ONLINE METHODS

Cases and samples. For exome sequencing, we selected ten individuals of self-reported European, Hispanic or mixed European and Haitian ancestry with Kabuki syndrome from ten unrelated families. Phenotypic data were collected from review of medical records, phone interviews and photographs. All participants provided written consent, and the Institutional Review Boards of Seattle Children's Hospital and the University of Washington approved all studies. The clinical characteristics of the 43 individuals in the validation cohort who had been diagnosed with Kabuki syndrome have been reported previously[7]. Subjective assessment and ranking of the Kabuki phenotype was based on pictures of each subject (**Supplementary Fig. 1**) and clinical information (**Supplementary Table 1**). Informed consent was obtained for publication of each of the facial photos shown.

Exome definition, array design and target masking. We targeted all protein-coding regions as defined by RefSeq 36.3. Entries were filtered for the following: (i) CDS as the feature type, (ii) transcript name starting with "NM_" or "-", (iii) reference as the group_label, (iv) not being on an unplaced contig (for example, 17|NT_113931.1). Overlapping coordinates were collapsed for a total of 31,922,798 bases over 186,040 discontiguous regions. A single custom array (Agilent, 1M features, aCGH format) was designed to have probes over these coordinates as previously described[3], except here, the maximum melting temperature (T_m) was raised to 73 °C.

The mappable exome was also determined as previously described[3] using this RefSeq exome definition instead. After masking for 'unmappable' regions, 30,923,460 bases were left as the mappable target.

Targeted capture and massive parallel sequencing. Genomic DNA was extracted from peripheral blood lymphocytes using standard protocols. Five micrograms of DNA from each of ten individuals with Kabuki syndrome was used for construction of a shotgun sequencing library as described previously[3] using paired-end adaptors for sequencing on an Illumina Genome Analyzer II (GAII). Each shotgun library was hybridized to an array for target enrichment; this was then followed by washing, elution and additional amplification. Enriched libraries were then sequenced on a GAII to get either single-end or paired-end reads.

Read mapping and variant analysis. Reads were mapped and processed largely as previously described[3]. In brief, reads were quality recalibrated using Eland and then aligned to the reference human genome (hg18) using Maq. When reads with the same start site and orientation were filtered, paired-end reads were treated like separate single-end reads; this method is overly conservative and hence the actual coverage of the exomes is higher than reported here. Sequence calls were performed using Maq and these calls were filtered to coordinates with ≥8× coverage and consensus quality ≥ 20.

Indels affecting coding sequences were identified as previously described[3], but we used phaster instead of cross_match and Maq. Specifically, unmapped reads from Maq were aligned to the reference sequence using phaster (version 1.100122a) with the parameters -max_ins:21 -max_del:21 -gapextend_ins:-1 -gapextend_del:-1 -match_report_type:1. Reads were then filtered for those with at most two substitutions and one indel. Reads that mapped to the negative strand were reverse complemented and, together with the other filtered reads, were remapped using the same parameters to reduce ambiguity in the called indel positions. These reads were then filtered for (i) having a single indel more than 3 bp from the ends and (ii) having no other substitutions in the read. Putative indels were then called per individual if they were supported by at least two filtered reads that started from different positions. An 'indel reference' was generated as previously described[3], and all the reads from each individual were mapped back to this reference using phaster with default settings and -match_report_type:1. Indel genotypes were called as previously described[3].

To determine the novelty of the variants, sequence calls were compared against 16 individuals for whom we had previously reported exome data[2,3] and 10 EGP exomes. Annotations of variants were based on NCBI and UCSC databases using an in-house server (SeattleSeqAnnotation). Loss-of-function variants were defined as nonsense mutations (premature stop) or frame-shifting indels. For each variant, we also generated constraint scores as implemented in GERP[10].

***Post hoc* ranking of candidate genes.** Candidate genes were ranked by summation of a case score and variant score. The case score was calculated by counting the total number of Kabuki exomes in which a variant was identified at a given gene, weighted for case rank from 1 to 10. For example, the top ranked case was weighted by a factor of 10, whereas the case ranked tenth was weighted by a factor of 1. The variant score was calculated by first counting the total number of nonsense, nonsynonymous and synonymous variants across the ten Kabuki exomes and assigning a prior probability of the occurrence of each variant type per gene based upon the target of 18,918 genes. Next, for each candidate gene shared among two or more Kabuki exomes, the scores for each newly identified variant were summed across the gene. The case score and variant score were summed as the candidate gene score.

Mutation validation. Sanger sequencing of PCR amplicons from genomic DNA was used to confirm the presence and identity of variants in the candidate gene identified via exome sequencing and to screen the candidate gene in additional individuals with Kabuki syndrome.

Array comparative genomic hybridization (CGH). Samples were hybridized to commercially available whole-genome tiling arrays consisting of one million oligonucleotide probes with an average spacing of 2.6 kb throughout the genome (SurePrint G3 Human CGH Microarray 1x1M, Agilent Technologies). Twenty-one probes on this array covered *MLL2* specifically. Data were analyzed using Genomics Workbench software according to the manufacturer's instructions.

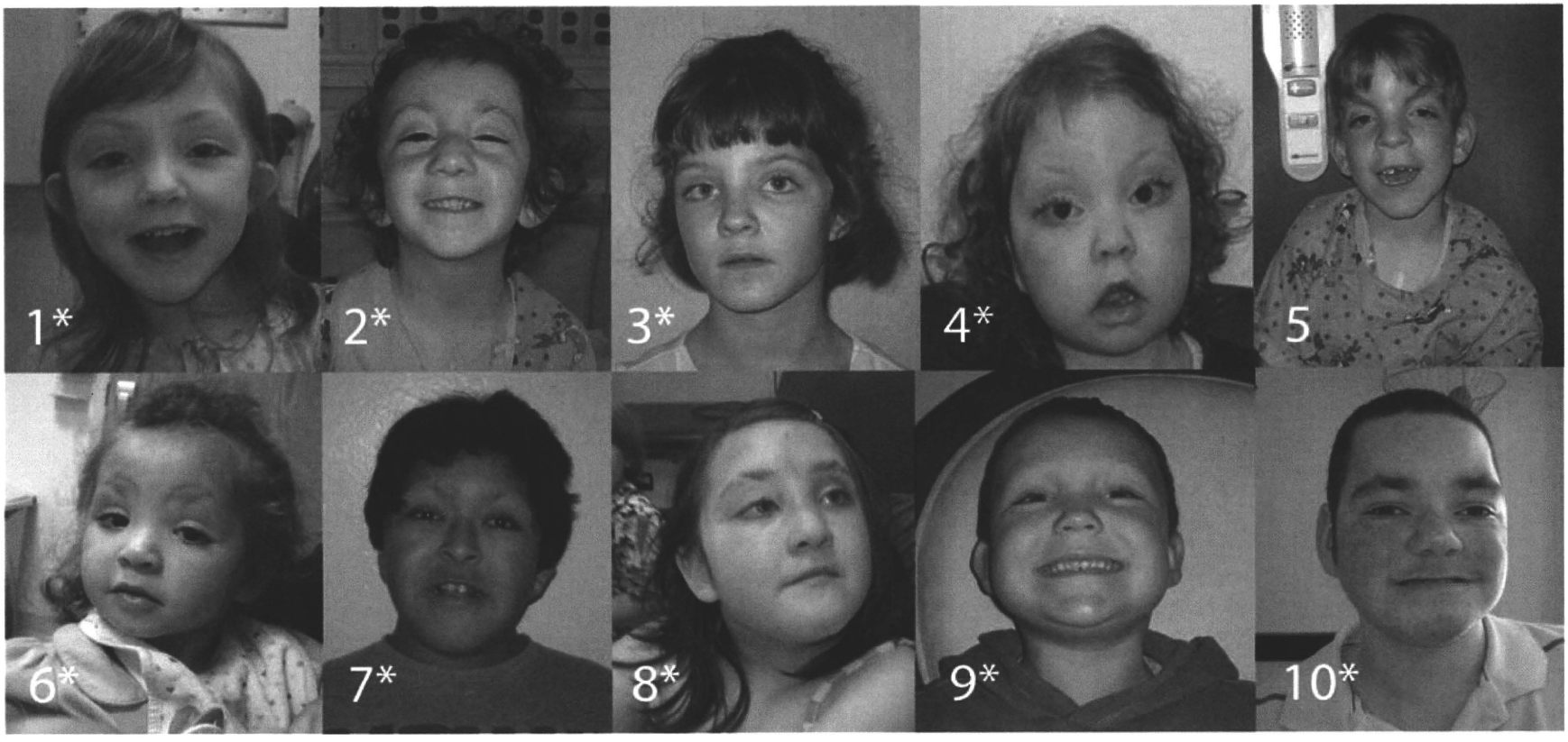
Exome sequencing identifies *MLL2* mutations as a cause of Kabuki syndrome

Sarah B. Ng[1*], Abigail W. Bigham[2*], Kati J. Buckingham[2], Mark C. Hannibal[2,3], Margaret McMillin[2], Heidi Gildersleeve[2], Anita E. Beck[2,3], Holly K. Tabor[2,3], Greg M. Cooper[1], Heather C. Mefford[2], Choli Lee[1], Emily H. Turner[1], Josh D. Smith[1], Mark J. Rieder[1], Koh-ichiro Yoshiura[4], Naomichi Matsumoto[5], Tohru Ohta[6], Norio Niikawa[6], Deborah A. Nickerson[1], Michael J. Bamshad[1,2,3†], Jay Shendure[1†]

Departments of [1]Genome Sciences and [2]Pediatrics, University of Washington, Seattle, Washington, USA. [3]Seattle Children's Hospital, Seattle, Washington, USA. [4]Department of Human Genetics, Nagasaki University Graduate School of Biomedical Sciences, Nagasaki, Japan. [5]Department of Human Genetics, Yokohama City University Graduate School of Medicine, Yokohama, Japan. [6]Research Institute of Personalized Health Sciences, Health Sciences University of Hokkaido, Hokkaido, Japan

