Molecular Autism

RESEARCH　　　　　　　　　　　　　　　　　　　　　　　　　　**Open Access**

Increased gene expression of *FOXP1* in patients with autism spectrum disorders

Wei-Hsien Chien[1], Susan Shur-Fen Gau[2,3,4*], Chun-Houh Chen[5], Wen-Che Tsai[3], Yu-Yu Wu[6], Po-Hsu Chen[7], Chi-Yung Shang[2,3] and Chia-Hsiang Chen[2,6*]

Abstract

Background: Comparative gene expression profiling analysis is useful in discovering differentially expressed genes associated with various diseases, including mental disorders. Autism spectrum disorders (ASD) are a group of complex childhood-onset neurodevelopmental and genetic disorders characterized by deficits in language development and verbal communication, impaired reciprocal social interaction, and the presence of repetitive behaviors or restricted interests. The study aimed to identify novel genes associated with the pathogenesis of ASD.

Methods: We conducted comparative total gene expression profiling analysis of lymphoblastoid cell lines (LCL) between 16 male patients with ASD and 16 male control subjects to screen differentially expressed genes associated with ASD. We verified one of the differentially expressed genes, *FOXP1*, using real-time quantitative PCR (RT-qPCR) in a sample of 83 male patients and 83 male controls that included the initial 16 male patients and male controls, respectively.

Results: A total of 252 differentially expressed probe sets representing 202 genes were detected between the two groups, including 89 up- and 113 downregulated genes in the ASD group. RT-qPCR verified significant elevation of the *FOXP1* gene transcript of LCL in a sample of 83 male patients (10.46 ± 11.34) compared with 83 male controls (5.17 ± 8.20, *P* = 0.001).

Conclusions: Comparative gene expression profiling analysis of LCL is useful in discovering novel genetic markers associated with ASD. Elevated gene expression of *FOXP1* might contribute to the pathogenesis of ASD.

Clinical trial registration: Identifier: NCT00494754

Keywords: Autism, FOXP1, Expression microarray, Genetics, Lymphoblastoid cell line

Background

Autism spectrum disorders (ASD) represent a constellation of childhood-onset neurodevelopmental disorders characterized by three core symptom domains: deficits in reciprocal social interactions, impairments in language development and verbal communication, and the presence of stereotyped, repetitive behaviors and restricted interests [1]. The clinical presentations of ASD are highly heterogeneous with varied expressivity and differences in the severity of clinical symptoms and functional impairments. The estimated prevalence of ASD in the general population worldwide is approximately 1 to 2.6%, with a male predominance [2,3]. The etiology of ASD is mainly attributable to genetic factors. The estimated heritability of ASD is more than 90%, and the genetic underpinnings of ASD are heterogeneous and complex, involving multiple genes, gene-gene interactions, and gene-environmental interactions [4,5]. Identification of genetic underpinnings can shed light on the pathogenesis of ASD, and facilitate the development of new treatments for ASD. The genetic risk factors for ASD identified so far range from common variants conferring a small clinical effect to rare mutations that confer a high clinical effect [6]. The high heterogeneity of ASD may account for the varied clinical presentations of patients with ASD. Several genes have been found to be associated with ASD, but more genes remain to be discovered.

---

* Correspondence: gaushufe@ntu.edu.tw; cchen3801@gmail.com
[2]Department of Psychiatry, National Taiwan University College of Medicine, No.1 Jen-Ai Rd. Section 1, Taipei, Taiwan
[6]Department of Psychiatry, Chang Gung Memorial Hospital at Linkou, 5 Fusing Street, Kueishan, Taoyuan, Taiwan
Full list of author information is available at the end of the article

Microarray-based gene expression profiling is a useful technology allowing simultaneous measurement of hundreds to thousands of gene transcripts, and is useful for large-scale gene discovery [7]. This technology has been used in several post-mortem brain studies of psychiatric disorders, including schizophrenia, bipolar disorder, and autism [8-11]. Since there are several limitations to the use of the post-mortem brain tissue in gene expression studies, and the study of fresh brain tissue from living psychiatric patients is impractical at the present time, several studies have reported the use of peripheral blood cells and lymphoblastoid cell lines (LCL) as surrogates for brain tissue in the gene expression studies of mental disorders [12-16]. In addition, a moderate correlation of gene expression between peripheral blood cells and brain tissue in humans has been reported, supporting the usefulness of peripheral blood cells in the gene expression studies for psychiatric research [17].

Comparative gene expression profiling analysis of LCL has been conducted in several autism studies [18-22]. For example, several differentially expressed genes important to the development, structure, and/or function of the nervous system were detected in monozygotic twins discordant for autism using this method [18]. Another genome-wide expression profiling of LCL identified shared pathways among different forms of autism [19]. Also, altered pathways in neural development and steroid biogenesis were detected in a study of total gene expression profiling analysis of LCL in autistic patients and their unaffected sib-pairs [20]. Another study found that the imbalance of cellular and mitochondrial glutathione redox in LCL was associated with autism [21]. LCL was also used successfully to investigate the role of microRNA in autism [22]. Taken together, these studies support the usefulness of comparative gene expression profiling analysis of LCL in the molecular genetic study of autism [23].

The above-mentioned findings attracted our interest in identifying novel differentially expressed genes associated with autism in our population using comparative gene expression profiling analysis of LCL. We first conducted a comparative gene expression profiling analysis of LCL in a sample of 16 male patients with ASD and 16 male control subjects as an initial screening of differentially expressed genes associated with autism. In the second stage, we verified the authenticity of a selected differentially expressed gene, *FOXP1*, in an expanded sample of 83 male patients and 83 male control subjects using real-time quantitative PCR (RT-qPCR).

