

ORIGINAL ARTICLE

# Mutation spectrum of and founder effects affecting the *PTS* gene in East Asian populations

Yen-Hui Chiu<sup>1,2</sup>, Ying-Chen Chang<sup>3</sup>, Yu-Hsin Chang<sup>1,16</sup>, Dau-Ming Niu<sup>4</sup>, Yan-Ling Yang<sup>5</sup>, Jun Ye<sup>6</sup>, Jianhui Jiang<sup>7</sup>, Yoshiyuki Okano<sup>8</sup>, Dong Hwan Lee<sup>9</sup>, Suthipong Pangkanon<sup>10</sup>, Chulaluck Kuptanon<sup>11</sup>, Ngu Lock Hock<sup>12</sup>, Mary Anne Chiong<sup>13</sup>, Barbra V Cavan<sup>14</sup>, Kwang-Jen Hsiao<sup>1,2,15</sup> and Tze-Tze Liu<sup>2,3</sup>

The enzyme 6-pyruvoyl-tetrahydropterin synthase (PTPS, gene symbol: *PTS*) is involved in the second step of the *de novo* biosynthesis of tetrahydrobiopterin (BH4), which is a vital cofactor of nitric oxide synthases and three types of aromatic amino acid hydroxylases; the latter are important enzymes in the production of neurotransmitters. We conducted a study of *PTS* mutations in East Asia, including Taiwan, Mainland China, Japan, South Korea, the Philippines, Thailand and Malaysia. A total of 43 mutations were identified, comprising 22 previously reported mutations and 21 new discovered mutations. Among these, the c.155A>G, c.259C>T, c. 272A>G, c.286G>A and c.84-291A>G mutations were the most common *PTS* mutations in East Asia, while the c.58T>C and c.243G>A mutations were, respectively, specific to Filipinos and Japanese originating from Okinawa. Further studies demonstrated that each of the mutations listed above was in linkage disequilibrium to a specific allele of polymorphic microsatellite marker, D11S1347. These results suggest the presence of founder effects that have affected these frequent mutations in East Asia populations. In this context, D11S1347 should become one of the most reliable polymorphic markers for use in prenatal diagnosis among PTPS deficient families, especially where mutations are yet to be identified.

*Journal of Human Genetics* (2012) 57, 145–152; doi:10.1038/jhg.2011.146; published online 12 January 2012

**Keywords:** founder effect; HPA; microsatellite marker; mutation; PTPS deficiency

## INTRODUCTION

Hyperphenylalaninemia (HPA) is caused by a disruption of hydroxylation step converting phenylalanine to tyrosine and involves either a deficiency of the enzyme phenylalanine hydroxylase (EC 1.14.16.1) or a deficiency of its cofactor, tetrahydrobiopterin (BH4).<sup>1</sup> Such depletion may lead to an increase in the individual's plasma phenylalanine concentration above 120 μM and this is the most common autosomal recessive disorder of amino acid metabolism. If mild form HPA is excluded, HPA patients show no clinical symptoms during neonatal period, but, in the absence of early treatment, then undergo impaired cognitive development and function that leads to mental retardation. This damage can be avoided by early biochemical diagnosis and appropriate treatment.<sup>2,3</sup> The estimated prevalence of HPA in Taiwan, Mainland China, Japan, Thailand, South Korea, the Philippines and Malaysia is 1/23 000,<sup>4</sup> 1/12 000,<sup>5</sup> 1/70 000,<sup>6</sup> 1/36 000 ~ 1/327 000,<sup>7</sup> 1/50 000, 1/114 000 and 1/450 000 (unpublished data),

respectively, whereas the prevalence in Caucasian populations is about 1 in 10 000.<sup>8</sup>

The synthesis of BH4 starts with GTP and is catalyzed by a series of enzymes, namely GTP cyclohydrolase I (GTPCH, EC 3.5.4.16), 6-pyruvoyl-tetrahydropterin synthase (PTPS; EC 4.6.1.10, gene symbol: *PTS*) and sepiapterin reductase (SR, EC 1.1.1.153).<sup>1</sup> Once formed as a cofactor, BH4 is converted to inactivated *q*-dihydrobiopterin, which can be regenerated by pterin-4 $\alpha$ -carbinolamine dehydrase (PCD; EC 4.2.1.96) and dihydropteridine reductase (DHPR; EC 1.6.99.7).<sup>1</sup> Depletions in any of the enzymes involved in BH4 biosynthesis or regeneration leads to BH4 deficiency.<sup>1</sup> Furthermore, BH4 is not just involved in the phenylalanine hydroxylation system; BH4 is also a vital cofactor for nitric oxide synthases, tyrosine hydroxylase and tryptophan hydroxylase, the latter two of which are involved in the biosynthesis of the neurotransmitters dopamine and serotonin,<sup>1</sup> respectively. As a result BH4-deficient patients without proper

<sup>1</sup>Institute of Genetics, National Yang-Ming University, Taipei, Taiwan; <sup>2</sup>Department of Education and Research, Taipei City Hospital, Taipei, Taiwan; <sup>3</sup>Genome Research Center, National Yang-Ming University, Taipei, Taiwan; <sup>4</sup>Department of Pediatrics, Taipei Veterans General Hospital, Taipei, Taiwan; <sup>5</sup>Department of Pediatrics, Peking University First Hospital, Beijing, China; <sup>6</sup>The Department of Endocrinology and Genetic Metabolism, Shanghai Institute for Pediatric Research, Xinhua Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China; <sup>7</sup>Guangzhou Neonatal Screening Center, Maternal and Neonatal Hospital of Guangzhou, Guangzhou, China; <sup>8</sup>Department of Pediatrics, Osaka City University Graduate School of Medicine, Osaka, Japan; <sup>9</sup>Department of Pediatrics, Soonchunhyang University Hospital, Seoul, South Korea; <sup>10</sup>Department of Pediatrics, Queen Sirikit National Institute of Child Health, College of Medicine, Rangsit University, Bangkok, Thailand; <sup>11</sup>Division of Medical Genetics, Department of Pediatrics, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand; <sup>12</sup>Paediatric Institute, Kuala Lumpur Hospital, Kuala Lumpur, Malaysia; <sup>13</sup>Institute of Human Genetics, National Institute of Health, University of the Philippines Manila, Manila, Philippine; <sup>14</sup>Children's Clinic and Human Genetics Information Resource Center, Cebu, Philippine and <sup>15</sup>Preventive Medicine Foundations, Taipei, Taiwan

<sup>16</sup>Current address: Department of Laboratory Medicine, Chang-Gung Medical Foundation, Taoyuan, Taiwan.

Correspondence: Dr T-T Liu, Genome Research Center, National Yang-Ming University, No.155, Sec.2, LiNong Street, Taipei 112, Taiwan.