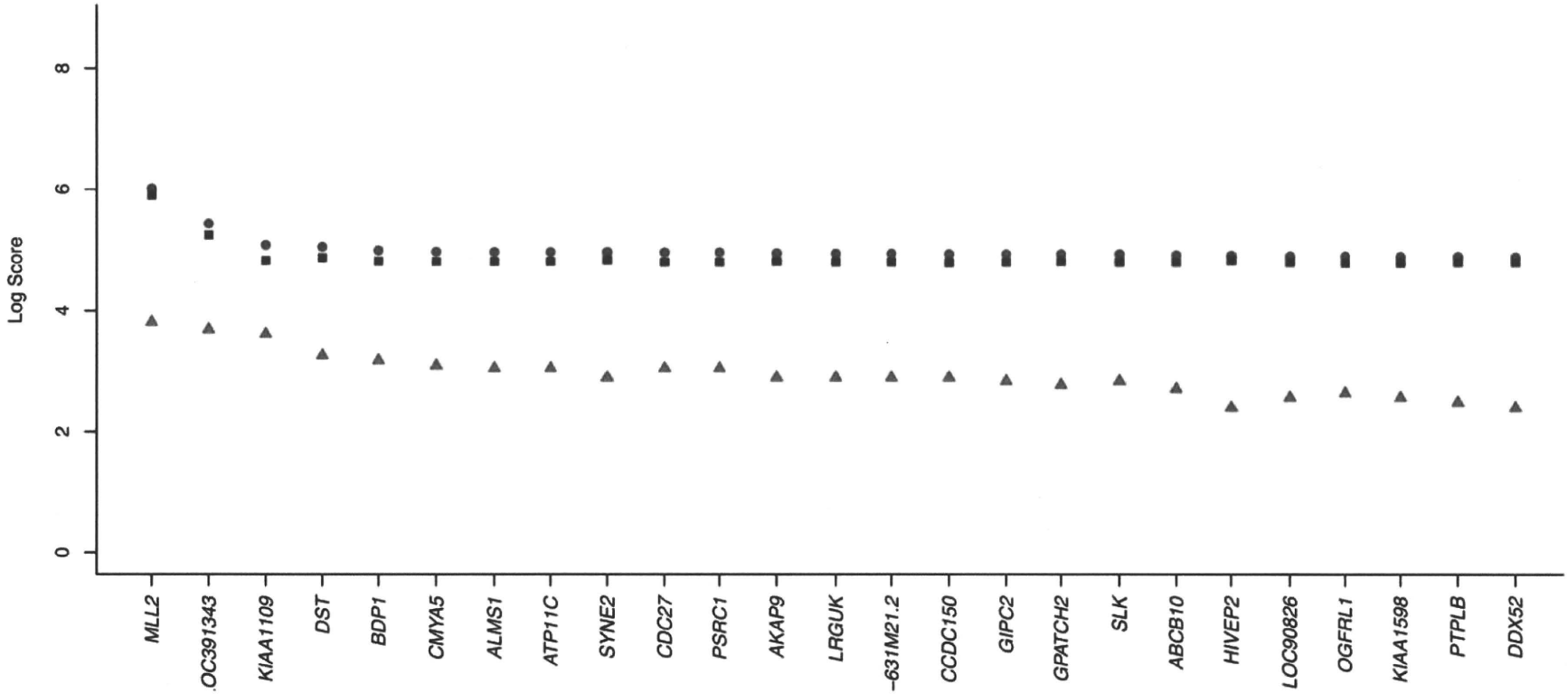
Supplementary Figures 1-3 and Supplementary Tables 1-3

Nature Genetics doi:10.1038/ng.646



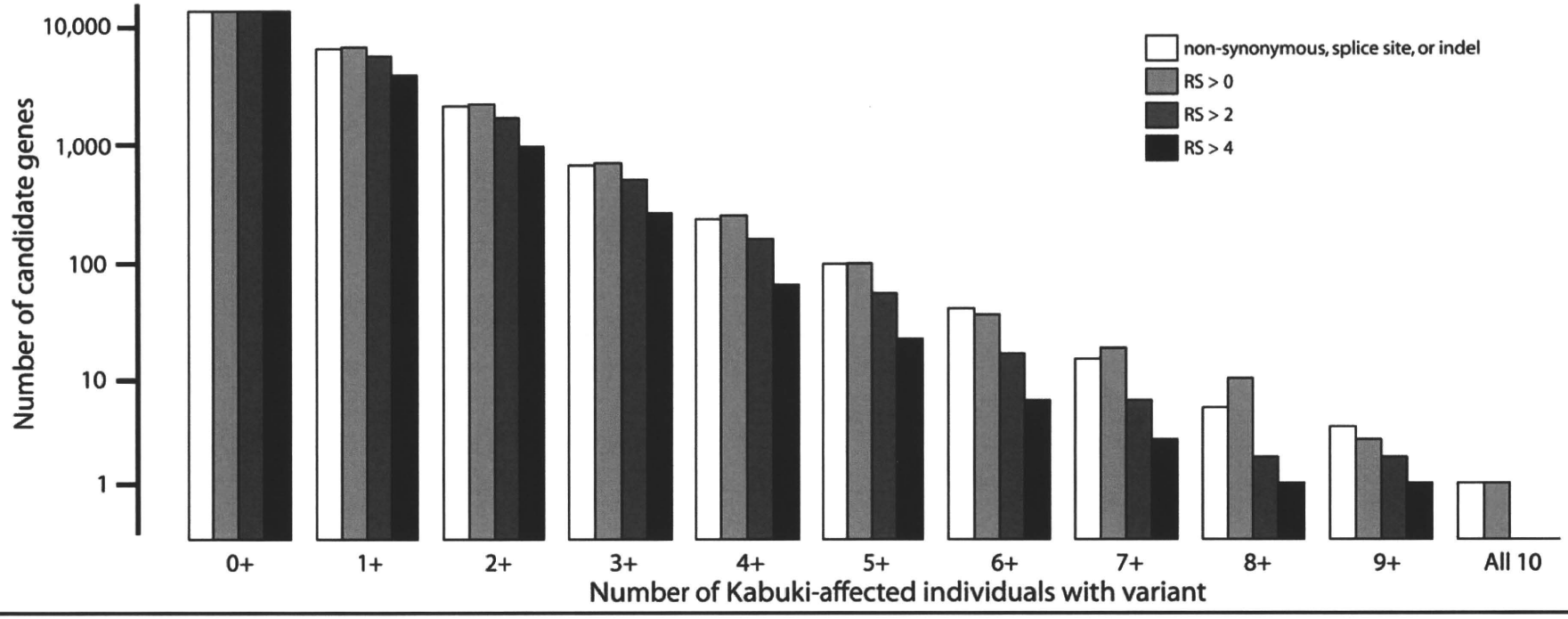
Supplementary Figure 1. Photographs of the facial characteristics used to determine the subjective ranking of Kabuki phenotypes. The phenotype ranking of the ten children with Kabuki syndrome is listed here from 1-10 based on similarity to the canonical phenotype of Kabuki syndrome. Asterisks indicate cases in which *MLL2* mutations were identified. Informed consent was obtained for publication of each of the facial photos shown.

Nature Genetics doi:10.1038/ng.646



Supplementary Figure 2. Ranking of candidate genes identified by exome sequencing based on shared variants and functional annotation. Plot of the log of the top twenty-five candidate gene scores (green circles) in 10 Kabuki cases ranked by sum of case score (red triangles) and variant score (blue squares).

Nature Genetics doi:10.1038/ng.646



	0+	1+	2+	3+	4+	5+	6+	7+	8+	9+	All 10
MLL2 rank :	3,732	3,732	1,232	399	136	47	14	6	3	NA	NA
Total (RS > 0) :	14,407	7,176	2,360	754	269	106	39	20	11	3	1

Supplementary Figure 3. Constraint scores enrich candidate gene pool for variants causing Kabuki syndrome. Number of candidate genes in which at least the given number of individuals with Kabuki syndrome has a rare variant that is functionally defined or with RS scores as indicated. Total number of candidate genes identified at RS > 0 and rank of *MLL2* among those genes are also given.

Nature Genetics doi:10.1038/ng.646

Supplementary Table 1. Clinical characteristics used to determine the subjective phenotype ranking of the 10 children with Kabuki syndrome.

	Cardiovascular	Spleen/Liver abnormality	Kidney abnormality	Kidney dysfunction	Hearing loss	Preauricular pits/tags	Cleft palate	High arched palate	Hypotonia	Developmental delay
1	ASD/VSD, aortic coarctation, bicuspid valves, dysrhythmia	np	X	np	X	X	X	X	X	X
2	aortic coarctation, bicuspid valves	-	X	X	X	X	X	X	X	X
3	ASD/VSD, aortic coarctation, bicuspid valves	-	-	-	X	np	np	X	X	X
4	VSD, aortic coarctation	X	np	np	X	-	X	X	X	X
5	ASD, VSD	np	np	np	np	np	X	X	X	X
6	np	np	np	np	-	np	np	X	X	X
7	np	-	X	-	X	np	np	X	X	X
8	np	X	X	np	np	X	X	X	X	X
9	np	np	np	np	np	np	np	np	np	X
10	ASD, VSD, dysrhythmia	X	X	X	X	np	np	X	X	X

X denotes abnormality present
np denotes abnormality not present
- denotes data not found in medical record

Nature Genetics doi:10.1038/ng.646

Supplementary Table 2. Lists of certain candidate genes as identified in Table 1 and 2.

Gene names given for cells in Table 1 where number of genes < 20

A. Subset analysis (any *x* of 10)	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
NS/SS/I	12,042	8,722	7,084	6,049	5,289	4,581	3,940	3,244	2,486	1,459
Not in dbSNP129 or 1000 genomes	7,419	2,697	1,057	488	288	192	128	88	60	34
Not in control exomes	7,827	2,865	1,025	399	184	90	50	22	7 MUC2, LIN37, MUC6, AHNAK2, UGT2B10, MUC16, MUC5B	2 MUC16, AHNAK2
Not in either	6,935	2,227	701	242	104	44	16 AHNAK2, MUC2, EVPL, LOC100128468, MUC6, MLL2, PCDHGB1, DSPP, LOC391343, MYCBP2, PCDHGA3, PCDHGA2, MUC16, PCDHGA1, LYST, PKHD1L1	6 MUC2, MUC6, DSPP, MUC16, AHNAK2, LYST	3 MUC16, AHNAK2, MUC2	1 MUC16
Is loss-of-function (nonsense / frameshift indel)	753	49	7 FLJ34443, GJB4, MLL3, MLL2, SQLE, LOC391343, ZNF598	3 MLL2, ZNF598, FLJ34443	2 MLL2, ZNF598	2 MLL2, ZNF598	1 MLL2	0	0	0

Nature Genetics doi:10.1038/ng.646

Supplementary Table 2 continued. Lists of certain candidate genes as identified in Table 1 and 2.

