dup(15q) by microarray analysis (Supplementary Material, Table S6). This gene is a particularly biologically important candidate, given its putative role in $GABA_B$ receptor expression (40) and microtubule networks (41).

**Table 1.** Genes dysregulated in both autism with FMR1-FM and dup(15q)

| Symbol | Gene product | Refseq ID | *P*-value[a] | FC (FM/C)[b] | FC (dup/C)[c] | Gene loci | Autism loci[d] | Reference |
|---|---|---|---|---|---|---|---|---|
| *G1P2* | Interferon alpha-inducible protein (clone IFI-15K) | NM_005101 | 2.2E-06 | 0.79 | 0.85 | 1p36.33 | | |
| *TNFRSF8* | Tumor necrosis factor receptor superfamily member 8 | NM_001243 | 1.5E-08 | 1.23 | 1.27 | 1p36.22 | 1p36.23–1q36.13 | 12 |
| *RIMS3* | Regulating synaptic membrane exocytosis 3 | NM_014747 | 1.9E-06 | 0.66 | 0.75 | 1p22.2 | | |
| *FAM46C* | Family with sequence similarity 46 member C | NM_017709 | 5.3E-06 | 0.76 | 0.82 | 1p12 | 1p13–1q12 | 17 |
| *HIST2H3C* | Histone 2 H3c | NM_021059 | 2.3E-06 | 0.77 | 0.84 | 1q21.3 | | |
| *C1orf61* | Chromosome 1 open reading frame 61 | NM_006365 | 7.5E-06 | 1.13 | 1.20 | 1q22 | | |
| *SLAMF1* | Signaling lymphocytic activation molecule family member 1 | NM_003037 | 3.4E-06 | 1.19 | 1.17 | 1q23.3 | | |
| *SELL* | Selectin L (lymphocyte adhesion molecule 1) | NM_000655 | 1.5E-06 | 0.70 | 0.79 | 1q24.2 | | |
| *DYRK3* | Dual-specificity tyrosine-phosphorylation regulated kinase 3 | NM_003582 | 9.5E-06 | 1.21 | 1.18 | 1q32.1 | | |
| *RRM2* | Ribonucleotide reductase M2 polypeptide | NM_001034 | 4.4E-07 | 0.82 | 0.81 | 2p25.1 | | |
| *RSNL2* | Restin-like 2 | NM_024692 | 1.4E-07 | 1.23 | 1.13 | 2p23.2 | | |
| *FAM82A* | Family with sequence similarity 82 member A | NM_144713 | 4.1E-07 | 1.47 | 1.35 | 2p22.2 | | |
| *HNRPLL* | Heterogeneous nuclear ribonucleoprotein L-like | NM_138394 | 2.8E-06 | 1.35 | 1.37 | 2p22.1 | | |
| *SLC20A1* | Solute carrier family 20 (phosphate transporter) member 1 | NM_005415 | 1.0E-07 | 1.13 | 1.14 | 2q13 | | |
| *BBS5* | Bardet–Biedl syndrome 5 | NM_152384 | 7.0E-06 | 1.11 | 1.16 | 2q31.1 | | |
| *GPR155* | G protein-coupled receptor 155 | NM_001033045 | 1.0E-05 | 0.82 | 0.83 | 2q31.1 | 2q31.1–2q33.3 | 6,8 |
| *WNT6* | Wingless-type MMTV integration site family member 6 | NM_006522 | 1.0E-05 | 0.86 | 0.87 | 2q35 | | |
| *CMKOR1* | Chemokine orphan receptor 1 | NM_020311 | 6.5E-08 | 1.40 | 1.57 | 2q37.3 | | |
| *LZTFL1* | Leucine zipper transcription factor-like 1 | NM_020347 | 3.4E-06 | 1.20 | 1.23 | 3p21.3 | | |
| *GCET2* | Germinal center expressed transcript 2 | NM_152785 | 5.8E-09 | 1.50 | 1.52 | 3q13.2 | | |
| *GHSR* | Growth hormone secretagogue receptor | NM_004122 | 2.9E-06 | 1.29 | 1.26 | 3q26.31 | | |
| *AREG* | Amphiregulin (schwannoma-derived growth factor) | NM_001657 | 9.0E-07 | 0.77 | 0.81 | 4q13.3 | | |
| *TCF7* | Transcription factor 7 (T-cell specific HMG-box) | NM_003202 | 1.3E-06 | 1.18 | 1.44 | 5q31.1 | 5q14.3–5q31.2 | 12 |
| *NRG2* | Neuregulin 2 | NM_004883 | 1.3E-05 | 1.36 | 1.57 | 5q31.2 | | |
| *NDFIP1* | Nedd4 family interacting protein 1 | NM_030571 | 1.7E-08 | 0.86 | 0.87 | 5q31.3 | | |
| *NR3C1* | Nuclear receptor subfamily 3 group C member 1 | NM_001018077 | 2.8E-07 | 1.13 | 1.14 | 5q31.3 | | |
| *NID67* | Putative small membrane protein NID67 | NM_032947 | 7.3E-07 | 1.31 | 1.21 | 5q33.1 | | |
| *LCP2* | Lymphocyte cytosolic protein 2 | NM_005565 | 3.4E-07 | 1.16 | 1.16 | 5q35.1 | | |
| *TXNDC5* | Thioredoxin domain containing 5 | NM_030810 | 5.9E-06 | 0.82 | 0.81 | 6p24.3 | | |
| *PHACTR1* | Phosphatase and actin regulator 1 | NM_030948 | 1.5E-07 | 1.26 | 1.23 | 6p24.1 | | |
| *PAQR8* | Progestin and adipoQ receptor family member VIII | NM_133367 | 1.6E-07 | 1.15 | 1.13 | 6p12.1 | | |
| *CHST12* | Carbohydrate (chondroitin 4) sulfotransferase 12 | NM_018641 | 1.8E-07 | 0.81 | 0.87 | 7p22 | | |
| *PSCD3* | Pleckstrin homology Sec7 and coiled-coil domains 3 | NM_004227 | 7.5E-06 | 1.13 | 1.27 | 7p22.1 | | |
| *C1GALT1* | Core 1 synthase galactosyltransferase 1 | NM_020156 | 3.9E-06 | 0.88 | 0.86 | 7p14–p13 | | |
| *RASA4* | RAS p21 protein activator 4 | NM_006989 | 5.0E-06 | 1.17 | 1.16 | 7q22.1 | 7q22.1–7q31.31 | 7 |
| *HIG2* | Hypoxia-inducible protein 2 | NM_013332 | 2.9E-08 | 1.24 | 1.12 | 7q32.2 | | |
| *CLDN23* | Claudin 23 | NM_194284 | 2.3E-07 | 0.79 | 0.82 | 8p23.1 | | |
| *TRAM1* | Translocation associated membrane protein 1 | NM_014294 | 7.1E-06 | 0.88 | 0.87 | 8q13.3 | | |
| *DENND3* | KIAA0870 protein | NM_014957 | 1.1E-05 | 1.16 | 1.22 | 8q24.3 | | |
| *TRAF1* | TNF receptor-associated factor 1 | NM_005658 | 1.4E-05 | 1.17 | 1.21 | 9q33–q34 | | |
| *KIAA0649* | KIAA0649 | NM_014811 | 4.8E-07 | 0.66 | 0.82 | 9q34.3 | | |
| *VIM* | Vimentin | NM_003380 | 3.0E-06 | 0.67 | 0.75 | 10p13 | | |
| *PTPLA* | Protein tyrosine phosphatase-like member a | NM_014241 | 1.1E-06 | 0.72 | 0.74 | 10p14–p13 | | |
| *RASSF4* | Ras association (RalGDS/AF-6) domain family 4 | NM_032023 | 1.6E-06 | 1.25 | 1.39 | 10q11.21 | | |
| *ACADSB* | Acyl-Coenzyme A dehydrogenase short/branched chain | NM_001609 | 3.0E-08 | 1.20 | 1.13 | 10q25–q26 | | |
| *APBB1* | Amyloid beta precursor protein-binding family B member 1 | NM_001164 | 1.2E-06 | 1.26 | 1.23 | 11p15 | | |
| *PRICKLE1* | Prickle-like 1 (*Drosophila*) | NM_153026 | 3.3E-07 | 0.82 | 0.81 | 12q12 | | |
| *SAV1* | Salvador homolog 1 (*Drosophila*) | NM_021818 | 1.6E-07 | 1.15 | 1.15 | 14q13–q23 | | |
| *ACTN1* | Actinin alpha 1 | NM_001102 | 4.3E-07 | 0.70 | 0.76 | 14q24.1–q24.2 | | |
| *JDP2* | Jun dimerization protein 2 | NM_130469 | 6.5E-06 | 1.19 | 1.20 | 14q24.3 | | |

*Continued*

**Table 1.** *Continued*

| Symbol | Gene product | Refseq ID | *P*-value[a] | FC (FM/C)[b] | FC (dup/C)[c] | Gene loci | Autism loci[d] | Reference |
|---|---|---|---|---|---|---|---|---|
| *BATF* | Basic leucine zipper transcription factor ATF-like | NM_006399 | 8.1E-06 | 1.11 | 1.17 | 14q24.3 | | |
| *KIAA1370* | Hypothetical protein FLJ10980 | NM_019600 | 2.0E-08 | 0.80 | 0.86 | 15q21.2–q21.3 | | |
| *SV2B* | Synaptic vesicle glycoprotein 2B | NM_014848 | 2.3E-07 | 1.51 | 1.42 | 15q26.1 | | |
| *CIITA* | MHC class II transactivator | NM_000246 | 7.0E-07 | 1.21 | 1.19 | 16p13 | 16p13.2–16p13.13 | 5 |
| *SNN* | Stannin | NM_003498 | 4.9E-07 | 1.19 | 1.21 | 16p13 | 16p13 | 17 |
| *KIAA0251* | KIAA0251 | NM_015027 | 4.1E-06 | 0.78 | 0.84 | 16p13.11 | | |
| *PRKCB1* | Protein kinase C beta 1 | NM_212535 | 6.3E-06 | 1.21 | 1.21 | 16p11.2 | 16p11.2 | 15 |
| *IL4R* | Interleukin 4 receptor | NM_000418 | 1.8E-07 | 1.13 | 1.15 | 16p12.1 | | |
| *IL21R* | Interleukin 21 receptor | NM_021798 | 1.5E-07 | 1.14 | 1.13 | 16p11 | | |
| *CCL17* | Chemokine (C-C motif) ligand 17 | NM_002987 | 3.6E-06 | 1.38 | 1.58 | 16q13 | | |
| *CCL22* | Chemokine (C-C motif) ligand 22 | NM_002990 | 1.5E-05 | 1.41 | 1.42 | 16q13 | | |
| *FLJ35773* | Hypothetical protein FLJ35773 | NM_152599 | 1.9E-06 | 1.23 | 1.16 | 17p13.1 | | |
| *RAB11FIP4* | RAB11 family interacting protein 4 (class II) | NM_032932 | 1.3E-05 | 1.11 | 1.17 | 17q11.2 | 17q11.2–17q12 | 11,12 |
| *RASL10B* | RAS-like family 10 member B | NM_033315 | 1.2E-08 | 0.88 | 0.84 | 17q12 | 17q11.2–17q12 | 12 |
| *ARHGAP23* | Rho GTPase activating protein 23 | XM_290799 | 4.2E-07 | 0.71 | 0.74 | 17q12 | 17q11.2–17q12 | 12 |
| *PITPNC1* | Phosphatidylinositol transfer protein cytoplasmic 1 | NM_012417 | 4.7E-06 | 1.23 | 1.41 | 17q24.2 | | |
| *SLC16A6* | Solute carrier family 16 member 6 | NM_004694 | 3.7E-08 | 0.84 | 0.85 | 17q24.2 | | |
| *MIA* | Melanoma inhibitory activity | NM_006533 | 2.3E-06 | 1.36 | 1.29 | 19q13.2 | 19q13.11–19q13.33 | 5 |

[a]*P*-value was calculated by one-way ANOVA using controls ($N = 15$), autism with FMR1-FM ($N = 8$) and autism with dup(15q) ($N = 7$).
[b]FC was calculated between mean values of controls ($N = 15$) and autism with FMR1-FM ($N = 8$).
[c]FC was calculated between mean values of controls ($N = 15$) and autism with dup(15q) ($N = 7$).
[d]Autism loci identified by other genetic studies were shown with references.