E-mail: tliu@ym.edu.tw

Received 17 September 2011; revised 21 November 2011; accepted 25 November 2011; published online 12 January 2012

treatment show an impairment in phenylalanine catabolism that is accompanied by a deficiency in neurotransmitters.<sup>3</sup> In contrast to the dietary therapy used for phenylalanine hydroxylase deficiency, the administration of BH4 and neurotransmitters is used for the treatment of most BH4-deficient patients and additional folic acid supplement is recommended for DHPR deficiency.<sup>3,9</sup> The incidence of BH4 deficiency is ~1-3% of HPA patients worldwide,<sup>10</sup> but is more frequent in East Asian, including Taiwan (17%),<sup>4</sup> Mainland China (9%),<sup>11</sup> Thailand (17%),<sup>7</sup> South Korea (10%), the Philippines (23%) and Malaysia (64%) (unpublished data) with the exception of Japan (3%) (unpublished data).

Depletion of the PTPS enzyme (MIM 261640), a highly heterogeneous disorder, is the most common cause of BH4 deficiency.<sup>3</sup> The correlation between phenotype and genotype is still unclear. In humans, the *PTS* gene contains six exons and spans ~8 kb on chromosome 11q22.3-11q23.3. More than 53 *PTS* mutations, including missense mutations, nonsense mutations, splicing error mutations and small insertion/deletions, which are spread over the six exons and the first three introns, have been described in various different populations.<sup>12</sup>

In this study, a molecular analysis of the *PTS* gene was performed on populations in East Asia, including the Han populations in Taiwan, Mainland China and Malaysia as well as the populations of Japan, South Korea, Thailand and the Philippines. A total of 43 different *PTS* gene mutations were observed. Four missense mutations c.155A>G, c.259C>T, c.272A>G and 286G>A together with one newly discovered intronic mutation c.84-291A>G were the most frequent alleles detected in these East Asian populations. A haplotype analysis of the common mutations in relation to a polymorphic microsatellite marker, D11S1347, suggested that these common mutations observed in Han populations were influenced by a founder effect. In the 137 PTPS-deficient families analyzed so far, no crossing over events were detected. We therefore proposed that each of the frequent mutations in East Asia is owing to the same original mutational event and that a haplotype analysis of D11S1347 should be included in the prenatal diagnostic test for PTPS deficiency, especially when the familial mutation is as yet unidentified.

## SUBJECTS AND METHODS

### Subjects

The patients were recruited from 176 families of whom 156 were ethnic Han Chinese populations (29 from Taiwan, 52 from southern Mainland China, 71 from northern Mainland China and four from Malaysia); the others were six Japanese, seven South Korean, three Thai and four Filipinos. Patients were all from non-consanguineous families. A total of 352 mutant alleles were analyzed. Southern and northern Mainland China refers to south and north of the Yangtze River in Mainland China, respectively.<sup>13</sup> All patients were confirmed as having PTPS deficiency by the urinary pterin test, the BH4 loading test and/or DHPR enzyme activity assay. Anonymous control samples were collected from apparently healthy individuals, 50 from Taiwan, 55 from southern Mainland China, 52 from northern Mainland China, 56 from Japan, 49 from South Korea, 50 from Thailand and 79 from the Philippines.

### Mutation analysis

Mutation analysis was mainly performed on dried blood spots stored on filter paper as previous reported.<sup>14</sup> In some cases, genomic DNA was extracted from blood samples, cultured fibroblasts or lymphoblasts by standard protocols. The DNA fragments for all patients as well as the normal controls were amplified by PCR followed by bidirectionally fluorescent sequencing on an ABI 377 or ABI 3730xl DNA analyzer (Life Technologies, Carlsbad, CA, USA) with BigDye chemistry (Life Technologies). The sequence data were analyzed by the Polyphred/Phrap/Consed system.<sup>15</sup> The coding regions and the flanking areas

of the *PTS* gene were analyzed first. Intronic sequences were also analyzed if only one mutant allele was found in flanking areas and genomic DNA was available. Mutation designation was according to the official mutation nomenclature (<http://www.hgvs.org/mutnomen/>). Reference accession numbers NM\_000317.2 and M97655.1 were, respectively, used as reference sequences for the human genomic DNA and human cDNA with the A at the ATG translational initiation codon being numbered 1. Disease-causing mutations were suggested as having no other variations in the coding regions of the *PTS* gene in the patient, no similar mutation in 50 normal controls, showing phenotypic association in families, involving a conserved amino acid residue in mammals, involving a change in chemical characteristics of the amino acid involved and/or concurrent appearance among more than one PTPS-deficient families.

### Splicing mutation analysis

The effects of the intronic mutations were analyzed when a patient's cells were available. Total RNA was isolated from cultured cells using RNAlarge reagent (Genepure Technology, Taichung, Taiwan) and the first strand cDNA was synthesized using the SuperScript Reverse Transcription System (Life Technologies) and the primer PTPSS (5'-AATTGAATTCAGATGAGCAG GAAGGTG-3'). Amplification of the *PTS* cDNA was performed with the primers PTPSS and PTPSAS (5'-AATGAATCAATAAATAGCACTCCA-3'). The cDNA was then subcloned into the vector pCRII using a TA Cloning Kit (Life Technologies) and analyzed by sequencing as described in previous section.

### Haplotype analysis

Genotypes were analyzed for the normal populations and the PTPS-deficient families as well as three control cells, CEPH1331-01, CEPH1331-02 and CEPH1347-02, which were obtained from the Foundation Jean Dausset-Centre d'Etude du polymorphisme Humain (CEPH). DNA fragments were amplified by PCR with fluorescence labeled primers, which was followed by analysis on an ABI 3730xl ABI 3730 XL DNA analyzer. The allele frequencies of the populations were calculated from the normal individuals. Linkage analysis of the microsatellite marker allele in relation to the various mutations was carried out for PTPS-deficient families. The difference in distribution of the microsatellite alleles among the populations and the association of specific alleles of the microsatellite marker with the common mutations was calculated by Fisher's two-tailed exact test. The  $\chi^2$  test with Yate's correction was used to assess the Hardy-Weinberg equilibrium status of the alleles within populations. A *P* value <0.05 was considered as significant.