Gene names given for cells in Table 2 where number of genes < 20

B. Sequential analysis	1	+2	+3	+4	+5	+6	+7	+8	+9	+10
NS/SS/I	5,282	3,850	3,250	2,354	2,028	1,899	1,772	1,686	1,600	1,459
Not in dbSNP129 or 1000 genomes	687	214	145	84	63	54	42	40	39	34
Not in control exomes	675	134	50	26	13 OBSCN, LIN37, MUC16, PCDHGB1, PCDHGA2, PCDHGA3, PCDHGA1, PCDHGA4, MUC2, MUC6, UGT2B10, AHNAK2, PKHD1L1	13 OBSCN, LIN37, MUC16, PCDHGB1, PCDHGA2, PCDHGA3, PCDHGA1, PCDHGA4, MUC2, MUC6, UGT2B10, AHNAK2, PKHD1L1	8 OBSCN, LIN37, MUC16, MUC2, MUC6, UGT2B10, AHNAK2, PKHD1L1	5 LIN37, MUC16, MUC2, MUC6, AHNAK2	4 MUC16, MUC2, MUC6, AHNAK2	2 MUC16, AHNAK2
Not in either	467	89	34	18 SRRM2, ATG2B, VPS13A, FNDC1, MLL2, MUC4, UBR1, MUC16, PCDHGB1, PCDHGA2, PCDHGA3, PCDHGA1, LYST, A26C1A, MUC2, MUC6, AHNAK2, PKHD1L1	9 MUC16, PCDHGB1, PCDHGA2, PCDHGA3, PCDHGA1, MUC2, MUC6, AHNAK2, PKHD1L1	8 MUC16, PCDHGB1, PCDHGA2, PCDHGA3, PCDHGA1, MUC2, MUC6, AHNAK2	4 MUC16, MUC2, MUC6, AHNAK2	4 MUC16, MUC2, MUC6, AHNAK2	3 MUC16, MUC2, AHNAK2	1 MUC16
Is loss-of-function (nonsense / frameshift indel)	25	1 MLL2	1 MLL2	1 MLL2	0	0	0	0	0	0

Nature Genetics doi:10.1038/ng.646

Supplementary Table 3. Annotation of all *MLL2* mutations found in 53 Kabuki cases screened.

Kindred	Indiv	Exome Sequenced	Mutation	Exon	Predicted Amino Acid Change	Confirmed as *de novo*	Position[a]
1		yes	c.G15195A	48	p.W5065X	+	chr12:47706821
2		yes	c.C6010T	28	p.Q2004X	+	chr12:47722238
3		yes	c.C12697T	39	p.Q4233X	+	chr12:47712058
4		yes	c.C8488T	34	p.R2830X	-	chr12:47718918
5		yes	--	--	--	--	--
6		yes	c.11794_11797delCAAC	39	p.Q3932SfsX46	+	chr12:47712958-61
7		yes	c.T15618G	48	p.Y5206X	+	chr12:47706398
8		yes	c.3585_3586insA	11	p.P1196TfsX11	+	chr12:47730053-54
9		yes	c.C6295T	31	p.R2099X	+	chr12:47721525
10		yes	c.6595delT	31	p.Y2199IfsX65	+	chr12:47721225
11		no	c.G15326T	48	p.C5109F	+	chr12:47706690
12		no	c.G15536A	48	p.R5179H	+	chr12:47706480
13		no	c.C11149T	39	p.Q3717X	+	chr12:47713606
14		no	c.15444_15445delTT	48	p.F5149CfsX9	-	chr12:47706571-72
15		no	c.C15217T	48	p.Q5073X	-	chr12:47706799
16		no	c.C9961T	34	p.R3321X	-	chr12:47717445
17		no	c.C14710T	48	p.R4904X	-	chr12:47707306
18		no	c.5875_5891dup17	28	p.E1965GfsX88	-	chr12:47722341-57
19		no	c.G15536A	48	p.R5179H	-	chr12:47706480
20		no	c.C12703T	39	p.Q4235X	-	chr12:47712152
21		no	c.C12241T	39	p.Q4081X	-	chr12:47712514
22		no	c.C13390T	39	p.Q4464X	+	chr12:47711365
23		no	c.G15641A	48	p.R5214H	-	chr12:47706375
24		no	--	--	--	--	--
25	1	no	c.A13580T	40	p.K4527X	*	chr12:47711035
	2	no	c.A13580T	40	p.K4527X	-	chr12:47711035
26		no	c.C16501T	53	p.R5501X	-	chr12:47702113
27	1	no	--	--	--	--	--
	2	no	--	--	--	--	--
28		no	--	--	--	--	--
29		no	c.C10738T	38	p.Q3580X	-	chr12:47714119
30		no	--	--	--	--	--
31		no	c.C16360T	52	p.R5454X	-	chr12:47702382
32		no	--	--	--	--	--
33		no	--	--	--	--	--
34		no	--	--	--	--	--
35		no	c.4956_4957insG	19	p.E1654X	-	chr12:47724800-01
36		no	c.10599_10630del32	38	p.V3534QfsX11	-	chr12:47714227-58
37		no	--	--	--	--	--
38		no	c.C13606T	40	p.R4536X	-	chr12:47711008
39		no	c.G16019T	50	p.R5340L	-	chr12:47704661
40		no	c.C4843T	19	p.R1615X	-	chr12:47724914
41		no	c.C16391T	52	p.T5464M	-	chr12:47702351
42		no	--	--	--	--	--
43		no	--	--	--	--	--
44		no	--	--	--	--	--
45		no	c.C16360T	52	p.R5454X	-	chr12:47702382
46		no	--	--	--	--	--
47		no	--	--	--	--	--
48		no	--	--	--	--	--
49		no	--	--	--	--	--
50		no	c.1324delC	10	p.P442HfsX487	-	chr12:47732409
51		no	--	--	--	--	--
52	1	no	c.C16391T	52	p.T5464M	*	chr12:47702351
	2	no	c.C16391T	52	p.T5464M	-	chr12:47702351
53		no	--	--	--	--	--

-- no mutation identified
+ confirmed *de novo*
- no parental samples available
X stop codon
fs frameshift
[a] chromosomal position was determined using March 2006 assembly from UCSC (hg18)
* confirmed as inherited

Kindreds 25, 27 and 52 show dominant transmission of Kabuki syndrome from parent to child. Both affected individuals are listed here.

Nature Genetics doi:10.1038/ng.646

RESEARCH ARTICLE

AMERICAN JOURNAL OF medical genetics PART A

Two Missense Mutations of the IRF6 Gene in Two Japanese Families With Popliteal Pterygium Syndrome

Noriko Matsuzawa,[1,2]* Shinji Kondo,[3] Kazuo Shimozato,[2] Toru Nagao,[1,2] Motoi Nakano,[4] Masayoshi Tsuda,[4,5] Akiyoshi Hirano,[4] Norio Niikawa,[6,7] and Koh-ichiro Yoshiura[5,7]

[1]Department of Oral and Maxillofacial Surgery, Okazaki City Hospital, Okazaki, Japan

[2]Department of Maxillofacial Surgery, Aichi-Gakuin University School of Dentistry, Nagoya, Japan

[3]Department of Clinical Pharmacy, Nagasaki University Graduate School of Biomedical Sciences, Nagasaki, Japan

[4]Department of Plastic and Reconstructive Surgery, Nagasaki University Graduate School of Biomedical Sciences, Nagasaki, Japan

[5]Department of Human Genetics, Nagasaki University Graduate School of Biomedical Sciences, Nagasaki, Japan

[6]Research Institute of Personalized Health Sciences, Health Sciences University of Hokkaido, Tobetsu, Japan

[7]Solution Oriented Research of Science and Technology (SORST), Japan Science and Technology Agency (JST), Tokyo, Japan

Received 27 February 2009; Accepted 3 January 2010

Mutations in the interferon regulatory factor 6 gene (*IRF6*) cause either popliteal pterygium syndrome (PPS) or Van der Woude syndrome (VWS), allelic autosomal dominant orofacial clefting conditions. To further investigate the *IRF6* mutation profile in PPS, we performed mutation analysis of patients from two unrelated Japanese families with PPS and identified mutations in *IRF6*: c.251G>T (R84L) and c.1271C>T (S424L). We also found R84L, which together with previous reports on R84 mutations, provided another line of evidence that both syndromes could result from the same mutation probably under an influence of a modifier gene(s). This supports the idea that the R84 residue in the DNA binding domain of *IRF6* is a mutational hot spot for PPS. A luciferase assay of the S424L protein in the other family demonstrated that the mutation decreased the *IRF6* transcriptional activity significantly to 6% of that of the wild-type. This finding suggests that the C-terminus region of IRF6 could have an important function in phosphorylation or protein interaction. To our knowledge, this is the first report of mutations observed in Japanese PPS patients.

Key words: *IRF6*; gene mutation; clinical spectrum; transcriptional activity

How to Cite this Article:
Matsuzawa N, Kondo S, Shimozato K, Nagao T, Nakano M, Tsuda M, Hirano A, Niikawa N, Yoshiura K-i. 2010. Two missense mutations of the IRF6 gene in two Japanese families with popliteal pterygium syndrome.
Am J Med Genet Part A 152A:2262–2267.

INTRODUCTION

Popliteal pterygium syndrome (PPS, OMIM #119500) and Van der Woude syndrome (VWS, OMIM #119300) are allelic, autosomal dominant orofacial cleft syndromes. Clinical manifestations of PPS are diverse, ranging from orofacial anomalies such as lower lip pits, cleft lip and/or palate and syngnathia to skin and genital abnormalities, including webbing of the lower limbs, syndactyly, hypoplasia of the labia majora and bifid or absent scrotum [Gorlin et al., 1968; Bixler et al., 1973]. VWS is characterized mainly by cleft lip and/or palate, and lower lip pits [Schinzel and Klausler, 1986]. The incidence of PPS is estimated to be 1/300,000 [Froster-Iskenius, 1990], while VWS is more frequent and its incidence in the Finnish population some 20 years ago was about 1/33,600 [Cheney et al., 1986]. Recently, Kondo et al. [2002] discovered mutations in the interferon regulatory factor 6 gene (*IRF6*) causing either PPS or

Grant sponsor: JST, Japan (SORST); Grant sponsor: Scientific Research (Priority Area "Applied Genomics"); Grant number: 17019055; Grant sponsor: Specially Promoted Research; Grant numbers: 17019056, 17790225; Grant sponsor: Ministry of Education, Culture, Sports, Science and Technology (MEXT) of Japan; Grant sponsor: Scientific Research from the Ministry of Health, Labour and Welfare.
Noriko Matsuzawa and Shinji Kondo contributed equally to this work.
*Correspondence to:
Noriko Matsuzawa, DDD, PhD, Department of Oral and Maxillofacial Surgery, Okazaki City Hospital, Koryuji-cho 3-1, Goshoai, Okazaki, Aichi 444-8553, Japan. E-mail: mnoriko51@yahoo.co.jp
Published online 18 August 2010 in Wiley Online Library (wileyonlinelibrary.com)
DOI 10.1002/ajmg.a.33338

