

Study of Gene Expression Profiles and Biological Mechanism of Cerebral Palsy Using a Monozygotic Twin Pair

Tiane Zhang,¹ Miqu Wang,¹ Ling Pan,¹ Weijun Ding,¹ J.G. Wang,² Liping Yang,¹ Ming Liu,¹ Weihong Li,¹ and Zhiyong Yan³

¹ Genetics Laboratory, Chengdu University of Traditional Chinese Medicine, Chengdu, China

² Centre for Protective Technology, National University of Singapore, Singapore

³ School of Pharmaceutical Sciences, Southwest Jiaotong University, Emei, Sichuan, China

The gene expression profile of a normal-suffering monozygotic twin pair is investigated to explore biological mechanisms of spastic type cerebral palsy. Main works include following three aspects: First, a cDNA microarray test is carried out to get the differentially expressed genes of the patient with cerebral palsy compared to her monozygotic twin sister. Second, these differentially expressed genes are searched for their bioinformation within 4 biological databases: FatiGO, FatiGOPlus, KEGG, and SOURCE. Third, a set of special genes and gene families are screened out from the spastic type cerebral palsy patient. These biological analyses reveal that those genes for cell junction are mostly down-regulated, while those genes for metabolism are mostly up-regulated. The individual genes, gene family, and their associated biological functions can reflect the pathological and physiological characteristics of the cerebral palsy.

Cerebral palsy (CP) is a chronic disability of brain origin. It behaves as aberrant control of movement or posture and appears in the early stage of life. The CP is not the result of any progressive disease, but is caused by many etiological factors. For instance, at the prenatal stage, risk factors include intrauterine growth retardation, intrauterine infection, and exposure to teratogens during pregnancy. Among preterm infants, CP incidences generally increase with decreasing gestational age. The origin of most patient cases can be traced to post- and peripartum periventricular leukomalacia and intraventricular or periventricular hemorrhage (Folkert, 2005). However, the cause for a large portion of cases still remains obscure. The CP may be originally caused by genome and some environmental incursion during the development of the central neural system (CNS; Kuban & Leviton, 1994).

CP is the most common disability in children. At present, approximately 1% to 4% live-births in the world are suffering from CP. Based on clinical classification, the CP can be divided into a number of

phenotypic groups. In China, an investigation in 2001 revealed that the total prevalence of CP was 1.92%, of which 53.56% were the spastic type, 13.39% the hypotonic type, 10.00% the ataxia type, 4.63% the athetosis type, 0.85% the tremor type, 0.55% of the tonic type, 12.44% the mixture type including two or more of the above types and 4.58% of unknown classification (Lin et al., 2001). Therefore, the spastic type is the most common CP in China. Most researchers estimate that the portion of CP cases with a genetic etiology is approximately 1% to 2% in all CP cases (Baraitser, 1985), and half of the ataxia type of CP is due to the autosomal recessive inheritance related to 9p12-q12, and the spastic type of CP is related to 2q24-25 (McHale et al., 1999, 2000).

Biological mechanisms of a wide range of complex diseases have been successfully explored by genomic analysis. For example, diabetes mellitus, coronary artery disease and hypertension were successfully explored by cDNA microarray technology. By means of globally monitoring gene expression profiling at transcript level, the cDNA microarray technology is of high capacity, high sensitivity, high efficiency and high automaticity. It is effective and facile to study complex diseases as they are the results of interactions between multiple genes and multiple environmental factors (Goldsmith & Dhanasekaran, 2004; Spijker et al., 2004; Weeraratna et al., 2004).

Twins, particularly monozygotic (MZ) twin pairs, are a good way to explore the biological characteristics of CP. CP is more common in twins than in singletons, especially in preterm twins (Bonellie et al., 2005). Further, the characteristics of coherence and nonsynchronization are observed in the development of disease in twins (Medici et al., 1999). So far only

Received 18 July, 2006; accepted 4 December, 2006.

Address for correspondence: Dr Jian-Guo Wang, Centre for Protective Technology, National University of Singapore, 10 Kent Ridge Crescent, Singapore 119260, Singapore. E-mail: cvewjg@nus.edu.sg

Table 1
Repetitive Frequency and Heterozygosity of 15 Polymorphic Microsatellite Loci

Loci	Frequency and heterozygosity		Loci	Frequency and heterozygosity	
	Twin 1	Twin 2		Twin 1	Twin 2
D8S1179	10	10	D2S1338	25	25
D21S11	30–32.2	30–32.2	D19S433	13	13
D7S820	9–11	9–11	VWA	14–18	14–18
CSF1P0	12	12	TPOX	8–9	8–9
D3S1358	15–16	15–16	D18S51	14–16	14–16
TH01	9	9	D5S818	9–11	9–11
D13S317	8–9	8–9	FGA	14–16	14–16
D16S539	11–12	11–12			

Note: Twin 1 and Twin 2 are female.

several mechanisms on the characteristics have been identified, including point mutations, extension of triplet repeat, chromosomal aberrations, altered X-chromosomal inactivation and aberrant DNA methylation (Petronis, 2001). cDNA microarray technology has been employed to study the gene expression profile in twins, but no investigations have been carried out on the biological mechanisms of CP in a MZ twin pair.

This study investigates a MZ twin pair, a patient with CP and her normal sister, through their gene expression profiles. First, a cDNA microarray test is carried out to get the differentially expressed gene profile between the patient and her normal sister. These genes are then data-mined through four well-known databases (FatiGO, FatiGOPlus, KEGG and SOURCE) for bioinformation. The molecular and biological mechanisms of spastic CP are explored. Finally, specific genes which may be potential causes or results of the spastic CP are discussed. This is the first attempt to explore molecular and biological mechanisms through comparing the microarray-based gene expression profiles of blood cells between a patient and her normal sister in a MZ twin pair.

Material and Method

A MZ Twin Pair and Their Diagnosis of CP

A 10-year-old patient (Twin 1) was diagnosed as CP when she was 2 months old, while her sister (Twin 2) was normal. This clinical result was obtained by professional physicians after a CP evaluation based on the child CP diagnostic criteria in the Code of China Health Medicine Diagnosis and Treatment Standard. The Code adopted the Gross Motor Function Measure (GMFM) and the Activity of Daily Living Scale (ADL) as scaling tables. The GMFM evaluation was divided into five function sections, a total of 88 items including a series of body gestures such as lying, turning, sitting, climbing, kneeling, standing, walking, running, and jumping. The ADL evaluation included nine aspects of movement tests, including controlling urination, taking food, changing clothes, defecating, using

apparatus, recognizing and communication, movement on the bed, shifting and walking. These two scaling tables were well recognized in the international community (Collin et al., 1988; Russell et al., 1989).