### Functional annotation revealed pathway dysregulation

In an attempt to uncover common fractions among the differentially expressed genes, we classified genes into gene ontology groups using DAVID (42). Table 3 shows the top three clusters identified by DAVID using the 68 genes dysregulated in autism with FMR1-FM and dup(15q) or the 52 genes selectively dysregulated in autism with FMR1-FM. The number of genes selectively dysregulated in autism with dup(15q) was too small to analyze using functional annotation clustering.

Genes related to cell communication ($P = 7.6 \times 10^{-6}$) and signal transduction ($P = 2.2 \times 10^{-5}$) were most significantly enriched in the 68 genes commonly dysregulated in autism with FMR1-FM and dup(15q). Genes related to immune response ($P = 3.7 \times 10^{-3}$) and defense response ($P = 7.3 \times 10^{-3}$) were also enriched in this gene set. Genes related to chaperone ($P = 2.6 \times 10^{-2}$) and protein folding ($P = 3.2 \times 10^{-2}$) were enriched in the 52 genes selectively dysregulated in autism with FMR1-FM. Genes related to RNA binding ($P = 1.2 \times 10^{-2}$) and mRNA metabolism ($P = 2.1 \times 10^{-2}$) were also enriched in this gene set, consistent with the FMRP protein's function as an RNA binding protein important in regulatory translation (43). Chaperones and folding proteins are commonly found to operate co-translationally, providing a potential link with FMRP function.

To provide a more refined functional classification of genes, we used Ingenuity Pathway Analysis (IPA) (44), a powerful tool for investigating the biological pathways represented by the genes commonly dysregulated in autism with FMR1-FM and dup(15q). IPA uses known protein–protein and gene–gene interactions that have been culled into a curated database and associates the list of differentially expressed genes with biological networks. IPA identified three statistically significant networks, each containing at least 10 genes (Table 4, Supplementary Material, Fig. S1). Principal functions associated with these networks were cell cycle ($P = 5.2 \times 10^{-8}$), cellular movement ($P = 1.3 \times 10^{-8}$) and cell-to-cell signaling and interaction ($P = 4.3 \times 10^{-8}$). The 'cell-to-cell signaling and interaction' was consistent with 'cell communication' and 'signal transduction' categories identified by DAVID. The identification of the 'molecular transport' pathway containing *JAKMIP1* was particularly salient, given this gene's known role in GABAR trafficking within neurons (41). There were also other important genes in this pathway, including *PSCD3*, an ADP-ribosylation factor of unknown CNS function, and *ACTN1*, a cytoskeletal anchoring protein. *JAKMIP1* may act along with these genes in the segregation of signaling complexes involved in neural transmission.

### Effect of downregulation of *FMR1* and up-regulation of *CYFIP1* in a neuronal cell on the expression of the dysregulated genes identified in lymphoblastoid cells

Although we identified dysregulated genes in autism with FMR1-FM and dup(15q) using lymphoblastoid cells, we were interested in whether the expression of these genes would also be dependent on *FMR1* and *CYFIP1* in neuronal cells. To examine the effect of *FMR1* and *CYFIP1* in neuronal cells, we used the well-characterized human neuronal cell line SH-SY5Y (45). *FMR1* and *CYFIP1* dependence in SH-SY5Y

**Table 2.** Genes selectively dysregulated in autism with FMR1-FM or dup(15q)

| Symbol | Gene product | Refseq ID | *P*-value[a] | FC (FM/C)[b] | Gene loci |
|---|---|---|---|---|---|
| Genes selectively dysregulated in autism with FMR1-FM | | | | | |
| *EDG1* | Endothelial differentiation sphingolipid G-protein-coupled receptor 1 | NM_001400 | 7.7E-06 | 0.84 | 1p21 |
| *LRRC8D* | Leucine rich repeat containing 8 family member D | NM_018103 | 3.7E-06 | 0.88 | 1p22.2 |
| *AK2* | Adenylate kinase 2 | NM_001625 | 5.4E-06 | 1.11 | 1p34 |
| *TRIT1* | tRNA isopentenyltransferase 1 | NM_017646 | 3.6E-07 | 1.15 | 1p35.3–p34.1 |
| *PSEN2* | Presenilin 2 (Alzheimer disease 4) | NM_000447 | 1.2E-06 | 0.86 | 1q31–q42 |
| *FAM89A* | Chromosome 1 open reading frame 153 | NM_198552 | 4.1E-06 | 1.23 | 1q42.2 |
| *PRKRA* | Protein kinase interferon-inducible dsRNA dependent activator | NM_003690 | 1.1E-06 | 1.14 | 2q31.2 |
| *KLF7* | Kruppel-like factor 7 (ubiquitous) | NM_003709 | 5.9E-06 | 1.26 | 2q32 |
| *APPL* | Adaptor protein containing pH PTB domain and leucine zipper motif 1 | NM_012096 | 1.7E-06 | 1.21 | 3p21.1–p14.3 |
| *KALRN* | Kalirin RhoGEF kinase | NM_001024660 | 7.8E-06 | 1.40 | 3q21.1–q21.2 |
| *ZCCHC4* | Zinc finger CCHC domain containing 4 | XM_376310 | 1.4E-05 | 1.13 | 4p15.2 |
| *JAKMIP1* | Janus kinase and microtubule interacting protein 1 | NM_144720 | 5.5E-06 | 1.98 | 4p16.1 |
| *AASDH* | 2-aminoadipic 6-semialdehyde dehydrogenase | NM_181806 | 8.2E-06 | 1.15 | 4q12 |
| *SCARB2* | Scavenger receptor class B member 2 | NM_005506 | 5.7E-06 | 0.78 | 4q21.1 |
| *ALPK1* | Alpha-kinase 1 | NM_025144 | 4.8E-06 | 1.22 | 4q25 |
| *SUB1* | Activated RNA polymerase II transcription cofactor 4 | NM_006713 | 3.8E-07 | 0.88 | 5p13.3 |
| *MAN2A1* | Mannosidase alpha class 2A member 1 | NM_002372 | 2.1E-06 | 0.74 | 5q21–q22 |
| *HNRPA0* | Heterogeneous nuclear ribonucleoprotein A0 | NM_006805 | 4.5E-06 | 1.12 | 5q31 |
| *DNAJC18* | DnaJ (Hsp40) homolog subfamily C member 18 | NM_152686 | 3.7E-06 | 1.20 | 5q31.2 |
| *HIST1H3D* | Histone 1 H3d | NM_003530 | 1.5E-06 | 0.78 | 6p21.3 |
| *HIST1H3H* | Histone 1 H3 h | NM_003536 | 6.5E-06 | 0.83 | 6p22–p21.3 |
| *CDC40* | Cell division cycle 40 homolog (yeast) | NM_015891 | 1.3E-06 | 1.12 | 6q21 |
| *SEC63* | SEC63-like (*S. cerevisiae*) | NM_007214 | 6.0E-06 | 1.11 | 6q21 |
| *RABGEF1* | RAB guanine nucleotide exchange factor (GEF) 1 | NM_014504 | 7.3E-06 | 1.12 | 7q11.21 |
| *STAG3* | Stromal antigen 3 | NM_012447 | 7.0E-06 | 1.39 | 7q22.1 |
| *STAR* | Steroidogenic acute regulator | NM_000349 | 4.2E-06 | 1.32 | 8p11.2 |
| *TRIM14* | Tripartite motif-containing 14 | NM_014788 | 5.1E-06 | 1.15 | 9q22.33 |
| *KIAA0368* | KIAA0368 | XM_001129450 | 8.4E-06 | 1.11 | 9q31.3 |
| *SEPHS1* | Selenophosphate synthetase 1 | NM_012247 | 5.7E-06 | 0.84 | 10p14 |
| *ANXA11* | Annexin A11 | NM_001157 | 9.6E-06 | 0.87 | 10q23 |
| *RGS10* | Regulator of G-protein signaling 10 | NM_001005339 | 9.0E-06 | 1.12 | 10q25 |
| *C10orf137* | Chromosome 10 open reading frame 137 | NM_015608 | 2.3E-06 | 0.88 | 10q26.13–q26.2 |
| *SPTY2D1* | Hypothetical protein FLJ39441 | NM_194285 | 2.7E-06 | 0.87 | 11p15.1 |
| *BIRC3* | Baculoviral IAP repeat-containing 3 | NM_001165 | 1.1E-05 | 1.19 | 11q22 |
| *HSP90B1* | Heat shock protein 90 kDa beta (Grp94) member 1 | NM_003299 | 2.7E-06 | 0.88 | 12q24.2–q24.3 |
| *RSN* | Reed–Steinberg cell-expressed intermediate filament-associated protein | NM_002956 | 8.3E-06 | 1.18 | 12q24.3 |
| *MGC7036* | Hypothetical protein MGC7036 | NM_145058 | 6.0E-06 | 0.88 | 12q24.31 |
| *GNG2* | Guanine nucleotide binding protein (G protein) gamma 2 | NM_053064 | 1.2E-05 | 1.15 | 14q21 |
| *SEL1L* | Sel-1 suppressor of lin-12-like (*C. elegans*) | NM_005065 | 1.8E-07 | 0.88 | 14q24.3–q31 |
| *KIAA1370* | Hypothetical protein FLJ10980 | NM_019600 | 4.8E-07 | 0.85 | 15q21.2–q21.3 |
| *MBTPS1* | Membrane-bound transcription factor protease site 1 | NM_003791 | 1.7E-09 | 0.88 | 16q24 |
| *MYO18A* | Myosin XVIIIA | NM_078471 | 8.3E-06 | 0.86 | 17q11.2 |
| *ZFP36* | Zinc finger protein 36 C3H type homolog (mouse) | NM_003407 | 7.0E-06 | 0.87 | 19q13.1 |
| *FTL* | Ferritin light polypeptide | NM_000146 | 1.0E-07 | 0.88 | 19q13.3–q13.4 |
| *ZNF160* | Zinc finger protein 160 | NM_033288 | 6.9E-06 | 1.13 | 19q13.41 |
| *HSPBP1* | hsp70-interacting protein | NM_012267 | 3.2E-06 | 0.87 | 19q13.42 |
| *CST3* | Cystatin C (amyloid angiopathy and cerebral hemorrhage) | NM_000099 | 1.1E-08 | 0.78 | 20p11.21 |
| *CST5* | Cystatin D | NM_001900 | 3.7E-07 | 0.83 | 20p11.21 |
| *C20orf35* | Chromosome 20 open reading frame 35 | NM_018478 | 6.2E-06 | 0.87 | 20q13.12 |
| *LRP5L* | Hypothetical protein DKFZp434O0213 | NM_182492 | 5.2E-06 | 1.11 | 22q11.23 |
| *OSM* | Oncostatin M | NM_020530 | 1.2E-05 | 0.86 | 22q12.2 |
| *FMR1* | Fragile X mental retardation 1 | NM_002024 | 4.1E-15 | 0.54 | Xq27.3 |
| Genes selectively dysregulated in autism with dup(15q) | | | | | |
| | | | | FC(dup/C)[c] | |
| *BRP44* | Brain protein 44 | NM_015415 | 1.6E-07 | 0.87 | 1q24 |
| *CIB4* | Calcium and integrin binding family member 4 | NM_001029881 | 4.0E-06 | 1.16 | 2p23.3 |
| *PTMA* | Prothymosin alpha | NM_002823 | 8.4E-07 | 0.87 | 2q35–q36 |