VWS. The gene product of *IRF6* belongs to the *IRF* family of transcriptional factors. All IRF proteins are characterized by two domains that consist of the highly conserved winged-helix DNA binding domain containing a penta-tryptophan motif and the less conserved protein binding domain termed SMAD/IRF (SMIR) or interferon association domain (IAD) [Eroshkin and Mushegian, 1999]. IRFs function not only to regulate interferon response after viral infection [Taniguchi et al., 2001], but also to regulate cell growth, differentiation, and apoptosis [Tanaka et al., 1994; Tamura et al., 1995; Holtschke et al., 1996; Harada et al., 1998; Lohoff et al., 2000; Bailey et al., 2008]. Among IRFs, the biological function of *IRF6* seems unique, but its involvement in viral infections during the early developmental stage, especially for PPS/VWS, remains unknown. Most *IRF6* mutations found in PPS are missense mutations located at R84 within the DNA binding domain or at surrounding amino acid residues [Kondo et al., 2002; Peyrard-Janvid et al., 2005], with one mutation, D430N, being reported in the SMIR domain. Mutations in VWS, however, occur across the entire coding region and consist of missense or truncation-type mutations [Kondo et al., 2002; Kayano et al., 2003; Wang et al., 2003; Matsuzawa et al., 2004, 2006; Peyrard-Janvid et al., 2005; Ye et al., 2005; Du et al., 2006; Tan et al., 2008; de Medeiros et al., 2008].

We studied two Japanese families, one of which had a patient with PPS and three other members with cleft lip with or without palate, and the other with a mother and son both affected with PPS. Here we report the results of mutation analysis in the two families and functional study of a mutated allele identified in one family. To our knowledge, this is the first report of mutations observed in Japanese PPS patients.

MATERIALS AND METHODS

Patients and Families

Family 1. Four PPS affected individuals appeared in three generations (Fig. 1A). The proband (III-2), a 1-year-old boy, had lip pits, bilateral cleft lip with palate, syngnathia, syndactyly as well as popliteal webbing, and atrophic testes (Fig. 2a–d). He had received a surgical treatment for syngnathia. His father had bilateral cleft lip with palate but no other signs of PPS. Both a paternal grandfather (I-1) and an aunt (II-2) had bifid uvula in addition to cleft palate but no lip pits. Variable expressivity was thus clinically evident in this family. Other individuals including his mother were phenotypically normal. There was no evidence that the mother had had a viral infection during early pregnancy with her son. The pedigree was constructed by physical examinations of the individuals concerned and by interviewing the parents of the proband.

Family 2. Two PPS affected members were present in two successive generations (Fig. 1B). The male proband (III-1) showed cleft palate, syngnathia, unclear scrotum, and syndactyly as well as mild popliteal webbing (Fig. 2e–g). His mother (II-1) was reported previously by Hirano et al. [1994], and had cleft lip and palate, syngnathia, syndactyly and hypoplasia of the labia majora (Fig. 2h). The proband had received an operation for the syngnathia. Other members of this family were phenotypically normal. The pedigree was constructed by physical examination of the individuals concerned and by interviewing a family member.

Blood samples were collected under informed consent from the proband and his parents in Family 1 and from the

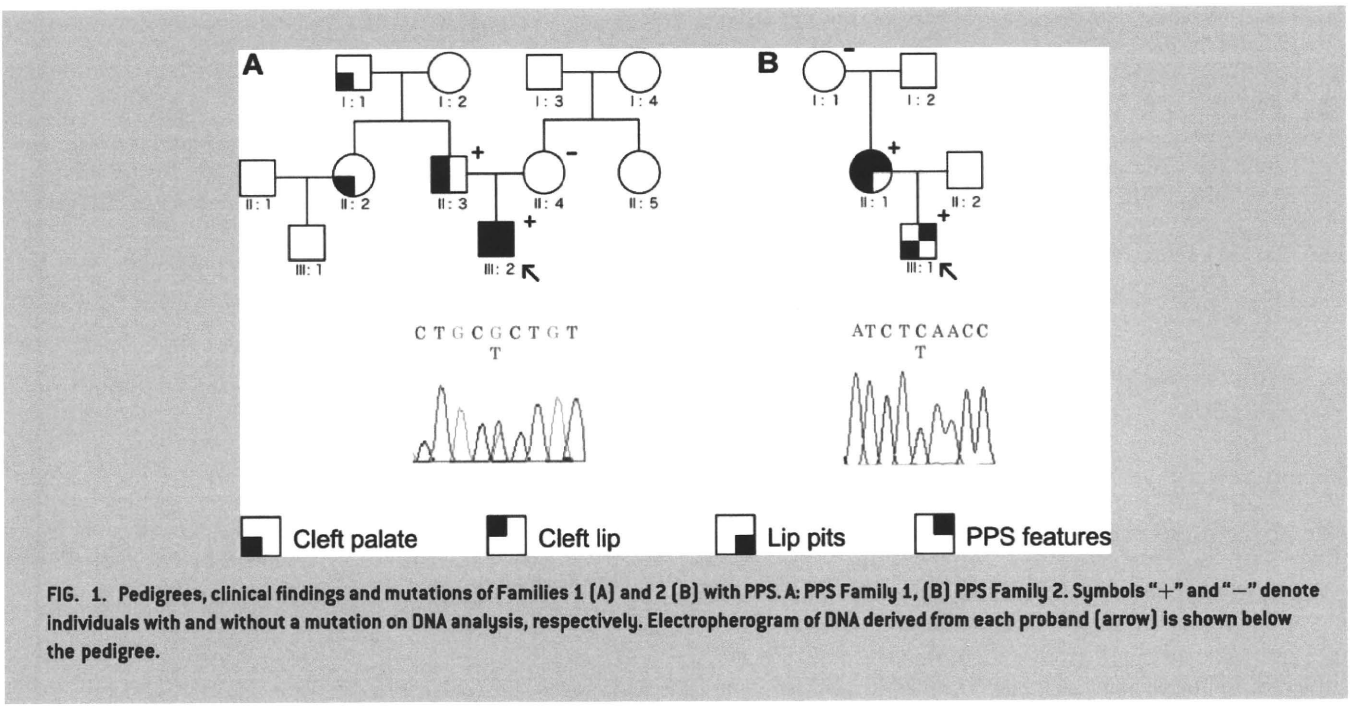


FIG. 1. Pedigrees, clinical findings and mutations of Families 1 (A) and 2 (B) with PPS. A: PPS Family 1, (B) PPS Family 2. Symbols "+" and "−" denote individuals with and without a mutation on DNA analysis, respectively. Electropherogram of DNA derived from each proband (arrow) is shown below the pedigree.

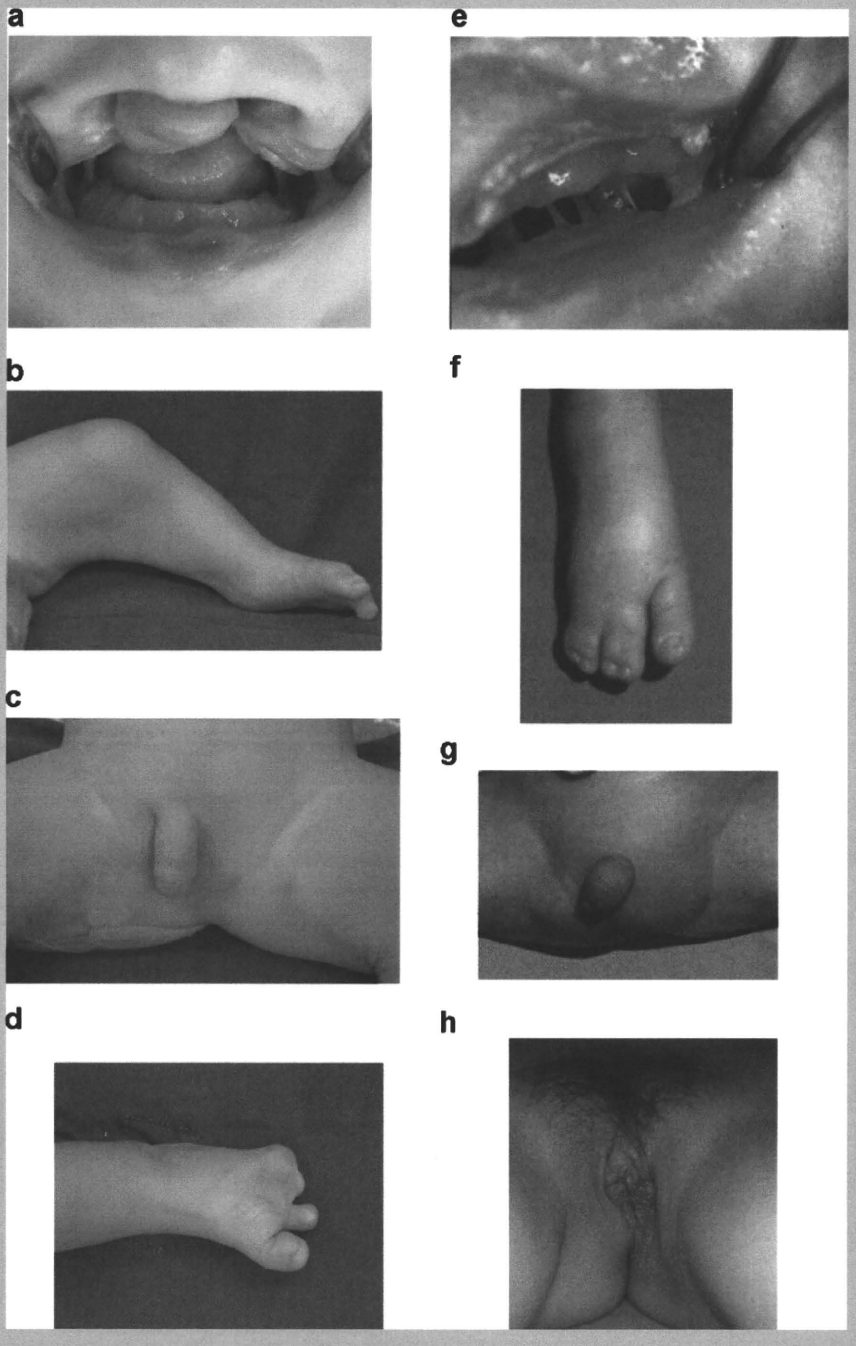


FIG. 2. Phenotypic appearance of the proband (a–d) from Family 1, and the proband (e–g) and his mother (h) from Family 2. a: Bilateral cleft lip and palate, lower lip pits, syngnathia; (b) popliteal web; (c) absent scrotum; (d) syndactyly; (e) syngnathia; (f) syndactyly; (g) unclear scrotum; (h) hypoplastic labia majora of mother.

proband, his parents and maternal grandmother in Family 2. All experimental procedures were approved by the Committees for the Ethical Issues on Human Genome and Gene Analysis in Aichi-Gakuin University and in Nagasaki University.