## RESULTS AND DISCUSSION

### Mutation spectrum in Han and other East Asian populations

The 43 autosomal mutations identified consisted of 22 previously reported mutations and 21 mutations that are first reported in this study; these were detected in a total of 352 mutant alleles (Table 1). (The number of mutant alleles observed in each population is listed in Supplementary Table 1). Among all mutations identified, there were 33 missense mutations, two nonsense mutations, one small deletion, one frameshift mutation, five confirmed/predicted splicing error mutations and one initiation codon change (Table 1). The mutations were distributed across all six exons as well as the first three introns. No polymorphisms other than the previously reported single-nucleotide polymorphism (SNP), c.163+14T/C (rs3819331),<sup>16</sup> were detected. The c.118-121del (p. F40Gfs\*17) was previously named c.116-119del.<sup>16</sup> It was possible to identify only one mutant allele in six patients after sequencing of the gene's coding regions, the flanking regions and the known intronic mutation, suggesting that more disease causing mutations may be located in the non-coding regions. However, mutation analysis on the non-coding regions in these patients was not possible because of the limited and/or poor quality of the samples.

Five common mutations, c.155A>G, c.259C>T, c.272A>G, c.286G>A and c.84-291A>G, were frequently observed in at least

**Table 1** P<sub>TPS</sub> gene mutations observed in East Asian populations

Mutation type	Mutations	Predicted effect on protein	Number of alleles	Key Reference(s)
Missense	c.26G>A	p.R9H	1	This study
	c.58T>C	p.F20L	6	This study
	c.68G>C	p.S23T	1	This study
	c.73C>G	p.R25G	7	Liu <i>et al.</i> <sup>14</sup>
	c.74G>A	p.R25Q	1	Thöny <i>et al.</i> <sup>39</sup>
	c.100G>A	p.E34K	1	This study
	c.120T>A	p.F40L	1	This study
	c.120T>G	p.F40L	1	Liu <i>et al.</i> <sup>16</sup>
	c.147T>G	p.H49Q	1	Vatanavicharn <i>et al.</i> <sup>40</sup>
	c.155A>G	p.N52S	55	Liu and Hsiao <sup>41</sup> , Leuzzi <i>et al.</i> <sup>17</sup>
	c.166G>A	p.V56M	7	Liu <i>et al.</i> <sup>14</sup>
	c.200C>A	p.T67K	1	This study
	c.200C>T	p.T67M	6	Oppliger <i>et al.</i> <sup>42</sup> Zekanowski <i>et al.</i> <sup>33</sup> Liu <i>et al.</i> <sup>16</sup>
	c.209T>A	p.V70D	1	Liu <i>et al.</i> <sup>14</sup>
	c.226C>T	p.L76F	1	Liu <i>et al.</i> <sup>16</sup>
	c.259C>T	p.P87S	132	Liu and Hsiao <sup>41</sup> Imamura <i>et al.</i> <sup>18</sup>
	c.260C>T	p.P87L	1	Oppliger <i>et al.</i> <sup>43</sup>
	c.272A>G	p.K91R	14	This study
	c.276T>A	p.N92K	1	This study
	c.277C>A	p.L93M	1	This study
	c.286G>A	p.D96N	37	Liu and Hsiao <sup>41</sup> Imamura <i>et al.</i> <sup>18</sup>
	c.293C>A	p.P98Q	1	This study
	c.293C>T	p.P98L	2	This study
	c.317C>T	p.T106M	14	Liu and Hsiao <sup>41</sup> Leuzzi <i>et al.</i> <sup>17</sup>
	c.331G>A	p.A111T	2	This study
	c.338A>G	p.Y113C	1	Thöny <i>et al.</i> <sup>12</sup>
	c.347A>G	p.D116G	1	Sherer-Oppliger <i>et al.</i> <sup>44</sup>
	c.379C>T	p.L127F	4	Qu <i>et al.</i> <sup>45</sup>
	c.382T>A	p.Y128N	1	This study
	c.393A>C	p.K131N	1	This study
	c.402A>C	p.E134D	1	This study
	c.407A>T	p.D136V	1	Oppliger <i>et al.</i> <sup>42</sup>
	c.430G>C	p.G144R	1	Liu <i>et al.</i> <sup>16</sup>
Nonsense	c.37C>T	p.Q13X	1	This study
	c.73C>T	p.R25X	1	This study
Small deletion	c.169_171del	p.V57del	1	Oppliger <i>et al.</i> <sup>42</sup> Liu <i>et al.</i> <sup>16</sup>
Frameshift mutations	c.118-121del	p.F40Gfs*17	4	Liu <i>et al.</i> <sup>16</sup>
Splicing error	c.84-291A>G	p.Y27Rfs*8	25	This study
	c.163+1G>A	p.=?	1	This study
	c.168G>A	p.=? (p.V56V)	1	This study
	c.186+1G>A	p.V55Dfs*2	1	Liu <i>et al.</i> <sup>16</sup>
	c.243G>A	p.631_81EdeI	3	Imamura <i>et al.</i> <sup>18</sup>
Initiation codon change	c.3G>A	p.M1?	1	This study
	ND	ND	6	

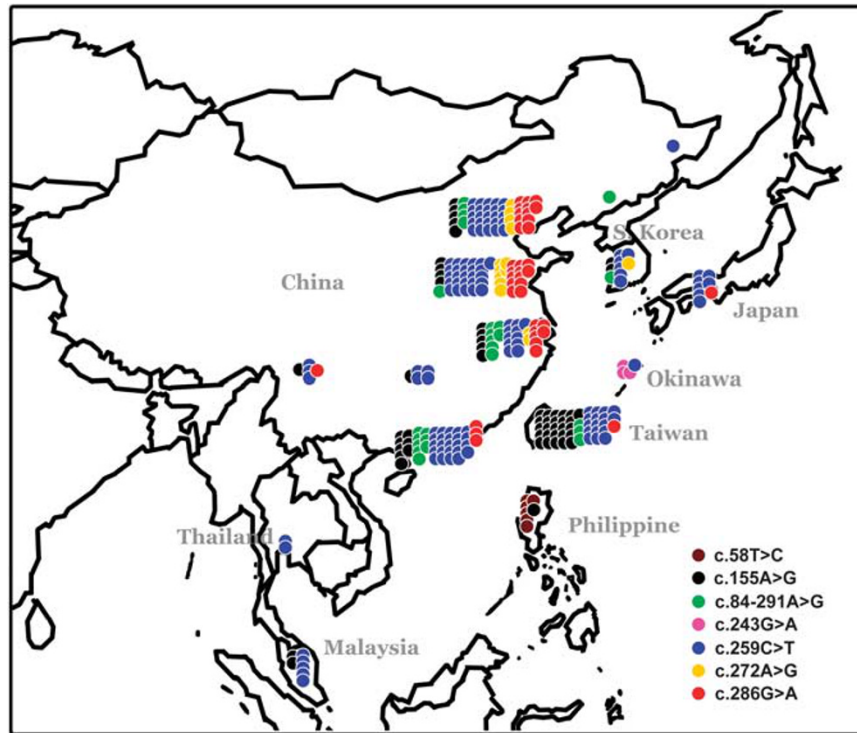
Abbreviation: ND, not determined.

three different East Asian populations and accounted for 15.6% (55/352), 37.5% (132/352), 4.0% (14/352), 10.5% (37/352) and 7.1% (25/352) of the alleles analyzed (Figure 1). The c.259C>T, c.272A>G and c.286G>A mutations have only been reported in Asian populations, whereas the c.155A>G mutation has been observed in western countries.<sup>12,17</sup> More than 89.7% (158/176) of the patients carried at least one of the five common mutations suggesting that screening for these common mutations should be an early step in the treatment of PTPS-deficient patients from East Asia. The distribution of the common mutations differs from population to population. The c.259C>T mutation was the most common mutation in Mainland China (94/246), whereas c.155A>G was the most frequent mutant allele found in Taiwan (26/58). The c.58T>C and c.243G>A mutations were only observed in the Philippines and in Okinawa, Japan,<sup>18</sup> respectively.