*Continued*

**Table 2.** *Continued*

| Symbol | Gene product | Refseq ID | *P*-value[a] | FC (FM/C)[b] | Gene loci |
|---|---|---|---|---|---|
| *FYTTD1* | Forty-two-three domain containing 1 | NM_001011537 | 4.8E-06 | 1.14 | 3q29 |
| *PSIP1* | PC4 and SFRS1 interacting protein 1 | NM_021144 | 8.8E-09 | 1.18 | 9p22.3 |
| *RCL1* | RNA terminal phosphate cyclase-like 1 | NM_005772 | 1.5E-07 | 1.12 | 9p24.1–p23 |
| *PPM1A* | Protein phosphatase 1A magnesium-dependent alpha isoform | NM_021003 | 5.8E-06 | 1.12 | 14q23.1 |
| *CYFIP1* | cytoplasmic FMR1 interacting protein 1 | NM_014608 | 1.1E-09 | 1.23 | 15q11.2 |
| *NIPA2* | Non-imprinted in Prader-Willi/Angelman syndrome 2 | NM_001008860 | 9.7E-08 | 1.14 | 15q11.2 |
| *GOLGA8E* | Golgi autoantigen golgin family member | NM_001012423 | 1.9E-06 | 1.17 | 15q11.2 |
| *UBE3A* | Ubiquitin protein ligase E3A | NM_130838 | 1.6E-09 | 1.18 | 15q11.2–15q12 |
| *ICAM1* | Intercellular adhesion molecule 1 human rhinovirus receptor | NM_000201 | 3.1E-06 | 1.13 | 19p13.3–p13.2 |

[a]*P*-value was calculated by one-way ANOVA using controls ($N = 15$), autism with FMR1-FM ($N = 8$) and autism with dup(15q) ($N = 7$).
[b]FC was calculated between mean values of controls ($N = 15$) and autism with FMR1-FM ($N = 8$).
[c]FC was calculated between mean values of controls ($N = 15$) and autism with dup(15q) ($N = 7$).

cells was assessed using a short hairpin RNA (shRNA) to reduce the expression of *FMR1* and a plasmid expression vector to induce the expression of *CYFIP1*, respectively. As shown in Figure 4A, the expression of *FMR1* was reduced to ~60% of its normal level in SH-SY5Y cells stably expressing *FMR1* shRNAs, whereas the expression of *CYFIP1* was significantly induced (11-fold) in SH-SY5Y cells stably transfected with the *CYFIP1* plasmid.

We were able to further demonstrate the effect of downregulation of *FMR1* and up-regulation of *CYFIP1* on the expression of two key downstream genes (Fig. 4B). In SH-SY5Y cells transfected with *FMR1* shRNA, the expression of *JAKMIP1* and *GPR155* was significantly reduced and induced, respectively. In SH-SY5Y cells over-expressing *CYFIP1*, the expression of *JAKMIP1* and *GPR155* was also reduced and induced, respectively. These findings demonstrated that the expression of *JAKMIP1* and *GPR155* was also dependent on *FMR1* and *CYFIP1* in SH-SY5Y cells and that reduction of *FMR1* and induction of *CYFIP1* can share common downstream effects on the expression of *JAKMIP1* and *GPR155*.

### The expression of JAKMIP1 protein was dependent on *FMR1* and *CYFIP1*

We next validated the effect of *FMR1* or *CYFIP1* on the protein expression of JAKMIP1 in the central nervous system (CNS). We examined the expression of the JAKMIP1 protein in the cortex of *Fmr1* knock-out (KO) and wild-type (WT) mice and SH-SY5Y cells transfected with the *CYFIP1* over-expression plasmid. The expression of Jakmip1 was reduced in the cortex of *Fmr1* KO mice (Fig. 5A) and SH-SY5Y cells over-expressing CYFIP1 (Fig. 5B). These results confirmed the *in vitro* findings that the expression of Jakmip1 was dependent on *Fmr1* in mouse brain, suggesting that at least some of the changes observed in lymphoblastoid cells reflect similar changes in the CNS.

### The expression of *JAKMIP1* and *GPR155* was significantly different between 27 male sib pairs discordant for idiopathic ASD

To determine the potential generalizability of these findings to idiopathic autism, we examined whether the expression of *JAKMIP1* and *GPR155* was also dysregulated in lymphoblastoid cells from idiopathic ASD cases. We selected 27 male sib pairs discordant for ASD from AGRE (Supplementary Material, Table S1). The 27 males with ASD did not have FMR1-FM or dup(15q) and had surrogate IQ markers (Raven's progressive matrices) >70. As shown in Figure 6, the expression of *JAKMIP1* and *GPR155* was significantly dysregulated in the 27 males with ASD when compared with their sibs without ASD. These results show that the dysregulation of *JAKMIP1* and *GPR155* is associated with ASD. The lack of general intellectual disability in this ASD group also shows that these dysregulations are not simply due to a non-specific cognitive impairment or intellectual disability observed in FXS and dup(15q). However, in both *in vitro* (SH-SY5Y cells) and *in vivo* (brain) CNS tissues, the directions of *JAKMIP1* and *GPR155* regulation were opposite to that observed in lymphoblastoid cells. The differences may reflect many facts, including immortalization or alternative regulatory signaling pathways in different tissues. However, these data are consistent between FMR1-FM and dup(15q) and indicate that expression of *JAKMIP1* and *GPR155* is regulated by both *FMR1* and *CYFIP1* levels, although differently between neural tissues and lymphoblastoid cells, providing potential common signaling pathways dysregulated in ASD (Fig. 7).

## DISCUSSION

Autism is a heterogeneous condition and likely results from the combined effects of multiple genetic changes including copy number variations and single nucleotide polymorphisms, interacting with environmental factors (1–4). Classification of autistic patients on the basis of genotypic and phenotypic information is one effective way to identify more homogeneous subgroups and hasten the identification of genes underlying autism (1–4). Approximately 3% of autistic children have either FMR1-FM or dup(15q); these patients comprise homogeneous populations for investigation.

In this study, we performed global mRNA expression profiling in males with autism carrying either FMR1-FM or dup(15q) and in control males. We found that these autistic individuals can be differentiated based on their genetic etiologies based on lymphoblast gene expression profiles. Interestingly, this analysis also revealed a common gene expression signature

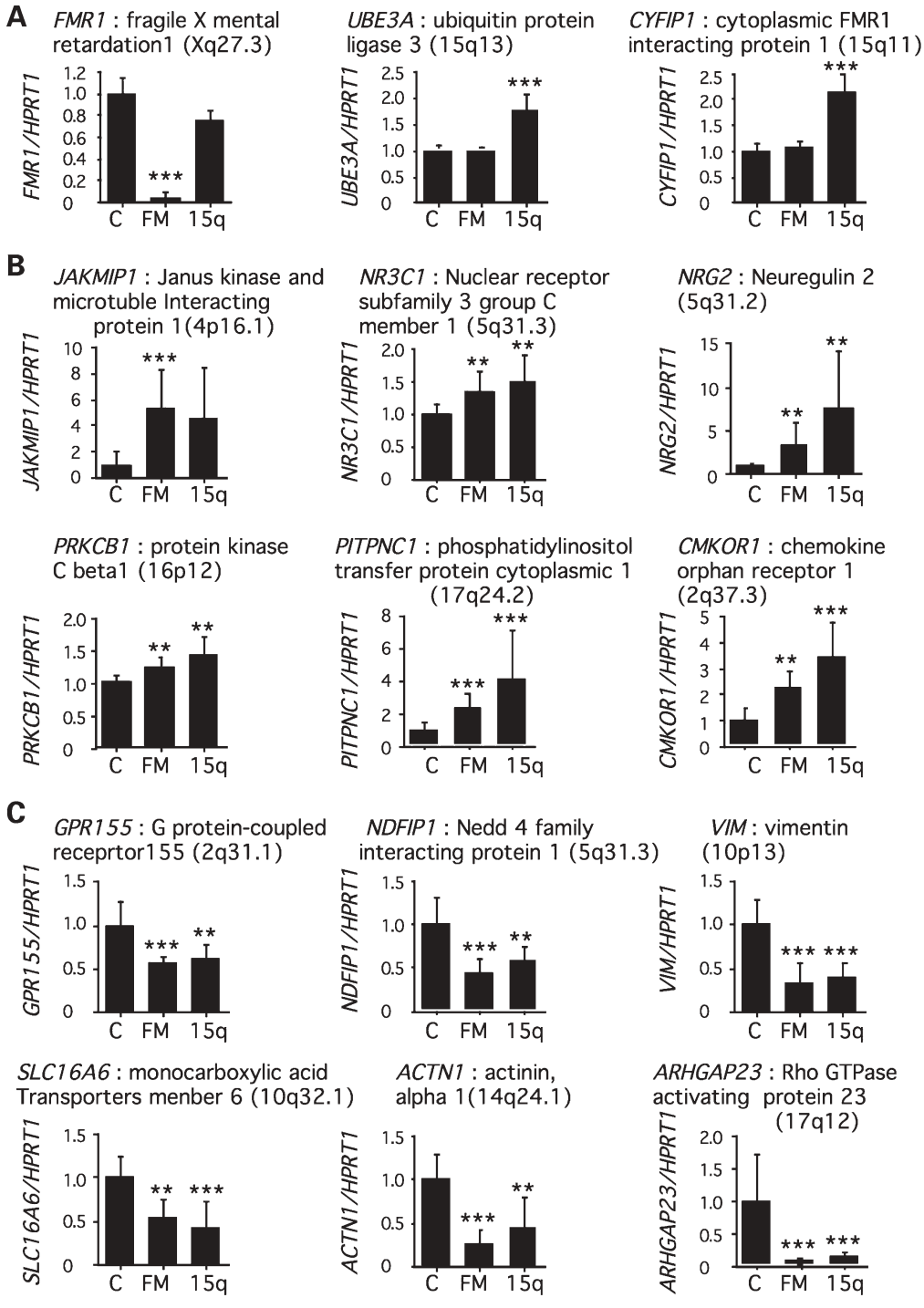


**Figure 3.** Confirmation of the differential gene expression by qRTPCR. Total RNA was extracted from lymphoblastoid cells with FMR1-FM, dup(15q) or control and qRTPCR was performed to confirm the differential expression identified by microarray analysis. (**A**) Genes specifically dysregulated in autism with FMR1-FM or dup(15q). (**B**) Genes up-regulated in both autism with FMR1-FM and dup(15q). (**C**) Genes downregulated in both autism with FMR1-FM and dup(15q). Results represent means ± SD of each group. The mean of the value of control subjects was set as 1. *P*-value was calculated by Mann–Whitney *U* test using control ($N = 15$) versus autism with FMR1-FM ($N = 8$) or autism with dup(15q) ($N = 7$). $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$.

across these two distinct genetic conditions that was significantly different from control profiles. We used the intersection of three different statistical tests to identify the most robustly differentially expressed genes (35,38,39), and the qRTPCR data confirmed this gene selection strategy.