IRF6 Mutation Search

Screening for mutations in *IRF6* was performed in individuals from the two families. All exons and exon–intron boundaries of *IRF6* were amplified by PCR using primer pairs designed from the

genomic sequence. PCR was performed in a 10 μl mixture containing 5 ng of genomic DNA, 1 μM each primer, 200 μM each dNTPs, 0.3 U TaKaRa ExTaq HS version (TaKaRa, Kyoto, Japan), and 10× PCR buffer supplied by TaKaRa. PCR conditions were as follows; initial incubation at 94°C for 2 min followed by 35 cycles of denaturation at 94°C for 30 sec, annealing at 60°C for 30 sec, elongation at 72°C for 30 sec, and final elongation at 72°C for 7 min. PCR products were treated with ExoSAP-IT (GE Healthcare, Buckinghamshire, England) following the supplier's instruction manual, and then sequenced directly using BigDye Terminator ver.3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, CA). Sequenced samples were purified with Sephadex G-50 (GE Healthcare) and run on an automated sequencer Model 3100 (Applied Biosystems). Sequence electropherograms were aligned using AutoAssembler software v2.1 (Applied Biosystems) to find a base alteration by visual inspection.

Construction of Plasmid Expressing *IRF6*

We generated plasmids expressing the wild-type (pS424) or mutant (pL424) *IRF6* by site-directed mutagenesis using the Genetailor Site-Directed Mutagenesis System (Invitrogen, Carlsbad, CA). To express a GAL4-IRF6 fusion protein, we fused the DNA sequence for IRF6 C-terminus portions (amino acid positions 114 to 467) to GAL4 DNA binding domain (DBD). We used pBIND-DEST expressing vector containing the GAL4-DNA binding domain in Gateway system (Invitrogen). Thus, each of the recombinant plasmids (pS424 and pL424) contained the sequence for the GAL4 -DBD and the IRF6 interaction domain with Ser or Leu at the 424th amino acid residue (Fig. 3a). All clones were verified for their integrity by sequencing, and their expressions were verified by western blot analysis with a polyclonal anti-IRF6 antibody (Active Motif, Carlsbad, CA) (data not shown).

Cell Culture, Transfection, and Luciferase Assay

We cultured 293T cells in Dulbecco's MEM (Sigma-Aldrich, St. Louis, MO) containing 10% FBS at 37°C in 5% CO_2. We transiently transfected expression plasmids (pS424 or pL424) to 293T cells with a luciferase-expression reporter plasmid driven by the GAL promoter, and a βgal-expression plasmid to normalize the transfection efficiency using Lipofectamine 2000 (Invitrogen). Twenty-four hours after transfection, soluble protein was extracted and luciferase activity was assayed by the Luciferase Assay System (Promega, Madison, WI) using Galacto-Light Plus (Applied Biosystems). Transcriptional activation by the GAL4DBD-IRF6 fusion protein was calculated as the ratio of luciferase activity in the sample to that of a positive control GAL4-DBD plasmid, after normalization to the βgal activity. Statistical analysis was carried out using StatView Version 5.0, and a P-value <0.05 was considered to be statistically significant.

RESULTS

The proband and his father from Family 1 showed a missense mutation, c.251G>T, in *IRF6*-exon 4 (Fig. 1A). This nucleotide substitution results in an amino acid change in the DNA binding domain from a polar charged basic arginine residue into a hydrophobic nonpolar leucine (R84L). No such mutation was present in either the mother or among 90 healthy Japanese individuals.

In the proband from Family 2, a missense mutation, c.1271C>T, was found in the last exon (exon 9) giving a predicted amino acid change from a polar uncharged serine to a hydrophobic nonpolar leucine (S424L) (Fig. 1B). This mutation was found in his affected mother, but neither in his unaffected maternal grandmother nor among 200 healthy Japanese individuals. The luciferase assay demonstrated that the S424L protein decreased the *IRF6* transcriptional promotion activity significantly to 6% of that of the wild-type S424 protein ($P < 0.05$) (Fig. 3b).

DISCUSSION

We have identified two mutations, R84L and S424L, in two Japanese families affected by PPS. Both mutations were hitherto undescribed, although R84H in two families and R84C in five families have previously been reported to be associated with PPS [Kondo

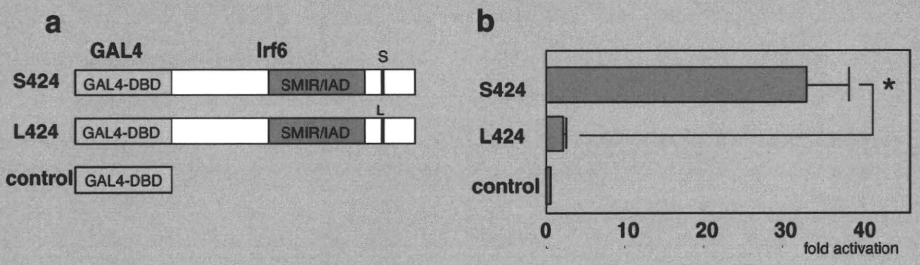


FIG. 3. Luciferase assay for transcriptional activity of mutated IRF6 protein by transient transfection of S424 into 293T cells. a: Illustration of mouse Irf6 protein (aa114–467) fused in frame to GAL4-DBD. The position of the residue 424 is indicated by S or L. GAL4-DBD, the GAL4 DNA binding domain; SMIR/IAD, SMAD/IRF or the interferon association domain. b: Luciferase value (fold activation) for each construct was measured as its ratio to a value measured for vector alone, after normalization to the co-tansfected βgal activity. The values represent the average of five experiments with standard deviation (error bars). *$P < 0.05$.

et al., 2002]. As R84 was assumed to be a critical site for DNA binding [Kondo et al., 2002], it is likely that R84L observed in Family 1 causes a change of DNA binding ability of *IRF6*, and the site R84 in the IRF6 protein is a mutational hot spot.

In Family 2, we identified another heterozygous base substitution, c.1271C>T, in the last exon of *IRF6*, and this mutation results in S424L at the C-terminus region of IRF6 protein. The S424L is a rare missense mutation found within the C-terminus domain, where only one mutation (D430N) in PPS [Kondo et al., 2002] and two mutations (P396S and R400W) in VWS [Kayano et al., 2003; Wang et al., 2003] have been reported. These findings suggest that the C-terminus region of IRF6 could have an important function, and the S424L identified in Family 2 could be associated with that function. A luciferase assay showed that the S424L mutation significantly reduces the transcriptional activity of IRF6, and two possible functions of the C-terminus are suggested. The region may act as an important domain for phosphorylation of *IRF6* via a kinase to allow correct localization as occurs in other IRFs. It has been suggested that S424 is a potential target residue of a kinase, although details are unknown [Bailey et al., 2005]. Studies on IRF3 and IRF7 have shown that phosphorylation at the C-terminus is required to unfold and dimerize them, and that the dimer is then translocated into the nucleus, where it binds with their targets and activates gene expression [Mamane et al., 1999; Sharma et al., 2003]. Alternatively, the C-terminus region could play a role in protein interaction or its regulation. This possibility is supported by a study showing an interaction of IRF6 to the mammary serine protease inhibitor (maspin) [Bailey et al., 2005]. We favor this second possibility, because a relationship between the phosphorylation state of IRF6 and subcellular localization was not supported either by immunofluorescent staining [Bailey et al., 2008] or by luciferase assay using GAL4-DBD [Little et al., 2009], and the present luciferase assay using GAL4-DBD suggests that the S424L mutation abolished transcriptional activation of mutant IRF6. However, why different mutations surrounding the SMIR domain result in either the PPS or the VWS phenotype still remains to be investigated.

The molecular mechanisms resulting in PPS and VWS and the genotype–phenotype correlations between them remain unclear. A possible explanation why mutations in the same gene cause two different syndromes is the nature of the mutations. One idea is that haploinsufficiency of *IRF6* causes VWS, while PPS develops by a dominant negative mechanism or by a situation of less than half insufficiency [Kondo et al., 2002]. According to this hypothesis, mutated IRF6 at R84 found in PPS would completely abolish the transcriptional activator function without any disturbance of localization or dimerization regulation. The R84 mutations found in Family 1 and in seven unrelated PPS families [Kondo et al., 2002] support this hypothesis. Likewise, mutations in the C-terminus region resulting in PPS, such as S424L seen in Family 2, would abolish completely the transcriptional activity. A quantitative assay reflecting intrinsic *IRF6* transcriptional activity should provide an answer. Alternatively, as the two disorders cannot always be differentiated clearly [Khan et al., 1986], they may fall into a wide clinical spectrum, resulting in little genotype–phenotype correlation.

Variable expressivity within a family is also evident, as seen in Family 1, where only the proband had typical clinical signs for PPS while other affected members had only cleft palate with or without cleft lip. Clinical severity is very different between typical PPS and VWS patients. Although these findings can be explained by an assumption that the two conditions represent each end of the spectrum, it is also plausible that various mutated IRF6 proteins would show consecutive intrinsic transcriptional activity under an influence of a modifier gene(s) and the PPS/VWS phenotypes would be distinctively categorized through thresholds according to the level of activity. All five mutations (R9W, R45Q, R84C, E349V, and P396S) reported in Japanese VWS patients [Kayano et al., 2003; Matsuzawa et al., 2004] are located at either the DNA binding domain or the SMIR domain, and there is no evidence for the existence of a Japanese-specific mutation spectrum. However, as we previously found R84C to be quite common [Matsuzawa et al., 2004], some genetic factors modifying IRF6 function might exist more frequently in the Japanese.

In conclusion, two mutations associated with PPS have been identified, c.251G>T (R84L) and c.1271C>T (S424L), one each in the two families studied. The nature of the mutations suggests R84 and S424 are biologically important sites for DNA binding ability and for transcriptional activation of IRF6, respectively. To our knowledge, this is the first report of mutations observed in Japanese PPS patients.