Methods
Subjects
The study was conducted in two tiers. In the first experiment, microarray-based total gene expression profiling analysis was performed with 20 male patients with ASD

and 20 male control subjects, but only 16 of each group passed the quality control of the experiment and they were subjected to further analysis. In the second experiment, the transcript level of *FOXP1* was compared between 83 male patients with ASD and 83 healthy male controls. The sample included the 20 male patients and 20 male control subjects from the first tier of the experiment. All the patients were recruited from the outpatient clinics of psychiatric departments in two medical centers (National Taiwan University Hospital [NTUH] in Taipei and Chang-Gung Memorial Hospital (CGMH in Kueishan). All subjects were Han Chinese. The clinical diagnoses of probands were made by senior board-certified child psychiatrists according to the DSM-IV diagnostic criteria for autistic disorder or Asperger's disorder, and were further confirmed by structural interviews using the Chinese version of the Autism Diagnostic Interview-Revised (ADI-R) [24-27]. The ADI-R [24] is a standardized, comprehensive, semi-structured, investigator-based interview scale that was designed for interviewing the caregivers/parents to obtain information on the developmental and behavioral aspects of ASD of children with a mental age from about 18 months to adulthood. The behavioral aspects of ASD assessed by the ADI-R include reciprocal social interactions, communication, and repetitive stereotyped behaviors/interests. Gau and colleagues have prepared the Chinese version of the ADI-R, which was approved by the World Psychological Association (WPS) in May 2007, for use in this study [25-27].

The case group in the first tier of the experiment consisted of 20 male probands, aged 9.1 ± 3.2 years (range 4 to 18 years old). These participants were selected from a family cohort of 393 probands with ASD (373 with autistic disorder and 20 with Asperger's disorder) [27] because they had the highest severity scores for problems in verbal communication and social reciprocity based on the ADI-R interview and clinical assessments. Probands in the original cohort with fragile X syndrome or Rett's disorder, and probands who had a previously identified chromosomal abnormality associated with autism, or had any other neurological or medical conditions, were excluded from the study. The control group consisted of 20 males (mean age: 37.15 ± 16.90 years; range 18 to 67 years) who were recruited among the cohort of individuals receiving regular medical checkups in a local medical center.

In the second tier of the experiment, a total of 83 male patients with ASD (including 20 from the first tier of the experiment), aged 10.2 ± 4.6 years (range 3 to 23 years), were recruited. The control group consisted of 83 males (including 20 from the first tier of the experiment), aged 43.5 ± 12.7 years (range 18 to 63 years) who were recruited from among individuals who received regular medical checkups in a local medical center. The mental

status and history of mental illness of each control subject were evaluated by a senior psychiatrist using the Mini International Neuropsychiatric Interview (MINI); subjects diagnosed with a DSM-IV axis 1 disorder including ASD were excluded. The Research Ethics Committee of the two research sites approved this study (IRB ID: 9561709027 for NTUH and C20060905 for CGMH). Written informed consent and child assent were obtained from the parents and the probands, respectively, after the purposes and procedures of the study were fully explained and confidentiality was ensured.

Preparation of lymphoblastoid cell line, total RNA and cDNA

Immortalized LCL from each subject was established by transforming lymphocytes with Epstein-Barr virus (EBV) following the procedures described elsewhere [16]. Total RNA from the LCL was extracted using TRIZOL Total RNA Isolation Reagent according to the manufacturer's instructions (Invitrogen Life Technologies, Carlsbad, California, USA). The cDNA was prepared using Superscript II RNase H⁻ Reverse Transcriptase (Invitrogen Life Technologies). The details were described in our previous report [16].

Microarray-based total gene expression experiment

The microarray experiment was performed at the Core Laboratory of National Taiwan University Hospital, Taipei Taiwan using the Affymetrix Human Genome U133 Plus 2.0 Array (Affymetrix Inc., Santa Clara, CA, USA). The experimental procedures followed the protocols provided by the manufacturer.

Identification of differentially expressed genes

The intensity files of all the subjects were input into the computer program GAP: Generalized Association Plots [28,29] (http://gap.stat.sinica.edu.tw/Software/GAP/index.htm) for quality control using visualization and descriptive statistics. The chips that passed the quality control were subjected to further differentially expressed gene analysis. The data of the chips were normalized using the Robust Multi-array Analysis (RMA) method [30]. In order to filter out probe sets with low variations and to reduce the impact of multiple comparisons, we kept only the 1,000 probe sets with the largest standard deviations. A differentially expressed probe set was defined as a P value smaller than 0.05 for a Student's t-test adjusted by the false discovery rate [31] of 5% and fold-change greater than 2.

Pathway analysis

We performed gene set enrichment analysis (GSEA) on the entire expression profile for all 32 gene chips using the GSEA software (v2.0.10) (http://www.broadinstitute.org/gsea/index.jsp) [32,33]. The phenotype of subject category (ASD and Control) was employed as phenotype annotation for GSEA.

In addition to GSEA, we also carried out pathway analysis using DAVID (Database for Annotation, Visualization and Integrated Discovery) [34] and verified the results from GSEA (v2.0.10) and DAVID (v6.7) with hypergeometric test using the runHyperKEGG procedure in EMA v1.3.1 (A R package for Easy Microarray data Analysis) [35]. Analysis with all three environments (GSEA, DAVID, and EMA) was based on the gene population of Affymetrix Human Genome U133 Plus 2.0 Array annotation data (chip hgu133plus2) with KEGG v3.1 gene symbol as the gene set and pathway database. The lists of up- and downregulated gene probe sets were fed to DAVID and runHyperKEGG.