Sample Preparation for cDNA Tests

Zygosity Diagnosis

In order to ascertain the zygosity of this twin pair by microsatellite polymorphism or STR (short string repetitive sequence), 5ml EDTA-2Na blood samples were collected from each of the twin sisters. Genomic DNA was then extracted by the Chelex-100 method (Walsh et al., 1991). Fifteen highly polymorphic microsatellite loci or human heredity markers were amplified by polymerase chain reaction (PCR) according to the specification of ABI Identifier Kit on an ABI9700 cycle. The PCR products were separated by capillary electrophoresis on ABI3100 heredity analyzer, followed by Genescan Analysis 317 and Genotyper 317 software. According to Mendel's Law, these 15 human heredity markers were used to identify kindred or zygosity in any two individuals.

Total RNA Sample Extraction

Ten ml of peripheral blood samples were collected from each of twin sisters at the same time for cDNA microarray tests. These peripheral blood samples were then harvested by means of the lysis of red blood cells with tris-buffered ammonium chloride to get mononuclear cells (PBMC). Total RNA of these PBMC was isolated using the Trizol (Invitrogen, Gaithersburg, MD, USA) one-step method, deposited and concentrated by isopropyl- alcohol, and purified with RNeasy mini spin column kit (Qiagen, Valencia, CA, USA). The quality and quantity of the total RNA was finally checked by formaldehyde denaturing agarose gel electrophoresis and spectrophotometer, respectively.

Microarray Experiment

Gene microarray chips with more than 20,000 human genes (CapitalBio Co. Ltd, Beijing, China) were used for microarray tests, including the following major five steps:

Table 2
Biological Processes for Down-/Up-Regulated Genes at Different Levels in FatIGO

GO No.	Biological process	Level	Down-regulated genes		Up-regulated genes		Unadjusted <i>p</i> value	Adjusted <i>p</i> value
			Number	(%)	Number	(%)		
GO:0009607	Response to biotic stimulus	3	48	7.55	116	14.81	1.843457e-05	6.386742e-03
GO:0008152	Metabolism	3	348	54.72	512	65.39	5.200361e-05	9.656872e-03
GO:0007154	Cell communication	3	195	30.66	171	21.84	1.934696e-04	2.048091e-02
GO:0009057	Macromolecule catabolism	4	186	30.19	336	43.64	2.787043e-07	3.540473e-04
GO:0050877	Neurophysiological process	4	42	6.82	15	1.95	7.689629e-06	2.930517e-03
GO:0006952	Defense response	4	48	7.79	110	14.29	1.682764e-04	1.886180e-02
GO:0044238	Primary metabolism	4	322	52.27	478	62.08	2.483520e-04	2.444026e-02
GO:0007267	Cell–cell signaling	4	39	6.33	18	2.34	3.191122e-04	2.966187e-02
GO:0044237	Cellular metabolism	4	337	54.71	493	64.03	5.080340e-04	4.302484e-02
GO:0007166	Cell surface receptor Linked signal transduction	5	84	14.09	52	6.98	2.427428e-05	7.116098e-03
GO:0006811	Ion transport	5	51	8.56	26	3.49	8.359075e-05	1.225247e-02
GO:0015031	Protein transport	5	20	3.36	60	8.05	2.821123e-04	2.687825e-02
GO:0051234	Establishment of protein localization	5	22	3.69	61	8.19	5.914207e-04	4.695634e-02
GO:0007186	G-protein coupled receptor protein signaling pathway	6	40	7.43	17	2.49	5.574683e-05	9.656872e-03
GO:0016070	RNA metabolism	6	13	2.42	49	7.16	1.214502e-04	1.596024e-02
GO:0007167	Enzyme linked receptor protein signaling pathway	6	17	3.16	4	0.58	6.407892e-04	4.884095e-02
GO:0030001	Metal ion transport	7	22	5	8	1.34	6.213554e-04	4.832623e-02

Note: Unadjusted *p* value from a Fisher's exact test and adjusted *p* value (after correction for multiple testing) based on False Discovery Rate (FDR) method are calculated. *p* value is automatically given by the FatIGO.

1. Label RNA sample by fluorescent dye. This labeling procedure includes four substeps:

- Synthesize cDNA: Total of 5 µg RNA was used as the templet and T7-Oligo (dT) 15 (5'-AAACGACGGCCAGTGAATTGTAATACGACTCACTATAGGCGCTTTTTTTTTTTTTTTTTT-3, where V may be either G, C or A from Bioasia Bio-Technology Ltd., Shanghai, China) was used as the Primer. The synthesized cDNA was purified by using QIAquick PCR Purification Kit (Qiagen).
- Transcribe and synthesize cRNA: In vitro, the cDNA was transcribed to synthesize cRNA with T7 RiboMAX Express Large Scale RNA Production System (Promega). It was then purified with RNeasy Mini Kit (Qiagen).
- Reverse transcription under Random Primer: Two µg of cRNA was used as the templet and 200 u/µl superscript II was used as reverse transcriptase. They were used along with 9 Random Primers (Bioasia Bio-Technology Ltd.) for reverse transcription. The resultant products were then purified by QIAquick PCR Purification Kit (Qiagen).
- Label cDNA: KLENOW catalyses were used for labeling. The concentration of reaction solutions was 120 µM for dATP, dGTP, dTTP, 60 µM for

dCTP (dNTP, Bioasia Bio-Technology Ltd.) and 40 µM for Cy5-dCTP, Cy3-dCTP (Amersham Pharmacia Biotech, Inc., Piscataway, NJ, USA; Random Primer DNA Labeling Kit, TaKaRa, Dalian, China). The labeled cDNA was purified by QIAquick PCR Purification Kit (Qiagen). The cDNA was labeled Cy3 for patient and Cy5 for normal sister.

- Hybridize and wash chips: The labeled cDNA was dissolved in 30 µl hybridization solution (3 × SSC, 0.2% SDS, 5 × Denhart's, 25% Formamide). Hybridization was going on overnight at 42°C. The hybridized chips were then washed in solutions of 2 × SSC and 0.2% SDS for 5 minutes each at 42°C. They were further washed in solution of 0.2 × SSC for 5 minutes each at room temperature and finally dried.
- Scan microarray chips: The hybridized microarray chips were scanned by a LuxScan 10K/A dual laser scanner (CapitalBio Ltd., Beijing, China) to detect the intensity of fluorescence from both Cy3 and Cy5 channels.
- Normalize image intensity: The signal intensities obtained from both Cy3 and Cy5 channels were normalized and analyzed by GenePix Pro 4.0 software (Axon Instruments Co. Ltd). The normalization was

so done that the mean Cy3/Cy5 ratio of all spots across the entire array was 1.0.

- Identify differentially expressed genes: A differentially expressed gene was identified if its ratio of Cy3 to Cy5 was out of range of (0.5, 2.0; Pan et al., 2006). The underlying information in differentially expressed genes was then explored through four public biological databases: FatiGO, FatiGOPlus, KEGG, and SOURCE.