A total of 21 of the mutations are newly reported, including 15 missense mutations, two nonsense mutations, three splicing error mutations and one initiation codon change. There is a supportive information that these new variations, which are only observed in a single patient, are pathogenic. First, initiation codon variants such as c.3G>A identified in this study, have been widely suggested to be a disease causing mutations.<sup>19</sup> Second, small in frame deletions usually cause protein misfolding and instability and this suggested that they will be disease causing mutations.<sup>20</sup> The c.169\_171del (p.V57del) mutation identified in this study might also lead to protein misfolding, because it is located between two conserved amino acids, one branched (valine) and one polar (threonine). Third, nonsense mutations have been widely suggested to be disease causing mutations.<sup>21</sup> The two new discovered nonsense mutations in this study, c.37C>T (p.Q13X) and c.73C>T (p.R25X), will result in proteins that loose approximately the last 3/4 and 1/2 of the protein sequence and are very likely to be non-functional. Finally, the remaining missense mutations are suggested to be disease-causing mutation, because they are absent from the normal population and NCBI single-nucleotide polymorphism database, they are conserved in primates/mammals and the alteration in amino acid chemical properties they produce is significant. Together these factors suggest that these rare variations, as described here, are associated with PTPS deficiency.

### Splicing error mutations

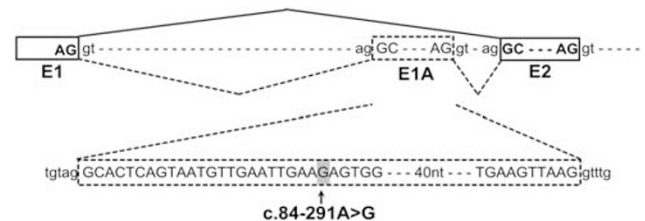
It has been estimated that 15% of inherited disease in humans is caused by mutations in splice sites or in splicing control sequences, such as exonic splicing enhancers (ESE) and intronic splicing enhancers, all of which lead to splicing errors.<sup>22-24</sup> The disruption of a consensus splice sequence is the most frequent type of splicing error mutation and leads to the skipping of exons.<sup>22</sup> Three such mutations, c.163+1G>A, c.186+1G>A and c.243G>A, were observed in this study. Two mutations, c.186+1G>A and c.243G>A, have been previously reported; they possibly lead to a splicing error by frame-shift skipping of exon 3 and in-frame deletion of exon 4, respectively.<sup>18,25</sup> The newly identified c.163+1G>A variant disrupts a consensus splice site sequence, which suggests that it causes a splicing error mutation. In human disease genes, there are numerous mutations in ESE control sequences that have been documented as causing aberrant exon skipping<sup>23</sup> or the creation of pseudoexon.<sup>24</sup> The c.168G>A mutation (p.V56V) was identified in a patient from southern Mainland China who also carried a single c.84-291A>G mutation. The c.168G>A (p.V56V) variation was initially suggested to be a new silent polymorphism. However, using ESE motif prediction tools, we found that this mutation represented a change in the SC35-binding site (Supplementary Table 2A), which, in turn, suggests



**Figure 1** The distribution of major common mutations in East Asia as determined by this study. The common mutations c.58T>C, c.155A>G, c.84-291A>G, c.243G>A, c.259C>T, c.272A>G and c.286G>A are represented in brown, black, green, pink, blue, orange and red, respectively. The mutation c.155A>G was predominantly observed in South East Asia, while c.286G>A and c.272A>G were mainly observed in North East Asia. The mutations c.58T>C and c.243G>A mutations were specifically observed in the Philippines and in Okinawa, Japan, respectively.

that c.168G>A causes a splicing error mutation. An analysis of the precise effects of the c.163+1G>A and c.168G>A mutations needs to be done using cells from a patient or from a carrier as they become available.

The intronic c.84-291A>G mutation is frequently observed in the Han populations. ESE finder predicted an alteration of the SRp40 splicing protein-binding site when the c.84-291A>G mutation is present (Supplementary Table 2B).<sup>26,27</sup> A 79 bp pseudoexon flanked by consensus splice-site can be observed in cells carrying the c.84-291A>G mutation (Figure 2). The same 79 bp pseudoexon can also be observed elsewhere in cells carrying the c.84-322A>T mutation in the *PTS* gene while no similar splicing pattern was identified in the control cells.<sup>28</sup> To quantify the effects of c.84-291A>G, cDNA from the skin fibroblasts of a patient heterozygous for the c.84-291A>G and c.259C>T mutations was analyzed by subcloning. As expected, 49% of the clones represented the c.259C>T mutation (40/82) and 6% of clones showed exon 3 skipping (5/82) as previously reported.<sup>16</sup> Intriguingly, only 9% of the clones contained the 79 bp pseudoexon (7/82), while the other 37% of clones represented a normal *PTS* cDNA (30/82). Although the phenotype of most patients carrying the c.84-291A>G was not available owing to early diagnosis by newborn screening and proper treatment, the low level of normal *PTS* cDNA correlates well with the very mild form of a PTPS-deficient patient, namely 115 in Liu TT *et al.*,<sup>25</sup> who is heterozygous for the c.84-291A>G and c.118-121del mutations (designated as 116-119del previously). The patient 115 was detected by newborn screening and treated with phenylalanine restricted dietary therapy before differential diagnosed as PTPS deficiency at 3 years and 8 months of age. This patient did not receive BH4 therapy nor neurotransmitters supplement and showed no neurological illness with IQ 109 until puberty.