## Gene expression profiles of lymphoblastoid cells carrying the FMR1-FM

We identified 120 genes differentially expressed in FMR1-FM carriers compared with controls. Among these genes, *NR3C1*

**Table 3.** Functional annotation clustering using the 68 genes dysregulated in both autism with FMR1-FM and dup(15q) or the 52 genes selectively dysregulated in autism with FMR1-FM

| Category | Gene ontology | | *P*-value[a] |
|---|---|---|---|
| Top 3 clusters enriched in the 68 genes [autism with FMR1-FM and dup(15q)] | | Hits in the 68 genes | |
| Cluster 1 | Enrichment score[b]: 3.53 | | |
| GO_BP | Cell communication | 27 | 7.6E-06 |
| GO_BP | Signal transduction | 25 | 2.2E-05 |
| GO_BP | Cellular process | 47 | 1.5E-04 |
| Cluster 2 | Enrichment score[b]: 1.47 | | |
| GO_BP | Immune response | 11 | 3.7E-03 |
| GO_BP | Defense response | 11 | 7.3E-03 |
| GO_BP | Response to biotic stimulus | 11 | 9.8E-03 |
| GO_BP | Response to pest, pathogen or parasite | 6 | 3.8E-02 |
| GO_BP | Response to other organism | 6 | 4.8E-02 |
| GO_BP | Response to stress | 8 | 6.9E-02 |
| GO_BP | Response to stimulus | 13 | 1.6E-01 |
| GO_BP | Organismal physiological process | 11 | 2.9E-01 |
| Cluster 3 | Enrichment score[b]: 1.33 | | |
| SP_PIR | Membrane | 21 | 4.1E-04 |
| SP_PIR | Transmembrane | 20 | 9.1E-04 |
| GO_MF | Signal transducer activity | 19 | 5.4E-03 |
| UP_SEQ | Transmembrane region | 20 | 1.9E-02 |
| GO_CC | Integral to membrane | 21 | 4.7E-02 |
| GO_CC | Intrinsic to membrane | 21 | 4.8E-02 |
| KEGG_PATHWAY | Cytokine–cytokine receptor interaction | 5 | 5.0E-02 |
| GO_CC | Membrane | 25 | 7.1E-02 |
| SP_PIR | Signal | 11 | 8.8E-02 |
| SP_PIR | Transmembrane protein | 5 | 1.1E-01 |
| SP_PIR | Glycoprotein | 13 | 1.2E-01 |
| UP_SEQ | Disulfide bond | 13 | 1.4E-01 |
| UP_SEQ | Signal peptide | 11 | 4.8E-01 |
| UP_SEQ | Glycosylation site:N-linked (GlcNAc) | 13 | 5.0E-01 |
| GO_CC | Cell | 41 | 7.5E-01 |
| Top 3 clusters enriched in the 52 genes (autism with FMR1-FM) | | Hits in the 52 genes | |
| Cluster 1 | Enrichment score[b]: 1.4 | | |
| SP_PIR | Chaperone | 3 | 2.6E-02 |
| GO_BP | Protein folding | 4 | 3.2E-02 |
| GO_MF | Unfolded protein binding | 3 | 7.7E-02 |
| Cluster 2 | Enrichment score[b]: 1.36 | | |
| GO_MF | RNA binding | 6 | 1.2E-02 |
| GO_BP | mRNA metabolism | 4 | 2.1E-02 |
| GO_BP | RNA metabolism | 5 | 2.7E-02 |
| GO_BP | RNA processing | 4 | 6.3E-02 |
| GO_BP | mRNA processing | 3 | 9.5E-02 |
| GO_CC | Ribonucleoprotein complex | 4 | 1.7E-02 |
| Cluster 3 | Enrichment score[b]: 1.31 | | |
| GO_CC | Membrane-enclosed lumen | 6 | 6.0E-03 |
| GO_CC | Organelle lumen | 6 | 6.0E-03 |
| GO_CC | Nucleoplasm | 4 | 2.6E-02 |
| GO_CC | Nuclear lumen | 4 | 6.2E-02 |
| SP_PIR | Calcium | 3 | 3.6E-02 |
| GO_MF | Calcium ion binding | 3 | 6.7E-01 |

GO, gene ontology; BP, biological process; MF; molecular function; CC, cellular component; SP, SWISS-PROT; PIR, Protein Information Resources. KEGG, Kyoto Encyclopedia of Genes and Genomes.
[a]*P*-value was calculated by Fisher Exact test.
[b]Enrichment score was the negative log of geometric mean of each member's *P*-values in the cluster.

and *VIM* were previously identified as target RNAs of FMRP (46), although the mRNA expression changes of these genes in FMR1-FM have not been reported. Brown *et al*. (32) previously identified 144 genes as differentially expressed in lymphoblasts with FMR1-FM using pooled fragile X lymphoblastoid cells and pooled normal lymphoblastoid cells. Because there was no overlap except for *FMR1* between these 144 genes and the 120 genes identified here with our most stringent analyses using ANOVA, SAM and RankProd, we used the larger gene list identified by either SAM and/or RankProd to compare with the 144 genes identified by Brown *et al*. We found that 13 genes were shared in these gene lists, including iduronate 2-sulfatase (*IDS*), hairy and enhancer of split 1 (*HES1*) and immunoglobulin

**Table 4.** Gene networks identified by IPA using the genes dysregulated in both autism with FMR1-FM and dup(15q)

| Genes in network[a] | Top functions (*P*-value[b]) |
|---|---|
| *ACTB*, ***APBB1***, *CTNNB1, CYP27A1, DBT, GAD1*, ***GHSR***, *HK2, IL8RA* ***ISG15***, ***JDP2***, *LAD1, LEP, MT1L*, ***NDFIP1***, *NEDD4*, ***NR3C1***, ***NRG2*** ***PHACTR1***, *PHKG1, PLEC1, PODXL, PTENRB1*, ***RRM2***, *RRM2B*, ***SLC20A1, SNN, TCF7***, *TNF*, ***TNFRSF8***, *TOB2, TP53*, ***TRAF1***, ***TRAM1*** | Cell cycle ($P = 5.2E-8$)<br>Cancer ($P = 2.3E-7$)<br>Gastrointestinal disease ($P = 2.3E-7$) |
| ***ACTN1***, *ADORA2A, ADRBK2, CCBP2*, ***CCL17***, ***CCL22***, *CCR4*, ***CIITA***, *CLEC11A*, ***CMKOR1***, *DOK2, EGF*, ***GCET2***, *IER3, IGFALS, IL13, IL13RA1 IL13RA2, IL1B*, ***IL21R, IL4R***, *JAK1*, ***JAKMIP1, MIA***, *NR4A3, PIP3-E, PLA2G2A, PLA2G4A, PLB1*, ***PSCD3***, ***SLAMF1***, ***SLC16A6***, *SPHK1, TYK2, ZYX* | Cellular movement ($P = 1.3E-8$)<br>Lipid metabolism ($P = 1.6E-8$)<br>Molecular transport ($P = 1.6E-8$) |
| *ANXA6*, ***AREG***, *BATF, CD53, CEACAM1, D830050J10RIK*, ***DYRK3***, *EPO, HIF1A*, ***HIG2***, ***HIST2H3C***, *HRAS, IL2, IL6*, ***LCP2***, *LDB3, MAZ, MSN, MYOD1 PODXL*, ***PRKCB1***, ***RASA4***, *RASGRP3*, ***RASSF4***, *RRAS, SCGB1A1*, ***SELL***, *SELPLG, SNAP23*, ***SV2B***, *TXK, TYK2, UPP1*, ***VIM***, ***WNT6*** | Cell-to-cell signaling and interaction ($P = 4.3E-8$)<br>Cell cycle ($P = 5.7E-7$)<br>Cancer ($P = 1.2E-6$) |

[a]Genes identified as differentially expressed by microarray analysis is shown in bold.
[b]*P*-value was calculated using the right-tailed Fisher's Exact Test.

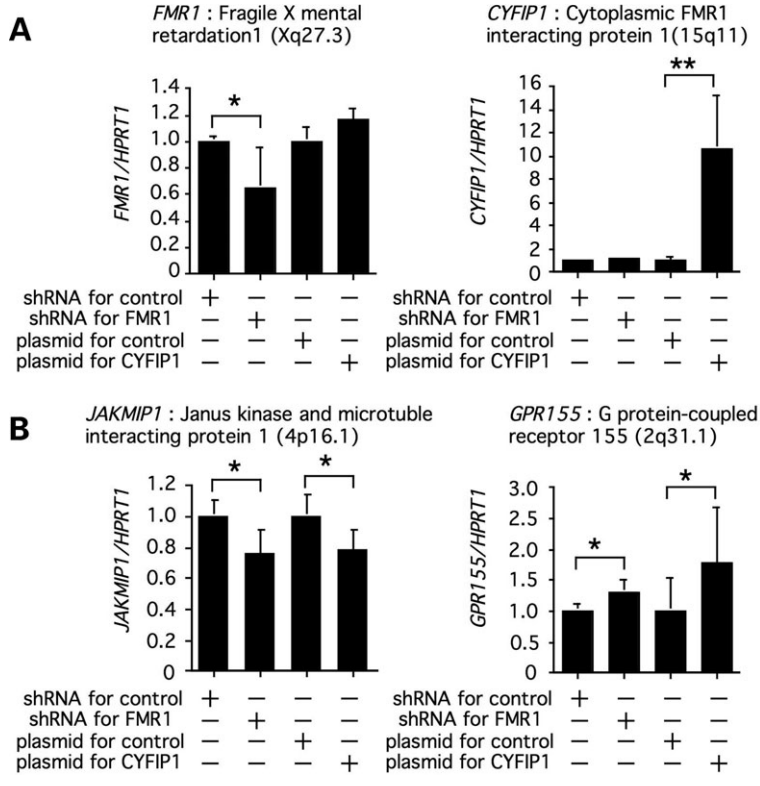


**Figure 4.** *JAKMIP1* and *GPR155* are dysregulated by reduction of *FMR1* and induction of *CYFIP1*. SH-SY5Y cells were stably transfected with (i) vector expressing shRNA for control, (ii) vector expressing shRNA for *FMR1*, (iii) empty expression vector or (iv) expression vector for *CYFIP1*. Total RNA was extracted from each and qRTPCR was performed to validate the effect of *FMR1* and *CYFIP1* on the expression of *JAKMIP1* and *GPR155*. (**A**) The expression of *FMR1* is significantly reduced in SH-SY5Y cells expressing shRNA for *FMR1*, whereas the expression of *CYFIP1* was significantly induced in SY5Y cells over-expressing *CYFIP1*. (**B**) The expression of *JAKMIP1* is significantly reduced in SH-SY5Y cells expressing shRNA for *FMR1* and over-expressing *CYFIP1*. The expression of *GPR155* is significantly induced in SH-SY5Y cells expressing shRNA for *FMR1* and over-expressing *CYFIP1*. Results represent means $\pm$ SD of each group. The mean of the value of each control was set as 1. Significance was calculated by the Mann–Whitney *U* test using SH-SY5Y cells expressing shRNA for control ($N = 4$) versus shRNA for *FMR1* ($N = 4$) or empty expression vector ($N = 8$) versus expression vector for *CYFIP1* ($N = 7$). $^{*}P < 0.05$, $^{**}P < 0.01$.

superfamily, member 3 (*IGSF3*) as up-regulated genes and CDK2-associated protein 2 (*CDK2AP2*), ubiquitin specific peptidase 8 (*USP8*), MAX-like protein X (*MLX*), ribosomal protein S5 (*RPS5*), C-terminal binding protein 1 (*CTBP1*), spleen tyrosine kinase (*SYK*), F-box protein 6 (*FBXO6*), mitogen-activated protein kinase kinase kinase 11 (*MAP3K11*), sorting nexin 15 (*SNX15*) and CD44 antigen (*CD44*) as downregulated genes. Although these genes have

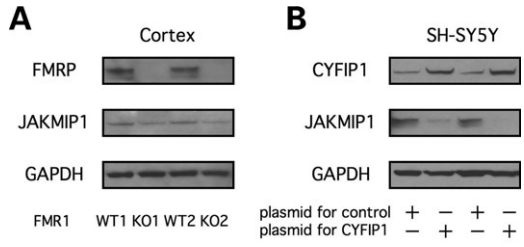


**Figure 5.** The expression of JAKMIP1 was dependent on *FMR1* and *CYFIP1* in mouse cortex and SH-SY5Y cells. Proteins were extracted from cortices of *Fmr1* WT or KO mice (**A**), or SH-SY5Y cells transfected with empty vector or *CYFIP1* cDNA (**B**). Western blotting was performed to validate the effect of the reduction of *FMR1* or induction of *CYFIP1* on the expression of *JAKMIP1* protein. The protein expression of Jakmip1 was reduced in cortex of Fmr1-KO mice (A) as well as SH-SY5Y cells transfected with shRNA for *FMR1* (data not shown) and SH-SY5Y over-expressing *CYFIP1* (B). Data shown in (A) and (B) were the representative of two independent experiments.

not been reported as associated with *FMR1* or autism, *HES1* was associated with attention-deficit hyperactive disorder (47), which is a symptom frequently seen in FXS (48) and overlapping with ASD (49). The relatively low overlap between the two gene lists could be due to the difference of clinical features of individuals (autism versus not specific for autism), experimental design (each individual versus pooled), microarray platforms (Agilent versus Affymetrix) and the statistical analysis used to find the differential expression between groups. The initial study (32) whose primary aim was to find FMRP ligand mRNPs was relatively underpowered to detect overall differences in gene expression and our study used very conservative statistical criteria. However, this core set of genes provides an interesting gene list for further investigation.