Real-time quantitative PCR

RT-qPCR was performed using the SYBR Green method and implemented in the StepOnePlus Real-Time PCR System according to the manufacturer's protocol (Applied Biosystems, Forster City, CA, USA). The detailed procedures were described in our previous report [16]. The relative standard curve method was used for quantification of mRNA expression (User Bulletin #2 ABI PRISM 7700 sequence detection system). With this method, serial dilutions of known amounts of RNA from a reference sample (pooled from 40 LCLs of male controls) were used to generate an external standard curve. For each unknown sample the relative amount was calculated using linear regression analysis from their respective standard curves. The expression level of FOXP1 mRNA was normalized by 18S rRNA. The18S rRNA reference gene was measured using pre-developed TaqMan assay reagents 18S rRNA MGB according to the manufacturer's protocol (Applied Biosystems). All experiments were performed in duplicate. Primer sequences for PCR amplification were designed using online Primer3 software (http://bioinfo.ut.ee/primer3/). All the sequences of primer sets used for RT-qPCR are listed in Table 1.

Results

Characteristics of patients with Autism Spectrum Disorders

The ADI-R interviews regarding the past behaviors revealed that the 16 patients, who underwent the microarray experiment, scored 24.31 ± 3.88 (range, 14 to 28) in the 'qualitative abnormalities in reciprocal social interaction' (cut-off = 10), 17.69 ± 2.98 (range, 14 to 28) in the 'qualitative abnormalities in communication, verbal' (cut-off = 8), 8.94 ± 2.72 (range, 4 to 14) in the 'qualitative abnormalities in communication, nonverbal' (cut-off = 7), and 8.69 ± 2.24 (range, 5 to 12) in the 'restricted, repetitive and stereotyped patterns of behaviors' (cutoff = 3). The average age at which the 16 patients said their first word was 27.5 ± 9.9 (range, 12 to 42) months and at which they said their first

Table 1 Sequences of primers used in the real-time quantitative PCR experiments in this study

Gene	Forward primer	Backward primer	Ta
FOXP1	5′-CAGATATTGCGCAGAACCAA-3′	5′-GCAAACATTCGTGTGAACCA-3′	60
FOXP2	5′-AGCAAATGCAGCAGATCCTT-3′	5′-TGCTGCAAAAGCTGAAGATG-3′	60
CNTNAP2	5′-CAGCTCCTCCTCCATCTCTG-3′	5′-TCACCCAATCTGAGCTGCTA-3′	60

Ta, annealing temperature. FOXP1[GenBank:NM_001012505.1], FOXP2[GenBank:NM_148898.3], CNTNAP2[GenBank:NM_014141.5].

phrase was 47.5 ± 12.4 (range, 30 to 78) months. Their current average intelligence quotients (IQ) were 82.13 ± 16.23 (range 62 to 117) for full-scale IQ, 89.13 ± 14.63 (range, 68 to 113) for performance IQ, and 78.44 ± 20.34 (range, 45 to 118) for verbal IQ assessed by the Wechsler Intelligence Scale for Children-3rd edition.

The ADI-R results from the parents of 83 patients showed that the 83 patients scored 24.07 ± 4.78 in the 'qualitative abnormalities in reciprocal social interaction' (cut-off = 10), 18.05 ± 3.74 in the 'qualitative abnormalities in communication, verbal' (cut-off = 8), 10.14 ± 2.41in the 'qualitative abnormalities in communication, nonverbal' (cut-off = 7), and 7.76 ± 2.41 in the 'restricted, repetitive and stereotyped patterns of behaviors' (cutoff = 3). The average age at which first word was said was 33.3 ± 16.6 (range, 12 to 84) months and that they had their first phrase was said was 49.9 ± 23.1 (range, 15 to 138) months. The average full-scale IQ, performance IQ, and verbal IQ of these 83 patients were 89.31 ± 21.52 (range, 62 to 130), 93.74 ± 19.61 (range, 68 to 140), and 86.89 ± 24.49 (range, 46 to 125), respectively, as assessed by the Wechsler Intelligence Scale for Children-3rd edition. All the probands had abnormal development evident before 30 months of age based on the ADI-R interview and validated by the medical records.

Identification of differentially expressed genes
The microarray data of 16 patients and 16 male controls passed the quality control. The data of 32 gene chips were then normalized using the RMA method [30]. The descriptive statistics for the distribution of standard deviations of all 54,675 probe sets were in the ranged from 0.025 to 3.668 with the 25th percentile as 0.157, median as 0.202, the 75th percentile as 0.298, and mean as 0.260; and the 98.17th (1,000 among 54,675) percentile was 0.835. Two hundred and fifty-two probe sets of 1,000 probe sets had standard deviations greater than 0.835 expressed differentially among the two groups with a P value less than 0.05 for a Student's t-test adjusted by the false discovery rate [31] of 5% and fold-change greater than 2. These 252 differentially expressed probe sets corresponded to 202 genes with 89 of them as upregulated and 113 as downregulated transcripts in the ASD group compared to the control group. The list of differentially expressed genes is shown in Additional file 1.
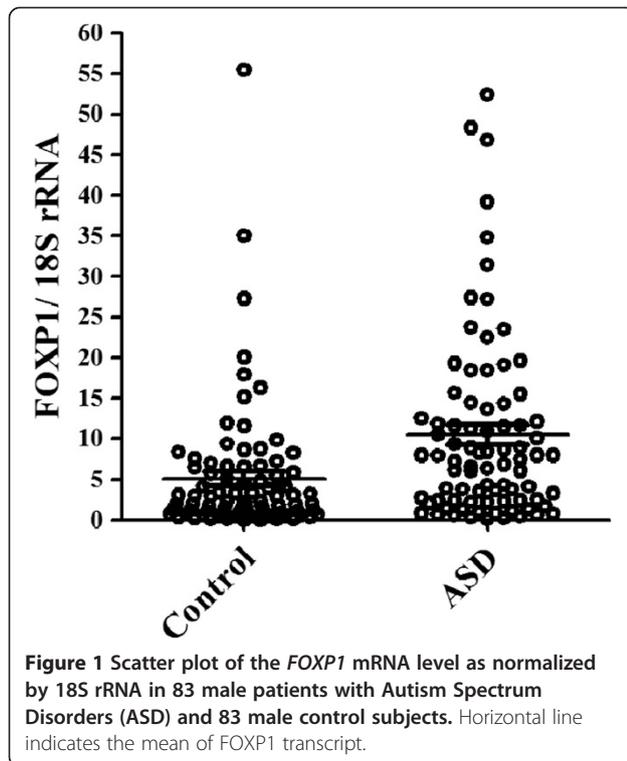
Pathway analysis
We performed gene set enrichment analysis on the entire expression profile of all 32 gene chips using the GSEA software (v2.0.10). The most significant 20 KEGG signaling pathways for both up- and downregulated gene sets in the significance order (size of P values) of GSEA are listed in the Additional file 2. The results of pathway analysis of DAVID and runHyperKEGG are also shown in the Additional file 2. Seven pathways were commonly found in three analyses, including the long-term depression, the vascular smooth muscle contraction, the arrhythmogenic right ventricular cardiomyopathy (ARVC), the glycerophospholipid metabolism, the allograft rejection, the Jak-STAT signaling pathway, the toll-like receptor signaling pathway, and the RIG-I-like receptor signaling pathways.