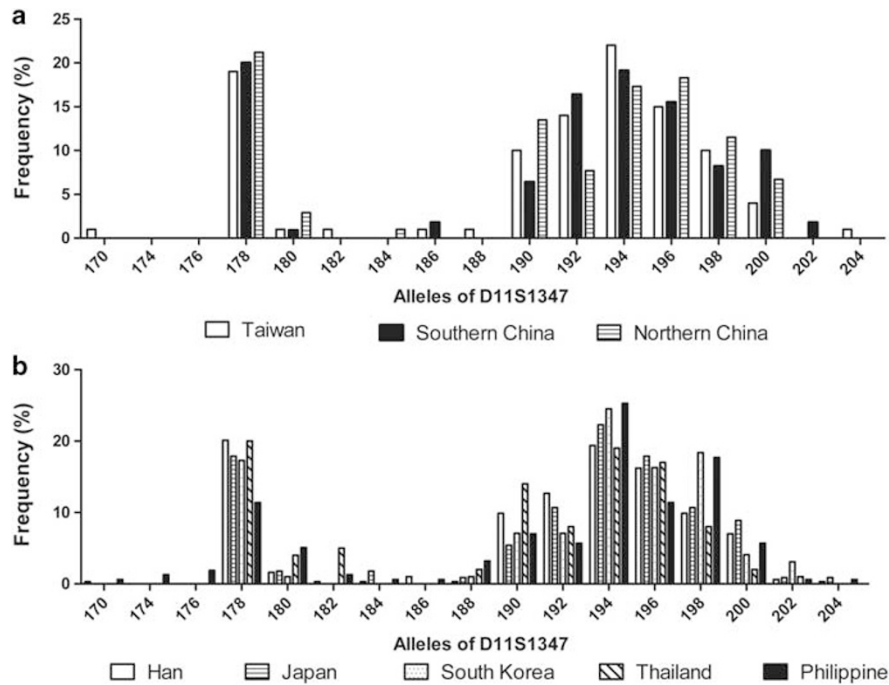


**Figure 2** Schematic representation of the splicing pattern caused by the c.84-291A>G mutation. The exon 1 and exon 2 of the *PTS* gene is enclosed in a box. The pseudoexon sequence is highlighted in a dashed box and the mutation is enclosed in a grey box. The black line represents the normal splicing and the dashed line indicates the abnormal splicing caused by the c.84-291A>G mutation.

It is unclear what the effect(s) of these mutations might be on cells in central nervous system, but c.84-291A>G may only create a weaker-binding site for the SR-splicing factors and allow a significant portion of normal splicing to be formed, resulting in mild form hyperphenylalaninemia.

#### Selection of a marker for linkage analysis

In order to identify a feasible marker for linkage analysis of the *PTS* gene, three polymorphic microsatellite markers around the *PTS* gene were analyzed. D11S1986, D11S1347 and D11S1987 are located at 0.9Mb downstream, 29 kb upstream and 0.58Mb upstream of the *PTS* gene, respectively. D11S1986 and D11S1987 were excluded because of their low heterozygosity in the Han population from Taiwan and the presence of recombination during a preliminary study of two PTPS-deficient families (data not shown). D11S1347 is a highly polymorphic



**Figure 3** Allele frequency of D11S1347 in East Asian populations. (a) No significant difference in D11S1347 allele frequency was seen between ethnic Han populations in Taiwan, southern Mainland China and northern Mainland China ( $P>0.05$  in all groups). Some rare D11S1347 alleles were only seen in the Han population in Taiwan. (b) The allele frequencies of D11S1347 among Han, Japanese, Korean and Thai populations were similar ( $P>0.05$  in all groups). However, Filipinos showed a very different allele frequency for D11S1347 compared with other nationalities except for Koreans ( $P>0.05$  with Koreans and  $P<0.05$  with other groups).

marker for which no recombination with the *PTS* gene was found in a preliminary study of 26 Taiwanese PTPS-deficient families. This suggests that D11S1347 might be a reliable polymorphic marker for linkage studies of *PTS* mutations. Furthermore, the data on D11S1347 in this study differed from that reported in CEPH. The control DNA, CEPH1331-01, CEPH1331-02 and CEPH1347-02, obtained from CEPH, were read in the original study as 178/196 bp, 178/190 bp and 194/194 bp, but were read as 178/197 bp, 178/191 bp and 195/195 bp in the present study, respectively (Supplementary Table 3). The difference might be the result of variations in methods and analysis used.

The allelic frequency of D11S1347 was first analyzed in general populations from Taiwan, southern Mainland China, northern Mainland China, Japan and South Korea in which the heterozygosities were found to be 0.88, 0.76, 0.87, 0.91, 0.96, 0.86 and 0.89, respectively (Figure 3). (The allele frequencies of D11S1347 in each population are listed in Supplementary Table 4). The data demonstrated a Hardy-Weinberg equilibrium in each population analyzed ( $\chi^2=0.02$ ,  $P>0.05$ ;  $\chi^2=0.84$ ,  $P>0.05$ ;  $\chi^2=0.00$ ,  $P>0.05$ ;  $\chi^2=0.16$ ,  $P>0.05$ ;  $\chi^2=0.56$ ,  $P>0.05$ ;  $\chi^2=0.00$ ,  $P>0.05$  and  $\chi^2=0.00$ ,  $P>0.05$ , respectively). Among the ethnic Han populations, no significant differences in D11S1347 allele frequency for Taiwan, southern Mainland China and northern Mainland China were seen (Figure 3a and Supplementary Table 4,  $P>0.05$  in all groups), although some rare alleles, such as 170 bp, 182 bp, 188 bp and 204 bp, were only observed in Taiwan. The Filipinos demonstrated a very different allele frequency pattern for D11S1347 compared with all other nationalities except for Koreans ( $P>0.05$  with South Koreans and  $P<0.05$  with other groups). The observation correlated with the distinct origin of Filipinos compared with other the populations that have been proposed in a previous study.<sup>29</sup> Excluding the Filipinos, no significant difference in the allele

frequency for D11S1347 was seen across the populations studied ( $P>0.05$  for all groups), although a distinction has been previously proposed for southern and northern populations in East Asia.<sup>30</sup>

### The founder events and mutations in East Asia

To examine the possibility of founder event effects involving the *PTS* common mutations observed in East Asia, the haplotype of mutations in conjunction with D11S1347 alleles were evaluated by DNA analysis for samples where samples from the parents of probands were available. There were 43 alleles of c.155A>G, 98 alleles of c.259C>T, 13 alleles of c.272A>G, 28 alleles of c.286G>A and 16 alleles of c.84-291A>G analyzed. The c.155A>G, c.259C>T, c.272A>G and c.286G>A mutations were in linkage disequilibrium with the 178 bp, 196 bp, 194 bp and 192 alleles of D11S1347 (Table 2,  $P<0.01$  for all groups). The c.84-291A>G mutation was statistical significant linked to the 198 bp allele ( $P=0.0008$ ), but not the 194 bp allele ( $P=0.0531$ ), although the number of these two haplotypes was equal ( $n=7$  in both types). None of the common mutations are located in CpG mutation hot spots, which suggest that the possibility of independent mutations is low. Thus the linkage disequilibrium data suggest that each of the common mutations observed in East Asia comes from a single ancestor.