### Gene expression profiles of lymphoblastoid cells with dup(15q)

We identified 80 genes differentially expressed in dup(15q) carriers compared with controls. Among these genes, four genes located in 15q11–q13 (the region of duplication) *UBE3A*, *CYFIP1*, *NIPA2* and *GOLGA8F* were all induced. It is important to note that five other genes located in the duplicated region, tubulin gamma complex associated protein 5 (*TUBGCP5*), *HERC2*, HERC2 pseudogene 2 (*HERC2P2*), *NIPA1* and *ATP10A* were also identified as up-regulated genes by at least one of the three different statistical analyses (Supplementary Material, Table S7). Five other genes in the duplicated region, gamma-aminobutyric acid A receptor (GABR) beta 3 (*GABRB3*), GABR alpha 5 (*GABRA5*), GABR gamma 3 (*GABRG3*), oculocutaneous albinism II (*OCA2*) and necdin homolog (*NDN*), were not expressed at detectable levels in the lymphoblastoid cells. It is important to emphasize that the 15q11–q13 region is subject to paternal imprinting. Three paternally imprinted genes, makorin ring finger protein 3 (*MKRN3*), MAGE-like 2 (*MAGEL2*) and SNRPN upstream reading frame (*SNURF*)-small nuclear ribonucleoprotein polypeptide N (*SNRPN*) were expressed in the lymphoblasts, but showed no significant changes relative to controls. These data are consistent with the fact that the duplicated region was maternally derived in all seven cases analyzed in this study. So, overall, these findings suggest that the genes located in the duplicated region were globally up-regulated except for the paternally imprinted genes. Global up-regulation due to gene-dosage has also been reported in Down syndrome (50,51).

Baron *et al*. (27) identified 81 known genes as differentially expressed in lymphoblastoid cells with dup(15q) (seven individuals) when compared with controls (eight individuals) using the Affymetrix platform. They identified up-regulation of *UBE3A*, *NIPA1*, *NIPA2* and *HERC2*, findings consistent with our results. We used the gene list identified by SAM and/or RankProd to compare with the 81 genes identified by Baron *et al*. and identified 11 other genes shared in the two gene lists, a significant overlap (the hypergeometric probability is 0.001). These genes were abhydrolase domain containing 6 (*ABHD6*), potassium channel, subfamily K, member 1 (*KCNK1*), hypothetical protein KIAA1147 and zinc finger, DHHC domain containing 14 (*ZDHCC14*) as up-regulated and Rho GTPase activating protein 25 (*ARHGAP25*), clone LOC387882, leukotriene B4 12-hydroxydehydrogenase (*LTB4DH*), clone MGC27165, PFTAIRE protein kinase 1 (*PFTK1*), zinc finger protein 43 (*ZNF43*) and ring finger protein 41 (*RNF41*) as downregulated. The relationships between these genes and autism remain unknown. Again, similar to the studies of FMR1-FM, these genes represent a set of independently replicated genes between two studies.

### Significant overlap of dysregulated genes in autism with FMR1-FM and dup(15q)

We identified 68 genes that were dysregulated in both autism with FMR1-FM and with dup(15q), a very significant result (the hypergeometric probability of this overlap is $1.2\times10^{-153}$). However, we cannot formally exclude the possibility that some of the 68 common dysregulated genes might be related to common pathways between FMR-FM and dup(15q) unrelated to ASD. Microarray analysis using lymphoblastoid cells with FMR1-FM or dup(15q), but without ASD is needed to exclude the possibility, as was done in tuberous sclerosis cases with and without autism (51).

We found that the expression of *CYFIP1* was significantly induced in autism with dup(15q). CYFIP1 has been shown to antagonize FMRP in the eye and nervous system of *Drosophila* (34). In FXS, the absence of FMRP, a binding partner to CYFIP1, results in excess free CYFIP1. Similarly, excess free CYFIP1 may be the outcome of dup(15q). Thus, antagonization of FMRP by over-expression of CYFIP1 and/or alternate actions of excess CYFIP1 may be common mechanistic links between FMR1-FM and dup(15q).

### Effect of *FMR1* and *CYFIP1* on the commonly dysregulated genes in SH-SY5Y and mouse brain

We validated the effect of downregulation of *FMR1* in SH-SY5Y cells and mouse brain and up-regulation of *CYFIP1* in SH-SY5Y cells on the expression of the commonly dysregulated genes identified in patient lymphoblastoid cells.

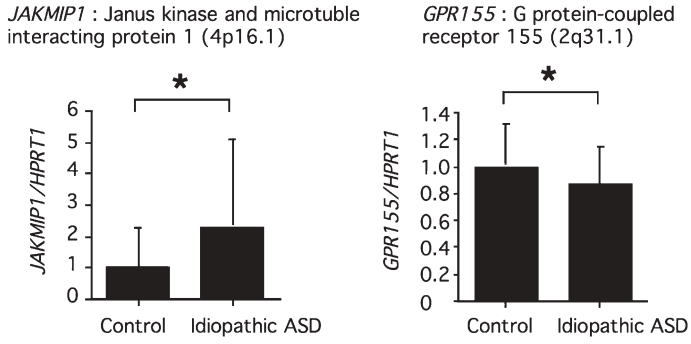


**Figure 6.** *JAKMIP1* and *GPR155* are dysregulated in the ASD proband in discordant male sib pairs. Total RNA was extracted from lymphoblastoid cells of 27 male sib pairs discordant for ASD and qRTPCR were performed to confirm the differential expression of *JAKMIP1* and *GPR155*. Results represent means ± SD of each group. The mean of the value of control subjects was set as 1. *P*-value was calculated by Wilcoxon rank-sum test using control ($N = 27$) versus ASD ($N = 27$). $^{*}P < 0.05$.

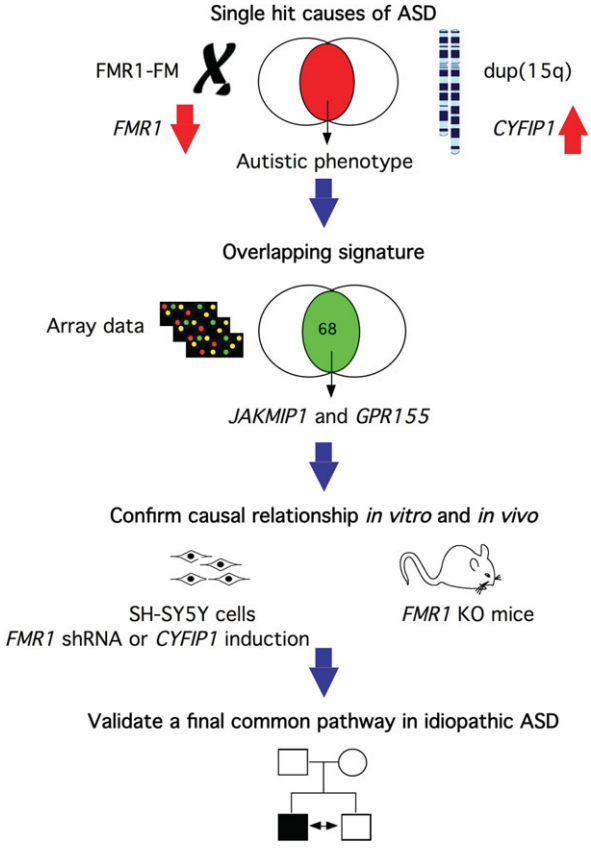


**Figure 7.** Molecular convergence of FMR1-FM, dup(15q) and idiopathic ASD. The mRNA expression profile in lymphoblastoid cells from autism with FMR1-FM or dup(15q) and control were compared using microarray analysis. Sixty-eight genes were dysregulated in both autism with FMR1-FM and dup(15q). Induction of *CYFIP1* in dup(15q) is a potential molecular link between FMR1-FM and dup(15q). Among the dysregulated genes, *JAKMIP1* and *GPR155* were further analyzed to confirm the causal relationship between *CYFIP1* and *FMR1* expression and their expression in neural cells or tissue and to validate the dysregulation of these genes in lymphoblastoid cells from subjects with idiopathic ASD.

We demonstrated that the expression of *JAKIMIP1* and *GPR155* was dysregulated by reduction of *FMR1* and induction of *CYFIP1* in SH-SY5Y cells. The Jakmip1 protein was also dysregulated by knock-out of *Fmr1* in mouse brain. Interestingly, the direction of changes observed in both of these genes was opposite in neural tissues (SH-SY5Y cells and brain) and lymphoblastoid cells. Such differences between brain and blood cells have been previously observed in other signaling pathways (31,52). It is likely that it is not the precise direction observed in lymphoblastoid cells that is most important, but the common dysregulation of *JAKMIP1* and *GPR155* downstream of these single gene defects, which is observed in idiopathic ASD.

*JAKMIP1* is associated with Janus kinases (41), microtubules (41) and $GABR_B$ receptors (40). The expression levels of *JAKMIP1* affect the intracellular levels of the $GABR_B$ receptor (40). Because the $GABR_B$ receptor could interact with the metabotropic glutamate receptor 1 (mGluR1) and increase the glutamate sensitivity of mGluR1 (53), *JAKMIP1* might affect mGluR1 signaling through $GABR_B$ receptors. It is important to note that mGluR signaling is exaggerated in *Fmr1* knockout mice (54) and that glutamergic and GABAergic systems have been reported to be abnormal in autism (55). *Jakmip1* is highly expressed throughout the mouse brain, especially in hippocampus, where $GABR_B$ receptors and mGluR1 are also highly expressed (56). Although the function of *Gpr155* is unknown, it is highly expressed in the limbic system in mouse brain (56), suggesting that *Gpr155* might have functions relevant to the limbic system.

It is also interesting to consider how the reduction of *FMR1* and the induction of *CYFIP1* might regulate the expression of *JAKMIP1* and *GPR155*. G-quadruplex motifs in RNA have been shown to play significant roles in FMRP binding (57). Using QGRS mapper (58), we found that human and mouse *JAKMIP1* each had two of the G-quadruplex ($G_2N_{2-4}G_2N_{2-4}G_2N_{2-4}G_2$) and that human and mouse *GPR155* had five and one of the G-quadruplex, respectively (data not shown). FMRP can also bind target RNAs through non-coding RNAs (59) or microRNAs (60). Using miRBase (61), we found putative microRNA binding sites in human and mouse *JAKIMIP1* and *GPR155* (data not shown). Further studies are required to clarify the functional importance of *JAKMIP1* and *GPR155* in autism and the mechanism of regulation of these genes by *FMR1* and *CYFIP1*. In this regard, the potential link with neuronal transmission is intriguing.