Results

Clinical Diagnosis and Medical Examination

For the clinical diagnosis, the procedure for a CP medical examination was strictly observed. Figure 1 shows the twins' standing gestures. Medical records showed that the twin sisters were preterm and low birthweight. After birth, Twin 1 had a series of unusual symptoms such as fidgeting, loss of appetite, a sucking disability, and so on. These were the early symptoms of brain injury. The CT image observed a right cerebral cyst, augmented side ventricle and hypogenetic periventricular white matter. Medical examination indicated that Twin 1 had mental retardation, dysphasia, and apperceive hindrance, while Twin 2 did not. Twin 1 had muscle spasms and stiffness of all four limbs. Further, Twin 1 graded 51.5 for GMFM (full marks are 264 according to the Rusell grading) and 26 for ADL (full marks are 100 according to the Barthel index number grading), while Twin 2 graded 252 for GMFM and 95 for ADL. Both GMFM and ADL grading systems imply that higher marks mean function that is closer to normal. Therefore, Twin 1 was a typical spastic CP while Twin 2 was a completely normal child.

Zygosity Examination

Table 1 shows that the 15 loci were completely identical in both frequency and heterozygosity which has



Figure 1

Comparison of standing position

Note: On the left is the normal child and on the right, the CP child.

right comparative probability of 99.999954% and total coincidence probability of 1.7×10^{-13} . Therefore, the twin sisters were MZ.

Online Analysis for Differentially Expressed Genes

From these twin pairs, 2703 differentially expressed genes were detected, in which 1272 genes were up-regulated and 1431 genes were down-regulated. These up- and down-regulated genes are searched for underlying bioinformation within the above-mentioned four biological databases through either their GenBank Accession Numbers or Gene Name. Sorting procedures and results are separately stated as follows.

First, the GenBank Accession Numbers of 1431 down-regulated genes and 1272 up-regulated genes are separately input into the FatiGO database, a web tool to find differential distributions of GO terms between two groups of genes. The search report showed that there were 505 unknown genes as well as 17 genes with more than one identifier in the down-regulated gene list (List 1) and 212 unknown genes as well as 10 genes with more than one identifier in up-regulated gene list (List 2). After removing these genes, the remaining genes for analysis had 909 in list 1 and 1050 in list 2. The search report indicated that these genes participated in 1750 biological processes, and annotated genes were 636 in List 1 and 783 in List 2. At present, FatiGO has nine levels for use. A higher level means more details available. However, detailed expressions may not master the global tendency of gene distribution. Table 2 lists the significant biological processes and their percentage for both down- and up-regulated genes in statistics at different levels. Table 3 lists the genes involved in 18 special biological processes which exist only in either down-regulated genes or up-regulated genes, among which 14 biological processes are relating to the down-regulated genes and only four biological processes are relating to the up-regulated genes.

Second, biochemical pathways of both up- and down-regulated genes are acquired in FatiGOPlus through genBank Accession Numbers. FatiGOPlus is an advanced version of FatiGO which can not only include the functions of FatiGO, but also provide a differential distribution of biological terms in a common pathway in two groups of genes (Al-Shahrour et al., 2006). It can compare the percentages of down-regulated genes (List 1) and up-regulated genes (List 2) which participate in the same pathway. Search results reported that KEGG found 134 pathways which involved 189 annotated genes in List 1 and 252 annotated genes in List 2. It was found that only the pathway, neuroactive ligand-receptor interaction attaching to cell junction, was significant in statistics. Figure 2 is the graphic presentation of cell junction obtained from KEGG, which includes 15 pathways relating to our genes. Table 4 compares the percentages and significant levels (*p* values) for up- and down-regulated genes in the 15 cell junction pathways. The independent sample *t* test was used to compare the percentages of down- and up-regulated

Table 3

Special Biological Processes Only Involved in Either Down- or Up-Regulated Genes

GB.Accession Num	Name	Locus	Cy3/Cy5 Ratio	Description
1 GO:0007530 sex determination				
NM_006557	DMRT2	9p24.3	0.3171	Doublesex and mab-3 related transcription factor 2
2 GO:0007548 sex differentiation				
NM_006942	SOX15	17p13	0.2671	SRY (sex determining region Y)-box 20
NM_006557	see 1			
NM_002191	INHHA	2q33-q36	0.4622	Inhibin, alpha
3 GO:0007611 learning and/or memory				
NM_001115	ADCY8	8q24	0.3362	Adenylate cyclase 8 (brain)
4 GO:0007631 feeding behavior				
NM_001524	HCRT	17q21	0.4220	Hypocretin (orexin) neuropeptide precursor
NM_001138	AGRP	16q22	0.3010	Agouti related protein homolog (mouse)
GO:0007638 mechanosensory behavior				
D10537	MPZ	1q22	0.2962	Myelin protein zero (Charcot-Marie-Tooth neuropathy 1B)
5 GO:0009791 post-embryonic development				
NM_002191	see 2			
6 GO:0019048 virus-host interaction				
NM_022740	HIPK2	7q32-q34	0.4381	Homeodomain interacting protein kinase 2
AK023750	No Gene Information was found matching Accession number query AK023750 in SOURCE			
7 GO:0019089 viral transmission				
AK023750	No Gene Information was found matching Accession number query AK023750 in SOURCE			
8 GO:0019098 tube development				
NM_014693	ECE2	3q27.1/1p36.1	0.3339	Endothelin converting enzyme 2
NM_003019	SFTPD	10q22.2-q23.1	0.4950	Surfactant, pulmonary-associated protein D
NM_004557	NOTCH4	6p21.3	0.3667	Notch homolog 4 (Drosophila)
9 GO:0035295 reproductive behavior				
NM_005958	MTNR1A	4q35.1	0.4604	Melatonin receptor 1A
10 GO:0043062 extracellular structure organization and biogenesis				
NM_001849	COL6A2	21q22.3	0.4442	Collagen, type VI, alpha 2
NM_004395	DBN1	5q35.3	0.4576	Drebrin 1
11 GO:0044419 interaction between organisms				
NM_022740	see6			
AK023750	No Gene Information was found matching Accession number query AK023750 in SOURCE			
12 GO:0048468 cell development				
NM_002385	MBP	18q23	0.4933	Myelin basic protein
NM_003551	NME5	5q31	0.2981	Non-metastatic cells 5, (nucleoside-diphosphate kinase)
D10537	repeat			
13 GO:0048511 rhythmic process				
NM_033282	OPN4	10q22	0.2135	Opsin 4 (melanopsin)
NM_012118	CCR4L	4q31.1	0.4083	CCR4 carbon catabolite repression 4-like (S. cerevisiae)
NM_005958	see 9			
NM_002191	see 2			
NM_003216	TEF	22q13	0.3594	Thyrotrophic embryonic factor
14 GO:0050795 regulation of behavior				
NM_016215	EGFL7	9q34.3	0.4029	NEU1 protein
15 GO:0050792 regulation of viral life cycle				
NM_021822	APOBEC3G	22q13.1-q13.2	3.5499	Phorbolin-like protein MDS019
NM_006058	TNIP1	5q32-q33.1	4.7260	Nef-associated factor 1