Two missense mutations, c.58T>C and c.243G>A, were demographically distinct, namely the Filipinos and the Japanese from Okinawa. Association of c.58T>C with the 180 bp allele of D11S1347 and c.243G>A with the 198 bp allele of D11S1347 was found (Table 3a,  $P=0.002$  for both groups). These results suggest each of these two demographically common mutations also involve a founder event.

The c.200C>T and c.317C>T mutations, on the other hand, were located in CpG mutational hot spots<sup>31</sup> and have been found in several

**Table 2** Numbers of specific D11S1347 alleles associated with c.155A>G, c.259C>T, c.272A>G, c.286G>A and c.84-291A>G mutations in *PTS* gene

Mutations	Populations	178 bp	188 bp	190 bp	192 bp	194 bp	196 bp	198 bp	200 bp
c.155A>G	Han in Taiwan	25							
	Han in SC	8							
	Han in NC	6							
	Korean	1				1			
	Malaysian	2							
c.259C>T	Han in Taiwan			1	1		10	2	
	Han in SC			2	2	1	24	1	
	Han in NC	1		5	3	2	21	1	
	Japanese						8		
	Korean						6		
	Thai						2		
	Malaysian			1			3		1
c.272A>G	Han in SC					2			
	Han in NC		1		1	7	1		
	Korean					1			
c.286G>A	Han in Taiwan				1				
	Han in SC			1	5	1	2		
	Han in NC				13	2	2		
	Japanese				1				
c.84-291A>G	Han in Taiwan							2	1
	Han in SC					6		4	1
	Han in NC					1			
	Korean							1	

Abbreviations: SC, southern Mainland China; NC, northern Mainland China.

independent areas. The c.200C>T mutation has been reported in Albania and Italy,<sup>32,33</sup> whereas the c.317C>T has been previously observed in the Caucasus region and Turkey.<sup>17</sup> No association of c.200C>T with any D11S1347 alleles was found, indicating that several independent mutations had occurred at this mutation hot spot (Table 3b). The diversity of haplotypes with c.317C>T also supported independent mutations, although a founder event might have occurred in some, if not all, of the eight alleles associated with 194 bp.

### The spread of the founder mutations

Among the common *PTS* mutations in East Asia, the c.259C>T mutation was the most widely distributed in northern and southern East Asia and had seven different haplotypes, which indicates that c.259C>T is likely to be the most ancient mutation. Moreover, six different haplotypes were identified in Mainland China, in contrast to four haplotypes in Taiwan, three in Malaysia and only one haplotype in Japan, Southern Korea and the Thailand. These observations suggest that the most common *PTS* mutation in East Asia comes from a Han ancestor. The mutations might have spread eastward to Korean and Japan and southward to Thailand via multiple migrations. In contrast, the c.155A>G mutation was predominantly observed in South East Asia and almost all mutant alleles (42/43) were linked to the 178 bp allele of D11S1347. These observations indicate that the c.155A>G mutation is a relatively young mutation that originated in South East Asia. The c.272A>G and c.286G>A mutations are mainly observed in North East Asia and are associated with four different D11S1347 alleles, which suggests that both mutations appeared in the period between the origin of c.259C>T and the origin of c.155A>G.

Several founder event mutations have been reported in Han populations previously, including c.1935C>A in the *GAA* gene,<sup>34</sup>

c.1517\_1525del in the *17OHD* gene<sup>35</sup> and c. 1199A>G in the *IVD* gene.<sup>36</sup> This study represents the largest number of founder event mutation profiles identified so far in Han populations. On the basis of the presumed population distribution and migration patterns of Han populations in East Asia as determined by population genetic studies<sup>30</sup> and the specific haplotypes of these founder mutations, we speculate that these common mutations in the *PTS* gene of East Asian individuals all originated from Han populations.

### The utility of D11S1347

Dried blood spots are a convenient resource for newborn screening;<sup>4,5,7</sup> however, this approach also limits sample size and quality. The DNA samples in this study were mainly extracted from dried blood spots and this led to limitations in terms of whole-gene analysis. To date, six of our PTPS-deficient patients have had only one mutant allele identified. Therefore, it is important to identify other polymorphic markers that are able to facilitate molecular analysis of this disorder. Microsatellite markers, in addition to genetic analysis, have been used for carrier and prenatal diagnosis in many studies.<sup>37</sup> D11S1347 is highly polymorphic in East Asian populations and can easily be analyzed. Furthermore, the analysis of microsatellite markers from a fetus can be done in such a way as to exclude contamination from maternal DNA.<sup>38</sup> No cross-over event has as yet been detected after the analysis of 139 PTPS-deficient families. The haplotype analysis of the mutant and D11S1347 alleles has been included in all of our prenatal diagnoses. Nonetheless, it is still important to recognize that the very low recombination rate must not be ignored. We suggested that haplotype analysis of D11S1347 in conjunction with the *PTS* mutant allele should be used in all prenatal diagnosis of PTPS-deficient families.

**Table 3** Number of the D11S1347 alleles associated with minor mutations, including c.58T>C, c.243G>A, c.200C>T and c.317C>T mutations

Mutations	Populations	178 bp	180 bp	194 bp	198 bp		
<b>(A)</b>							
c.58T>C	Filipino	1	3	1			
c.243G>A	Japanese				3		
Mutations	Populations	178 bp	188 bp	192 bp	194 bp	196 bp	198 bp
<b>(B)</b>							
c.200C>T	Han in NC			1			
	Thai	2			1		
c.317C>T	Han in Taiwan			1			
	Han in SC				1		
	Han in NC				7	1	
	Korean			1			1

Abbreviations: SC, southern Mainland China; NC, northern Mainland China.

## Conclusions

Here, we reported the largest study of PTPS deficiency in East Asia up to the present. The mutations c.155A>G, c.259C>T, c.272A>G, c.286G>A and c.84-291A>G were the most frequent mutant alleles in Han populations. Strong linkage was found with a specific allele of the polymorphic microsatellite marker D11S1347 and this indicates that each of the common mutations widely observed in East Asia is likely to be the result of a founder mutation in a Han population. Furthermore, the mutations c.58T>C and c.243G>A, which dominated in the Philippines and Okinawa Japan, respectively, also represent founder events that are restricted to certain isolated geographic regions. Based on the very low recombination rate detected in this study, we proposed that D11S1347 is likely to be a highly reliable polymorphic marker for linkage analysis of the PTS gene and it therefore should be applied for prenatal diagnosis in families where the specific mutation involved in the disease has not yet been identified.