### The expression of *JAKMIP1* and *GPR155* was also dysregulated in 27 males with idiopathic ASD

The findings in autism with FMR1-FM and dup(15q) suggest that *JAKMIP1* and *GPR155* may be involved more generally in idiopathic ASD, since their dysregulation is observed in neural cells and brain. We tested whether dysregulation of these genes were more generalizable in an independent sample of idiopathic ASD cases. To attempt to reduce the heterogeneity of idiopathic ASD and extend these findings beyond those with mental retardation or intellectual disability, we used an IQ surrogate based on Raven's Progressive Matrices, which is highly correlated with IQ defined by other measures (62). We selected 27 ASD males with an IQ score of more than 70. These data demonstrated that the expression of *JAKMIP1* and *GPR155* was significantly dysregulated in lymphoblastoid cells from idiopathic ASD when compared with controls. These results based on independent data on lymphoblastoid cell gene expression from ASD subjects with FMR1-FM, or dup(15q), as well as idiopathic ASD, suggest that *JAKMIP1* and *GPR155* may be useful as biomarkers for ASD. In this sample overall, *JAKMIP1* appeared to be more robustly dysregulated than *GPR155*, although the expression differences appeared heterogeneous. How the dysregulation of these genes relates to distinct autism subtypes will also be important to determine in the future.

The mechanism for the opposite regulation of *JAKMIP1* and *GPR155* in lymphobastoid cells and neural cells remains unclear. There are several previous reports of genes showing the opposite expression between lymphoblastoid cells and brains in neuropsychiatric disease. One example is inositol monophosphatase 2 (*IMPA2*), which has been identified as a plausible locus for bipolar disorder (63–65). The expression of *IMPA2* was reduced and induced in lymphoblastoid cells and brains, respectively, in patients with bipolar disorder (66). A genetic association between *IMPA2* promoter polymorphism and bipolar disorder has been confirmed (67,68). In this regard, it is notable that *GPR155* is located on 2q31.1, 300 kb from D2S2188, which has shown strong linkage to autism in studies by two independent groups (6,69). Association analyses for *GPR155* and *JAKMIP1* are ongoing using the large AGRE cohort. These data provide the first identification and independent validation of the roles of *JAKMIP1* and *GPR155* dysregulation in ASD (Fig. 7). Further work is needed to understand the functional consequences of these changes in the developing brain and to assess the general utility of these and other genes as potential biomarkers.

## MATERIALS AND METHODS

### Individuals and lymphoblastoid cells analyzed in this study

We have analyzed individuals diagnosed with ASD using standard validated measures, including the Autism Diagnostic Interview (ADI-R) (70) and Autism Diagnostic Observation Schedule (ADOS) (71). Eight males with FMR1-FM and three males with dup(15q) were drawn from AGRE (33) (http://www.agre.org/). An additional four males with dup(15q) were obtained from N.C.S. Twenty-seven males without autism, FMR1-FM or dup(15q) were drawn from the AGRE for controls. In addition, another 27 males with idiopathic ASD and with unaffected male siblings were chosen from AGRE for a comparison sample (Supplementary Material, Table S1). Surrogate IQ scores (using the Raven Progressive Matrices) were available. FMR1-FM and dup(15q) were examined by PCR and fluorescence *in situ* hybridization, respectively. The 15q11–q13 duplicated regions in the seven males analyzed in this study were all maternally derived. We also used 14 other individuals from AGRE for common reference (pool) in microarray analysis (data not shown). Lymphoblastoid cell lines (human Epstein-Barr virus transformed lymphocytes) from these individuals were available from AGRE.

The lymphoblastoid cells of the subjects were grown in RPMI 1640 medium with 2 mM L-glutamine and 25 mM HEPES (Invitrogen, Carlsbad, CA, USA), 10% fetal bovine serum, and 1 × Antibiotic-Antimycotic solution (Invitrogen) at 37°C in a humidified 5% $CO_2$ chamber. Cells were grown to a density of $6 \times 10^5$/ml. Special attention was given to maintain all the cell lines in the same conditions to minimize environmental variation.

### Microarray experiments

A total of $9 \times 10^6$ of lymphoblastoid cells were seeded in a T-75 flask in 30 ml of fresh medium. After 24 h, total RNA was extracted from the cells using an RNeasy Mini Kit with DNase treatment (Qiagen, Valencia, CA, USA) according to the manufacturer's protocol. The RNA quantity and quality were measured by ND-100 (Nanodrop, Wilmington, DE, USA) and 2100 Bioanalyzer (Agilent, Santa Clara, CA, USA), respectively.

Target preparation was performed using a Low RNA Input Fluorescent Linear Amplification Kit (Agilent) according to the manufacturer's protocol. We extracted total RNA from lymphoblastoid cells from each individual and made target and labeled it with Cy5 fluorescence. We also made reference target by using pooled total RNA from the 14 individuals for reference and labeled it with Cy3 fluorescence. The generated targets were mixed and subjected to hybridization to the Whole Human Genome Array G4112A (Agilent) according to the manufacturer's protocol. Scanning of the microarrays were done by a DNA microarray scanner (Agilent).

Scanner output image files were normalized and filtered using Feature Extraction Software v8.5 (Agilent). Normalization was performed so that overall intensity ratio of Cy5 to Cy3 was equal to one. Probes with signal-to-noise ratio $>2.7$ in both Cy3 and Cy5 in at least 14 of 15 controls were used for further analysis.

### Statistical analysis of microarray data

ANOVA was performed by MeV3.1 (72). *P*-values were calculated based on 1000 permutations. Hierarchical clustering using Spearman's rank correlation with average linkage clustering was performed by MeV3.1. PCA was performed using GeneSpring GX7.3 (Agilent). SAM (38) and RankProd (39) were performed using Bioconductor (73) packages Siggene and RankProd, respectively. For cross-validation in SAM and RankProd, respectively, 100 and 1000 permutations

were performed. We used three different statistical tests to conservatively identify the most robustly differentially expressed genes. Numerous feature selection methods have been applied to the identification of differentially expressed genes in microarray data (74). The genes commonly identified by ANOVA, SAM and RankProd are likely to be differentially expressed, given the relative robustness of these statistical approaches (35,38,39,74). Functional Annotation Clustering was performed by DAVID (42) with medium classification stringency. The clustering algorithm is based on the hypothesis that similar annotations should have similar gene members. The Functional Annotation Clustering uses two different statistics to measure the degree of the common genes between two annotations and to classify the groups with similar annotations. The Group Enrichment Score is the geometric mean (in $-\log$ scale) of a member's *P*-values in a corresponding annotation cluster. IPA was used to find significant pathways related to the genes commonly dysregulated in autism with FMR1-FM and dup(15q). The Ingenuity Pathway Knowledge Base builds gene networks based upon known protein and gene interactions (44). IPA determines a statistical score for each network according to the probability of the network given in the gene list. The Ingenuity Pathway Knowledge Base provides pathways with biological function based upon the scientific literature. The significance value associated with Functions and Pathways measures how likely it is that genes from the data set file participate in that biological function. The significance was expressed as a *P*-value, which is calculated using the right-tailed Fisher's exact test.

### qRTPCR analysis

One microgram of total RNA was used to make cDNA by SuperScript III First-Strand Synthesis SuperMix (Invitrogen). The qRTPCR was done by ABI Prism 7900 (Applied Biosystems, Foster City, CA, USA) using Platinum SYBR Green qPCR SuperMix UDG with ROX (Invitrogen). Thermal cycling consisted of an initial step at 50°C for 2 min followed by another step at 95°C for 2 min and 50 cycles of 95°C for 15 s and 60°C for 30 s. The qRTPCR was performed for 16 genes. The primers used in this study are shown in Supplementary Material, Table S8. TaqMan probe (Hs00327005_m1, Applied Biosystems) was used to measure *JAKMIP1* expression in lymphoblastoid cells. Data were normalized by the quantity of hypoxanthine phosphoribosyltransferase 1 (*HPRT1*). *HPRT1* was selected rather than beta-actin, glyceraldehyde-3-phosphate dehydrogenase or other possible internal controls because it was shown to be the most stable RNA species from the lymphoblastoid cell lines. This allowed us to account for the variability in the conversion efficiency of the reverse transcription reaction.

### Transduction of retroviral shRNAs

To construct retrovirus vectors expressing shRNAs, oligonucleotides encoding stem-loop shRNAs for *FMR1* (Supplementary Material, Table S8) and negative control were ligated into the BamHI and EcoRI site of the pSIREN-RetroQ (BD Clontech, Mountain View, CA, USA). PT67 cells (BD Clontech) were transfected for retrovirus production. A total of $6 \times 10^6$ of SH-SY5Y cells were seeded in a T-75 flask in 20 ml of fresh medium of DMEM (Invitrogen) with 10% FBS. After 1 day, SH-SY5Y cells were infected with retroviruses in the presence of 5 μg/ml of polybrene. After 2 days, the SH-SY5Y cells were treated with 10 μg/ml of puromycin (Sigma, St Louis, MO, USA). Cells that survived after 4 weeks were collected, and this population of cells was used for further experiments. Total RNA was extracted from the cells using RNeasy Mini Kit with DNase treatment (Qiagen) according to the manufacturer's protocol. We compared SH-SY5Y cells expressing *FMR1* shRNA ($n = 4$) and SH-SY5Y cells expressing shRNAs for negative control ($n = 4$) to examine the effect of reduction of *FMR1* on the expression of *JAKMIP1* and *GPR155*.

### Transfection of *CYFIP1*

The human *CYFIP1* coding region (amino acid 1–1254) obtained by PCR using IMAGE clone 10625411 (ATCC, Manassas, VA, USA) was subcloned into the *EcoRV* and *Not*I sites of the plasmid vector pIRES-neo3 (BD Clontech). The sequence of the construct was confirmed by automated DNA sequencing.

A total of $6 \times 10^6$ of SH-SY5Y cells were seeded in a T-75 flask in 20 ml of fresh medium of DMEM (Invitrogen) with 10% FBS. After 1 day, SH-SY5Y cells were transfected with 120 μl of lipofectamine 2000 (Invitrogen) diluted with 3 ml of OptiMEM (Invitrogen) and 24 μg of plasmid (pIRES-CYFIP1 or pIRES) diluted with 3 ml of OptiMEM (Invitrogen). After 5 min at room temperature, they were combined and incubated for 20 min. The reaction mixture was added with 16 ml of DMEM with 10% FBS. The cell culture medium was replaced by this solution. After 2 days, the SH-SY5Y cells were treated with 500 μg/ml of G418 (Invitrogen). Cells that survived after 3 weeks were collected, and this population of cells was used for further experiments. Total RNA was extracted from the cells using RNeasy Mini Kit with DNase treatment (Qiagen) according to the manufacturer's protocol. We compared SH-SY5Y cells stably transfected with the expression vector for *CYFIP1* ($n = 7$) and SH-SY5Y cells transfected with the empty expression vector ($n = 8$) to examine the effect of induction of *CYFIP1* on the expression of *JAKMIP1* and *GPR155*. Protein was also extracted using Cellytic M (Sigma) with proteinase inhibitors (Sigma) according to the manufacturer's protocol.

### Animals and tissue collection

WT and *Fmr1* KO mice were raised at the Emory University animal facility and treated in accordance with National Institue of Health regulations and under approval of the Emory University Institutional Animal Care and Use Committee. WT and *Fmr1* KO littermates were produced by breeding heterozygous females with *FMR1* KO males in a congenic background of C57BL/6. The genotype of each animal was confirmed by PCR. For tissue collection, cortices were dissected and protein was isolated using Cellytic M (Sigma) with proteinase inhibitors (Sigma) according to the manufacturer's protocol.

### Immunoblot analysis

Proteins extracted from SH-SY5Y cells or cortices of *Fmr1* WT and KO mice were subjected to SDS–PAGE using NuPAGE Novex 4–20% Bis–Tris gel and MOPS buffer (Invitrogen) according to the manufacturer's protocol. After electrophoresis, gels were electroblotted onto PVDF membranes (Millipore, Bedford, MA, USA). After electroblotting, membranes were blocked in SuperBlock blocking buffer (Pierce Biotechnology, Rockford, IL, USA). Membranes were probed in the blocking solution at 4°C overnight with the following antibodies: FMRP (Chemicon, Temecula, CA, USA), CYFIP1 (Upstate, Temecula, CA, USA), JAKMIP1 (41) or GAPDH. Membranes were washed 3× in PBS supplemented with 0.05% Tween 20 (PBS-T) and incubated with the appropriate horseradish peroxidase-conjugated secondary antibody in the blocking solution for 1 h at room temperature. Membranes were again washed 3× in PBS-T and developed with SuperSignal West Pico Chemiluminescent (Pierce Biotechnology). Membranes were stripped by Restore Western Blot Stripping Buffer (Pierce Biotechnology) and used for different antibodies.