ACKNOWLEDGMENTS

The participation of the patients in this study is highly appreciated. We thank Ms. Y. Noguchi, K. Miyazaki, C. Hayashida, and M. Ohga for their technical assistance. SORST from JST, Japan for N. N.; Grants-in-Aid for Scientific Research (Priority Area "Applied Genomics," No. 17019055) for N. N., and for Specially Promoted Research (Nos. 17019056 and 17790225) for S. K. from the Ministry of Education, Culture, Sports, Science and Technology (MEXT) of Japan. K.Y. was supported in part by Grants-in-Aid for Scientific Research from the Ministry of Health, Labour and Welfare.

REFERENCES

Bailey CM, Khalkhali-Ellis Z, Kondo S, Margaryan NV, Seftor REB, Wheaton WM, Amir S, Pins MR, Schutte BC, Hendrix MJC. 2005. Mammary serine protease inhibitor (maspin) binds directly to interferon regulatory factor 6. J Biol Chem 280:34210–34217.

Bailey CM, Abbott DE, Margaryan NV, Khalkhali-Ellis Z, Hendrix MJC. 2008. Interferon regulatory factor 6 promotes cell cycle arrest and is regulated by the proteasome in a cell cycle-dependent manner. Mol Cell Biol 28:2235–2243.

Bixler D, Poland C, Nance WE. 1973. Phenotypic variation in the popliteal pterygium syndrome. Clin Genet 4:220–228.

Cheney ML, Cheny WR, Lejeune FE. 1986. Familial incidence of labial pits. Am J Otolaryngol 7:311–313.

de Medeiros F, Hansen L, Mawlad E, Eiberg H, Asklund C, Tommerup N, Jakobsen LP. 2008. A novel mutation in IRF6 resulting in VWS-PPS spectrum disorder with renal aplasia. Am J Med Genet Part A 146A: 1605–1608.

Du X, Tang W, Tian W, Li S, Li X, Liu L, Zheng X, Chen X, Lin Y, Tang Y. 2006. Novel IRF6 mutations in Chinese patients with Van der Woude syndrome. J Dent Res 85:937–940.

Eroshkin A, Mushegian A. 1999. Conserved transactivation domain shared by interferon regulatory factors and Smad morphogens. J Mol Med 77:403–405.

Froster-Iskenius UG. 1990. Popliteal pterygium syndrome. J Med Genet 27:320–326.

Gorlin RJ, Sedano HO, Cervenka J. 1968. Popliteal pterygium syndrome. A syndrome comprising cleft lip-palate, popliteal and intercrural pterygia, digital and genital anomalies. Pediatrics 41:503–509.

Harada H, Taniguchi T, Tanaka N. 1998. The role of interferon regulatory factors in the interferon system and cell growth control. Biochimie 80:641–650.

Hirano A, Iio Y, Murakami R, Fujii T. 1994. Recurrent trismus: Twenty-year follow-up result. Cleft Palate Craniofac J 31:309–312.

Holtschke T, Lohler J, Kanno Y, Fehr T, Giese N, Rosenbauer F, Lou J, Knobeloch KP, Gabriele L, Waring JF, Bachmann MF, Zinkernagel RM, Morse HC III, Ozato K, Horak I. 1996. Immunodeficiency and chronic myelogenous leukemia-like syndrome in mice with a targeted mutation of the ICSBP gene. Cell 87:307–317.

Kayano S, Kure S, Suzuki Y, Kanno K, Aoki Y, Kondo S, Schutte BC, Murray JC, Yamada A, Matsubara Y. 2003. Novel IRF6 mutations in Japanese patients with Van der Woude syndrome: Two missense mutations (R45Q and P396S) and a 17-kb deletion. J Hum Genet 48:622–628.

Khan SN, Hufnagle KG, Pool R. 1986. Intrafamilial variability of popliteal pterygium syndrome: A family description. Cleft Palate 23:233–236.

Kondo S, Schutte BC, Richardson RJ, Bjork BC, Knight AS, Watanabe Y, Howard E, de Lima RL, Daack-Hirsch S, Sander A, McDonald-McGinn DM, Zackai EH, Lammer EJ, Aylsworth AS, Ardinger HH, Lidral AC, Pober BR, Moreno L, Arcos-Burgos M, Valencia C, Houdayer C, Bahuau M, Moretti-Ferreira D, Richieri-Costa A, Dixon MJ, Murray JC. 2002. Mutations in IRF6 cause Van der Woude and popliteal pterygium syndromes. Nat Genet 32:285–289.

Little HJ, Rorick NK, Su LI, Baldock C, Malhotra S, Jowitt T, Gakhar L, Subramanian R, Schutte BC, Dixon MJ, Shore P. 2009. Missense mutation that cause Van der Woude syndrome and popliteal pterygium syndrome affect the DNA-binding and transcriptional activation functions of IRF6. Hum Mol Genet 18:535–545.

Lohoff M, Duncan GS, Ferrick D, Mittrücker HW, Bischof S, Prechtl S, Röllinghoff M, Schmitt E, Pahl A, Mak TW. 2000. Deficiency in the transcription factor interferon regulatory factor (IRF)-2 leads to severely compromised development of natural killer and T helper type 1 cells. J Exp Med 192:325–336.

Mamane Y, Heylbroeck C, Genin P, Algarté M, Servant MJ, LePage C, DeLuca C, Kwon H, Lin R, Hiscott J. 1999. Interferon regulatory factors: The next generation. Gene 237:1–14.

Matsuzawa N, Yoshiura K, Machida J, Nakamura T, Niimi T, Furukawa H, Toyoda T, Natsume N, Shimozato K, Niikawa N. 2004. Two missense mutations in the IRF6 gene in two Japanese families with Van der Woude syndrome. Oral Surg Oral Med Oral Pathol Oral Radiol Endod 98: 414–417.

Matsuzawa N, Shimozato K, Natsume N, Niikawa N, Yoshiura K. 2006. A novel missense mutation in Van der Woude syndrome: Usefulness of fingernail DNA for genetic analysis. J Dent Res 85:1143–1146.

Peyrard-Janvid M, Pegelow M, Koillinen H, Larsson C, Fransson I, Rautio J, Hukki J, Larson O, Karsten AL, Kere J. 2005. Novel and de novo mutations of the IRF6 gene detected in patients with Van der Woude or popliteal pterygium syndrome. Eur J Hum Gene 13:1261–1267.

Schinzel A, Klausler M. 1986. The Van der Woude syndrome (dominantly inherited lip pits and cleft). J Med Genet 23:291–294.

Sharma S, tenOever BR, Grandvaux N, Zhou GP, Lin R, Hiscott J. 2003. Triggering the interferon antiviral response through an IKK-related pathway. Science 300:1148–1151.

Tamura T, Ishihara M, Lamphier MS, Tanaka N, Oishi I, Aizawa S, Matsuyama T, Mak TW, Taki S, Taniguchi T. 1995. An IRF-1-dependent pathway of DNA damage-induced apoptosis in mitogen-activated T lymphocytes. Nature 376:596–599.

Tan EC, Lim EC, Yap SH, Lee ST, Cheng J, Por YC, Yeow V. 2008. Identification of IRF6 gene variants in three families with Van der Woude syndrome. Int J Mol Med 21:747–751.

Tanaka N, Ishihara M, Kitagawa M, Harada H, Kimura T, Matsuyama T, Lamphier MS, Aizawa S, Mak TW, Taniguchi T. 1994. Cellular commitment to oncogene-induced transformation or apoptosis is dependent on the transcription factor IRF-1. Cell 77:829–839.

Taniguchi T, Ogasawara K, Takaoka A, Tanaka N. 2001. IRF family of transcription factors as regulators of host defense. Annu Rev Immunol 19:623–655.

Wang X, Liu J, Zhang H, Xiao M, Li J, Yang C, Lin X, Wu Z, Hu L, Kong X. 2003. Novel mutations in the IRF6 gene for Van der Woude syndrome. Hum Genet 113:382–386.

Ye XO, Jin HX, Shi LS, Fan MW, Song GT, Fan HL, Bian Z. 2005. Identification of novel mutations of IRF6 gene in Chinese families with Van der Woude syndrome. Int J Mol Med 16:851–856.

PRENATAL DIAGNOSIS
Prenat Diagn 2010; **30**: 849–861.
Published online 27 July 2010 in Wiley Online Library
(wileyonlinelibrary.com) **DOI:** 10.1002/pd.2570

The possibility of microarray-based analysis using cell-free placental mRNA in maternal plasma

Kiyonori Miura[1]*, Shoko Miura[1], Kentaro Yamasaki[1], Takako Shimada[1], Akira Kinoshita[2], Norio Niikawa[3], Koh-ichiro Yoshiura[2] and Hideaki Masuzaki[1]

[1]*Department of Obstetrics and Gynecology, Nagasaki University Graduate School of Biomedical Sciences, Nagasaki, Japan*
[2]*Department of Human Genetics, Nagasaki University Graduate School of Biomedical Sciences, Nagasaki, Japan*
[3]*Research Institute of Personalized Health Sciences, Health Sciences University of Hokkaido, Hokkaido, Japan*

Objective The purpose of this study is to investigate a possibility of overall assessment of cell-free (CF) placental mRNAs in maternal plasma.

Methods First, placenta-predominantly expressed transcripts were selected by the analysis of GeneChip using three sets of placental tissues and corresponding maternal blood cells. Subsequently, a custom cDNA array panel of placenta-predominantly expressed transcripts was designed and used to compare the RNA profiles of maternal plasma collected from 12 preeclamptic and 12 uncomplicated pregnancies. Scatter plots for the signal intensities of the comparative cDNA hybridization revealed either unchanged or aberrant patterns.

Results We selected top 50 placenta-predominantly expressed transcripts that were >2500 times higher in placental tissues than in corresponding whole blood samples. A custom cDNA array analysis detected the aberrant pattern in five preeclamptic women with severe hypertension but not in seven preeclamptic women with mild hypertension ($P < 0.05$, Fisher's direct method). The aberrant pattern of above RNA transcripts in maternal plasma was validated by quantitative real-time reverse transcription-polymerase chain reaction. The mean (range) value of coefficient of variations in this custom array quantification was 9.4% (3.0–16.2%).

Conclusion Our custom cDNA array is expected to be useful for overall assessment of CF placental mRNAs in maternal plasma in a single experiment. Copyright © 2010 John Wiley & Sons, Ltd.