Verification by real-time quantitative PCR
The authenticity of the differential expression of FOXP1 was first verified between the 20 ASD and 20 control subjects using RT-qPCR. The average expression level of FOXP1 in the 20 patients with ASD (4.57 ± 3.04) was significantly higher than that of the control subjects (1.89 ± 2.64, P = 0.005). The upregulation of FOXP1 gene expression was further supported by an expanded sample size of 83 male ASD patients and 83 male control subjects. This expanded sample size included the initial 20 patients and 20 control subjects. The average transcript level of FOXP1 in the patient group (10.46 ± 11.34) was significantly higher than that in the control group (5.17 ± 4.20, P = 0.001). The distribution and mean of the FOXP1 transcript level of LCL in 83 patients and 83 controls are shown in Figure 1.

Discussion
Aberrant language development is one of the clinical characteristics of autism. Mutations of FOXP2 have been linked to speech and language disorders and ASD [36-38]. FOXP1 is a member of subfamily P of the forkhead box (FOX) transcription factor family and forms a heterodimer with FOXP2 [39,40]. The elevated mRNA level of FOXP1 in patients with ASD in the initial expression microarray experiment attracted our attention; hence, it was selected for further validation in an expanded sample size using RT-qPCR. The present study showed significant elevation of the mRNA level of FOXP1 in patients with ASD compared to that in the control group. Our finding

Figure 1 Scatter plot of the *FOXP1* mRNA level as normalized by 18S rRNA in 83 male patients with Autism Spectrum Disorders (ASD) and 83 male control subjects. Horizontal line indicates the mean of FOXP1 transcript.

suggests that increased *FOXP1* expression may be involved in the pathogenesis of ASD. However, the interpretation of the result should be cautious as there was a significant difference in the mean age between the two groups, which might be a confounder affecting *FOXP1* gene expression. We were not able to obtain blood from the children and young adolescents as the age-matched controls because of ethical considerations and patient protection, which is a limitation of this study. In addition, the findings from LCL can only be considered as indirect evidence reflecting what may happen in the brain.

The *FOXP1* gene encodes a member of the forkhead box transcription factor family that contains a DNA-binding domain and a protein-protein interaction domain. The *FOXP1* gene functions as a transcription repressor [39,41,42], and is widely expressed in the developing and mature brain. The gene has been suggested to be involved in the development and function of the brain [41,43]. In the literature, three subjects with mental retardation and significant language and speech deficits were detected to have heterozygous deletions overlapping the *FOXP1* gene [44]. Two mentally retarded individuals with autistic features were detected to have a *de novo* intergenic deletion and a *de novo* nonsense mutation in the *FOXP1* gene, respectively [45]. In an exome sequencing study of 20 sporadic ASD patients (simplex ASD), a *de novo* single-base insertion in *FOXP1* that introduces a frameshift and a premature stop codon was identified in a severely affected

patient [46]. These data suggest that haploinsufficiency or hypomorphic mutations of *FOXP1* with reduced expression or deficient activity of *FOXP1* are associated with syndromic or non-syndromic ASD. However, these *FOXP1* mutations associated with ASD are rare; they may not apply to the pathogenesis of autism in general. In the present study, we found that increased *FOXP1* gene expression was associated with autism in general. Our finding may expand our understanding about the relationship of *FOXP1* with autism.

FOXP1 protein was known to interact with another subfamily member FOXP2 to form a heterodimer and co-expressed with FOXP2 in several brain regions, suggesting close functional collaboration between the two proteins [39,40]. Contactin-associated protein-like 2 (*CNTNAP2*), a neurexin family protein that functions as a neuronal adhesion molecule and receptor, was found to be a direct neural target bound by human FOXP2 protein [47]. Mutations of *FOXP2* and *CNTNAP2* were linked to speech and language disorders and ASD [36-38,48-51]. Taken together, these data indicate that interactions among *FOXP1*, *FOXP2*, and *CNTNAP2* genes may play an essential role underlying the pathogenesis of syndromic and non-syndromic ASD.