## SUPPLEMENTARY MATERIAL

Supplementary Material is available at HMG Online.

## ACKNOWLEDGEMENTS

We gratefully acknowledge the resources provided by the AGRE consortium and sincerely thank the AGRE families and the IsoDicentric 15 Exchange Advocacy and Support (IDEAS) group who have participated in the study. AGRE is a program of the Cure Autism Now foundation (CAN) and is supported, in part, by grant MH64547 from the National Institute of Mental Health (NIMH) to D.H.G. (PI). This work was supported by NIMH grant MH64547 (to D.H.G.), The Gassin Family Foundation (to Y.N. and D.H.G.), The Boler Company Foundation (to Y.N. and D.H.G.), NICHD grant HD020S21 (to S.T.W.) and NIH U19-HD35470 CPEA grant (to N.C.S., Marian Sigman, PI). Y.N. is a recipient of a Young Investigator Award from CAN. We are also grateful to Giovanni Coppola and Brett Abrahams for valuable discussions; Gena Konopka and Michelle Stofa for critical readings of the manuscript; Jeffrey Gregg, Stephenie Liu, Barb Malone and Jen Driscoll for providing lymphoblastoid cells with dup(15q): Tamika Malone for providing brains of FMR1-KO and WT mice; and Lauren Kawaguchi for her help as laboratory manager.

*Conflict of Interest statement*. D.H.G. provides consulting service to AGRE, a non-profit arm of Cure Autism Now, as Chief Scientific Officer.

## APPENDIX

The AGRE Consortium: D.H.G., M.D., Ph.D., UCLA, Los Angeles, CA; Maja Bucan, Ph.D., University of Pennsylvania, Philadelphia, PA; W. Ted Brown, M.D., Ph.D., F.A.C.M.G., N.Y.S. Institute for Basic Research in Developmental Disabilities, Staten Island, NY; Rita M. Cantor, Ph.D., UCLA School of Medicine, Los Angeles, CA; John N. Constantino, M.D., Washington University School of Medicine, St Louis, MO; T. Conrad Gilliam, Ph.D., University of Chicago, Chicago, IL; Martha Herbert, M.D., Ph.D., Harvard Medical School, Boston, MA; Clara Lajonchere, Ph.D, Cure Autism Now, Los Angeles, CA; David H. Ledbetter, Ph.D., Emory University, Atlanta, GA; Christa Lese-Martin, Ph.D., Emory University, Atlanta, GA; Janet Miller, J.D., Ph.D., Cure Autism Now, Los Angeles, CA; Stanley F. Nelson, M.D., UCLA School of Medicine, Los Angeles, CA; Gerard D. Schellenberg, Ph.D., University of Washington, Seattle, WA; Carol A. Samango-Sprouse, Ed.D., George Washington University, Washington, D.C.; S.J.S., M.D., Ph.D., NIMH, Bethesda, MD; Matthew State, M.D., Ph.D., Yale University, New Haven, CT; Rudolph E. Tanzi, Ph.D., Massachusetts General Hospital, Boston, MA.

## REFERENCES

1. Folstein, S.E. and Rosen-Sheidley, B. (2001) Genetics of autism: complex aetiology for a heterogeneous disorder. *Nat. Rev. Genet.*, **2**, 943–955.
2. Belmonte, M.K., Cook, E.H., Jr., Anderson, G.M., Rubenstein, J.L., Greenough, W.T., Beckel-Mitchener, A., Courchesne, E., Boulanger, L.M., Powell, S.B., Levitt, P.R. *et al.* (2004) Autism as a disorder of neural information processing: directions for research and targets for therapy. *Mol. Psychiatry*, **9**, 646–663.
3. Veenstra-Vanderweele, J., Christian, S.L. and Cook, E.H., Jr. (2004) Autism as a paradigmatic complex genetic disorder. *Annu. Rev. Genomics Hum. Genet.*, **5**, 379–405.
4. Muhle, R., Trentacoste, S.V. and Rapin, I. (2004) The genetics of autism. *Pediatrics*, **113**, e472–e486.
5. Liu, J., Nyholt, D.R., Magnussen, P., Parano, E., Pavone, P., Geschwind, D., Lord, C., Iversen, P., Hoh, J., Ott, J. *et al.* (2001) A genomewide screen for autism susceptibility loci. *Am. J. Hum. Genet.*, **69**, 327–340.
6. IMGSAC (2001) A genomewide screen for autism: strong evidence for linkage to chromosomes 2q, 7q, and 16p. *Am. J. Hum. Genet.*, **69**, 570–581.
7. IMGSAC (2001) Further characterization of the autism susceptibility locus AUTS1 on chromosome 7q. *Hum. Mol. Genet.*, **10**, 973–982.
8. Buxbaum, J.D., Silverman, J.M., Smith, C.J., Kilifarski, M., Reichert, J., Hollander, E., Lawlor, B.A., Fitzgerald, M., Greenberg, D.A. and Davis, K.L. (2001) Evidence for a susceptibility gene for autism on chromosome 2 and for genetic heterogeneity. *Am. J. Hum. Genet.*, **68**, 1514–1520.
9. Alarcon, M., Cantor, R.M., Liu, J., Gilliam, T.C. and Geschwind, D.H. (2002) Evidence for a language quantitative trait locus on chromosome 7q in multiplex autism families. *Am. J. Hum. Genet.*, **70**, 60–71.
10. Yonan, A.L., Alarcon, M., Cheng, R., Magnusson, P.K., Spence, S.J., Palmer, A.A., Grunn, A., Juo, S.H., Terwilliger, J.D., Liu, J. *et al.* (2003) A genomewide screen of 345 families for autism-susceptibility loci. *Am. J. Hum. Genet.*, **73**, 886–897.
11. Stone, J.L., Merriman, B., Cantor, R.M., Yonan, A.L., Gilliam, T.C., Geschwind, D.H. and Nelson, S.F. (2004) Evidence for sex-specific risk alleles in autism spectrum disorder. *Am. J. Hum. Genet.*, **75**, 1117–1123.
12. Alarcon, M., Yonan, A.L., Gilliam, T.C., Cantor, R.M. and Geschwind, D.H. (2005) Quantitative genome scan and Ordered-Subsets Analysis of autism endophenotypes support language QTLs. *Mol. Psychiatry*, **10**, 747–757.
13. Cantor, R.M., Kono, N., Duvall, J.A., Alvarez-Retuerto, A., Stone, J.L., Alarcon, M., Nelson, S.F. and Geschwind, D.H. (2005) Replication of autism linkage: fine-mapping peak at 17q21. *Am. J. Hum. Genet.*, **76**, 1050–1056.
14. McCauley, J.L., Li, C., Jiang, L., Olson, L.M., Crockett, G., Gainer, K., Folstein, S.E., Haines, J.L. and Sutcliffe, J.S. (2005) Genome-wide and Ordered-Subset linkage analyses provide support for autism loci on 17q and 19p with evidence of phenotypic and interlocus genetic correlates. *BMC Med. Genet.*, **6**, 1.