Table 3 (CONTINUED)

Special Biological Processes Only Involved in Either Down- or Up-Regulated Genes

GB.Accession Num	Name	Locus	Cy3/Cy5 Ratio	Description
16 GO:0051098 regulation of binding				
NM_001641	APEX	14q11.2-q12	2.8479	APEX nuclease (multifunctional DNA repair enzyme)
17 GO:0006944 membrane fusion				
AL122075	GSA7	3p25.3	2.5298	Ubiquitin activating enzyme E1-like protein
NM_003764	STX11	6q24.2	2.7529	Syntaxin 11
NM_006370	VT11B	14q24.1	8.2650	Vesicle-associated soluble NSF attachment protein receptor (v-SNARE; homolog of <i>S. cerevisiae</i> VT11)
AF038202	STX6	1q25.3	2.8475	Homo sapiens clone 23570 mRNA sequence
18 GO:0007498 mesoderm development				
NM_006060	ZNFN1A1	7p13-p11.1	3.9652	Zinc finger protein, subfamily 1A, 1 (Ikaros)
NM_002110	HCK	20q11-q12	3.8957	Hemopoietic cell kinase
NM_000061	BTK	Xq21.33-q22	6.2190	Bruton agammaglobulinemia tyrosine kinase

genes involved in 15-cell junction pathways. Finally, the *p* value of .021 was obtained. Therefore, the down-regulated genes exceeded significantly the up-regulated genes as a whole.

Third, the number of genes distributing subfamily pathways was enumerated in KEGG. We obtained the *p* value through the statistical disposition with two independent sample tests of nonparametric tests. Table 5 summarizes the number of differentially expressed genes in each biological pathway and its significant level in KEGG database. During the analysis, we picked out a series of gene families with diversity in biological function (not listed here). We also looked

into the distribution of up- and down-regulated genes which participated in the same pathway. Only two typical pathways were given — the neuroactive ligand-receptor interaction in Figure 3 and the regulation of actin cytoskeleton in Figure 4 where down-regulated genes concentrate in the upstream and up-regulated genes distribute in the downstream. The genes relating to actin polymerization and depolymerization are almost up-regulated.

Although the four above databases possess different bioinformation in quantity and searching mode, their search outcomes had high consistency in tendency.

Table 4Percentage and *p* Value for Down- and Up-Regulated Genes in Cell Junction

Entry	Pathway name	Down-regulated genes (%) (List 1)	Up-regulated genes (%) (List 2)	Unadjusted <i>p</i> value	Adjusted <i>p</i> value
1. hsa04010	MAPK signaling pathway	7.94	5.95	.44818	1
2. hsa04020	Calcium signaling pathway	6.88	3.97	.19757	1
3. hsa04060	Cytokine-cytokine receptor interaction	8.47	5.95	.34876	1
4. hsa04070	Phosphatidylinositol signaling system	2.12	4.37	.28895	1
5. hsa04080	Neuroactive ligand-receptor interaction	9.52	1.19	.00005	.03311*
6. hsa04110	Cell cycle	4.23	2.38	.28665	1
7. hsa04210	Apoptosis	2.12	3.97	.41146	1
8. hsa04310	Wnt signaling pathway	4.76	4.37	1	1
9. hsa04350	TGF-beta signaling pathway	2.12	1.19	.46827	1
10. hsa04510	Focal adhesion	10.58	7.14	.23140	1
11. hsa04512	ECM-receptor interaction	5.82	0.79	.00288	.38824
12. hsa04520	Adherens junction	4.76	3.17	.45722	1
13. hsa04530	Tight junction	4.76	3.57	.62878	1
14. hsa04540	Gap junction	4.76	1.19	.03487	1
15. hsa04810	Regulation of actin cytoskeleton	10.05	6.35	.15956	1
Total		88.89	55.55		