In a more recent study, forkhead box protein p1 was found to function as a transcriptional repressor of immune signaling in the mouse brain, and was involved in the pathophysiology of Huntington's disease [52]. The study suggested that Foxp1-regulated pathways might be important mediators of neuronal-glial cell communication. Thus, the increased *FOXP1* gene expression in the LCL of ASD patients as found in this study may offer a new insight that dysregulated immune signaling in the brain contributes to the pathogenesis of ASD.

A previous study demonstrated that FOXP2 protein directly binds to intron 1 of the *CNTNAP2* gene and significantly downregulates the mRNA levels of *CNTNAP2* [47]. This interaction was further supported by the findings in a study of Foxp2 R553H knock-in mice [53], which showed a significant increase in the Cntnap2 mRNA level in the cerebellum of the transgenic animal [53]. Thus, we inferred that increased *FOXP1* gene expression may lead to increased expression of the *FOXP2* gene through a feedback mechanism, which may in turn reduce the gene expression of *CNTNAP2* in patients with ASD. We hypothesized that there might be significant differences in the gene expression of *FOXP2* and *CNTNAP2* in LCL between patients with ASD and control subjects. To address this issue, we attempted to measure the mRNA of *FOXP2* and *CNTNAP2* in LCL in our subjects using RT-qPCR. However, the mRNA levels of these two genes were too low to be detected in our system.

The present study may have several implications. First, as LCL is derived from the EBV transformation of lymphocytes of peripheral blood, it would be interesting to

know whether there is also an increased *FOXP1* expression in the peripheral blood cells. If so, the increased *FOXP1* expression can be considered as a potential biomarker of ASD in future studies. Second, recent reports indicate that the expression of FOXP1 is subject to regulation by cis-acting SNPs, methylation, and microRNA [54-57], and thus, it would be interesting to explore the underlying mechanisms that may lead to increased expression of the *FOXP1* gene in ASD. Third, recent studies have demonstrated the feasibility of establishing induced neuronal cells from induced pluripotent stem cells (iPSC) [58] or directly from non-neuronal somatic cells [59]. Hence, it would be interesting to verify whether the increased expression of *FOXP1* in LCL of ASD patients can be observed in the induced neuronal cells. Furthermore, it would also be interesting to measure the expression levels of *FOXP2* and *CNTNAP2* in the induced neuronal cells that cannot be detected in LCL and investigate their interactions with *FOXP1*.

In a recent review of studies of genome-wide expression analysis in LCL or blood cells in ASD patients, Voineagu reported that not much of a convergent pathway was found in terms of dysregulated pathways from several studies he reviewed [60]. He provided several explanations for such discrepant findings, including differences in study designs, sample sizes, and the clinical heterogeneity of patients with ASD [60]. Nevertheless, given no overlap of specific pathways and genes among different gene expression studies of peripheral tissues in autism patients, he reported that an upregulation of several genes involved in immune response is common to several studies, which is compatible with the findings from several total gene expression studies of brain tissue in autism patients [60]. In the present study, when we carried out pathway analysis using three platforms, we found several immune-related pathways showing significant differences between the ASD patients and controls, including the RIG-I-like receptor signaling, the toll-like receptor signaling, the allograft rejection, the primary immunodeficiency, the cytokine-cytokine receptor interaction, the chemokine signaling, the intestinal immune network for IgA production, and the Jak-STAT signaling pathways. Our findings are compatible with several previous studies and support the notion that dysregulated immunological pathways may be involved in the pathogenesis of ASD [60]. In addition, we found several other pathways showing significant differences between the ASD patients and controls in three platforms, such as the long-term depression, the vascular smooth muscle contraction, the arrhythmogenic right ventricular cardiomyopathy, and the glycerophospholipid metabolism pathways. The detection of the long-term depression pathway is consistent with the finding from the study by Kong and colleagues [61]. Taken together, our findings in pathway analysis support the idea that autism is a complex genetic disorder.

Conclusions

Our study suggests that comparative gene expression profiling analysis of LCL is a useful tool to discover novel genes associated with autism. In this study, we found a significant elevation of the gene transcript of *FOXP1* in patients with ASD, suggesting that increased *FOXP1* gene expression might be involved in the pathogenesis of ASD. Nevertheless, limitations of this study, such as the small sample size, a significant difference in the mean age between the two groups, and the non-brain tissue of the study, should be taken into consideration while assessing these findings.

Additional files

Additional file 1: List of the differentially expressed probe sets and genes identified in this study.

Additional file 2: List of pathways showing differences between the Autistic Spectrum Disorder patients and controls using GSEA, DAVID, and EMA in this study.

Author details
[1]Department of Occupational Therapy, College of Medicine, Fu-Jen Catholic University, No. 510, Zhongzheng Rd, New Taipei City, Taiwan. [2]Department of Psychiatry, National Taiwan University College of Medicine, No.1 Jen-Ai Rd. Section 1, Taipei, Taiwan. [3]Department of Psychiatry, National Taiwan

University Hospital, No. 7, Chung-Shan South Rd, Taipei, Taiwan. ⁴Department of Psychology and School of Occupational Therapy, Graduate Institute of Brain and Mind Sciences, Graduate Institute of Epidemiology and Preventive Medicine, National Taiwan University, Taipei, Taiwan. ⁵Institute of Statistical Science, Academia Sinica, No. 128, Academia Rd. Nankang, Taipei, Taiwan. ⁶Department of Psychiatry, Chang Gung Memorial Hospital at Linkou, 5 Fusing Street, Kueishan, Taoyuan, Taiwan. ⁷Department of Statistics, Ohio State University, 1958 Neil Ave, Columbus, OH, USA.