KEY WORDS: microarray-based analysis; cell-free placental mRNA; maternal plasma; placental status; preeclampsia; noninvasive diagnosis

INTRODUCTION

Assessment of the fetal/placental status during pregnancy is currently carried out by fetal cardiotocography, ultrasonography and/or tests using biological marker molecules, but they often produce false-positive results (Bobby, 2003). To obtain direct information regarding the fetal/placental status, conventional prenatal diagnostic procedures such as amniocentesis, cordocentesis or fetal scalp blood sampling may remain useful. However, these invasive procedures always involve the risk of serious complications, for example fetal loss, rupture of the membrane and infections. Therefore, noninvasive detection of fetus or placenta-derived molecular markers in pregnant women is desirable for accurate monitoring of the fetal/placental status. The recent discovery of cell-free (CF) placental mRNA in maternal plasma has provided possibilities for exploring placental dysfunction (Lo and Chiu, 2007; Maron and Bianchi, 2007). CF placental transcripts can be detected in maternal plasma by week 4 of gestation and have a median half-life of 14 min (Chiu *et al.*, 2006). The pregnancy specificity of CF placental mRNA has been demonstrated by its rapid clearance from the maternal plasma after delivery. Therefore, its detection in the maternal circulation appears to be a promising approach for the development of sex- and polymorphism-independent fetal/placental molecular markers for prenatal gene expression profiling. CF placental mRNA in maternal plasma is measurable by quantitative real-time reverse transcription-polymerase chain reaction (RT-PCR) (Miura *et al.*, 2008), and placental transcripts are more readily detectable in plasma than in whole blood (Heung *et al.*, 2009).

The placental functions during pregnancy are regulated by a large variety of genes. Preeclampsia is a serious complication of pregnancy, and its pathogenesis is known to be associated with multiple factors, such as abnormal placentation, reduced placental perfusion, endothelial cell dysfunction and systemic vasospasm. The plasma concentrations of circulating CF placental mRNA including transcripts for corticotrophin-releasing hormone, placenta-specific gene 1 and selectin P were reported to be higher for pregnant women with preeclampsia than for normal pregnant women (Lo and Chiu, 2007; Maron and Bianchi, 2007; Purwosunu *et al.*, 2007). These findings suggest that certain different kinds of CF placental mRNA in maternal plasma can be used as markers for preeclampsia. However, overall assessment of CF placental mRNA by a single quantitative

*Correspondence to: Dr Kiyonori Miura, Department of Obstetrics and Gynecology, Nagasaki University Graduate School of Biomedical Sciences, 1-7-1 Sakamoto, Nagasaki 852-8501, Japan.
E-mail: kiyonori@nagasaki-u.ac.jp

Received: 23 February 2010
Revised: 13 April 2010
Accepted: 10 May 2010
Published online: 27 July 2010

real-time RT-PCR analysis is difficult because of the limitation of the blood sampling volume. To resolve this technical issue, the development of a high-throughput microarray-based approach is promising. In addition, microarray comparative genomic hybridization analysis of CF fetal DNA in amniotic fluid already provides rapid screening for chromosome abnormalities with copy-number changes (Larrabee *et al.*, 2004; Miura *et al.*, 2006), though amniotic fluid cannot be obtained noninvasively. The application of a microarray-based method using CF placental mRNA in maternal plasma is feasible and may allow a noninvasive overall assessment of the placental status in a single experiment.

In this study, we initially identified placenta-predominantly expressed transcripts in maternal blood by comparisons between their expression levels in the placenta and maternal blood (Tsui *et al.*, 2004). Subsequently, we generated a custom cDNA microarray panel composed of the cDNAs of these genes. Finally, we adopted a custom array-based comparative cDNA hybridization using CF mRNA in maternal plasma to establish an overall assessment of CF placental mRNA levels in pregnant women.

MATERIALS AND METHODS

Subjects and sample collection

The study subjects comprised 12 pregnant women with preeclampsia and 12 normal pregnant women whose gestational weeks were matched with those of the preeclamptic women. CF plasma samples (12 mL) were prepared from maternal blood by a double centrifugation method. All these women attended the Department of Obstetrics and Gynecology at the Nagasaki University Hospital. Preeclampsia was defined as gestational hypertension (systolic pressure $\geq$140 mm Hg and/or diastolic blood pressure $\geq$90 mm Hg on at least two occasions during measurements after 20 weeks of gestation) with proteinuria ($\geq$0.3 g/day). Severe preeclampsia was defined by the presence of one or more of the following findings: severe gestational hypertension (systolic pressure $\geq$160 mm Hg and/or diastolic blood pressure $\geq$110 mm Hg on two occasions at least 6 h apart while the patient was on bed rest); severe proteinuria ($\geq$5 g of protein in a 24-h urine specimen, or 3 or greater in two random urine samples collected at least 4 h apart). Women in whom the onset of preeclampsia was before 32 weeks of gestation were defined as having the early-onset type, while those with onset after 32 weeks of gestation were defined as having the late-onset type.

Placental tissue samples were obtained from three normal pregnant women immediately after termination of pregnancy during the first (7–12 weeks of pregnancy), second (18–21 weeks) or third (37–39 weeks) trimesters (Figure 1), immediately placed in RNAlater™ (Ambion, Austin, TX) and stored at −80 °C until RNA extraction. Blood samples (6 mL) were collected from the women into PAXgene™ blood RNA tubes (PreAnalytiX, Hombrechtikon, Switzerland) before the termination of pregnancy at each trimester. All the samples

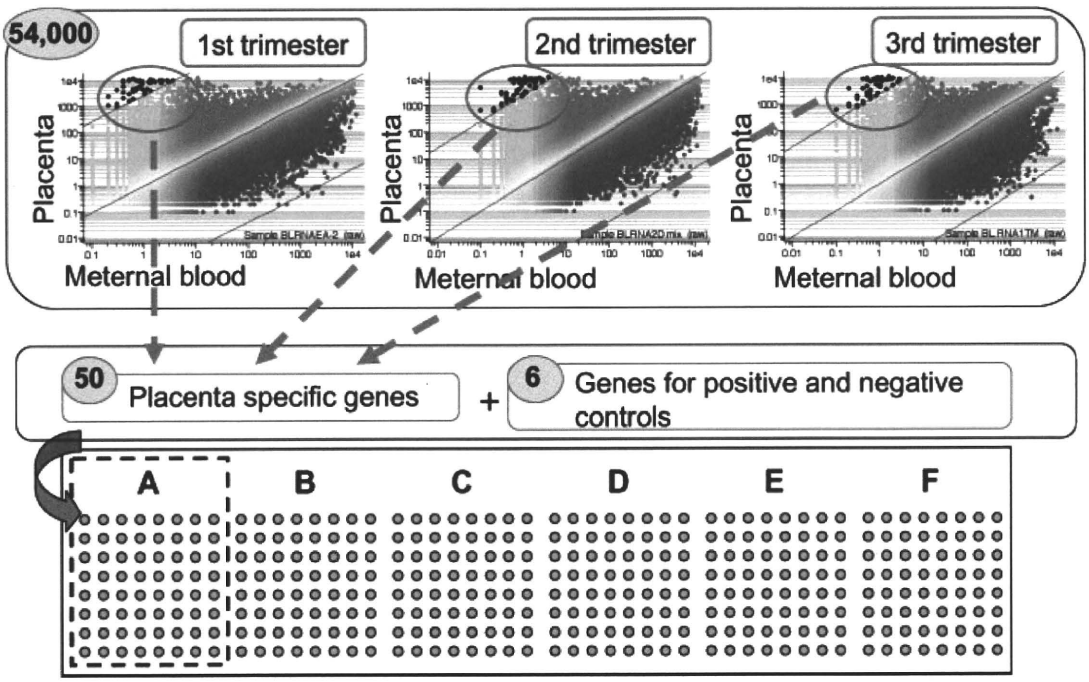


Figure 1—Outline of the strategy used for the development of a custom microarray panel of placenta-specific genes in maternal blood. Upper panel: gene expression profiles of three paired samples of placenta tissues and corresponding whole maternal blood (at the first, second and third trimesters). Black dots indicate placenta-specific transcripts exhibiting signal intensities that are >2500 times higher in the placenta in every trimester compared with the corresponding whole maternal blood samples. Middle panel: collection of 50 placenta-specific genes plus 6 positive and negative control genes. Lower panel: a custom cDNA microarray panel where 6 blocks containing a total of 56 gene-specific sequences were robotically spotted onto a single slide. The 50 placenta-specific genes were spotted in the order of their high expression levels in the placenta

Prenat Diagn 2010; **30**: 849–861.
DOI: 10.1002/pd

were obtained after receiving written informed consent, and the study protocol was approved by the Institutional Review Board for Ethical, Legal and Social Issues at the Nagasaki University.

Identification of placenta-predominantly expressed transcripts in the maternal circulation

Total RNA from the placental tissue samples and the peripheral blood were extracted as described, respectively, by Tsui *et al.* (2004) and Okazaki *et al.* (2007). The same amount of RNA from the placental or corresponding maternal blood sample was subjected to GeneChip® analysis (Affymetrix, Santa Clara, CA). Because the expression pattern of placental transcripts is changing during pregnancy, three sets (placental tissue and corresponding maternal blood sample) at each trimester were analyzed. For each sample, the extracted RNA (1 μg) was labeled by the one-cycle target labeling method, and biotinylated cDNA (15 μg) was hybridized to a GeneChip® Human Genome U133 Plus 2.0 (Affymetrix) according to the manufacturer's instructions. After the hybridization, the arrays were washed and stained in a GeneChip Fluidics Station 450 (Affymetrix). The chips were scanned with a GeneChip Scanner 3000 (Affymetrix) and analyzed using GeneChip Microarray Suite 5.0 (Affymetrix). The raw intensity data were analyzed with the GeneChip Operating Software (Affymetrix), and data mining was performed with the GeneSpring software (Agilent Technologies, Palo Alto, CA). The corrected signal intensities for the placental tissues and blood samples were expressed as scatter plots (Figure 1). Three sets of placental tissue and maternal blood cell samples across the three trimesters were used to look for the placenta-predominantly expressed genes that were detected in one or more out of the three profiles. The top 50 placenta-predominantly expressed transcripts exhibiting raw signal intensities that were >2500 times higher in the placental tissues than in the corresponding whole blood samples were selected according to a previously reported strategy for systematic identification of pregnancy-specific placental mRNA markers expressed in the maternal circulation (Tsui *et al.*, 2004).