## ACKNOWLEDGEMENTS

The work was financially supported by the grants NSC87-2314-B-010-085-032 from the National Science Council and VGH-384 from the Taipei Veterans General Hospital, Taiwan.

- Thöny, B., Auerbach, G. & Blau, N. Tetrahydrobiopterin biosynthesis, regeneration and functions. *Biochem. J.* **347**, 1–16 (2000).
- Aoki, K., Ohwada, M. & Kitagawa, T. Long-term follow-up study of patients with phenylketonuria detected by the newborn screening programme in Japan. *J. Inher. Metab. Dis.* **30**, 608 (2007).
- Longo, N. Disorders of biopterin metabolism. *J. Inher. Metab. Dis.* **32**, 333–342 (2009).
- Niu, D. M., Chien, Y. H., Chiang, C. C., Ho, H. C., Hwu, W. L., Kao, S. M. *et al.* Nationwide survey of extended newborn screening by tandem mass spectrometry in Taiwan. *J. Inher. Metab. Dis.* **33**, S295–S305 (2010).
- Zhan, J. Y., Qin, Y. F. & Zhao, Z. Y. Neonatal screening for congenital hypothyroidism and phenylketonuria in China. *World J. Pediatr.* **5**, 136–139 (2009).
- Aoki, K. Long term follow-up of patients with inborn errors of metabolism detected by the newborn screening program in Japan. *Southeast Asian J. Trop. Med. Public Health* **34**(Suppl 3), 19–23 (2003).
- Pangkanon, S., Charoensiriwatana, W., Janejai, N., Boonwanich, W. & Chaisomchit, S. Detection of phenylketonuria by the newborn screening program in Thailand. *Southeast Asian J. Trop. Med. Public Health* **40**, 525–529 (2009).
- Steinfeld, R., Kohlschütter, A., Ullrich, K. & Lukacs, Z. Efficiency of long-term tetrahydrobiopterin monotherapy in phenylketonuria. *J. Inher. Metab. Dis.* **27**, 449–453 (2004).
- Scriver, C. R., Beaudet, A. L., Sly, W. S., Valle, D., in (eds Blau, N. Thöny, B. Cotton, R. & Hyland, K.) 1725–1776 (McGraw-Hill, New York, 2001).
- Blau, N., Barnes, I. & Dhondt, J. L. International database of tetrahydrobiopterin deficiencies. *J. Inher. Metab. Dis.* **19**, 8–14 (1996).
- Ye, J., Qiu, W., Han, L., Zhou, J., Gao, X. & Gu, X. The investigation of differential diagnostic development and incidence of tetrahydrobiopterin deficiency. *Chin. J. Prev. Med.* **43**, 128–131 (2009).
- Thöny, B. & Blau, N. Mutations in the BH4-metabolizing genes GTP cyclohydrolase I, 6-pyruvoyl-tetrahydropterin synthase, sepiapterin reductase, carbinolamine-4a-dehydratase, and dihydropteridine reductase. *Hum. Mutat.* **27**, 870–878 (2006).
- Wen, B., Li, H., Lu, D., Song, X., Zhang, F., He, Y. *et al.* Genetic evidence supports demic diffusion of Han culture. *Nature* **431**, 302–305 (2004).
- Liu, T. T., Hsiao, K. J., Lu, S. F., Wu, S. J., Wu, K. F., Chiang, S. H. *et al.* Mutation analysis of the 6-pyruvoyl-tetrahydropterin synthase gene in Chinese hyperphenylalaninemia caused by tetrahydrobiopterin synthesis deficiency. *Hum. Mutat.* **11**, 76–83 (1998).
- Ewing, B. & Green, P. Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Res.* **8**, 186–194 (1998).
- Liu, T. T., Chiang, S. H., Wu, S. J. & Hsiao, K. J. Tetrahydrobiopterin-deficient hyperphenylalaninemia in the Chinese. *Clin. Chim. Acta.* **313**, 157–169 (2001).
- Leuzzi, V., Carducci, C. A., Carducci, C. L., Pozzessere, S., Burlina, A., Cerone, R. *et al.* Phenotypic variability, neurological outcome and genetics background of 6-pyruvoyl-tetrahydropterin synthase deficiency. *Clin. Genet.* **77**, 249–257 (2010).
- Imamura, T., Okano, Y., Shintaku, H., Hase, Y. & Isshiki, G. Molecular characterization of 6-pyruvoyl-tetrahydropterin synthase deficiency in Japanese patients. *J. Hum. Genet.* **44**, 163–168 (1999).
- Kets, C. M., Hoogerbrugge, N., van Krieken, J. H., Goossens, M., Brunner, H. G. & Ligtenberg, M. J. Compound heterozygosity for two MSH2 mutations suggests mild consequences of the initiation codon variant c.1A>G of MSH2. *Eur. J. Hum. Genet.* **17**, 159–164 (2009).
- Douna, V., Papassotiriou, I., Metaxotou-Mavrommati, A., Stamoulakotou, A., Liapi, D., Kampourakis, D. *et al.* Further identification of the hyperunstable alpha-globin chain variant Hb Heraklion [codons 36/37 (-CCC); Pro->O (alpha1)] in Greek cases with co-inherited alpha+-thalassemia mutations. *Hemoglobin* **32**, 379–385 (2008).
- Nytofte, N. S., Serrano, M. A., Monte, M. J., Gonzalez-Sanchez, E., Tumer, Z., Ladefoged, K. *et al.* A homozygous nonsense mutation (c.214C->A) in the biliverdin reductase alpha gene (BLVRA) results in accumulation of biliverdin during episodes of cholestasis. *J. Med. Genet.* **48**, 219–225 (2011).
- Mount, S. M. Genomic sequence, splicing, and gene annotation. *Am. J. Hum. Genet.* **67**, 788–792 (2000).
- Collin, R. W., de Heer, A. M., Oostrik, J., Pauw, R. J., Plantinga, R. F., Huygen, P. L. *et al.* Mid-frequency DFNB8/12 hearing loss caused by a synonymous TECTA mutation that affects an exonic splice enhancer. *Eur. J. Hum. Genet.* **16**, 1430–1436 (2008).
- Homolova, K., Zavadakova, P., Doktor, T. K., Schroeder, L. D., Kozich, V. & Andresen, B. S. The deep intronic c.903+469T>C mutation in the MTRR gene creates an SF2/ASF binding exonic splicing enhancer, which leads to pseudoxon activation and causes the cbIE type of homocystinuria. *Hum. Mutat.* **31**, 437–444 (2010).