Note: Independent Sample Test for all genes: *p* = .021

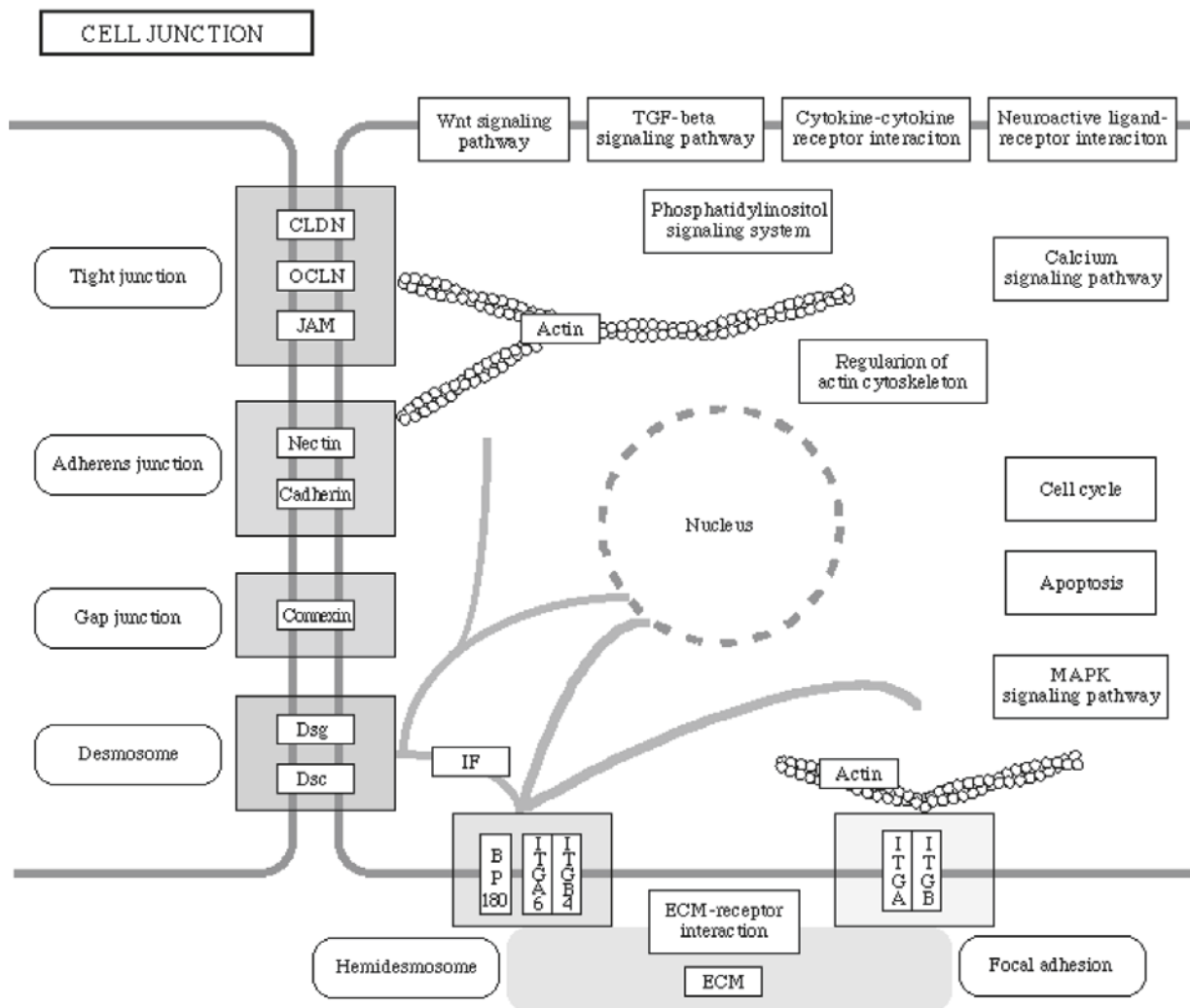


Figure 2
Cell junction pathways in KEGG database.

Therefore, the above search results are cross-validated and reliable.

Discussion

Advantage of MZ Twin Models

MZ twin models have sole advantage in exploring the biological mechanism of CP. A MZ twin pair has the same DNA structures after birth but different gene expressions in each individual due to their cellular environment, the smallest unit to express the interaction between individual and environment. A MZ twin model can rule out genetic heterogeneity to its maximum and focus only on cellular environment, thus being superior to those models without identically genetic backgrounds (Kohsaka et al., 1996; Poulsen et al., 2003; Reuss et al., 2002). The current study is thus a good example of the identification of the differentially expressed gene profile between CP PBMC and normal PBMC.

The analysis for differentially expressed genes is to understand the global distribution and tendency of a large quantity of gene expressions. The differentially expressed genes of this twin pair mainly concentrate on metabolism, cell communication or junction, and response to biotic stimulus (see Table 2). Biological pathway and process are mainly related to cell communication for cellular processes, signal transduction for environmental information processing, metabolism, and human diseases (see Table 5). Further analysis revealed that the down-regulated genes, which participate in 15 cell junction pathways, were the absolute majority in all genes and they mainly concentrate on the upstream, while the up-regulated genes mainly concentrate on the downstream, or down-/up-regulated genes alternatively distribute. Such a distribution may imply a compensation mechanism in the CP patient. That is, upstream down-regulated genes may activate downstream genes to generate an

Table 5

Distribution of Differentially Expressed Genes in KEGG Database

The property of pathway	Down-regulated gene number	Up-regulated gene number	NPar Tests <i>p</i> value
1. Metabolism	104/208 (50%)	191/298 (64%)	.000***
1.1 Carbohydrate metabolism	13	38	.006***
1.2 Energy metabolism	7	18	.280
1.3 Lipid metabolism	11	17	.387
1.4 Nucleotide metabolism	17	21	.221
1.5 Amino acid metabolism	21	32	.284
1.6 Metabolism of other amino acids	3	8	.057
1.7 Glycan biosynthesis and metabolism	9	20	.154
1.8 Metabolism of cofactors and vitamins	7	16	.042*
1.9 Biosynthesis of secondary metabolites	1	4	.155
1.10 Xenobiotics biodegradation and metabolism	15	17	.160
2. Genetic information processing	11/208 (5%)	57/298 (19%)	.001***
2.1 Transcription	2	9	.121
2.2 Translation	2	19	.102
2.3 Replication and repair	1	3	.317
2.4 Folding, sorting and degradation	6	26	.026**
3. Environmental information processing	125/208 (60%)	112/298 (38%)	.713
3.1 Membrane transport	2	0	.317
3.2 Signal transduction	71	74	.704
3.3 Signaling molecules and interaction	52	38	.386
4. Cellular Processes	165/208 (79%)	195/298 (65%)	.310
4.1 Cell motility	20	17	.317
4.2 Cell growth and death	14	19	.221
4.3 Cell communication	39	34	.381
4.4 Endocrine system	25	18	.663
4.5 Immune system	36	91	.002**
4.6 Nervous system	13	8	.121
4.7 Sensory system	7	2	.317
4.8 Development	10	5	1
4.9 Behavior	1	1	1
5. Human diseases	47/208 (23%)	48/298 (16%)	.596
5.1 Neurodegenerative disorders	13	11	.432
5.2 Infectious diseases	6	12	.034*
5.3 Metabolic disorders	3	4	1
5.4 Cancers	25	21	.294

Note: * denotes < .05; ** denotes < .01, and *** denotes < .001.

overexpressed profile, making the CP individual obtain homeostasis. At present, it is known that 294 genes participate in neuroactive ligand-receptor interaction. The receptors belong to G-protein-coupled receptors (GPCRs) while the ligands include protein and peptide hormones, serotonin and gamma-aminobutyric acid. Our study found 16 down- and four up-regulated genes (see Figure 3). Whether those exceptional genes in transcription are the results or the causes of CP is worthy of further investigation.

Gene Family Relating to CP

Some gene families which repeatedly appeared were found in the 15 cell junction pathways, including

integrin- α 3,10, β 4; laminin- α 5, β 4, γ 3; CLDN-3,5,7,10; WASL and WASF-2,3; ADCY-4,6,8; FGF-7,8,14, R1, R4, 22; and DFNA-1,20,26. Integrins are heterodimers comprised of alpha and beta subunits and participate in focal adhesion, ECM-receptor interaction, and the regulation of actin cytoskeleton. Particularly, integrin α 3 is a receptor for fibronectin, laminin, collagen, epiligrin, thrombospondin and CSPG4, while integrin α 10 and integrin β 4 are the receptors for collagen and laminins. Laminins, a family of extracellular matrix glycoproteins, are the major noncollagenous constituents of basement membranes. They regulate a wide variety of biological processes including cell