Development of a custom microarray panel of placenta-predominantly expressed genes

Preparation of probe cDNAs

The 50 placenta-predominantly expressed transcripts (Nos. 1–50) and six control transcripts (Nos. 51–56) were selected as probe cDNAs to be spotted onto a custom microarray panel (Table 1, Figure 1). The control transcripts included five placenta-nonspecific genes as positive controls (Nos. 51–55) and the plasmid sequence of the pBluescriptII SK(−) vector as a negative control (No. 56). The 50 primer sets for the placenta-specific genes and 5 primer sets for the placenta-nonspecific genes (TaqMan® Gene Expression Assays) were purchased from Applied Biosystems (Foster City, CA, USA) (each assay ID is listed in Table 1). RT-PCR amplifications of the placenta mRNA samples were carried out, and the PCR products were sequenced to confirm the accuracy of each gene-specific primer set. As the primer sets for *CSH1* and *HERV-FRD* failed to amplify their exact cDNA sequences, custom primers were made for these genes as described previously (Ng *et al.*, 2003b; Okahara *et al.*, 2004). Each placental cDNA aliquot (2 μL) was subjected to 30 cycles of PCR in a total volume of 20 μL. The PCR product was cloned into the TOPO II vector (Invitrogen (Carlsbad, CA, USA)), and the plasmid DNA was extracted. Several clones for each transcript were sequenced using an ABI3100 to confirm the sequence integrity, and a total of 55 clones, 1 for each from the placenta-specific and nonspecific transcripts, were selected.

As probe DNA resources for the microarray panel, the sequences integrated into the TOPO II vector were amplified with primers that were modified to have sequences for an amino residue ($-NH_2$) at their 5′ ends. The primer sequences were as follows: topo-forward: 5′-agtgtgctggaattcgccct-3′; topo-reverse: 5′-gatatctgcagaattcgccct-3′. After purification using a Microcon YM-10/30/100 (Millipore Corporation, Billerica, MA), the concentration of the collected DNA was 138–580 ng/μL and each volume was 50–100 μL. The PCR amplifications involved 30 cycles of 94 °C for 30 s, 60 °C for 30 s and 72 °C for 30 s in a 100-μL mixture containing 1 ng of cDNA, 10 pM of primers (topo-forward and topo-reverse), 250 μM dNTP, 0.5 U of Ex Taq polymerase (TaKaRa Bio Inc., Tokyo, Japan) and 10× PCR buffer (TaKaRa Bio Inc.). PCR amplification was performed four times for each cDNA clone.

DNA spotting

The PCR products were dissolved in distilled water, and an equal volume of spotting solution DSP0050 (Matsunami, Osaka, Japan) was added. The final concentration of the PCR products was 0.15 μg/μL. The resulting DNA samples (0.15 μg/μL at 200 pL/spot) were robotically spotted using an inkjet printing technique (NGK, Nagoya, Japan) in six blocks onto CodeLink™ activated slides (Amersham Biosciences, Piscataway, NJ). The design of the custom cDNA microarray is shown in Figure 1.

Microarray-based comparative cDNA hybridization between samples from preeclamptic women and control pregnant women

cDNA preparation from the women and T7-based CF mRNA amplification

CF mRNA was extracted as described previously (Zhong *et al.*, 2008) and amplified by a T7-based RNA amplification method (Abe *et al.*, 2003), because the concentration of CF mRNA in maternal plasma was too low for direct use in a microarray analysis (Figure 2). A total of 56 specific primers [RNA amplification primers (RAPs)]

Table 1—Fifty genes highly expressed in the placenta compared with the maternal blood

	Gene symbol (name)	GenBank accession number	Assay ID	Chromosome localization	Gene expression[a]	Gene function	CV value[b] (custom array) (%)	CV value (real-time RT-PCR) (%)
1	LUM (lumican)	NM_002345.3	Hs00158940_m1	12q21.3-q22	P	A member of the small leucine-rich proteoglycan family	12.0	1.7
2	RAI14 (retinoic acid induced 14)	NM_015577.1	Hs00210238_m1	5p13.3-p13.2	P	Retinoic acid regulated gene	14.9	3.3
3	CDH1 [cadherin 1, type 1, E-cadherin (epithelial)]	NM_004360.2	Hs00170423_m1	16q22.1	P	Calcium ion-dependent cell adhesion molecule	7.8	5.8
4	CSRP2 (cysteine and glycine-rich protein 2)	NM_001321.1	Hs00426717_m1	12q21.1	P	Member of the CRP (cysteine- and glycine-rich protein) family of LIM (Lin/Isl/Mec) domain proteins	13.2	4.8
5	ERVWE1 [endogenous retroviral family W, env(C7), member 1, syncytin 1]	NM_014590.3	Hs00205893_m1	7q21-q22	P	Part of an HERV provirus, which is expressed in the placental syncytiotrophoblast and is involved in fusion of the cytotrophoblast cells to form the syncytial layer of the placenta	10.2	3.1
6	INHBA (inhibin, beta A)	NM_002192.2	Hs00170103_m1	7p15-p13	P	Growth/differentiation hormone	17.2	4.7
7	PSG5 (pregnancy-specific beta-1-glycoprotein 5)	NM_002781.2	Hs00818332_m1	19q13.2	P	A group of molecules that are mainly produced by the placental syncytiotrophoblasts during pregnancy	12.1	10.6
8	TFPI (tissue factor pathway inhibitor)	NM_001032281.2	Hs00196731_m1	2q31-q32.1	P	A protease inhibitor that regulates the tissue factor-dependent pathway of blood coagulation	6.9	6.0
9	INSL4 (insulin-like 4)	NM_002195.1	Hs00171411_m1	9q24	P	A member of the insulin superfamily	9.1	4.6
10	LEP (leptin)	NM_000230.1	Hs00174877_m1	7q31.3	P	Secreted by adipocytes and by placenta	13.0	9.1
11	TFPI2 (tissue factor pathway inhibitor 2)	NM_006528.2	Hs00197918_m1	7q22	P	A placental glycoprotein that inhibits plasmin, trypsin and thrombin	7.3	8.9
12	GH1 (growth hormone 1, transcript variant 1)	NM_000515.3	Hs00236859_m1	17q24.2	P	A member of the somatotrophin/prolactin family of hormones which play an important role in growth control	13.1	2.6
13	ADAM12 (a disintegrin and metalloproteinase domain 12)	NM_003474	Hs01106104_m1	10q26.3	P	Candidate regulator of trophoblast fusion	12.7	5.2
14	ANGPT2 (angiopoietin 2)	NM_001147.1	Hs00169867_m1	8p23.1	P	An antagonist of angiopoietin 1 (ANGPT1) and endothelial TEK (endothelial-specific receptor tyrosine kinase) tyrosine kinase (TIE-2 (tunica internal endothelial cell kinase 2), TEK)	10.9	6.1
15	GH1 (growth hormone 1), transcript variant 4	NM_022561.2	Hs00737955_m1	17q24.2	P	A member of the somatotrophin/prolactin family of hormones that play an important role in growth control	11.4	7.6
16	PSG9 (pregnancy-specific beta-1-glycoprotein 9)	NM_002784.2	Hs00358192_m1	19q13.2	P	A group of molecules that are mainly produced by the placental syncytiotrophoblasts during pregnancy	16.2	7.7
17	CGA (glycoprotein hormones, alpha polypeptide)	NM_000735.2	Hs00174938_m1	6q12-q21	P	The alpha subunit and belongs to the glycoprotein hormones alpha chain family	12.0	5.2
18	KISS1 (KiSS-1 metastasis suppressor)	NM_002256.2	Hs00158486_m1	1q32	P	A putative role in the regulation of events downstream of cell-matrix adhesion	8.9	2.8

Prenat Diagn 2010; **30**: 849–861.
DOI: 10.1002/pd

Table 1—(*Continued*)

	Gene symbol (name)	GenBank accession number	Assay ID	Chromosome localization	Gene expression[a]	Gene function	CV value[b] (custom array) (%)	CV value (real-time RT-PCR) (%)
19	CAPN6 (calpain 6)	NM_014289.2	Hs00560073_m1	Xq23	P	Ubiquitous, well-conserved family of calcium-dependent, cysteine proteases	12.5	6.1
20	PSG6 (pregnancy-specific beta-1-glycoprotein 6)	NM_001031850.1	Hs00747417_m1	19q13.2	P	A group of molecules that are mainly produced by the placental syncytiotrophoblasts during pregnancy	9.9	6.3
21	TIMP3 (TIMP metallopeptidase inhibitor 3)	NM_000362.4	Hs00165949_m1	22q12.1-q13.2	P	Inhibitors of the matrix metalloproteinases	5.8	3.3
22	FBLN1 (fibulin 1)	NM_001996.2	Hs00242546_m1	22q13.3	P	A secreted glycoprotein that becomes incorporated into a fibrillar extracellular matrix	13.6	9.4
23	PRG2 (plasticity-related gene 2)	NM_002728.4	Hs00794928_m1	19p13.3	P	Lipid phosphate phosphatase family	8.4	3.0
24	CYP19A1 (cytochrome P450, family 19, subfamily A, polypeptide 1)	NM_031226.1	Hs00240671_m1	15q21.1	P	A member of the cytochrome P450 superfamily of enzymes	8.4	4.9
25	PSG3 (pregnancy-specific beta-1-glycoprotein 3)	NM_021016.3	Hs00360732_m1	19q13.2	P	A group of molecules that are mainly produced by the placental syncytiotrophoblasts during pregnancy	6.0	2.8
26	PPAP2B (phosphatidic acid phosphatase type 2B)	NM_177414.1	Hs00170359_m1	1pter-p22.1	P	A member of the phosphatidic acid phosphatase family	12.6	0.6
27	P11 (26 serine protease, placental protein 11)	NM_006025.2	Hs00195731_m1	12q13.1	P	A serine protease specifically expressed in the syncytiotrophoblast	6.9	4.5
28	PAGE4 (P antigen family, member 4)	NM_007003.2	Hs00199655_m1	Xp11.23	P	Expressed in a variety of tumors and in some fetal and reproductive tissues including placenta, a member of the GAGE (G antigen) family	9.6	9.3
29	SMARCA1 (SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 1)	NM_003069.2	Hs00161922_m1	Xq25	P	A member of the SWI/SNF (switch/sucrose nonfermentation) family of proteins	4.7	4.9
30	COL1A2 (collagen, type I, alpha 2)	NM_000089.3	Hs00164099_m1	7q22.1	P	The pro-alpha 2 chain of type I collagen whose triple helix comprises two alpha 1 chains and one alpha 2 chain	8.6	0.9
31	GULP1 (GULP, engulfment adaptor PTB (phosphotyrosine binding) domain containing 1)	NM_016315.2	Hs00169604_m1	2q32.3-q33	P	An evolutionarily conserved adaptor protein required for efficient engulfment of apoptotic cells by phagocytes	8.4	6.2
32	PLEKHC1 (pleckstrin homology domain containing, family C member 1)	NM_006832.1	Hs00235033_m1	14q22.1	P	A component of ECM (cell-extracellular matrix) structures in mammalian cells	6.9	1.8

Prenat Diagn 2010; **30**: 849–861.
DOI: 10.1002/pd