- 25 Liu, T. T., Chang, Y. H., Chiang, S. H., Yang, Y. L., Yu, W. M. & Hsiao, K. J. Identification of three novel 6-pyruvoyl-tetrahydropterin synthase gene mutations (226C>T, IVS3+1G>A, 116-119delTGTT) in Chinese hyperphenylalaninemia caused by tetrahydrobiopterin synthesis deficiency. *Hum. Mutat.* **18**, 83 (2001).
- 26 Cartegni, L., Wang, J., Zhu, Z., Zhang, M. Q. & Krainer, A. R. ESEfinder: A web resource to identify exonic splicing enhancers. *Nucleic. Acids. Res.* **31**, 3568–3571 (2003).
- 27 Smith, P. J., Zhang, C., Wang, J., Chew, S. L., Zhang, M. Q. & Krainer, A. R. An increased specificity score matrix for the prediction of SF2/ASF-specific exonic splicing enhancers. *Hum. Mol. Genet.* **15**, 2490–2508 (2006).
- 28 Meili, D., Kralovicova, J., Zagalak, J., Bonafé, L., Fiori, L., Blau, N. *et al.* Disease-causing mutations improving the branch site and polypyrimidine tract: pseudoexon activation of LINE-2 and antisense Alu lacking the poly(T)-tail. *Hum. Mutat.* **30**, 823–831 (2009).
- 29 Delfin, F., Salvador, J. M., Calacal, G. C., Perdigon, H. B., Tabbada, K. A., Villamor, L. P. *et al.* The Y-chromosome landscape of the Philippines: extensive heterogeneity and varying genetic affinities of Negrito and non-Negrito groups. *Eur. J. Hum. Genet.* **19**, 224–230 (2011).
- 30 Chu, J. Y., Huang, W., Kuang, S. Q., Wang, J. M., Xu, J. J., Chu, Z. T. *et al.* Genetic relationship of populations in China. *Proc. Natl Acad. Sci. USA* **95**, 11763–11768 (1998).
- 31 Bird, A. P. DNA methylation and the frequency of CpG in animal DNA. *Nucleic Acids Res.* **8**, 1499–1504 (1980).
- 32 Dudesek, A., Röschinger, W., Muntau, A. C., Seidel, J., Leupold, D., Thöny, B. *et al.* Molecular analysis and long-term follow-up of patients with different forms of 6-pyruvoyl-tetrahydropterin synthase deficiency. *Eur. J. Pediatr.* **160**, 267–276 (2001).
- 33 Zekanowski, C., Nowacka, M., Cabalska, B., Sendecka, E., Stowik, M., Gizewska, M. *et al.* Mutations causing hereditary hyperphenylalaninemia. *Med. Wieku. Rozwoj.* **3**, 55–66 (1999).
- 34 Shieh, J. J. & Lin, C. Y. Frequent mutation in Chinese patients with infantile type of GSD II in Taiwan: evidence for a founder effect. *Hum. Mutat.* **11**, 306–312 (1998).
- 35 Yang, J., Cui, B., Sun, S., Shi, T., Zheng, S., Bi, Y. *et al.* Phenotype-genotype correlation in eight Chinese 17 $\alpha$ -hydroxylase/17,20 lyase-deficiency patients with five novel mutations of CYP17A1 gene. *J. Clin. Endocrinol. Metab.* **91**, 3619–3625 (2006).
- 36 Lee, H. H., Lee, R. S., Lai, C. K., Yuen, Y. P., Siu, T. S., Chan, A. Y. *et al.* A novel duplication at the putative DNA polymerase alpha arrest site and a founder mutation in Chinese in the IVD gene underlie isovaleric acidaemia. *Hong. Kong. Med. J.* **16**, 219–222 (2010).
- 37 Harraway, J. R., Smith, M. P. & George, P. M. A highly informative, multiplexed assay for the indirect detection of hemophilia A using five-linked microsatellites. *J. Thromb. Haemost.* **4**, 587–590 (2006).
- 38 Winsor, E. J., Akoury, H., Chitayat, D., Steele, L. & Stockley, T. L. The role of molecular microsatellite identity testing to detect sampling errors in prenatal diagnosis. *Prenat. Diagn.* **30**, 746–752 (2010).
- 39 Thöny, B., Leimbacher, W., Blau, N., Harvie, A. & Heizmann, C. W. Hyperphenylalaninemia due to defects in tetrahydrobiopterin metabolism: molecular characterization of mutations in 6-pyruvoyl-tetrahydropterin synthase. *Am. J. Hum. Genet.* **54**, 782–792 (1994).
- 40 Vatanavicharn, N., Kuptanon, C., Liammongkolkul, S., Liu, T. T., Hsiao, K. J., Ratanarak, P. *et al.* Novel mutation affecting the pterin-binding site of PTS gene and review of PTS mutations in Thai patients with 6-pyruvoyltetrahydropterin synthase deficiency. *J. Inherit. Metab. Dis.* (e-pub ahead of print).
- 41 Liu, T. T. & Hsiao, K. J. Identification of a common 6-pyruvoyl-tetrahydropterin synthase mutation at codon 87 in Chinese phenylketonuria caused by tetrahydrobiopterin synthesis deficiency. *Hum. Genet.* **98**, 313–316 (1996).
- 42 Oppliger, T., Thöny, B., Kluge, C., Matasovic, A., Heizmann, C. W., Ponzone, A. *et al.* Identification of mutations causing 6-pyruvoyl-tetrahydropterin synthase deficiency in four Italian families. *Hum. Mutat.* **10**, 25–35 (1997).
- 43 Oppliger, T., Thöny, B., Nar, H., Bürgisser, D., Huber, R., Heizmann, C. W. *et al.* Structural and functional consequences of mutations in 6-pyruvoyltetrahydropterin synthase causing hyperphenylalaninemia in humans. Phosphorylation is a requirement for in vivo activity. *J. Biol. Chem.* **270**, 29498–29506 (1995).
- 44 Scherer-Oppliger, T., Matasovic, A., Laufs, S., Levy, H. L., Quackenbush, E. J., Blau, N. *et al.* Dominant negative allele (N47D) in a compound heterozygote for a variant of 6-pyruvoyltetrahydropterin synthase deficiency causing transient hyperphenylalaninemia. *Hum. Mutat.* **13**, 286–289 (1999).
- 45 Qu, Y. J., Song, F., Jin, Y. W. & Wang, H. Mutation analysis and one novel mutation detection of 6-pyruvoyl tetrahydropterin synthase gene in children with tetrahydrobiopterin deficiency. *Zhongguo Yi Xue Ke Xue Yuan Xue Bao.* **30**, 170–174 (2008).

Supplementary Information accompanies the paper on Journal of Human Genetics website (<http://www.nature.com/jhg>)