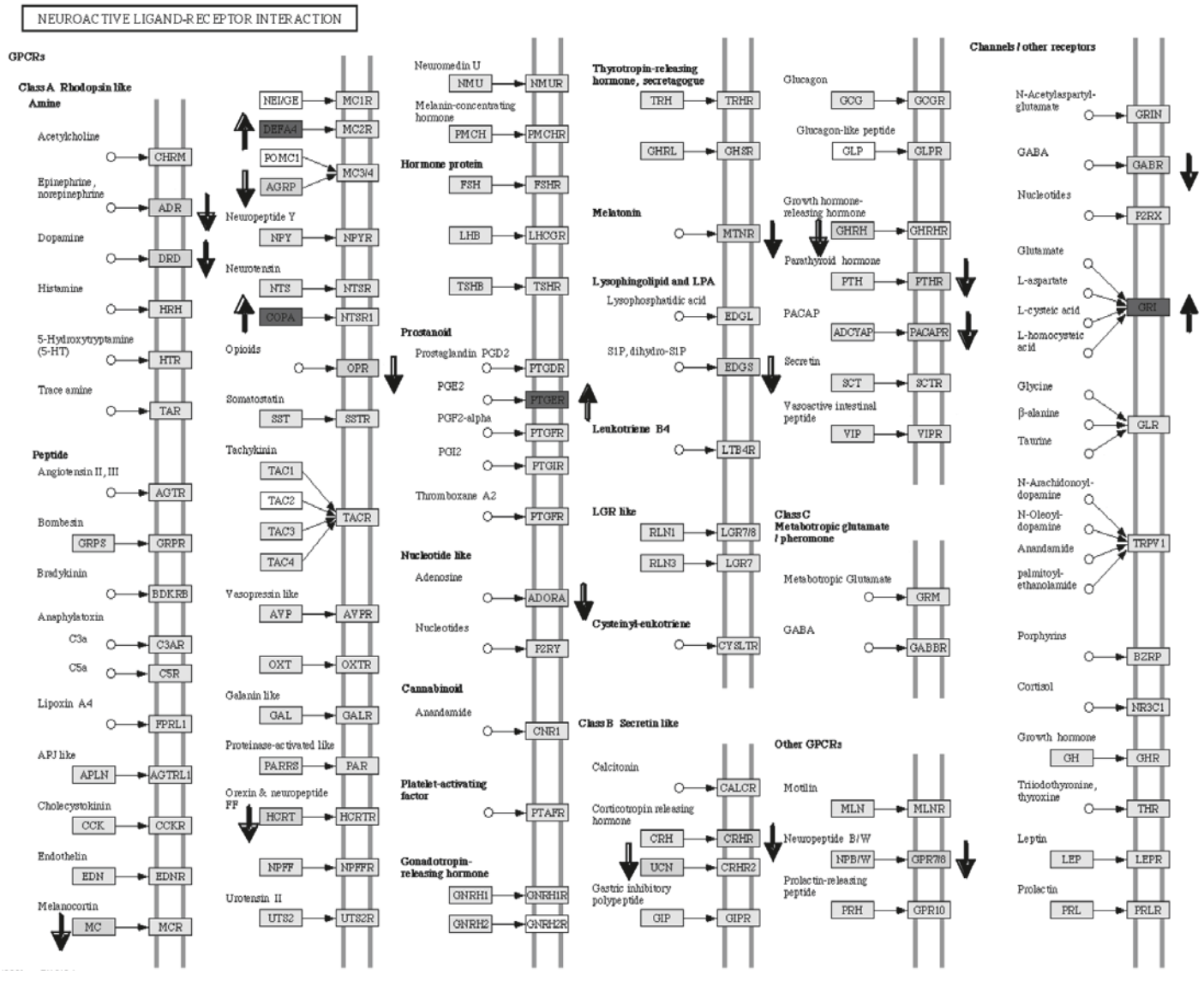


Figure 3 Distribution of down- and up-regulated genes in neuroactive ligand-receptor interaction obtained from KEGG (where a down-arrow denotes a down-regulated gene, an up-arrow denotes an up-regulated gene, and the rest denote inherent genes).

adhesion, differentiation, migration, signaling, neurite outgrowth, and metastasis. LAMC3 was reported to have neural and muscle-associated expression and function (Manuel et al., 1999). CLDN-3,5,7,10 genes which encode proteins, are the members of the claudin family. Claudins are integral membrane proteins and the components of tight junction strands which serve as a physical barrier to prevent solutes and water from passing freely through the paracellular space between epithelial or endothelial cell sheets. The CLDN3 haploinsufficiency may be the cause of certain cardiovascular and musculo-skeletal abnormalities observed in Williams-Beuren syndrome, a rare developmental disorder (Paperna et al., 1998). WASL and WASF-2,3 genes encode the member of the Wiskott-

Aldrich syndrome protein (WASP) family. The WASP, WASP-like (WASL), and WASF2 are among the downstream effector molecules involved in the transmission of signals from tyrosine kinases and small GTPases to actin cytoskeleton. The actin cytoskeleton plays critical roles in cell morphologic changes and motility (Dahl et al., 2003). These gene families, which have low expression but play important roles in cell communication, may express to some extent the biological traits of CP.

Special Individual Genes Found in CP patient

A set of genes was screened out to reflect pathological and physiological characteristics of the CP in this study. For example, SOX15 and INHA were