15. Philippi, A., Roschmann, E., Tores, F., Lindenbaum, P., Benajou, A., Germain-Leclerc, L., Marcaillou, C., Fontaine, K., Vanpeene, M., Roy, S. *et al.* (2005) Haplotypes in the gene encoding protein kinase c-beta (PRKCB1) on chromosome 16 are associated with autism. *Mol. Psychiatry*, **10**, 950–960.
16. Bartlett, C.W., Goedken, R. and Vieland, V.J. (2005) Effects of updating linkage evidence across subsets of data: reanalysis of the autism genetic resource exchange data set. *Am. J. Hum. Genet.*, **76**, 688–695.
17. Chen, G.K., Kono, N., Geschwind, D.H. and Cantor, R.M. (2006) Quantitative trait locus analysis of nonverbal communication in autism spectrum disorder. *Mol. Psychiatry*, **11**, 214–220.
18. Ylisaukko-oja, T., Alarcon, M., Cantor, R.M., Auranen, M., Vanhala, R., Kempas, E., von Wendt, L., Jarvela, I., Geschwind, D.H. and Peltonen, L. (2006) Search for autism loci by combined analysis of Autism Genetic Resource Exchange and Finnish families. *Ann. Neurol.*, **59**, 145–155.
19. Brown, W.T., Jenkins, E.C., Cohen, I.L., Fisch, G.S., Wolf-Schein, E.G., Gross, A., Waterhouse, L., Fein, D., Mason-Brothers, A., Ritvo, E. *et al.* (1986) Fragile X and autism: a multicenter survey. *Am. J. Med. Genet.*, **23**, 341–352.
20. Verkerk, A.J., Pieretti, M., Sutcliffe, J.S., Fu, Y.H., Kuhl, D.P., Pizzuti, A., Reiner, O., Richards, S., Victoria, M.F., Zhang, F.P. *et al.* (1991) Identification of a gene (FMR-1) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome. *Cell*, **65**, 905–914.
21. Rogers, S.J., Wehner, D.E. and Hagerman, R. (2001) The behavioral phenotype in fragile X: symptoms of autism in very young children with fragile X syndrome, idiopathic autism, and other developmental disorders. *J. Dev. Behav. Pediatr.*, **22**, 409–417.
22. Goodlin-Jones, B.L., Tassone, F., Gane, L.W. and Hagerman, R.J. (2004) Autistic spectrum disorder and the fragile X premutation. *J. Dev. Behav. Pediatr.*, **25**, 392–398.
23. Bailey, A., Phillips, W. and Rutter, M. (1996) Autism: towards an integration of clinical, genetic, neuropsychological, and neurobiological perspectives. *J. Child Psychol. Psychiatry*, **37**, 89–126.
24. Sutcliffe, J.S., Nurmi, E.L. and Lombroso, P.J. (2003) Genetics of childhood disorders: XLVII. Autism, part 6: duplication and inherited susceptibility of chromosome 15q11–q13 genes in autism. *J. Am. Acad. Child Adolesc. Psychiatry*, **42**, 253–256.
25. Wang, N.J., Liu, D., Parokonny, A.S. and Schanen, N.C. (2004) High-resolution molecular characterization of 15q11–q13 rearrangements by array comparative genomic hybridization (array CGH) with detection of gene dosage. *Am. J. Hum. Genet.*, **75**, 267–281.
26. Thomas, N.S., Roberts, S.E. and Browne, C.E. (2003) Estimate of the prevalence of chromosome 15q11–q13 duplications. *Am. J. Med. Genet. A*, **120**, 596–598.
27. Baron, C.A., Tepper, C.G., Liu, S.Y., Davis, R.R., Wang, N.J., Schanen, N.C. and Gregg, J.P. (2006) Genomic and functional profiling of duplicated chromosome 15 cell lines reveal regulatory alterations in UBE3A-associated ubiquitin-proteasome pathway processes. *Hum. Mol. Genet.*, **15**, 853–869.
28. Sebat, J., Lakshmi, B., Troge, J., Alexander, J., Young, J., Lundin, P., Maner, S., Massa, H., Walker, M., Chi, M. *et al.* (2004) Large-scale copy number polymorphism in the human genome. *Science*, **305**, 525–528.
29. Horike, S., Cai, S., Miyano, M., Cheng, J.F. and Kohwi-Shigematsu, T. (2005) Loss of silent-chromatin looping and impaired imprinting of DLX5 in Rett syndrome. *Nat. Genet.*, **37**, 31–40.
30. Meloni, I., Muscettola, M., Raynaud, M., Longo, I., Bruttini, M., Moizard, M.P., Gomot, M., Chelly, J., des Portes, V., Fryns, J.P. *et al.* (2002) FACL4, encoding fatty acid-CoA ligase 4, is mutated in nonspecific X-linked mental retardation. *Nat. Genet.*, **30**, 436–440.
31. Iwamoto, K., Kakiuchi, C., Bundo, M., Ikeda, K. and Kato, T. (2004) Molecular characterization of bipolar disorder by comparing gene expression profiles of postmortem brains of major mental disorders. *Mol. Psychiatry*, **9**, 406–416.
32. Brown, V., Jin, P., Ceman, S., Darnell, J.C., O'Donnell, W.T., Tenenbaum, S.A., Jin, X., Feng, Y., Wilkinson, K.D., Keene, J.D. *et al.* (2001) Microarray identification of FMRP-associated brain mRNAs and altered mRNA translational profiles in fragile X syndrome. *Cell*, **107**, 477–487.
33. Geschwind, D.H., Sowinski, J., Lord, C., Iversen, P., Shestack, J., Jones, P., Ducat, L. and Spence, S.J. (2001) The autism genetic resource exchange: a resource for the study of autism and related neuropsychiatric conditions. *Am. J. Hum. Genet.*, **69**, 463–466.
34. Schenck, A., Bardoni, B., Langmann, C., Harden, N., Mandel, J.L. and Giangrande, A. (2003) CYFIP/Sra-1 controls neuronal connectivity in *Drosophila* and links the Rac1 GTPase pathway to the fragile X protein. *Neuron*, **38**, 887–898.
35. Cui, X. and Churchill, G.A. (2003) Statistical tests for differential expression in cDNA microarray experiments. *Genome Biol.*, **4**, 210.
36. Sutcliffe, J.S., Nelson, D.L., Zhang, F., Pieretti, M., Caskey, C.T., Saxe, D. and Warren, S.T. (1992) DNA methylation represses FMR-1 transcription in fragile X syndrome. *Hum. Mol. Genet.*, **1**, 397–400.
37. Herzing, L.B., Cook, E.H., Jr. and Ledbetter, D.H. (2002) Allele-specific expression analysis by RNA-FISH demonstrates preferential maternal expression of UBE3A and imprint maintenance within 15q11–q13 duplications. *Hum. Mol. Genet.*, **11**, 1707–1718.
38. Tusher, V.G., Tibshirani, R. and Chu, G. (2001) Significance analysis of microarrays applied to the ionizing radiation response. *Proc. Natl Acad. Sci. USA*, **98**, 5116–5121.
39. Breitling, R., Armengaud, P., Amtmann, A. and Herzyk, P. (2004) Rank products: a simple, yet powerful, new method to detect differentially regulated genes in replicated microarray experiments. *FEBS Lett.*, **573**, 83–92.
40. Couve, A., Restituito, S., Brandon, J.M., Charles, K.J., Bawagan, H., Freeman, K.B., Pangalos, M.N., Calver, A.R. and Moss, S.J. (2004) Marlin-1, a novel RNA-binding protein associates with GABA receptors. *J. Biol. Chem.*, **279**, 13934–13943.
41. Steindler, C., Li, Z., Algarte, M., Alcover, A., Libri, V., Ragimbeau, J. and Pellegrini, S. (2004) Jamip1 (marlin-1) defines a family of proteins interacting with janus kinases and microtubules. *J. Biol. Chem.*, **279**, 43168–43177.
42. Dennis, G., Jr., Sherman, B.T., Hosack, D.A., Yang, J., Gao, W., Lane, H.C. and Lempicki, R.A. (2003) DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol.*, **4**, 3. P.
43. Bagni, C. and Greenough, W.T. (2005) From mRNP trafficking to spine dysmorphogenesis: the roots of fragile X syndrome. *Nat. Rev. Neurosci.*, **6**, 376–387.
44. Ingenuity Pathway Analysis. http://www.ingenuity.com/.
45. Millar, J.K., Pickard, B.S., Mackie, S., James, R., Christie, S., Buchanan, S.R., Malloy, M.P., Chubb, J.E., Huston, E., Baillie, G.S. *et al.* (2005) DISC1 and PDE4B are interacting genetic factors in schizophrenia that regulate cAMP signaling. *Science*, **310**, 1187–1191.
46. Miyashiro, K.Y., Beckel-Mitchener, A., Purk, T.P., Becker, K.G., Barret, T., Liu, L., Carbonetto, S., Weiler, I.J., Greenough, W.T. and Eberwine, J. (2003) RNA cargoes associating with FMRP reveal deficits in cellular functioning in Fmr1 null mice. *Neuron*, **37**, 417–431.
47. Brookes, K., Xu, X., Chen, W., Zhou, K., Neale, B., Lowe, N., Aneey, R., Franke, B., Gill, M., Ebstein, R. *et al.* (2006) The analysis of 51 genes in DSM-IV combined type attention deficit hyperactivity disorder: association signals in DRD4, DAT1 and 16 other genes. *Mol. Psychiatry*, **10**, 934–953.
48. Hagerman, R.J. (2006) Lessons from fragile X regarding neurobiology, autism, and neurodegeneration. *J. Dev. Behav. Pediatr.*, **27**, 63–74.
49. Todd, R.D. (2001) Probing the nature of child psychopathology. *Child Adolesc. Psychiatr. Clin. N. Am.*, **10**, 209–224.
50. Mao, R., Zielke, C.L., Zielke, H.R. and Pevsner, J. (2003) Global up-regulation of chromosome 21 gene expression in the developing Down syndrome brain. *Genomics*, **81**, 457–467.
51. Tang, Y., Schapiro, M.B., Franz, D.N., Patterson, B.J., Hickey, F.J., Schorry, E.K., Hopkin, R.J., Wylie, M., Narayan, T., Glauser, T.A. *et al.* (2004) Blood expression profiles for tuberous sclerosis complex 2, neurofibromatosis type 1, and Down's syndrome. *Ann. Neurol.*, **56**, 808–814.
52. Middleton, F.A., Pato, C.N., Gentile, K.L., McGann, L., Brown, A.M., Trauzzi, M., Diab, H., Morley, C.P., Medeiros, H., Macedo, A. *et al.* (2005) Gene expression analysis of peripheral blood leukocytes from discordant sib-pairs with schizophrenia and bipolar disorder reveals points of convergence between genetic and functional genomic approaches. *Am. J. Med. Genet. B Neuropsychiatr. Genet.*, **136**, 12–25.
53. Tabata, T., Araishi, K., Hashimoto, K., Hashimotodani, Y., van der Putten, H., Bettler, B. and Kano, M. (2004) Ca2+ activity at GABAB receptors constitutively promotes metabotropic glutamate signaling in the absence of GABA. *Proc. Natl Acad. Sci. USA*, **101**, 16952–16957.
54. Bear, M.F., Huber, K.M. and Warren, S.T. (2004) The mGluR theory of fragile X mental retardation. *Trends Neurosci.*, **27**, 370–377.

55. Polleux, F. and Lauder, J.M. (2004) Toward a developmental neurobiology of autism. *Ment. Retard. Dev. Disabil. Res. Rev.*, **10**, 303–317.
56. Allen Brain Atlas. http://www.brain-map.org/.
57. Darnell, J.C., Jensen, K.B., Jin, P., Brown, V., Warren, S.T. and Darnell, R.B. (2001) Fragile X mental retardation protein targets G quartet mRNAs important for neuronal function. *Cell*, **107**, 489–499.
58. Kikin, O., D'Antonio, L. and Bagga, P.S. (2006) QGRS Mapper: a web-based server for predicting G-quadruplexes in nucleotide sequences. *Nucleic Acids Res.*, **34**, W676–W682.
59. Zalfa, F., Giorgi, M., Primerano, B., Moro, A., Di Penta, A., Reis, S., Oostra, B. and Bagni, C. (2003) The fragile X syndrome protein FMRP associates with BC1 RNA and regulates the translation of specific mRNAs at synapses. *Cell*, **112**, 317–327.
60. Jin, P., Alisch, R.S. and Warren, S.T. (2004) RNA and microRNAs in fragile X mental retardation. *Nat. Cell Biol.*, **6**, 1048–1053.
61. Griffiths-Jones, S., Grocock, R.J., van Dongen, S., Bateman, A. and Enright, A.J. (2006) miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Res.*, **34**, D140–D144.
62. Mottron, L. (2004) Matching strategies in cognitive research with individuals with high-functioning autism: current practices, instrument biases, and recommendations. *J. Autism Dev. Disord.*, **34**, 19–27.
63. Yoshikawa, T., Turner, G., Esterling, L.E., Sanders, A.R. and Detera-Wadleigh, S.D. (1997) A novel human myo-inositol monophosphatase gene, IMP.18p, maps to a susceptibility region for bipolar disorder. *Mol. Psychiatry*, **2**, 393–397.
64. Nothen, M.M., Cichon, S., Rohleder, H., Hemmer, S., Franzek, E., Fritze, J., Albus, M., Borrmann-Hassenbach, M., Kreiner, R., Weigelt, B. *et al.* (1999) Evaluation of linkage of bipolar affective disorder to chromosome 18 in a sample of 57 German families. *Mol. Psychiatry*, **4**, 76–84.
65. Lin, P.I., McInnis, M.G., Potash, J.B., Willour, V.L., Mackinnon, D.F., Miao, K., Depaulo, J.R. and Zandi, P.P. (2005) Assessment of the effect of age at onset on linkage to bipolar disorder: evidence on chromosomes 18p and 21q. *Am. J. Hum. Genet.*, **77**, 545–555.
66. Yoon, I.S., Li, P.P., Siu, K.P., Kennedy, J.L., Cooke, R.G., Parikh, S.V. and Warsh, J.J. (2001) Altered IMPA2 gene expression and calcium homeostasis in bipolar disorder. *Mol. Psychiatry*, **6**, 678–683.
67. Sjoholt, G., Ebstein, R.P., Lie, R.T., Berle, J.O., Mallet, J., Deleuze, J.F., Levinson, D.F., Laurent, C., Mujahed, M., Bannoura, I. *et al.* (2004) Examination of IMPA1 and IMPA2 genes in manic-depressive patients: association between IMPA2 promoter polymorphisms and bipolar disorder. *Mol. Psychiatry*, **9**, 621–629.
68. Ohnishi, T., Yamada, K., Ohba, H., Iwayama, Y., Toyota, T., Hattori, E., Inada, T., Kunugi, H., Tatsumi, M., Ozaki, N. *et al.* (2007) *Neuropsychopharmacology*, in press.
69. Romano, V., Cali, F., Seidita, G., Mirisola, M., D'Anna, R.P., Gambino, G., Schinocca, P., Romano, S., Ayala, G.F., Canziani, F. *et al.* (2005) Suggestive evidence for association of D2S2188 marker (2q31.1) with autism in 143 Sicilian (Italian) TRIO families. *Psychiatr. Genet.*, **15**, 149–150.
70. Lord, C., Rutter, M. and Le Couteur, A. (1994) Autism Diagnostic Interview-Revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *J. Autism Dev. Disord.*, **24**, 659–685.
71. Lord, C., Leventhal, B.L. and Cook, E.H., Jr. (2001) Quantifying the phenotype in autism spectrum disorders. *Am. J. Med. Genet.*, **105**, 36–38.
72. Saeed, A.I., Sharov, V., White, J., Li, J., Liang, W., Bhagabati, N., Braisted, J., Klapa, M., Currier, T., Thiagarajan, M. *et al.* (2003) TM4: a free, open-source system for microarray data management and analysis. *Biotechniques*, **34**, 374–378.
73. Gentleman, R.C., Carey, V.J., Bates, D.M., Bolstad, B., Dettling, M., Dudoit, S., Ellis, B., Gautier, L., Ge, Y., Gentry, J. *et al.* (2004) Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol.*, **5**, R80.
74. Jeffery, I.B., Higgins, D.G. and Culhane, A.C. (2006) Comparison and evaluation of methods for generating differentially expressed gene lists from microarray data. *BMC Bioinformatics*, **7**, 359.