References

1. America Psychiatric Association: *Diagnostic and Statistical Manual of Mental Disorders (DSM-IV)*. Washington, DC: American Psychiatric Association Press Inc; 1994.
2. Kogan MD, Blumberg SJ, Schieve LA, Boyle CA, Perrin JM, Ghandour RM, Singh GK, Strickland BB, Trevathan E, van Dyck PC: **Prevalence of parent-reported diagnosis of autism spectrum disorder among children in the US, 2007.** *Pediatrics* 2009, **124**:1395–1403.
3. Kim YS, Leventhal BL, Koh YJ, Fombonne E, Laska E, Lim EC, Cheon KA, Kim SJ, Kim YK, Lee H, Song DH, Grinker RR: **Prevalence of autism spectrum disorders in a total population sample.** *Am J Psychiatry* 2011, **168**:904–912.
4. Holt R, Monaco AP: **Links between genetics and pathophysiology in the autism spectrum disorders.** *EMBO Mol Med* 2011, **3**:438–450.
5. Eapen V: **Genetic basis of autism: is there a way forward?** *Curr Opin Psychiatry* 2011, **24**:226–236.
6. State MW, Levitt P: **The conundrums of understanding genetic risks for autism spectrum disorders.** *Nat Neurosci* 2011, **14**:1499–1506.
7. Schena M, Shalon D, Heller R, Chai A, Brown PO, Davis RW: **Parallel human genome analysis: microarray-based expression monitoring of 1000 genes.** *Proc Natl Acad Sci USA* 1996, **93**:10614–10619.
8. Mistry M, Gillis J, Pavlidis P: **Genome-wide expression profiling of schizophrenia using a large combined cohort.** *Mol Psychiatry* 2013, **18**:215–225.
9. Ginsberg SD, Hemby SE, Smiley JF: **Expression profiling in neuropsychiatric disorders: emphasis on glutamate receptors in bipolar disorder.** *Pharmacol Biochem Behav* 2012, **100**:705–711.
10. Purcell AE, Jeon OH, Zimmerman AW, Blue ME, Pevsner J: **Postmortem brain abnormalities of the glutamate neurotransmitter system in autism.** *Neurology* 2001, **57**:1618–1628.
11. Konradi C: **Gene expression microarray studies in polygenic psychiatric disorders: applications and data analysis.** *Brain Res Brain Res Rev* 2005, **50**:142–155.
12. Woelk CH, Singhania A, Perez-Santiago J, Glatt SJ, Tsuang MT: **The utility of gene expression in blood cells for diagnosing neuropsychiatric disorders.** *Int Rev Neurobiol* 2011, **101**:41–63.
13. Gardiner E, Beveridge NJ, Wu JQ, Carr V, Scott RJ, Tooney PA, Cairns MJ: **Imprinted DLK1-DIO3 region of 14q32 defines a schizophrenia-associated miRNA signature in peripheral blood mononuclear cells.** *Mol Psychiatry* 2012, **17**:827–840.
14. Yasuda Y, Hashimoto R, Yamamori H, Ohi K, Fukumoto M, Umeda-Yano S, Mohri I, Ito A, Taniike M, Takeda M: **Gene expression analysis in lymphoblasts derived from patients with autism spectrum disorder.** *Mol Autism* 2011, **2**:9.
15. Padmos RC, Hillegers MH, Knijff EM, Vonk R, Bouvy A, Staal FJ, de Ridder D, Kupka RW, Nolen WA, Drexhage HA: **A discriminating messenger RNA signature for bipolar disorder formed by an aberrant expression of inflammatory genes in monocytes.** *Arch Gen Psychiatry* 2008, **65**:395–407.
16. Huang CH, Chen ML, Tsai YL, Tsai MT, Chen CH: **Elevated adrenomedullin mRNA in lymphoblastoid cells from schizophrenic patients.** *Neuroreport* 2004, **15**:1443–1446.
17. Sullivan PF, Fan C, Perou CM: **Evaluating the comparability of gene expression in blood and brain.** *Am J Med Genet B Neuropsychiatr Genet* 2006, **141B**:261–268.
18. Hu VW, Frank BC, Heine S, Lee NH, Quackenbush J: **Gene expression profiling of lymphoblastoid cell lines from monozygotic twins discordant in severity of autism reveals differential regulation of neurologically relevant genes.** *BMC Genomics* 2006, **7**:118.
19. Nishimura Y, Martin CL, Vazquez-Lopez A, Spence SJ, Alvarez-Retuerto AI, Sigman M, Steindler C, Pellegrini S, Schanen NC, Warren ST, Geschwind DH: **Genome-wide expression profiling of lymphoblastoid cell lines distinguishes different forms of autism and reveals shared pathways.** *Hum Mol Genet* 2007, **16**:1682–1698.
20. Hu VW, Nguyen A, Kim KS, Steinberg ME, Sarachana T, Scully MA, Soldin SJ, Luu T, Lee NH: **Gene expression profiling of lymphoblasts from autistic and nonaffected sib pairs: altered pathways in neuronal development and steroid biosynthesis.** *PLoS One* 2009, **4**:e5775.
21. James SJ, Rose S, Melnyk S, Jernigan S, Blossom S, Pavliv O, Gaylor DW: **Cellular and mitochondrial glutathione redox imbalance in lymphoblastoid cells derived from children with autism.** *FASEB J* 2009, **23**:2374–2383.
22. Talebizadeh Z, Butler MG, Theodoro MF: **Feasibility and relevance of examining lymphoblastoid cell lines to study role of microRNAs in autism.** *Autism Res* 2008, **1**:240–250.
23. Baron CA, Liu SY, Hicks C, Gregg JP: **Utilization of lymphoblastoid cell lines as a system for the molecular modeling of autism.** *J Autism Dev Disord* 2006, **36**:973–982.
24. Lord C, Rutter M, Le Couteur A: **Autism diagnostic interview-revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders.** *J Autism Dev Disord* 1994, **24**:659–685.
25. Gau SS, Chou MC, Lee JC, Wong CC, Chou WJ, Chen MF, Soong WT, Wu YY: **Behavioral problems and parenting style among Taiwanese children with autism and their siblings.** *Psychiatry Clin Neurosci* 2010, **64**:70–78.
26. Chien WH, Wu YY, Gau SSF, Huang YS, Soong WT, Chiu YN, Chen CH: **Association study of the SLC25A12 gene and autism in Han Chinese in Taiwan.** *Prog Neuropsychopharmacol Biol Psychiatry* 2010, **34**:189–192.
27. Chien Y-L, Gau SS-F, Gadow KD: **Sex difference in the rates and co-occurring conditions of psychiatric symptoms in incoming college students in Taiwan.** *Compr Psychiatry* 2011, **52**:195–207.
28. Chen CH: **Generalized association plots: information visualization via iteratively generated correlation matrices.** *Stat Sin* 2002, **12**:7–30.
29. Wu HM, Tien YJ, Chen CH: **GAP: a graphical environment for matrix visualization and cluster analysis.** *Comput Stat Data Anal* 2010, **54**:767–778.
30. Irizarry RA, Bolstad BM, Collin F, Cope LM, Hobbs B, Speed TP: **Summaries of affymetrix GeneChip probe level data.** *Nucleic Acids Res* 2003, **31**:e15.
31. Yoav B, Yosef H: **Controlling the false discovery rate: a practical and powerful approach to multiple testing.** *J R Stat Soc Series B Stat Methodol* 1995, **57**:289–300.
32. Mootha VK, Lindgren CM, Eriksson KF, Subramanian A, Sihag S, Lehar J, Puigserver P, Carlsson E, Ridderstrale M, Laurila E, Houstis N, Daly MJ, Patterson N, Mesirov JP, Golub TR, Tamayo P, Spiegelman B, Lander ES, Hirschhorn JN, Altshuler D, Groop LC: **PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes.** *Nat Genet* 2003, **34**:267–273.
33. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP: **Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles.** *Proc Natl Acad Sci USA* 2005, **102**:15545–15550.
34. da Huang W, Sherman BT, Lempicki RA: **Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources.** *Nat Protoc* 2009, **4**:44–57.
35. Servant N, Gravier E, Gestraud P, Laurent C, Paccard C, Biton A, Brito I, Mandel J, Asselain B, Barillot E, Hupé P: **EMA - A R package for easy microarray data analysis.** *BMC research notes* 2010, **3**:277.
36. Lai CS, Fisher SE, Hurst JA, Vargha-Khadem F, Monaco AP: **A forkhead-domain gene is mutated in a severe speech and language disorder.** *Nature* 2001, **413**:519–523.
37. Newbury DF, Monaco AP: **Genetic advances in the study of speech and language disorders.** *Neuron* 2010, **68**:309–320.
38. Mikhail FM, Lose EJ, Robin NH, Descartes MD, Rutledge KD, Rutledge SL, Korf BR, Carroll AJ: **Clinically relevant single gene or intragenic deletions encompassing critical neurodevelopmental genes in patients with developmental delay, mental retardation, and/or autism spectrum disorders.** *Am J Med Genet A* 2011, **155A**:2386–2396.
39. Wang B, Lin D, Li C, Tucker P: **Multiple domains define the expression and regulatory properties of Foxp1 forkhead transcriptional repressors.** *J Biol Chem* 2003, **278**:24259–24268.
40. Li S, Weidenfeld J, Morrisey EE: **Transcriptional and DNA binding activity of the Foxp1/2/4 family is modulated by heterotypic and homotypic protein interactions.** *Mol Cell Biol* 2004, **24**:809–822.
41. Tamura S, Morikawa Y, Iwanishi H, Hisaoka T, Senba E: **Expression pattern of the winged-helix/forkhead transcription factor Foxp1 in the developing central nervous system.** *Gene Expr Patterns* 2003, **3**:193–197.