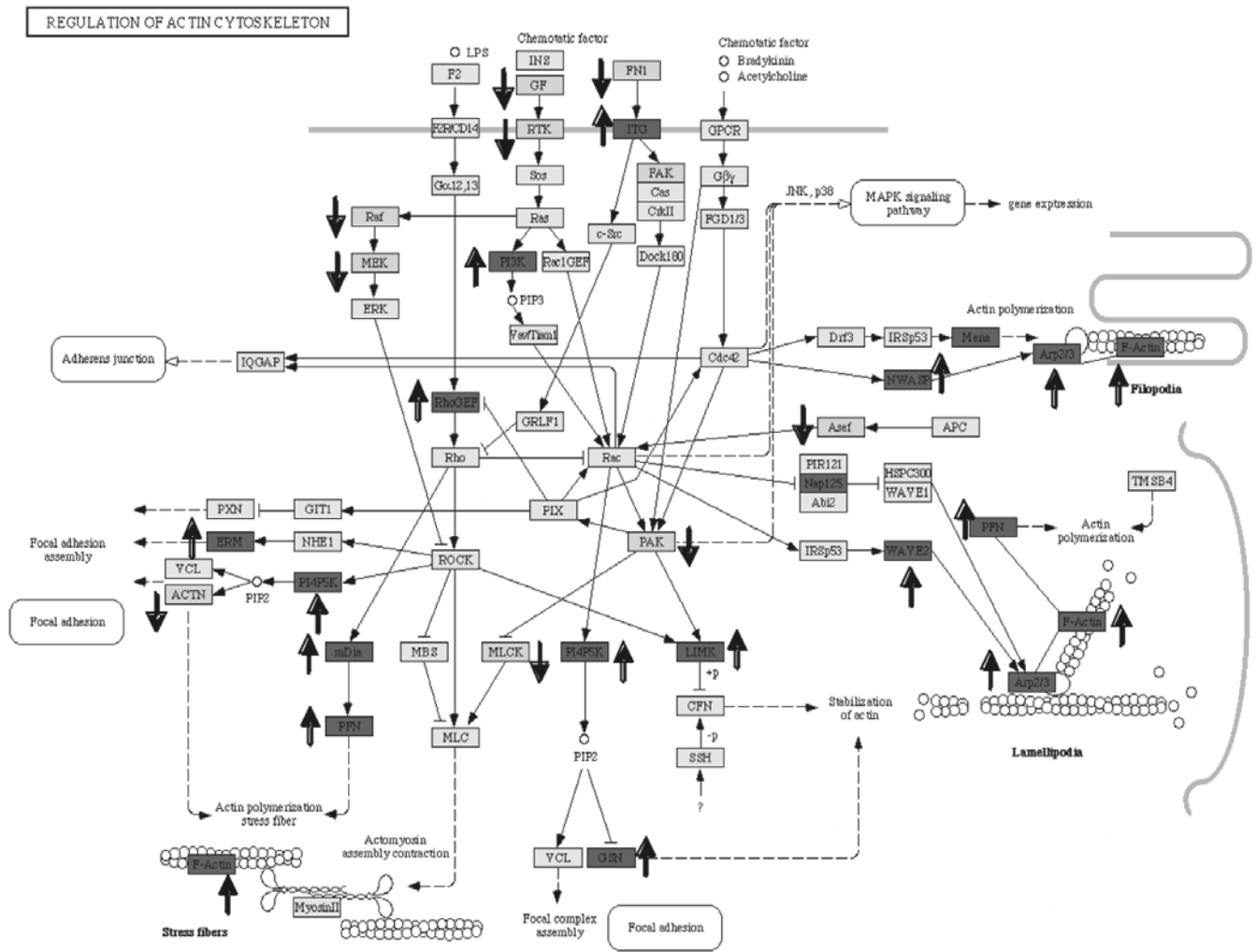


Figure 4

Pathway in regulation of actin cytoskeleton obtained from KEGG (where a down-arrow denotes a down-regulated gene, an up-arrow denotes an up-regulated gene, and the rest denote inherent genes).

down-regulated by approximately 5-fold and 2.5-fold, respectively. These two genes participate in the biological process of sex differentiation. SOX 15 is a member of the SOX (SRY-related HMG-box) family of transcription factors involved in the regulation of embryonic development and in the determination of cell fate. It plays a role in the regulation of skeletal muscle myogenesis (Lee et al., 2004). INHA (inhibin) is normally produced by ovarian granulosa cells and inhibits the secretion of follitropin by the pituitary gland. The inhibin is involved in the regulation of a number of diverse functions such as hypothalamic and pituitary hormone secretion, gonadal hormone secretion, germ cell development and maturation, erythroid differentiation, insulin secretion, nerve cell survival, embryonic axial development or bone growth. At present, inhibin is supposed to be the growth/differentiation factor and hormone (Farnworth et al., 2006; Kobayashi et al., 2006). Medical examination indi-

cates that the 10-year-old girl with CP is obviously stunted and has amyotrophy compared with her twin sister. The down-regulation of SOX15, INHA genes and the other relational genes listed in Table 3 may be partial reasons for this. This should be further validated by further tests such as assay of the level of blood serum hormone and western blot for the relevant protein.

Special Genes for Memory and Behavior

The differentially expressed genes for memory and behaviors were investigated. These genes, being all down-regulated, include ADCY8 relating to memory, HCRT and AGRP relating to feeding behaviors, and MPZ relating to mechanosensory behaviors. The ADCY8 (adenylate cyclase 8, brain) is a membrane-bound enzyme that catalyzes the formation of cyclic AMP from ATP and participates in purine metabolism, the calcium signaling pathway, and gap junction. Recent researches have revealed that

ADCY8 is one of seven genes that are important to human memory (de Quervain et al., 2006). Its down-regulation may imply the basic embodiment of central nerve injury. MPZ, which belongs to the myelin P0 protein family, is the major structural protein of peripheral myelin, accounting for more than 50% of the protein present in the sheath of peripheral nerves. Fifty-six distinct mutations in the MPZ gene had been identified in association with hereditary peripheral neuropathy. For example, charcot-marie-tooth diseases are a series of neuropathy that result from the MPZ mutation [MIM: 118200, 607677, 607736, 607791]. These diseases take on the following clinical symptoms: slow progressive distal muscle atrophy and weakness, absent deep tendon reflexes, reduced nerve conduction velocities and gait ataxia (Nelis et al., 1999). Whether MPZ occurred to the mutation or not has not been confirmed in our patient, but the outcome from the cDNA microarray assay shows that the down-regulation of MPZ above 3-fold may be another biological basis of the CP. HCRT and AGRP share similar functions, neuropeptides regulating food intake, participant of neuroactive ligand-receptor interaction, and locating to hypothalamus. The HCRT plays an important role in orchestrating the sleep-wake cycle (Hagan et al., 1999), but the AGRP plays a role in weight homeostasis, regulating body weight via central melanocortin receptors. Being associated with the history of the patient such as low birthweight, eating little and sucking disability in infancy, and the body weight of the patient being lower than her twin sister, the down-regulations of HCRT and AGRP were inferred to be one of the biological mechanisms for the clinic symptom of CP.

In this study, some genes and gene families were found which reflect the pathological and physiological characteristics of the CP. That is, the genes relating to cell communication and response to biotic stimulus take on low expression levels while those relating to metabolism take on high expression levels. Of course, these results and conclusions are preliminary. More experimental data are necessary for their further verifications.

Acknowledgments

The study was supported by Natural Science Foundation of China (granted No: 90209013, 30672579) and '973' Project of the Ministry of Science and Technology in China.