42. Banham AH, Beasley N, Campo E, Fernandez PL, Fidler C, Gatter K, Jones M, Mason DY, Prime JE, Trougouboff P, Wood K, Cordell JL: **The FOXP1 winged helix transcription factor is a novel candidate tumor suppressor gene on chromosome 3p.** *Cancer Res* 2001, **61**:8820–8829.

43. Ferland RJ, Cherry TJ, Preware PO, Morrisey EE, Walsh CA: **Characterization of Foxp2 and Foxp1 mRNA and protein in the developing and mature brain.** *J Comp Neurol* 2003, **460**:266–279.

44. Horn D, Kapeller J, Rivera-Brugues N, Moog U, Lorenz-Depiereux B, Eck S, Hempel M, Wagenstaller J, Gawthrope A, Monaco AP, Bonin M, Riess O, Wohlleber E, Illig T, Bezzina CR, Franke A, Spranger S, Villavicencio-Lorini P, Seifert W, Rosenfeld J, Klopocki E, Rappold GA, Strom TM: **Identification of FOXP1 deletions in three unrelated patients with mental retardation and significant speech and language deficits.** *Hum Mutat* 2010, **31**:E1851–E1860.

45. Hamdan FF, Daoud H, Rochefort D, Piton A, Gauthier J, Langlois M, Foomani G, Dobrzeniecka S, Krebs MO, Joober R, Lafrenière RG, Lacaille JC, Mottron L, Drapeau P, Beauchamp MH, Phillips MS, Fombonne E, Rouleau GA, Michaud JL: **De novo mutations in FOXP1 in cases with intellectual disability, autism, and language impairment.** *Am J Hum Genet* 2010, **87**:671–678.