References

- Al-Shahrour, F., Minguéz, P., Tárraga, J., Montaner, D., Alloza, E., Vaquerizas, J.M.M., Conde, L., Blaschke, C., Vera, J., & Dopazo, J. (2006). BABELOMICS: A systems biology perspective in the functional annotation of genome-scale experiments. *Nucleic Acids Research*, 34, 472–476.
- Baraitser, M. (1985). *The genetics of neurological disorders*. Oxford: Oxford University Press.
- Bonellie, S. R., Currie, D., & Chalmers, J. (2005). Comparison of risk factors for cerebral palsy in twins and singletons. *Developmental Medicine and Child Neurology*, 47, 587–591.
- Collin, C., Wade, D. T., Davies, S., & Horne, V. (1988). The Barthel ADL Index: A reliability study. *International Disability Studies*, 10, 61–63.
- Dahl, J. P., Wang-Dunlop, J., Gonzales, C., Goad, M. E., Mark, R. J., & Kwak, S. P. (2003). Characterization of the WAVE1 knock-out mouse: implications for CNS development. *Journal Neuroscience*, 23, 3343–3352.
- de Quervain, D. J.-F., & Papassotiropoulos, A. (2006). Identification of a genetic cluster influencing memory performance and hippocampal activity in humans. *Proceedings of the National Academy Science USA*, 103, 4270–4274.
- Farnworth, P. G., Wang, Y., Leembruggen, P., Ooi, G. T., Harrison, C., Robertson, D. M., & Findlay, J. K. (2006). Rodent adrenocortical cells display high affinity binding sites and proteins for inhibin A, and express components required for autocrine signalling by activins and bone morphogenetic proteins. *Journal of Endocrinology*, 188, 451–465.
- Goldsmith, Z. G., & Dhanasekaran, N. (2004). The microevolution: Applications and impacts of microarray technology on molecular biology and medicine. *International Journal of Molecular Medicine*, 13, 483–495.
- Folkerth, R. D. (2005). Neuropathologic substrate of cerebral palsy. *Journal of Child Neurology*, 20, 940–949.
- Hagan, J. J., Leslie, R. A., Patel, S., Evans, M. L., Wattam, T. A., Holmes, S., Benham, C. D., Taylor, S. G., Routledge, C., Hemmati, P., Munton, R. P., Ashmeade, T. E., Shah, A. S., Hatcher, J. P., Hatcher, P. D., Jones, D. N. C., Smith, M. I., Piper, D. C., Hunter, A. J., Porter, R. A., & Upton, N. (1999). Orexin A activates locus coeruleus cell firing and increases arousal in the rat. *Proceedings of the National Academy Science USA*, 96, 10911–10916.
- Kobayashi, Y., Jimenez-Crasser, F., Ireland, J. J., & Smith, G. W. (2006). Evidence of a local negative role for cocaine and amphetamine regulated transcript (CART), inhibits and low molecular weight insulin like growth factor binding proteins in regulation of granulosa cell estradiol production during follicular waves in cattle. *Reproductive Biology and Endocrinology*, 4, 22–34.
- Kohsaka, H., Carson, D. A., Rassenti, L. Z., Ollier, W. E., Chen, P. P., Kipps, T. J., & Miyasaka, N. (1996). The human immunoglobulin V(H) gene repertoire is genetically controlled and unaltered by chronic autoimmune stimulation. *Journal of Clinical Investigation*, 98, 2794–2800.
- Kuban, K. C. K., & Leviton, A. (1994). Cerebral palsy. *New England Journal of Medicine*, 330, 188–195.

- Lee, H. -J., Goring, W., Ochs, M., Muhlfeld, C., Steding, G., Paprotta, I., Engel, W., & Adham, I. M. (2004). Sox15 is required for skeletal muscle regeneration. *Molecular and Cellular Biology*, *24*, 8428–8436.
- Lin, Q., Li, S., Liu, J. M., Zhang, Sh. X., Hong, Sh. X., Jiang, M. F., Wang, T. M., Zhu, Y., Zhao, F. L., Zhong, Y. Q., Liang, Y. L., Shi, R. F., Zhao, P., & Li. (2001). Prevalence and clinical classification of cerebral palsy of children in 6 provinces or autonomous region in China. *China Journal Pediatric*, *39*, 613–615 (in Chinese).
- Koch, M., Olson, P. F., Albus, A., Jin, W., Hunter, D. D., Brunken, W. J., Burgeson, R. E., & Champlaud, M. F. (1999). Characterization and expression of the laminin gamma3 chain: A novel, non-basement membrane-associated, laminin chain. *Journal of Cell Biology*, *145*, 605–617.
- McHale, D. P., Jackson, A. P., Campbell, Levene, M. I., Corry, P., Woods, C. G., Lench, N. J., Mueller, R. F., & Markham, A. F. (2000). A gene for ataxic cerebral palsy maps to chromosome 9p12-q12. *European Journal of Human Genetics*, *8*, 267–272.
- McHale, D. P., Mitchell, S., Bunday, S., Moynihan, L., Campbell, D. A., Woods, C. G., Lench, N. J., Mueller, R. F., & Markham, A. F. (1999). A gene for autosomal recessive symmetrical spastic cerebral palsy maps to chromosome 2q24–25. *American Journal of Human Genetics*, *64*, 526–532.
- Medici, F., Hawa, M., Ianari, A., Pyke, D. A., & Leslie, R. D. (1999). Concordance rate for type II diabetes mellitus in monozygotic twins: Actuarial analysis. *Diabetologia*, *42*, 125–127.
- Nelis, E., Haites, N., & Van Broeckhoven, C. (1999). Mutations in the peripheral myelin genes and associated genes in inherited peripheral neuropathies. *Human Mutation*, *13*, 11–28.
- Pan, L., Wang, M. Q., Wang, J. G., Wu, B., & Hui, K. M. (2006). Clinical and molecular evaluation of warming and tonic herb treatment for sibling patients of a typical kidney-yang deficiency family. *American Journal of Chinese Medicine*, *34*, 387–400.
- Paperna, T., Peoples, R., Wang, Y. K., Kaplan, P., & Francke, U. (1998). Genes for the CPE receptor (CPETR1) and the human homolog of RVP1 (CPETR2) are localized within the Williams-Beuren syndrome deletion. *Genomics*, *54*, 453–459.
- Petronis, A. (2001). Human morbid genetics revisited: Relevance of epigenetics. *Trends in Genetics*, *17*, 142–146.
- Poulsen, P., Andersen, G., Fenger, M., Hansen, T., Echwald, S. M., Volund, A., Beck-Nielsen, H., Pedersen, O., & Vaag, A. (2003). Impact of two common polymorphisms in the PPARgamma gene on glucose tolerance and plasma insulin profiles in monozygotic and dizygotic twins: Thrifty genotype, thrifty phenotype, or both? *Diabetes*, *52*, 194–198.
- Reuss, E., Fimmers, R., Kruger, A., Becker, C., Rittner, C., & Hohler, T. (2002). Differential regulation of interleukin-10 production by genetic and environmental factors: A twin study. *Genes and Immunity*, *3*, 407–413.
- Russell, D., Rosenbaum, P., Cadman, D., Gowland, C., Hardy, S., & Jarvis, S. (1989). The gross motor function measure: A means to evaluate the effects of physical therapy. *Developmental Medicine and Child Neurology*, *31*, 341–352.
- Spijker, S., van de Leemput, J. C. H., Hoekstra, C., Boomsma, D. I., & Smit, A. B. (2004). Profiling gene expression in whole blood samples following an in vitro challenge. *Twin Research*, *7*, 564–570.
- Walsh, B. S., Petzger, D. A., & Higuchi, R. (1991). Chelex-100 as medium for simple extraction of DNA for PCR-based typing from forensic material. *Biotechniques*, *10*, 506–513.
- Weeraratna, A. T., Nagel, J. E., Mello, C. V., & Taub, D. D. (2004). Gene expression profiling: From microarrays to medicine. *Journal of Clinical Immunology*, *24*, 213–224.