46. O'Roak BJ, Deriziotis P, Lee C, Vives L, Schwartz JJ, Girirajan S, Karakoc E, Mackenzie AP, Ng SB, Baker C, Rieder MJ, Nickerson DA, Bernier R, Fisher SE, Shendure J, Eichler EE: **Exome sequencing in sporadic autism spectrum disorders identifies severe de novo mutations.** *Nat Genet* 2011, **43**:585–589.

47. Vernes SC, Newbury DF, Abrahams BS, Winchester L, Nicod J, Groszer M, Alarcon M, Oliver PL, Davies KE, Geschwind DH, Monaco AP, Fisher SE: **A functional genetic link between distinct developmental language disorders.** *N Eng J Med* 2008, **359**:2337–2345.

48. Alarcon M, Abrahams BS, Stone JL, Duvall JA, Perederiy JV, Bomar JM, Sebat J, Wigler M, Martin CL, Ledbetter DH, Nelson SF, Cantor RM, Geschwind DH: **Linkage, association, and gene-expression analyses identify CNTNAP2 as an autism-susceptibility gene.** *Am J Hum Genet* 2008, **82**:150–159.

49. Arking DE, Cutler DJ, Brune CW, Teslovich TM, West K, Ikeda M, Rea A, Guy M, Lin S, Cook EH, Chakravarti A: **A common genetic variant in the neurexin superfamily member CNTNAP2 increases familial risk of autism.** *Am J Hum Genet* 2008, **82**:160–164.

50. Bakkaloglu B, O'Roak BJ, Louvi A, Gupta AR, Abelson JF, Morgan TM, Chawarska K, Klin A, Ercan-Sencicek AG, Stillman AA, Tanriover G, Abrahams BS, Duvall JA, Robbins EM, Geschwind DH, Biederer T, Gunel M, Lifton RP, State MW: **Molecular cytogenetic analysis and resequencing of contactin associated protein-like 2 in autism spectrum disorders.** *Am J Hum Genet* 2008, **82**:165–173.

51. Gregor A, Albrecht B, Bader I, Bijlsma EK, Ekici AB, Engels H, Hackmann K, Horn D, Hoyer J, Klapecki J, Kohlhase J, Maystadt I, Nagl S, Prott E, Tinschert S, Ullmann R, Wohlleber E, Woods G, Reis A, Rauch A, Zweier C: **Expanding the clinical spectrum associated with defects in CNTNAP2 and NRXN1.** *BMC Med Genet* 2011, **12**:106.

52. Tang B, Becanovic K, Desplats PA, Spencer B, Hill AM, Connolly C, Masliah E, Leavitt BR, Thomas EA: **Forkhead box protein p1 is a transcriptional repressor of immune signaling in the CNS: implications for transcriptional dysregulation in Huntington disease.** *Hum Mol Genet* 2012, **21**:3097–3111.

53. Fujita E, Tanabe Y, Momoi MY, Momoi T: **Cntnap2 expression in the cerebellum of Foxp2(R552H) mice, with a mutation related to speech-language disorder.** *Neurosci Lett* 2012, **506**:277–280.

54. Morley M, Molony CM, Weber TM, Devlin JL, Ewens KG, Spielman RS, Cheung VG: **Genetic analysis of genome-wide variation in human gene expression.** *Nature* 2004, **430**:743–747.

55. Datta J, Kutay H, Nasser MW, Nuovo GJ, Wang B, Majumder S, Liu CG, Volinia S, Croce CM, Schmittgen TD: **Methylation mediated silencing of MicroRNA-1 gene and its role in hepatocellular carcinogenesis.** *Cancer Res* 2008, **68**:5049–5058.

56. Cicek MS, Koestler DC, Fridley BL, Kalli KR, Armasu SM, Larson MC, Wang C, Winham SJ, Vierkant RA, Rider DN, Block MS, Klotzle B, Konecny G, Winterhoff BJ, Hamidi H, Shridhar V, Fan JB, Visscher DW, Olson JE, Hartmann LC, Bibikova M, Chien J, Cunningham JM, Goode EL: **Epigenome-wide ovarian cancer analysis identifies a methylation profile differentiating clear cell histology with epigenetic silencing of HERG K+ channel.** *Hum Mol Genet* 2013 [Epub ahead of print].

57. Otaegi G, Pollock A, Hong J, Sun T: **MicroRNA miR-9 modifies motor neuron columns by a tuning regulation of FoxP1 levels in developing spinal cords.** *J Neurosci* 2011, **31**:809–818.

58. Brennand KJ, Simone A, Jou J, Gelboin-Burkhart C, Tran N, Sangar S, Li Y, Mu Y, Chen G, Yu D, McCarthy S, Sebat J, Gage FH: **Modelling schizophrenia using human induced pluripotent stem cells.** *Nature* 2011, **473**:221–225.

59. Pang ZP, Yang N, Vierbuchen T, Ostermeier A, Fuentes DR, Yang TQ, Citri A, Sebastiano V, Marro S, Sudhof TC, Wernig M: **Induction of human neuronal cells by defined transcription factors.** *Nature* 2011, **476**:220–223.

60. Voineagu I: **Gene expression studies in autism: moving from the genome to the transcriptome and beyond.** *Neurobiol Dis* 2012, **45**:69–75.

61. Kong SW, Collins CD, Shimizu-Motohashi Y, Holm IA, Campbell MG, Lee IH, Brewster SJ, Hanson E, Harris HK, Lowe KR, Saada A, Mora A, Madison K, Hundley R, Egan J, McCarthy J, Eran A, Galdzicki M, Rappaport L, Kunkel LM, Kohane IS: **Characteristics and predictive value of blood transcriptome signature in males with autism spectrum disorders.** *PLoS One* 2012, **7**:e49475.